

HOÀNG TRỌNG PHÁN (chủ biên)  
TRƯƠNG THỊ BÍCH PHƯƠNG  
TRẦN QUỐC DUNG

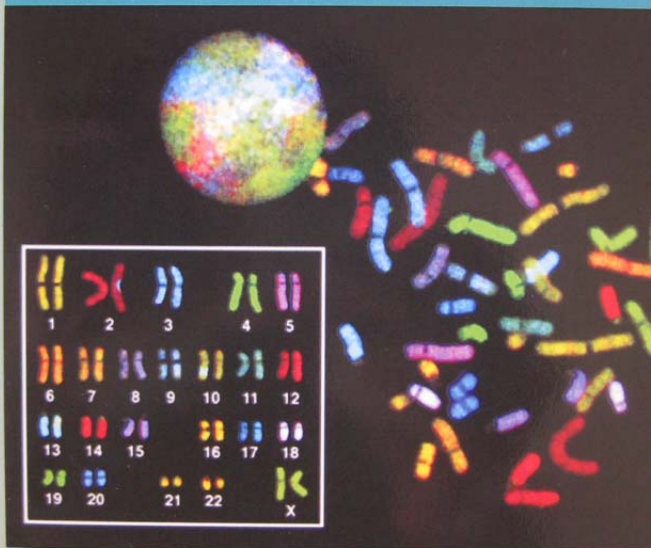
GIÁO TRÌNH

SƯ TÂM & TỔNG HỢP

DOCTOR PLUS CLUB

# DI TRUYỀN HỌC

<https://doctorplus.com> - <https://doctoplus.com>



NHÀ XUẤT BẢN ĐẠI HỌC HUẾ



Cảm ơn bạn đã tải sách từ **Doctor Plus Club**.

Tất cả ebook được **Doctor Plus Club** sưu tầm & tổng hợp từ nhiều nguồn trên internet, mạng xã hội. Tất cả sách **Doctor Plus Club** chia sẻ vì đích duy nhất là để đọc, tham khảo, giúp sinh viên, bác sĩ Việt Nam tiếp cận, hiểu biết nhiều hơn về y học.

Chúng tôi không bán hay in ấn, sao chép, không thương mại hóa những ebook này (nghĩa là quy đổi ra giá và mua bán những ebook này).

Chúng tôi sẵn sàng gỡ bỏ sách ra khỏi website, fanpage khi nhận được yêu cầu từ tác giả hay những người đang nắm giữ bản quyền những sách này.

Chúng tôi không khuyến khích các cá nhân hay tổ chức in ấn, phát hành lại và thương mại hóa các ebook này nếu chưa được sự cho phép của tác giả.

**Nếu có điều kiện các bạn hãy mua sách gốc từ nhà sản xuất để ủng hộ tác giả.**

Mọi thắc mắc hay khiếu nại xin vui lòng liên hệ chúng tôi qua email: [support@doctorplus.club](mailto:support@doctorplus.club)

Website của chúng tôi: <https://doctorplus.club>

Fanpage của chúng tôi: <https://www.facebook.com/doctorplus.club/>

Like, share là động lực để chúng tôi tiếp tục phát triển hơn nữa

Chân thành cảm ơn. Chúc bạn học tốt!



SƯU TẦM &amp; TỔNG HỢP

DOCTOR PLUS CLUB

<https://doctorplus.club/> - <https://facebook.com/doctorplus.club/>

## Lời nói đầu

Đến nay, di truyền học ra đời chỉ mới hơn một trăm năm song nó đã phát triển với một tốc độ hết sức nhanh chóng. Đặc biệt là, trong vòng 50 năm lại đây kể từ ngày James Watson và Francis Crick khám phá ra cấu trúc phân tử DNA, 25/4/1953. Sự hoàn thành việc giải mã di truyền bởi hai nhóm nghiên cứu của Marshall Nirenberg và Har Gobind Khorana vào tháng 6 năm 1966 và sự ra đời của Kỹ thuật Di truyền và Công nghệ DNA tái tổ hợp vào giữa thập niên 1970 là hai sự kiện nổi bật nhất kể từ sau khi sinh học phân tử ra đời. Kế đó, sự hoàn tất của Dự án Bộ gene Người vào tháng 4 năm 2003 được xem là một trong những kỳ công thám hiểm vĩ đại nhất của loài người. Lần đầu tiên con người có thể đọc được một cách đầy đủ toàn bộ trình tự 3.164.700.000 cặp base trong bộ gene của mình. Tất cả những sự kiện nổi bật này minh chứng một điều rằng: Sự phát triển cùng với những thành tựu đạt được của di truyền học trong thời gian qua quả là vô cùng to lớn!

Để góp phần đổi mới nội dung giáo trình Di truyền học theo hướng cập nhật kiến thức cũng như phương pháp dạy và học bộ môn ở bậc Đại học, chúng tôi đã tham cứu nhiều tài liệu khác nhau và nỗ lực biên soạn giáo trình trên tinh thần ấy. Chúng tôi hy vọng rằng giáo trình này sẽ đáp ứng được phần nào nhu cầu giảng dạy và học tập của giảng viên và sinh viên, và cũng có thể sử dụng như một tài liệu tham khảo bổ ích cho giáo viên Sinh học các trường THPT trong bối cảnh đổi mới giáo dục hiện nay.

Nội dung giáo trình gồm phần Mở đầu cộng với 12 chương bao quát các kiến thức đại cương của một giáo trình Di truyền học. Các chương 1-4 đề cập chủ yếu nội dung thuộc Di truyền học cổ điển, các chương 5-10 tập trung vào phần Di truyền học phân tử và chương 12 được xem là phần nhập môn của Di truyền học quần thể, còn chương 11 là sự kết hợp giữa các kiến thức di truyền cổ điển và hiện đại trên đối tượng là con người. Cuối mỗi chương đều có các phần Câu hỏi và Bài tập và Tài liệu Tham khảo để bạn đọc tiện ôn tập và tra cứu.

Giáo trình Di truyền học được ra đời trong khuôn khổ của Dự án Giáo dục thuộc Đại học Huế, vì vậy một số kiến thức nâng cao sẽ được đề cập trong một giáo trình riêng, như: Di truyền Vi sinh vật và Ứng dụng, và Công nghệ DNA Tái tổ hợp. Bên cạnh đó, một số thuật ngữ khoa học được thống nhất sử dụng bằng tiếng Anh để giúp người học dễ dàng hơn trong việc tiếp cận với thông tin qua sách báo nước ngoài hoặc internet.





## I. Khái niệm di truyền học

Theo quan niệm của Bateson (1906), *di truyền học* (genetics) là khoa học nghiên cứu các đặc tính di truyền và biến dị vốn có của mọi sinh vật cùng với các nguyên tắc và phương pháp điều khiển các đặc tính đó. Ở đây, *tính di truyền* (heredity) được biểu hiện ở sự giống nhau giữa con cái với cha mẹ; và *tính biến dị* (variability) biểu hiện ở sự sai khác giữa cha mẹ và con cái cũng như giữa các con cái với nhau.

Cần lưu ý rằng, *gene* là khái niệm căn bản của di truyền học cho nên nội dung của khái niệm gene không ngừng được phát triển cùng với sự phát triển của di truyền học.

## II. Lược sử phát triển của di truyền học

Sự ra đời và phát triển của di truyền học gắn liền với công trình nghiên cứu của Gregor Mendel năm 1865. Tuy nhiên, trước thời Mendel đặc biệt là từ thế kỷ XVII có một số sự kiện quan trọng sau đây: (1) Sự ra đời của kính hiển vi sơ khai bởi A.van Leuvenhook (1632-1723); và (2) Sinh học bắt đầu phát triển mạnh vào thế kỷ XIX với sự ra đời thuyết tế bào của M.Schleiden và T.Schwann (1838,1839) và các thuyết tiến hóa của J.B.Lamarck (1809) và đặc biệt là của R.C.Darwin (1859). Nhìn chung, quan niệm phổ biến thời bấy giờ vẫn là sự *di truyền các tính trạng tập nhiễm* (inheritance of acquired characters) do Lamarck đề xuất và sự *di truyền hòa hợp* (blending inheritance), nghĩa là sự pha lẫn "tinh cha huyết mẹ" ở con cái.

### 1. Sự ra đời và phát triển của di truyền Mendel



Từ đậu Hà Lan (*Pisum sativum*), với ý tưởng và phương pháp nghiên cứu độc đáo, năm 1865 Gregor Mendel (Hình 1) đã phát hiện ra các quy luật di truyền cơ sở đầu tiên và qua đó suy ra sự tồn tại tất yếu của các đơn vị di truyền đặc thù - *nhân tố di truyền* (genetic factor) - quy định các tính trạng được truyền từ thế hệ này sang thế hệ khác mà sau này gọi là *gene*. Tuy nhiên, giới khoa học đương thời không hiểu và do đó không thể đánh giá tầm vóc vĩ đại của phát minh này.

**Hình 1 G. Mendel** Mãi đến năm 1900, ba nhà thực vật học là Carl Correns (Germany), Hugo de Vries (Netherlands) và Erich von Tschermak (Austria) độc lập nhau khám phá lại các quy luật di truyền của Mendel. Và di truyền học chính thức ra đời từ đây mà người sáng lập là Mendel.

## SƯU TÂM & TỔNG HỢP

### DOCTOR PLUS PLUS

<https://www.doctorplus.com>

Trong những năm đầu thế kỷ XX, nhờ ứng dụng di truyền Mendel, các nhà chọn giống đã phát hiện thêm các hiện tượng như: trội không hoàn toàn, đồng trội, gene gây chết, đa allele, các kiểu tương tác gene... Ở giai đoạn này, ngoài thuyết đột biến của H.de Vries năm 1901, còn có hai sự kiện liên quan đến sự ra đời của thuyết di truyền nhiễm sắc thể và di truyền học quần thể sau này, đó là sự khởi xướng "thuyết nhiễm sắc thể" bởi Walter Sutton và Theodor Boveri năm 1902 và việc thiết lập quy luật Hardy-Weinberg năm 1908.

Một số thuật ngữ thông dụng cũng được đề xuất trong giai đoạn này, như: *di truyền học* (genetics) bởi W.Bateson năm 1906, *gene*, *kiểu gene* (genotype) và *kiểu hình* (phenotype) bởi W.Johannsen năm 1909.

### 2. Sự ra đời và phát triển của thuyết di truyền nhiễm sắc thể



Từ 1910, Thomas Hunt Morgan (Hình 2) cùng với ba cộng sự là Alfred H.Sturtevant, Calvin Bridges và Herman J. Muller đã xây dựng thành công *thuyết di truyền nhiễm sắc thể* (chromosome theory of inheritance) dựa trên đối tượng nghiên cứu là ruồi giấm *Drosophila melanogaster*. Học thuyết này xác nhận rằng gene là đơn vị cơ sở của tính di truyền nằm trên nhiễm sắc thể (ở trong nhân); trên đó các gene sắp xếp theo đường thẳng

#### Hình 2 T.H.Morgan

tạo thành nhóm liên kết. Những đóng góp đáng kể của các môn đệ xuất sắc của Morgan đó là: xây dựng bản đồ di truyền (Sturtevant 1913), chỉ ra cơ chế xác định các kiểu hình giới tính ở ruồi giấm (Bridges 1916) và phát triển phương pháp gây đột biến bằng tia X (Muller 1927). Với đóng góp to lớn đó Morgan đã được trao giải Nobel năm 1933 và Muller năm 1946.



Năm 1931, Barbara McClintock (Hình 3) và Harriet Creighton thu được bằng chứng vật lý trực tiếp về tái tổ hợp ở ngô. Sau đó, hiện tượng này cũng được C. Stern quan sát ở *Drosophila*. Như vậy tái tổ hợp có thể được phát hiện cả về mặt vật lý lẫn di truyền ở động vật cũng như ở thực vật. Đến 1944, McClintock phát hiện các *yếu tố di truyền vận động* (transposable genetic elements), và bà đã được trao giải Nobel năm 1983 về khám phá này.

#### Hình 3 B.McClintock

### 3. Sự ra đời và phát triển của di truyền học phân tử

Sự ra đời của *di truyền học phân tử* (molecular genetics) gắn liền với các khám phá về DNA (deoxyribonucleic acid) từ giữa thế kỷ XX trên đối

# SƯU TÂM & TỔNG HỢP

## DOCTOR PLUS CLUB

<https://doctorplus.club/> - <https://facebook.com/doctorplus.club/>

trọng nghiên cứu chủ yếu là các vi sinh vật. Tuy nhiên, trước đó Friedrich Miescher (1869) đã khám phá ra một hỗn hợp trong nhân tế bào gọi là nuclein mà thành phần chính của nó sau này được biết là DNA.



**Hình 4** Beadle, Tatum, Jacob và Monod (từ trái sang)

Về mối quan hệ giữa gene và protein, từ 1902 Archibald Garrod qua nghiên cứu bệnh alcaptonuria ở người đã gợi ý rằng đây là một tính trạng lặn Mendel, có thể liên quan tới sự sai hỏng một enzyme. Bằng các thí nghiệm gây đột biến các gene liên quan đến các con đường sinh hóa trên nấm mốc *Neurospora*, năm 1941 George Beadle và E.L. Tatum (Hình 4) xác nhận mỗi gene kiểm soát sự tổng hợp một enzyme đặc thù. Chính *giả thuyết một gene-một enzyme* (one gene-one enzyme hypothesis) nổi tiếng này đã mở đường cho sự ra đời của di truyền hóa-sinh, và hai ông đã được trao giải Nobel cùng với Joshua Lederberg năm 1958. Về sau, giả thuyết này được chính xác hóa là một gene xác định chỉ một chuỗi polypeptid - cấu trúc sơ cấp của các protein, trong đó có các enzyme.



**Hình 5** O.T. Avery

Vậy bản chất của gene là gì? Năm 1944, Oswald Avery (Hình 5) và các cộng sự là MacLeod và McCarty bằng thí nghiệm biến nạp *in vitro* đã chứng minh rằng DNA là vật chất mang thông tin di truyền. Năm 1949, Erwin Chargaff công bố các kết quả đầu tiên về thành phần hóa học của DNA một số loài.



**Hình 6** R.Franklin (trái) và M.Wilkins

Việc nghiên cứu cấu trúc phân tử DNA được bắt đầu từ 1951 với các dẫn liệu nhiễu xạ tia X của Rosalind Franklin và Maurice Wilkins (Hình 6). Các số liệu hóa học và vật lý này là cơ sở mà từ đó James Watson và Francis Crick (Hình 7) đã xây dựng thành công mô hình cấu trúc phân tử DNA năm 1953, còn gọi là *chuỗi*



*xoắn kép* (double helix). Phát minh vĩ đại này mở ra kỷ nguyên mới cho sự phát triển của di truyền học và sinh học nói chung. Với phát minh đó, Watson và Crick cùng với Wilkins được trao giải Nobel năm 1962.

Bài báo nhan đề "Một cấu trúc cho Deoxyribose Nucleic Acid" của Watson và Crick đăng trên tạp chí *Nature* ngày 25/4/1953 được đánh giá

### Hình 7 J.D.Watson (trái) và F.H.C.Crick

là một bài báo không bình thường. Chỉ vồn vẹn có 128 dòng nhưng đằng sau bài báo là cả một bước tiến lịch sử vĩ đại của di truyền học mà mỗi dòng là một câu chuyện. Thật vậy, sau cấu trúc chuỗi xoắn kép là hàng loạt các khám phá mới. Năm 1958 Matthew Meselson và Franklin Stahl chứng minh sự tái bản bán bảo toàn của DNA; và năm 1961



Seymour Benzer hoàn tất công trình nghiên cứu cấu trúc tinh vi của gene; Francois Jacob và Jacques Monod (Hình 4) tìm ra cơ chế điều hòa sinh tổng hợp protein (giải Nobel 1965 với Andre Lwoff); S.Brenner, Jacob và Meselson khám phá ra RNA thông tin; S.Brenner và F.Crick chứng minh mã di truyền là mã bộ ba; công trình giải mã

### Hình 8 H.G.Khorana (trái) và M.Nirenberg

di truyền này được hoàn thành vào tháng 6 năm 1966 bởi hai nhóm nghiên cứu của Marshall Nirenberg và Har Gobind Khorana (giải Nobel năm 1968; Hình 8).

#### 4. Sự ra đời và phát triển của công nghệ DNA tái tổ hợp

Có thể nói, nền tảng của công nghệ DNA tái tổ hợp (recombinant DNA technology) được thành lập từ 1972 khi Paul Berg (Hình 10) tạo ra phân tử DNA tái tổ hợp đầu tiên trong ống nghiệm (recombinant DNA *in vitro*). Một năm sau Herbert Boyer và Stanley Cohen (Hình 10) lần đầu tiên sử dụng plasmid để tạo dòng DNA. Lĩnh vực ứng dụng mới này của sinh học phân tử đã tạo ra một cuộc cách mạng mới trong sinh học. Đóng góp đáng kể trong lĩnh vực này là khám phá về các enzyme giới hạn (restriction enzyme) từ 1961-1969 của Werner Arber, Daniel Nathans và



## SƯU TÂM & TỔNG HỢP

### DOCTOR PLUS CLUB

<https://doctorplus.club/> - <https://facebook.com/doctorplus.club/>

Hamilton Smith (giải Nobel 1978; Hình 10); đề xuất các phương pháp xác định trình tự base trong các nucleic acid năm 1977 bởi P.Berg, W.Gilbert



**Hình 9** Các nhà khoa học đoạt giải Nobel y học liên quan kỹ thuật gene. Từ trái sang: D.Nathans, H.Smith, W.Arber, P.Sharp và R.Robert.



**Hình 10** Các nhà khoa học đoạt giải Nobel hóa học liên quan kỹ thuật gene. Từ trái sang: H.Boyer, S.Cohen, P.Berg, W.Gilbert, F.Sanger và K.Mullis.

và Frederick Sanger (giải Nobel hóa học 1980; Hình 10); sự khám phá ra các *gene phân đoạn* (split gene) năm 1977 bởi Phillip Sharp và Richard Robert (giải Nobel 1993; Hình 9); sự phát minh ra phương pháp PCR (*polymerase chain reaction*) của Kary B.Mullis năm 1985 (Hình 10) và phương pháp *gây đột biến định hướng* (site-directed mutagenesis) của Michael Smith từ 1978-1982 (giải Nobel hóa học 1993)...

Cùng với những thành tựu ứng dụng ly kỳ trong sản xuất và đời sống xã hội, như việc sản xuất các chế phẩm y-sinh học bằng công nghệ DNA tái tổ hợp, sử dụng *liệu pháp gene* (gene therapy) trong điều trị bệnh di truyền, tạo các giống sinh vật mới bằng con đường biến đổi gene (genetically modified organisms = GMOs), *dự án bộ gene người* (Human Genome Project = HGP)... gây ra không ít hoài nghi, tranh cãi xung quanh các vấn đề về *đạo lý sinh học* (bioethics) và *an toàn sinh học* (biosafety).

### III. Đối tượng và các lĩnh vực nghiên cứu của di truyền học

Trong giai đoạn đầu, đối tượng của di truyền học là các thực vật, động vật, người và các vi sinh vật. Từ đó dẫn tới sự hình thành các lĩnh vực nghiên cứu tương ứng là di truyền học thực vật, động vật, người và di truyền học vi sinh vật, trong đó di truyền học tế bào là cơ sở. Giai đoạn

này kéo dài từ thời Mendel cho đến thập niên 1940, với đặc trưng là nghiên cứu quy luật di truyền các tính trạng qua các thế hệ. Vì vậy nó thường được gọi là giai đoạn *di truyền học Mendel* hay *di truyền học cổ điển* (Mendelian or classical genetics).

Từ thập niên 1950 đến nay, với sự ra đời của *di truyền học phân tử* (molecular genetics), đối tượng nghiên cứu là tổ chức, cấu trúc, chức năng và cơ chế hoạt động của các *bộ gene* (genomes), các gene và các sản phẩm của chúng ở mức phân tử. Đặc biệt với sự ra đời của các *kỹ thuật tạo dòng gene* (gene-cloning techniques) từ thập niên 1970, việc nghiên cứu cơ bản cũng như ứng dụng trở nên hết sức thuận lợi. Sự phân hóa thành các chuyên ngành lúc này là vô cùng phong phú và đa dạng, như: *di truyền học bệnh* (genetics of disease), *di truyền học ung thư* (genetics of cancer), *di truyền học phát triển* (developmental genetics), *công nghệ sinh học* (biotechnology)... Gần đây còn xuất hiện một số lĩnh vực nghiên cứu mới như *genomics*, *DNA chip technology*, *DNA microarray technology*...

Một hướng khác của di truyền học chuyên nghiên cứu cấu trúc di truyền của các quần thể và sự biến đổi di truyền bên trong quần thể và giữa các quần thể. Đó là nhánh *di truyền học quần thể* (population genetics), mà nguyên lý cơ sở của nó được G.Hardy (England) và W.Weinberg (Germany) độc lập đưa ra năm 1908. Tuy nhiên, lĩnh vực nghiên cứu này chỉ thực sự bắt đầu từ thập niên 1930 với các công trình của R.A.Fisher, J.B.S.Haldane và Sewall Wright. Ngày nay các nhà di truyền học không còn thiên về nghiên cứu sự biến đổi ở mức kiểu hình mà tập trung vào sự biến đổi phân tử trong một quần thể nhằm tìm hiểu ý nghĩa tiến hóa của các biến đổi đó. Và như thế di truyền học quần thể trở thành nền tảng cho các thuyết tiến hóa hiện đại.

Qua phân tích ở trên cho thấy ba nhánh chính của di truyền học, đó là: di truyền học Mendel, di truyền học phân tử và di truyền học quần thể.

#### **IV. Các phương pháp nghiên cứu của di truyền học**

Việc nghiên cứu di truyền học được tiến hành bởi nhiều phương pháp khác nhau. Bên cạnh các phương pháp kinh điển đặc thù còn có sự phát triển và tích hợp của các phương pháp từ toán học, tin học, vật lý và hóa học đặc biệt là trong lĩnh vực sinh học phân tử.

##### *1. Các phương pháp kinh điển đặc thù của di truyền học*

Trước hết đó là phương pháp tự thụ phân dùng để chọn các dòng thuần làm bố mẹ trong các phép lai, các phương pháp lai một hoặc đồng thời nhiều tính trạng giữa các bố mẹ do Mendel đề xuất. Trong đó bao gồm cả các hình thức lai thuận nghịch và *lai phân tích* (testcross) nhằm rút



ra quy luật, kiểm tra kiểu gene hoặc dùng để thiết lập bản đồ di truyền.

Đối với nghiên cứu di truyền người và một số vật nuôi *giao phối cận huyết* (consanguineous), đặc trưng là phương pháp *phân tích phả hệ* (pedigree analysis) nhằm xác định đặc điểm di truyền trội-lặn của một tính trạng hoặc bệnh tật; nghiên cứu trẻ sinh đôi cùng trứng và khác trứng nhằm xác định *hệ số di truyền* (heritability) của tính trạng; phương pháp *gây đột biến* (mutagenesis) kết hợp với lai hữu tính dùng trong nghiên cứu và chọn tạo giống, đặc biệt là ở thực vật...

## 2. Phương pháp toán học

Trong nghiên cứu khoa học nói chung và di truyền học nói riêng, việc áp dụng các công cụ toán thống kê và lý thuyết xác suất để phân tích định lượng và lý giải các kết quả nghiên cứu là rất thiết yếu. Chính điều đó làm cho di truyền học trở thành một khoa học chính xác và mang tính dự báo. Điều này thể hiện rõ trong các công trình nghiên cứu của Mendel, Morgan cũng như trong các nghiên cứu di truyền số lượng và di truyền quần thể.

## 3. Các phương pháp vật lý và hóa học

Trong nghiên cứu di truyền, đặc biệt là di truyền phân tử và tế bào đòi hỏi phải sử dụng các kỹ thuật và phương pháp của vật lý và hóa học. Chẳng hạn, trong nghiên cứu hình thái nhiễm sắc thể không thể thiếu các loại kính hiển vi quang học và điện tử, các *kỹ thuật nhuộm băng* (banding techniques); nghiên cứu thành phần hóa học và cấu trúc DNA đòi hỏi các phương pháp sắc ký và nhiễu xạ tia X...

## 4. Các phương pháp tế bào học

Trong nghiên cứu di truyền tế bào có rất nhiều phương pháp được sử dụng để quan sát hình thái nhiễm sắc thể và thiết lập *kiểu nhân* (karyotype) của các loài cũng như để chẩn đoán các bệnh tật liên quan đến sự biến đổi số lượng và cấu trúc nhiễm sắc thể.

Bên cạnh các kỹ thuật nuôi cấy mô và chuẩn bị nhiễm sắc thể là các kỹ thuật nhuộm băng khác nhau, kỹ thuật *lai huỳnh quang tại chỗ* (fluorescence *in situ* hybridization = FISH), kỹ thuật miễn dịch tế bào học... (xem Verma và Babu 1995).

## 5. Các phương pháp của sinh học phân tử - kỹ thuật di truyền

Sự tiến bộ nhanh chóng gần đây của sinh học phân tử và công nghệ sinh học là nhờ sự phát triển mạnh mẽ của các kỹ thuật tái tổ hợp DNA, lai phân tử, sử dụng các mẫu dò và phương pháp đánh dấu khác nhau, phương pháp PCR, các phương pháp xác định trình tự nucleic acid cũng như các phương pháp biến đổi vật liệu di truyền mới.

Với các công cụ kỹ thuật mới này đã cho phép đi sâu nghiên cứu tổ

chức của các gene và bộ gene, các cơ chế điều hòa và biến đổi di truyền ở mức phân tử cũng như các thành tựu ứng dụng mới trong y-sinh học, nông nghiệp và các lĩnh vực khác của đời sống-xã hội.

## **V. Các nguyên tắc nghiên cứu và phương pháp học tập di truyền học**

### *1. Các nguyên tắc nghiên cứu của di truyền học*

Trong nghiên cứu di truyền học và sinh học nói chung có các nguyên tắc chung cần tuân thủ như là phương pháp luận, sau đây: (1) Lấy tế bào làm đơn vị nghiên cứu; (2) Thông tin di truyền chứa trong bộ gene tế bào chi phối mọi biểu hiện sống của nó mà các gene là đơn vị di truyền cơ sở; (3) Sự hoạt động của các gene trong quá trình phát triển cá thể là đặc trưng cho từng gene trong từng giai đoạn cụ thể; (4) Các quá trình trong các hệ thống sống phải được điều hòa và kiểm soát để đảm bảo cho sự tồn tại của nó là liên tục, trong đó phổ biến là sự tự điều chỉnh bằng các *cơ chế phản hồi thông tin* (feed-back mechanism); (5) Sự thống nhất giữa cấu trúc và chức năng biểu hiện ở tất cả các mức độ tổ chức khác nhau của sự sống; (6) Tất cả các tổ chức và quá trình sống đều tuân theo các quy luật vật lý và hóa học; (7) Sự sống trên trái đất trải qua quá trình tiến hóa khoảng 3,5 tỷ năm qua, vì vậy khi so sánh, những nét tương đồng giữa chúng cho thấy tính thống nhất về mặt nguồn gốc và những nét dị biệt cho thấy tính phát triển, sự phân hóa đa dạng tất yếu của chúng.

### *2. Phương pháp học tập môn di truyền học*

Cũng như bất kỳ môn học nào khác, việc học tập môn di truyền đòi hỏi phải nắm vững lịch sử môn học, đối tượng, nhiệm vụ, phương pháp nghiên cứu và hệ thống kiến thức căn bản của nó. Bên cạnh các nguyên tắc nói trên vốn rất cần cho tư duy trong học tập, dưới đây nêu một số điểm chính liên quan phương pháp học tập đặc thù của bộ môn.

(1) Nắm vững các kiến thức liên môn (như tế bào học, hóa sinh học, vi sinh học, học thuyết tiến hóa...) và liên ngành (như các khái niệm và nguyên lý cơ bản của toán thống kê-xác suất, vật lý và hóa hữu cơ).

(2) Nắm vững hệ thống khái niệm cơ bản cũng như các thuật ngữ mới không ngừng nảy sinh. Trong đó gene là khái niệm căn bản có nội hàm không ngừng phát triển, đặc biệt là trong ba thập niên lại đây.

(3) Hiểu rõ bản chất của các nguyên lý di truyền trong từng chủ đề cũng như mối liên quan giữa chúng để có thể giải thích và vận dụng trong giải quyết các bài toán hoặc tình huống của đời sống và thực tiễn sản xuất.

(4) Để nắm kiến thức và phát triển các kỹ năng tư duy một cách vững chắc đòi hỏi phải biết vận dụng kiến thức vào giải bài tập cũng như các kỹ năng thực hành thí nghiệm.

(5) Di truyền học là một khoa học thực nghiệm, nên thông tin thu được là nhờ các quan sát từ thế giới tự nhiên, và phương pháp khoa học chính là công cụ để hiểu biết các quan sát đó. Nói đến phương pháp nghiên cứu khoa học là nói đến các bước tiến hành theo một trình tự tổ chức công việc chặt chẽ sau đây: Quan sát → Giả thuyết → Dự đoán → Thực nghiệm (để kiểm tra giả thuyết đặt ra) → Đề xuất giả thuyết mới.

Cũng cần lưu ý rằng, đặc tính của khoa học là hoài nghi, đòi hỏi phải có bằng chứng xác thực, là kết hợp giữa logic và trí tưởng tượng, là giải thích và dự đoán, không có sự độc đoán; người nghiên cứu hay nhà khoa học phải nhận biết sáng tỏ, trung thực, vô tư, và nói chung là chịu sự chi phối của các nguyên tắc đạo đức đã được thừa nhận rộng rãi.

(6) Trong khi học giáo trình, bạn nên làm ít nhất một tiểu luận về một vấn đề cập nhật mình yêu thích. Điều đó rất lý thú và bổ ích. Bởi công việc này đòi hỏi sự say mê tìm tòi các thông tin mới, đặc biệt là trên mạng để viết một bài tổng luận có tính khoa học và trình bày trong một seminar.

(7) Bởi di truyền học là một ngành khoa học non trẻ nhưng phát triển với tốc độ cực nhanh, nên khối lượng kiến thức mới tích lũy được là vô cùng phong phú và đa dạng. Để có thể cập nhật thông tin về môn học đòi hỏi phải tăng cường khả năng sử dụng tiếng Anh và internet. Điều quan trọng là phải tạo cho mình một hoài bão học tập, một khả năng và phương pháp tự học và thành lập cho được một thư mục tra cứu. Trong đó đáng kể là các trang web (được giới thiệu trong từng chương), hoặc có thể sử dụng ngay các từ khóa (key words) được cho ở từng chủ đề để tìm kiếm với công cụ có thể nói mạnh nhất hiện nay là *Google*.

## **VI. Di truyền học với công nghệ sinh học, tin học và các vấn đề xã hội**

Như đã đề cập, sự phát triển hết sức nhanh chóng của di truyền học trong vài thập niên qua, đặc biệt là sự tiến bộ của *công nghệ sinh học* (biotechnology) nói chung đã có những tác động mạnh mẽ lên nhiều ngành khoa học và trên mọi mặt của đời sống, kinh tế, chính trị và xã hội ở phạm vi toàn cầu. Di truyền học được hình dung ở vị trí trung tâm và giao thoa với sinh học, hóa sinh học, kỹ nghệ, y-dược, nông nghiệp, sinh thái học, kinh tế học, luật, xã hội học và triết học.

Giáo sư danh dự môn hóa học ở Đại học Havard, F.H. Westheimer, bình luận về sinh học phân tử như sau: "Cuộc cách mạng trí tuệ vĩ đại nhất của 40 năm qua đã xảy ra trong sinh học. Liệu có thể tìm ra một người nào đó có học ngày nay mà không hiểu biết chút gì về sinh học phân tử?" (dẫn theo Weaver và Hedrick 1997, tr.15).

Các thành tựu đạt được nhờ ứng dụng di truyền học trong nông nghiệp

là vô cùng to lớn, góp phần tạo nên cuộc "cách mạng xanh lần thứ hai" với sự ra đời của hàng loạt các giống vật nuôi-cây trồng có ưu thế lai vượt trội, các sinh vật biến đổi gene (GMOs) mang những đặc tính hoàn toàn mới lạ.

Trong y học, đó là sự ra đời của hàng loạt các dược phẩm được sản xuất bằng kỹ thuật di truyền dùng cho điều trị bệnh và cải biến trí thông minh của con người; đó là các phương pháp chẩn đoán và điều trị bệnh ở mức phân tử... Sự thành công của dự án bộ gene người (HGP) vào tháng 4 năm 2003 cho phép chúng ta lần đầu tiên đọc được toàn bộ trình tự khoảng 3,2 tỷ cặp nucleotide trong bộ gene con người (*Homo sapiens*). HGP là một trong những kỳ công thám hiểm vĩ đại nhất trong lịch sử nhân loại (NHGRI 2005). Theo ước tính mới nhất được công bố ngày 21/10/2004 trên tạp chí *Nature*, bộ gene chúng ta chứa số lượng gene mã hóa protein thấp một cách đáng kinh ngạc, khoảng 20.000 đến 25.000 chứ không phải là 50.000 đến 140.000 gene như dự đoán ban đầu hoặc 35.000 theo dự đoán trong vài ba năm lại đây (NHGRI 2005).

Chính sự kết hợp tin học và máy tính trong nghiên cứu sinh học phân tử dẫn tới sự ra đời một ngành mới là *sinh-tin học* (bioinformatics) cho phép thu thập, tổ chức và phân tích số lượng lớn các số liệu sinh học nhờ sử dụng mạng máy tính và các nguồn dữ liệu (databases). Và một số lĩnh vực nghiên cứu mới khác cũng ra đời như: *genomics* - phân tích toàn bộ genome của một sinh vật được chọn, *DNA microchip technology* - xác định các đột biến trong các gene, *DNA microarray technology* - nghiên cứu cách thức một số lượng lớn các gene tương tác lẫn nhau và cơ chế mạng lưới điều hòa của tế bào kiểm soát đồng thời số lượng cực kỳ lớn các gene...

Những thách thức cho tương lai của nghiên cứu *khoa học về các bộ gene* (genomics) đối với sinh học, vấn đề sức khỏe và xã hội cũng được đặt ra (Collins và cs 2003). Sự hoàn tất của HGP tự nó không có nghĩa là đã xong mà đúng hơn là điểm khởi đầu cho công cuộc nghiên cứu thậm chí còn hứng thú hơn. Các nhà nghiên cứu hiện giờ đang cố gắng làm sáng tỏ một số quá trình phức tạp nhất của sinh học, đó là: một đĩa bé phát triển từ một tế bào đơn lẻ bằng cách nào, các gene phối hợp chức năng của các mô và cơ quan như thế nào, sự tiền định bệnh tật xảy ra như thế nào và bộ não người làm việc ra sao (NHGRI 2005).

Cùng với những thành tựu ứng dụng ly kỳ trong sản xuất và đời sống xã hội nói trên, nhiều vấn đề mới được đặt ra cho các ngành giáo dục, luật, triết học, xã hội học và đã gây không ít hoài nghi, tranh cãi xung quanh các vấn đề về đạo lý sinh học, an toàn sinh học và môi sinh.

## Tài liệu Tham khảo

### Tiếng Việt

- Hồ Huỳnh Thùy Dương. 1997. *Sinh học Phân tử*. NXB Giáo Dục.  
 Phạm Thành Hồ. 2000. *Di truyền học*. Tái bản lần II, NXB Giáo Dục.  
 Phan Cự Nhân (chủ biên), Nguyễn Minh Công, Đặng Hữu Lanh. 1999. *Di truyền học*. NXB Giáo Dục.

### Tiếng Anh

- Collins FS, Green ED, Guttmacher AE, Guyer MS. 2003. A vision for the future of genomics research. *Nature*, Vol. 422, No. 6934, p.835-847.
- McKusick V. 1998. *Mendelian Inheritance in Man*. 12th ed., Johns Hopkins University Press, Baltimore.
- National Human Genome Research Institute (NHGRI)/ National Institute of Health (NIH). 2005: <http://www.genome.gov/>  
<http://www.ncbi.nlm.nih.gov/>  
 Online Mendelian Inheritance in Man (OMIM™):  
<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>
- Tamarin RH. 1999. *Principles of Genetics*. 6<sup>th</sup> ed, McGraw-Hill, Inc, NY.
- Verma RS and Babu A. 1995. *Human chromosome : principles and techniques*. 2nd ed, international edition, McGraw-Hill, Inc., New York.
- Watson JD and Crick FHC. 1953. A structure for Deoxyribose Nucleic Acid. April 25, 1953, *Nature*, Vol.171, page 737.
- Weaver RF, Hedrick PW. 1997. *Genetics*. 3<sup>rd</sup> ed, McGraw-Hill Companies, Inc. Wm.C.Brown Publishers, Dubuque, IA.
- Zinnen T. 2005. JD Watson and FHC Crick - A structure for Deoxyribose Nucleic Acid - *Nature*, 25 April 1953. (<http://www.accessexcellence.org>)

## Chương 1

# Cơ sở của Di truyền học Mendel

Cho đến đầu thế kỷ XX, mọi người còn chưa hiểu được cơ chế của sự di truyền, mặc dù vẫn biết rằng con cái sinh ra thường giống bố mẹ. Quan niệm phổ biến cho đến giữa thế kỷ XIX được gói gọn trong cái gọi là *thuyết di truyền hòa hợp* (theory of blending inheritance) nhằm giải thích sự kiện con cái mang các đặc điểm của cả hai bố mẹ. Tuy nhiên, đến năm 1866 Gregor Mendel đã đưa ra *thuyết di truyền gián đoạn* (theory of particulate inheritance), với gợi ý rằng: Đơn vị di truyền đặc thù kiểm soát một tính trạng được truyền từ thế hệ này sang thế hệ khác tồn tại dưới dạng hạt, ngày nay ta gọi là *gene*.

Các khám phá quan trọng của Mendel đặt nền móng cho sự ra đời của di truyền học sau này. Như Thomas Hunt Morgan đã nhận định: "Trong mười năm nghiên cứu ở cây đậu Hà Lan trong ngôi vườn của tu viện, G. Mendel đã làm nên sự khám phá vĩ đại nhất trong sinh học đã đạt được trong năm trăm năm qua".

### I. Tiểu sử Mendel - Cha đẻ của Di truyền học

Gregor Mendel sinh năm 1822, lớn lên ở trang trại của cha mình tại một tỉnh của Austria (Áo). Vì gia đình nghèo nên ông phải vào tu viện để tiếp tục việc học của mình. Trong khoảng thời gian này, Mendel nghiên cứu vật lý và toán là những môn học giúp ông nhiều trong các thí nghiệm di truyền sau này. Ông đã được gọi tới Đại học Vienna để thi lấy bằng giáo viên chính thức, nhưng không đỗ và quay về tu viện "dạy học" trong nhiều năm.



**Hình 1.1** Gregor Mendel trong ngôi vườn của tu viện Brno.

Khi còn ở trang trại của cha mình, Mendel đã quan tâm tới các cây cối và con vật, và thường giữ lại những cái hoa, con ong và chuột. Sau này ở tu viện Brno ông tập trung vào các cây đậu Hà Lan (*Pisum sativum*).

Mendel đã xác định được các nguyên lý di truyền cơ sở từ các thí nghiệm chọn giống thực vật. Các kết quả nghiên cứu này đã được Mendel trình bày năm 1865 trước Hội Nghiên cứu Khoa học Tự nhiên Brno và công bố năm 1866 ở Germany trong một bài báo nhan đề là *Các thí nghiệm lai ở thực vật* (Experiments on Plant Hybrids). Bài báo này nhanh chóng có mặt ở nhiều thư viện, nhưng những người cùng thời ông không



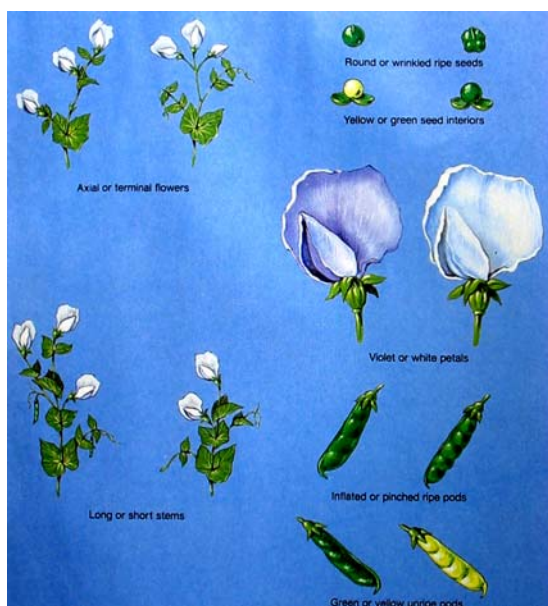
hiểu được các phát hiện của ông, có lẽ một phần là do ông sử dụng toán học để lý giải các kết quả của mình. Ngoài ra, hầu hết các nhà nghiên cứu đương thời do tiến hành nhiều tính trạng đồng thời dẫn tới các kết quả rối bời nên không thể nhận ra được các nguyên lý di truyền cơ sở.

Mendel trở thành tu viện trưởng từ năm 1868 và không công bố thêm một kết quả nào về di truyền nữa kể từ sau kiệt tác năm 1866.

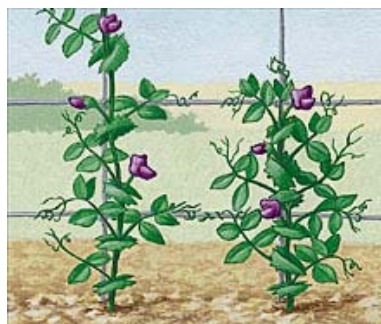
Mendel qua đời năm 1884 trước khi công trình của ông được giới khoa học thấu hiểu. Mãi đến năm 1900, công trình của ông mới được ba nhà thực vật học độc lập nhau khám phá lại, đó là Carl Correns của Germany, Hugo de Vries của Netherlands và Erich von Tschermak của Austria. Đây là mốc khởi đầu cho các nghiên cứu di truyền học hiện đại. Ngày nay, phương pháp thí nghiệm của Mendel được xem là thí dụ kinh điển về sự nghiên cứu khoa học được lập kế hoạch cẩn thận và bài báo của ông là sự minh họa tuyệt vời của một thiên tài khoa học.

## II. Đối tượng và phương pháp thí nghiệm của Mendel

### 1. Đối tượng



(a)



(b)

**Hình 1.2 (a) Bảy tính trạng tương phản ở đậu Hà Lan được Mendel nghiên cứu; dạng trội nằm bên trái của mỗi trường hợp. (b) Cấu tạo hoa đậu, phương pháp thụ phấn chéo và cây đậu Hà Lan.**

Mendel chọn đậu Hà Lan (*Pisum sativum*) làm đối tượng nghiên cứu vì chúng có hai đặc điểm cơ bản là sai khác nhau về nhiều tính trạng tương phản dễ quan sát (hình 1.2a) và sinh sản bằng lối tự thụ phấn. Ngoài ra, đậu có hoa khá lớn nên thao tác dễ dàng (hình 1.2b); có khả năng cho số lượng đời con nhiều; và nhiều giống đậu lúc bấy giờ có giá trị kinh tế cao.

## 2. Phương pháp

Tính chất độc đáo của phương pháp nghiên cứu Mendel thể hiện ở chỗ: (1) Chọn các dòng thuần (pure lines) khác nhau bằng cách cho tự thụ phấn liên tiếp nhiều thế hệ dùng làm dạng bố mẹ trong các phép lai; (2) Theo dõi trước tiên kết quả di truyền riêng biệt của từng tính trạng qua vài thế hệ, trong đó thế hệ cây lai thứ nhất hay  $F_1$  sinh ra do giao phấn giữa hai dạng bố mẹ thuần chủng khác nhau, còn thế hệ cây lai thứ hai hay  $F_2$  sinh ra từ sự tự thụ phấn của các cây lai  $F_1$ , rồi sau đó mới tiến hành nghiên cứu sự di truyền đồng thời của hai hoặc nhiều tính trạng; (3) Khái quát và lý giải các kết quả thí nghiệm thu được bằng toán thống kê và xác suất; và (4) Kiểm tra lại một cách cẩn thận các giả thuyết khoa học bằng các phép lai thuận nghịch (reciprocal matings) và lai phân tích (testcross).

## III. Lai một tính và nguyên lý phân ly

### 1. Kết quả thí nghiệm lai một tính (monohybrid cross)

Mendel đã tiến hành bảy phép lai một tính khác nhau và các kết quả thu được được trình bày ở Bảng 1.1.

**Bảng 1.1 Các kết quả lai một tính của Mendel**

TT	Kiểu hình P	$F_1$	$F_2$	Tỷ lệ $F_2$
1	Hạt trơn × nhăn	Trơn	5474 trơn : 1850 nhăn	2,96:1
2	Hạt vàng × xanh	Vàng	6022 vàng : 2001 xanh	3,01:1
3	Hoa đỏ tía × trắng	Đỏ tía	705 đỏ tía : 224 trắng	3,15:1
4	Quả phồng × tóp	Phồng	882 phồng : 299 tóp	2,95:1
5	Quả xanh × vàng	Xanh	428 xanh : 152 vàng	2,82:1
6	Hoa dọc thân × đỉnh	Dọc thân	651 dọc thân : 207 đỉnh	3,14:1
7	Thân cao × thấp	Cao	787 cao : 277 thấp	2,84:1

Từ tất cả các phép lai trên cho thấy: Khi bố mẹ ở thế hệ xuất phát (P) thuần chủng khác nhau về một cặp tính trạng tương phản, thì ở thế hệ  $F_1$  tất cả con lai đều biểu hiện chỉ một tính trạng của bố hoặc mẹ, tính trạng đó được gọi là tính trạng trội (dominant) và tính trạng kia không quan sát được gọi là tính trạng lặn (recessive). Sau đó cho các con lai  $F_1$  tự thụ phấn thì ở thế hệ  $F_2$  ông thu được cả hai kiểu hình (phenotype) của bố mẹ ban đầu với tỷ lệ xấp xỉ 3/4 trội và 1/4 lặn.

Ngoài ra, Mendel cũng cho các cây  $F_2$  tự thụ phấn riêng rẽ và theo dõi sự phân ly ở thế hệ  $F_3$ . Kết quả cho thấy 1/4 cây của  $F_2$  sinh ra kiểu hình



$F_1$  đều có kiểu hình trội hạt vàng. Khi bước vào giảm phân, các cá thể  $F_1$  dị hợp tử (Aa) sẽ cho hai loại giao tử (A và a) với tỷ lệ tương đương. Đó cũng là thực chất của nguyên lý phân ly (principle of segregation), hay quy luật thứ nhất của Mendel (Mendel's first law). Kết quả của sự tự thụ tinh ngẫu nhiên giữa các loại giao tử đực và cái của  $F_1$  này sẽ cho ra bốn kiểu tổ hợp ở  $F_2$ , với tỷ lệ phân ly theo kiểu gene là 1AA: 2Aa: 1aa và tỷ lệ kiểu hình tương ứng là 3 vàng (A-): 1 xanh (aa). Lưu ý rằng thông thường người ta sử dụng *khung Punnett* (Punnett square) do nhà di truyền học người Anh R.C.Punnett đưa ra năm 1905 để xác định kết quả di truyền của các phép lai. Sơ đồ biểu diễn kết quả lai một tính được nêu ở hình 1.3.

Để khẳng định nguyên lý phân ly, Mendel đã tiến hành hai thí nghiệm: Một là, cho các cá thể dị hợp tử  $F_1$  tự thụ phấn như đã nói ở trên; và hai là, cho  $F_1$  lai ngược trở lại với bố hoặc mẹ có kiểu hình lặn. Phép lai với một cá thể lặn như thế được gọi là *lai phân tích* (testcross), thay vì gọi là *lai ngược* (backcross) bởi vì nó cho phép kiểm tra kiểu gene của một cá thể trội là dị hợp hay đồng hợp. Cơ sở di truyền của kiểu lai này như sau:

Bố mẹ	Aa (hạt vàng)	×	aa (hạt xanh)
Giao tử	50% A : 50% a		100% a
Đời con	50% Aa (vàng)	:	50% aa (xanh)

### 3. Nguyên lý phân ly và tính phổ biến của nó

Sau khi các nguyên lý di truyền của Mendel được khám phá lại năm 1900, tính phổ biến của các nguyên lý Mendel nói chung và nguyên lý phân ly nói riêng đã được các nhà nghiên cứu khẳng định bằng cách lặp lại các phép lai của ông (chẳng hạn giữa đậu hạt vàng và hạt xanh; bảng 1.2) cũng như tiến hành các phép lai tương tự ở các động vật và thực vật khác.

**Bảng 1.2 Các kết quả lai lặp lại ở đậu Hà Lan**

Nhà nghiên cứu	Vàng	Nhãn	Tỷ lệ $F_2$
Mendel (1866)	6.022	2.001	3,01:1
Correns (1900)	1.394	453	3,08:1
Tschermak (1900)	3.580	1.190	3,01:1
Bateson (1905)	11.902	3.903	3,05:1
Darbishire (1909)	109.060	36.186	3,01:1
Tính toàn bộ	131.958	43.733	3,02:1

Nội dung chính của nguyên lý hay quy luật phân ly có thể tóm lược như sau: *Các allele là những dạng khác nhau của cùng một gene; trong các thể dị hợp tử chúng phân ly về các giao tử với tỷ lệ tương đương.*

## IV. Lai hai tính và nguyên lý phân ly độc lập

### 1. Kết quả thí nghiệm lai hai tính (dihybrid cross)

Để xác định sự di truyền đồng thời của nhiều cặp tính trạng, Mendel đã tiến hành nhiều thí nghiệm khác nhau. Bảng 1.3 giới thiệu kết quả lai hai tính giữa các giống đậu thuần chủng hạt vàng-trơn và hạt xanh-nhăn.

**Bảng 1.3 Các kết quả lai hai tính của Mendel**

Thế hệ	Kiểu hình hạt	Số lượng	Tỷ lệ F <sub>2</sub> (quan sát)	Tỷ lệ F <sub>2</sub> (kỳ vọng)
P <sub>tc</sub>	Vàng-trơn × xanh-nhăn	–	–	–
F <sub>1</sub>	Vàng-trơn	–	–	–
F <sub>2</sub>	Vàng-trơn	315	9,84	9
	Vàng-nhăn	101	3,16	3
	Xanh-trơn	108	3,38	3
	Xanh-nhăn	32	1,0	1
Tổng =		556		

Với phép lai này, tất cả con lai F<sub>1</sub> đều có kiểu hình trội kép là hạt vàng và trơn. Khi cho F<sub>1</sub> tự thụ phấn, ở F<sub>2</sub> xuất hiện 4 kiểu hình là vàng-trơn, vàng-nhăn, xanh-trơn và xanh-nhăn với tỷ lệ xấp xỉ 9:3:3:1.

## 2. Giải thích và nội dung nguyên lý phân ly độc lập

Nếu xét tỷ lệ phân ly của từng tính trạng ở F<sub>2</sub>, ta có: 315 + 101 = 416 vàng và 108 + 32 = 140 xanh, xấp xỉ 3 vàng : 1 xanh. Tương tự, về hình dạng hạt, ta có 315 + 108 = 423 trơn và 101 + 32 = 133 nhăn, xấp xỉ 3 trơn : 1 nhăn. Điều đó chứng tỏ mỗi tính trạng đều tuân theo quy luật phân ly 3 trội : 1 lặn.

Bằng cách áp dụng quy tắc nhân xác suất của các biến cố độc lập (xem mục VI), ta dễ dàng chứng minh được rằng sự phân ly của hai tỷ lệ này là hoàn toàn độc lập nhau như dự đoán ban đầu. Thật vậy, (3 vàng : 1 xanh)(3 trơn : 1 nhăn) = 9 vàng-trơn : 3 vàng-nhăn : 3 xanh-trơn : 1 xanh-nhăn.

Cần lưu ý là tỷ lệ 9:3:3:1 này cũng được Mendel tìm thấy trong khi lặp lại thí nghiệm với các tính trạng khác. Từ đó ông mới xây dựng nên *nguyên lý phân ly độc lập* (principle of independent assortment), còn gọi là *quy luật thứ hai của Mendel* (Mendel's second law). Nội dung của nguyên lý này phát biểu rằng: *Các allele của các gene khác nhau thì phân ly một cách độc lập với nhau trong quá trình hình thành giao tử (kết quả là tạo ra tỷ lệ 9:3:3:1 ở thế hệ F<sub>2</sub> từ một phép lai hai tính).*

Để minh họa cho những điều trình bày trên đây, ta quy ước: A - vàng, a - xanh, B - trơn, b - nhăn.

*Lưu ý:* Để kiểm tra lại giả thuyết phân ly độc lập, Mendel đã tiến hành lai phân tích giữa các cây vàng-trơn F<sub>1</sub> (AaBb) với cây xanh-nhăn (aabb). Kết quả thu được gồm 55 vàng-trơn : 49 vàng-nhăn : 51 xanh-trơn : 53

xanh-nhãn, tương đương với tỷ lệ  $1:1:1:1 = (1:1)(1:1)$ . Điều đó chứng tỏ các cây  $F_1$  đã cho bốn loại giao tử với tỷ lệ ngang nhau, nghĩa là chứa hai cặp gene dị hợp phân ly độc lập.

$P_{tc}$  Kiểu hình      vàng, trơn      ×      xanh, nhãn  
 Kiểu gene      AAbb           aabb  
 Giao tử      AB           ab

$F_1$  Kiểu gene      AaBb  
 Kiểu hình      vàng, trơn  
 Giao tử       $\frac{1}{4} AB : \frac{1}{4} Ab : \frac{1}{4} aB : \frac{1}{4} ab$

Khung Punnett

	$\frac{1}{4} AB$	$\frac{1}{4} Ab$	$\frac{1}{4} aB$	$\frac{1}{4} ab$
$\frac{1}{4} AB$	AABB	AABb	AaBB	AaBb
$\frac{1}{4} Ab$	AABb	AAbb	AaBB	Aabb
$\frac{1}{4} aB$	AaBB	AaBb	aaBB	aaBb
$\frac{1}{4} ab$	AaBb	Aabb	aaBb	aabb

$F_2$       Tỷ lệ kiểu gene      Tỷ lệ kiểu hình

$\frac{1}{16} AABB + \frac{2}{16} AaBB +$   
 $\frac{2}{16} AABb + \frac{4}{16} AaBb$       =  $\frac{9}{16}$  vàng, trơn  
 $\frac{1}{16} AAbb + \frac{2}{16} Aabb$       =  $\frac{3}{16}$  vàng, nhãn  
 $\frac{1}{16} aaBB + \frac{2}{16} aaBb$       =  $\frac{3}{16}$  xanh, trơn  
 $\frac{1}{16} aabb$       =  $\frac{1}{16}$  xanh, nhãn

**Hình 1.4** Cơ sở di truyền học của nguyên lý phân ly độc lập.

## V. Sự di truyền Mendel ở người

Cũng như ở đậu Hà Lan, ruồi giấm hay mèo, ở người có rất nhiều tính trạng di truyền theo các quy luật Mendel. Chẳng hạn, theo thống kê của Victor A. McKusick năm 1994, có 6.678 tính trạng và các bệnh đơn gene. Cho đến ngày 8/2/2005, các số liệu về số lượng gene và kiểu hình được nêu ở bảng 1.4 (OMIM 2005).

Nói chung, việc xác định phương thức di truyền ở người là tương đối khó khăn, vì mỗi gia đình có ít con, thường không quá 10 người. Để khắc phục điều đó người ta sử dụng phương pháp *phân tích phả hệ* (pedigree analysis). Dưới đây nêu một số tính trạng trội và lặn ở người mà không phân tích đặc điểm của các kiểu di truyền đó (xem chương 11).



**Bảng 1.4 Số lượng các mục**

	NST thường	Liên kết X	Liên kết Y	DNA ty thể	Tổng
Gene có trình tự đã biết	9517	423	48	37	10.025
Gene có trình tự và kiểu hình đã rõ	360	38	0	0	398
Mô tả kiểu hình, cơ sở ph.tử đã rõ	1512	137	2	27	1.678
Kiểu hình hay locus Mendel, cơ sở phân tử chưa biết	1326	134	4	0	1.464
Các kiểu hình chính yếu khác có cơ sở Mendel còn khả nghi	2150	153	2	0	2.305
<b>Tổng</b>	<b>14.865</b>	<b>885</b>	<b>56</b>	<b>64</b>	<b>15.870</b>

### 1. Các tính trạng lặn (*recessive traits*)

Ở người, hầu hết các rối loạn di truyền là lặn (xem bảng 1.5). Đa số những người mắc các bệnh này thường có bố mẹ đều bình thường về kiểu hình, nhưng lại mang gene bệnh ở trạng thái dị hợp.

**Bảng 1.5 Một số rối loạn đơn gene thuộc nhiễm sắc thể thường ở người**  
(phỏng theo Campbell và Reece 2001; Lewis 2003)

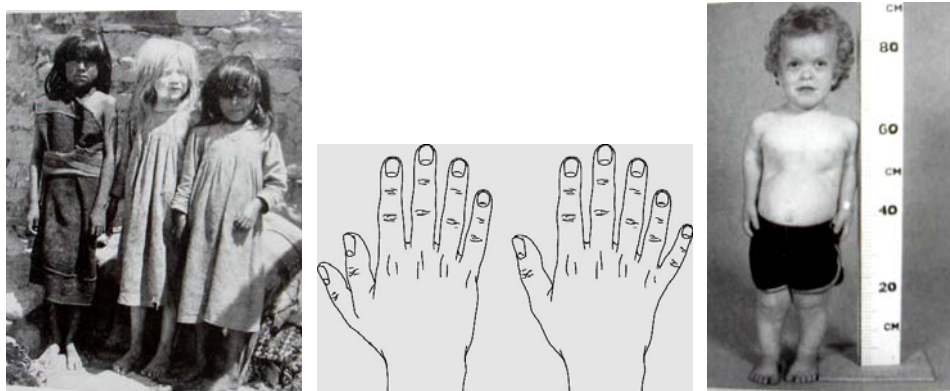
<b>Rối loạn</b>	<b>Các triệu chứng (và nguy cơ mắc phải)</b>
<b><i>Các rối loạn lặn</i></b>	
Bạch tạng	Thiếu sắc tố ở da, tóc và mắt (1/22.000)
Hóa xơ nang	Thừa chất nhầy ở phổi, dịch tiêu hóa, gan
Galactosemia	Tích lũy galactose ở các mô; trì độn; tổn thương mắt và gan. (1/100.000; xử lý bằng kiêng galactose).
Gaucher	Gan và lách sưng phồng, thiếu máu, xuất huyết nội
Hemochromatosis	Cơ thể giữ lại sắt; nguy cơ lây nhiễm cao, tổn thương gan, tim và tụy, thừa sắc tố da
Phenylketonuria	Tích lũy phenylalanyl trong máu; thiếu sắc tố da bình thường; trì độn, hôn mê, chết ở tuổi nhỏ (1/10.000)
Bệnh hồng cầu liềm (đồng hợp tử)	Tổn thương lách và nhiều cơ quan; nguy cơ lây nhiễm cao (1/500 ở người Mỹ gốc châu Phi; đồng trội)
Bệnh Tay-Sachs	Tích lũy lipid trong các tế bào não gây suy thoái thần kinh; mù màu; chết thời thơ ấu
<b><i>Các rối loạn trội</i></b>	
Achondroplasia	Lùn với tứ chi ngắn, đầu và thân bình thường
Bệnh Alzheimer	Thoái hóa tâm thần; thường xảy ra muộn màng
Bệnh Huntington	Thoái hóa tâm thần và các cử động mất kiểm soát
Hypercholesterolemia	Thừa cholesterol trong máu; bệnh tim; 1/500 là dị hợp
Không dung nạp lactose	Không có khả năng phân giải đường lactose, gây ra tình trạng chuột rút sau khi ăn loại đường này
Hội chứng Marfan	Tứ chi dài như vượn, ngực lõm, hồng thủy tinh thể, các ngón tay mảnh khảnh, động mạch chủ suy yếu

Bệnh thận đa nang	Các bọng trong các quả thận, urê trong máu, huyết áp cao, đau bụng dưới
Tật nhiều ngón	Thừa các ngón tay, ngón chân

Hai bệnh lặn điển hình đó là bạch tạng và hóa xơ nang. Những người bị *bạch tạng* (albino) là do thiếu hụt sắc tố melanin, nên da dễ trắng bạch, tóc và tròng đen của mắt trở nên nhạt khác thường (hình 1.5)... *Hóa xơ nang* (cystic fibrosis) là một bệnh di truyền gây chết phổ biến nhất ở Mỹ (USA). Bệnh lặn này phổ biến nhất ở những người Mỹ da trắng gốc Capca (Caucasians), với tần số chung là 1/1.800, nghĩa là trung bình cứ 25 người có một người mang allele lặn này (Campbell và Reece 2001) hay đối với trẻ sơ sinh là 1/2.500 (Weaver và Hedrick 1997). Người bị bệnh này có đặc điểm là tiết ra một lượng dư thừa chất nhầy dày ở phổi, tụy và các cơ quan khác. Các chất nhầy này có thể làm nhiều loạn sự thở, tiêu hóa và chức năng của gan và làm cho người bệnh rơi vào nguy cơ viêm phổi và các bệnh truyền nhiễm khác. Nếu không được điều trị, hầu hết trẻ em mắc bệnh này sẽ bị chết ở độ tuổi lên năm. Theo thống kê của McKusick năm 1994, có 1.730 mục cho các locus lặn.

## 2. Các tính trạng trội (dominant traits)

Mặc dù hầu hết các allele có hại là allele lặn, nhưng một số các rối loạn ở người là do các allele trội (xem bảng 1.5). Trong số đó có một vài allele không gây chết, chẳng hạn như tật thừa các ngón tay và chân (hình 1.5), hoặc có màng da giữa các ngón tay và chân. Các tính trạng như có tàn nhang, mái tai thông cũng như các khả năng uốn lưỡi hình ống và gập ngược lưỡi lên trên đều do các gene trội đơn khác nhau kiểm soát. Theo thống kê của McKusick năm 1994, có 4.458 mục cho các locus trội.



**Hình 1.5 Một số ví dụ về di truyền Mendel ở người.**

Từ trái sang là bạch tạng, tật nhiều ngón và dạng lùn phổ biến (achondroplasia).

Thí dụ điển hình về rối loạn trội nghiêm trọng đó là dạng lùn phổ biến do thoái hóa sụn gọi là *achondroplasia* (hình 1.5), đầu và thân mình phát triển bình thường nhưng tay chân ngắn một cách bất thường; tỷ lệ mắc bệnh này là khoảng 1 trên 25.000 người. Chỉ những người dị hợp tử mới bị rối loạn này; còn kiểu gen đồng hợp tử trội gây chết phôi. Trường hợp khác là *bệnh Huntington* (Huntington's disease), một dạng rối loạn do sự suy thoái của hệ thần kinh thường xảy ra từ sau độ tuổi trung niên. Khi bệnh tiến triển, nó làm cho các cử động trên mọi phần của cơ thể mất khả năng kiểm soát. Sự mất mát các tế bào não dẫn tới mất trí nhớ và khả năng suy xét, góp phần đẩy nhanh sự suy thoái. Cuối cùng, mất luôn các kỹ năng vận động làm cho không nuốt và nói năng được. Cái chết thường xảy ra sau khi các triệu chứng đó bắt đầu biểu hiện khoảng 10-20 năm (Campbell và Reece 2001).

## VI. Lý thuyết xác suất trong dự đoán và phân tích di truyền học

Để hiểu rõ các phát hiện của Mendel và các nguyên lý của di truyền học nói chung, cũng như để vận dụng các kiến thức này một cách có hiệu quả vào trong học tập và thực tiễn đời sống-sản xuất, chúng ta cần nắm vững một vài khái niệm và nguyên lý xác suất cơ bản sau đây.

### 1. Một số khái niệm và tính chất cơ bản của xác suất

(1) Một cách đơn giản, *xác suất* (probability) được định nghĩa bằng số lần xảy ra một *biến cố* hay *sự kiện* (event) cụ thể chia cho tổng số cơ may mà biến cố đó có thể xảy ra. Nếu ta ký hiệu xác suất của biến cố A là  $P(A)$ , m là số lần xuất hiện của A và n là tổng số phép thử hay toàn bộ số khả năng có thể có, khi đó:  $P(A) = m / n$ ; trong đó  $0 \leq m \leq n$ , và  $n > 0$ . Như vậy,  $0 \leq P(A) \leq 1$ .

(2) *Phép thử* là việc thực hiện một nhóm các điều kiện xác định, ví dụ một thí nghiệm tung đồng xu hay gieo hạt xúc xích hoặc một phép lai cụ thể. Các kết cục khác nhau có thể có từ phép thử gọi là các *biến cố*, được ký hiệu bằng các chữ cái in hoa A, B, C... Ví dụ: Khi tung một đồng xu, sự kiện xảy ra chỉ có thể là mặt sấp (S) hoặc ngửa (N) với xác suất tương đương là 0,5. Tương tự, kiểu gene dị hợp Aa có thể tạo ra hai loại giao tử mang A và a với xác suất ngang nhau là 0,5 trong khi các kiểu gene đồng hợp như aa chẳng hạn chỉ cho một loại giao tử mang a; vì vậy đời con của phép lai phân tích  $Aa \times aa$  có tỷ lệ kỳ vọng là 0,5 Aa : 0,5 aa.

Khi thực hiện phép thử có thể xuất hiện một trong các loại biến cố sau:

-*Biến cố ngẫu nhiên* (A) là sự kiện có thể xảy ra nhưng cũng có thể không,  $0 \leq P(A) \leq 1$ .

-*Biến cố chắc chắn* ( $\Omega$ ) là sự kiện nhất thiết xảy ra,  $P(\Omega) = 1$ .

-*Biến cố không thể có* ( $\emptyset$ ) là sự kiện nhất thiết không xảy ra,  $P(\emptyset) = 0$ .

-*Biến cố xung khắc*: Hai biến cố A và B gọi là đôi xung khắc với nhau nếu tích của chúng là một biến cố không thể có:  $A \cap B = \emptyset \Rightarrow P(A \cap B) = 0$  và  $P(A \cup B) = P(A) + P(B)$ .

-*Biến cố đối lập*: "không A" ( $\bar{A}$ ) được gọi là biến cố đối lập của biến cố A khi  $\bar{A} = \Omega \setminus A$  và  $\bar{A} \cup A = \Omega$ . Khi đó  $P(\bar{A}) = 1 - P(A)$ .

-*Nhóm đầy đủ các biến cố* hay *không gian biến cố sơ cấp* ( $\Omega$ ) là tập hợp toàn bộ các biến cố sơ cấp ( $\omega$ ) của một phép thử mà khi được thực hiện thì nhất thiết một trong chúng phải xảy ra, và có hiện tượng xung khắc từng đôi. Ví dụ, dãy các biến cố  $B_1, B_2, \dots, B_n$  lập thành một nhóm đầy đủ nếu thỏa mãn hai điều kiện: (i) Tổng của chúng là một biến cố chắc chắn:  $B_1 \cup B_2 \cup \dots \cup B_n = \Omega$ ; và (ii) Chúng xung khắc từng đôi một:  $B_i B_j = \emptyset$ ;  $i \neq j$  ( $i, j = 1, 2, \dots, n$ ).

## 2. Một số nguyên lý xác suất cơ bản

### 2.1. Quy tắc cộng

Quy tắc cộng phát biểu rằng: *xác suất kết hợp của hai (hoặc nhiều) sự kiện xung khắc từng đôi xảy ra là tổng các xác suất riêng rẽ của chúng.*

(i) Với A, B là hai biến cố bất kỳ, ta có  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(ii) Nếu A và B là hai biến cố xung khắc, thì:  $P(A \cup B) = P(A) + P(B)$

(iii) Hệ quả: Nếu  $\bar{A}$  là biến cố đối lập của A, thì  $P(\bar{A}) = 1 - P(A)$ .

Ví dụ, với kiểu gene Rr, nếu ta ký hiệu xác suất của loại giao tử mang allele R là P(R) và của loại giao tử mang allele a là P(r), theo lý thuyết ta có:  $P(R \text{ hoặc } r) = P(R) + P(r) = 1/2 + 1/2 = 1$ .

### 2.2. Xác suất điều kiện

*Định nghĩa*: Nếu A, B là hai biến cố bất kỳ và  $P(A) > 0$ , thì xác suất điều kiện của biến cố B với điều kiện biến cố A đã xảy ra là:

$$P(B/A) = P(A \cap B) : P(A).$$

*Ví dụ*: Từ kết quả thí nghiệm lai giữa hai thứ đậu thuần chủng hạt vàng và hạt xanh của Mendel, hãy tìm xác suất để một cây đậu hạt vàng ở  $F_2$  là thể dị hợp (Vv). Bằng lập luận thông thường ta thấy trong số bốn kiểu tổ hợp ở  $F_2$  có 3 kiểu tổ hợp cho kiểu hình hạt vàng (V-) nhưng chỉ có 2 kiểu là dị hợp (Vv). Vì vậy xác suất cần tìm là 2/3.

Nếu giải theo định nghĩa xác suất, ta ký hiệu: A là sự kiện hạt vàng ở  $F_2$  và B là sự kiện hạt vàng dị hợp. Theo lý thuyết, ở  $F_2$  có 4 sự kiện đồng khả năng với tỷ lệ là 1VV: 2Vv: 1vv. Ở đây  $P(A) = 3/4$  và  $P(A.B) = P(B) = 2/4 \Rightarrow P(B/A) = P(A \cap B) : P(A) = 2/4 : 3/4 = 2/3$ .

### 2.3. Quy tắc nhân

Từ định nghĩa về xác suất điều kiện, ta đi đến định nghĩa về *hai biến cố độc lập* như sau: Hai biến cố A và B được gọi là độc lập với nhau nếu  $P(B/A) = P(B)$  hoặc  $P(A/B) = P(A)$ . Nghĩa là sự xảy ra hay không xảy ra của biến cố này không ảnh hưởng đến sự xảy ra của biến cố kia.

Khi đó, quy tắc nhân được phát biểu như sau: *Xác suất trùng hợp của cả hai biến cố độc lập bằng tích các xác suất riêng rẽ của chúng*. Nghĩa là, nếu A và B là các biến cố độc lập thì:  $P(A \cap B) = P(A) \cdot P(B)$

*Ví dụ:* Khi gieo một đồng xu hai lần liên tiếp hoặc gieo hai đồng xu cân đối và đồng chất cùng một lượt thì có 4 khả năng xảy ra là (SS), (SN), (NS), (NN) với xác suất ngang nhau. Thật vậy, xác suất xuất hiện cả hai mặt sấp hoặc ngửa là:  $P(S,S) = P(N,N) = 1/2 \times 1/2 = 1/4$ ; còn xác suất xuất hiện một mặt sấp và một mặt ngửa là  $P(S,N) = 2 \times 1/2 \times 1/2 = 1/2$ ; và tổng các xác suất của tất cả các sự kiện đó đương nhiên bằng 1.

Từ đây ta có thể xây dựng cơ sở của "phép thử độc lập", nguyên tắc giải bài toán lai hai tính và các phương pháp biểu diễn kết quả lai như sau:

(i) Mệnh đề "*phép thử độc lập*" dưới đây được thiết lập dựa trên quy tắc nhân xác suất, chủ yếu dùng để kiểm tra xem quy luật di truyền chi phối đồng thời cả hai tính trạng nào đó trong một phép lai là độc lập hay liên kết (nếu là liên kết thì liên kết hoàn toàn hay không hoàn toàn), hoặc một tính trạng nào đó có liên kết với giới tính hay không. Mệnh đề này được phát biểu như sau: *Nếu các gene phân ly độc lập và tổ hợp tự do, thì tỷ lệ phân ly đồng thời của cả hai tính trạng bằng tích các tỷ lệ phân ly riêng rẽ của các tính trạng đó, và ngược lại*.

*Ví dụ 1:* Kết quả của một phép lai hai tính có tỷ lệ phân ly là 3:3:1:1 và các tỷ lệ của từng tính là 3:1 và 1:1. Ta nói rằng các tính trạng đó tuân theo quy luật phân ly độc lập, vì:  $3:3:1:1 = (3:1)(1:1)$ .

*Ví dụ 2:* Kết quả sự tự thụ tinh của một cá thể  $F_1$  (hoặc tạp giao giữa các cá thể  $F_1$  sinh ra từ các bố mẹ thuần chủng khác nhau về từng cặp gene allele) về một tính trạng nào đó mà có tỷ lệ phân ly là 9:3:3:1 hoặc là một biến dạng của công thức phân ly đó. Ta có thể khẳng định rằng tính trạng này tuân theo quy luật tương tác giữa các gene không allele phân ly độc lập. Thật vậy, từ kết quả trên cho thấy ở đời lai có  $9+3+3+1 = 16$  kiểu tổ hợp với tỷ lệ ngang nhau, chứng tỏ (các) cá thể  $F_1$  đó đã cho ít nhất 4 loại giao tử với tỷ lệ tương đương, nghĩa là dị hợp ít nhất 2 cặp gene phân ly độc lập. Nghĩa là tính trạng đó tuân theo quy luật tương tác gene.

(ii) Hệ quả: *Nếu tích các tỷ lệ phân ly riêng rẽ của các tính trạng khác với tỷ lệ phân ly đồng thời của cả hai tính, chứng tỏ các tính trạng đó tuân*

theo quy luật di truyền liên kết.

*Vi dụ:* Kết quả của một phép lai cho thấy tỷ lệ phân ly của cả hai tính là 1A-bb:2A-B-:1aabb hoặc 3A-B-:1aabb, trong khi tỷ lệ phân ly của mỗi tính trạng vẫn là 3:1. Ta dễ dàng thấy rằng tích  $(3:1)(3:1) \neq 1:2:1$  hoặc 3:1, chứng tỏ các tính trạng này tuân theo quy luật liên kết hoàn toàn, và kiểu gene của bố mẹ chúng đối với trường hợp đầu là Ab/aB  $\times$  Ab/aB hoặc AB/ab  $\times$  Ab/aB, và trường hợp sau là AB/ab  $\times$  AB/ab.

(iii) Nguyên tắc tổng quát giải một bài toán lai hai hoặc nhiều tính

Từ những điều trình bày trên đây cho thấy rằng thực ra toán lai gồm hai dạng tổng quát: lai một tính và lai nhiều tính, và việc giải một bài toán lai hai tính trở lên nói chung được thực hiện theo ba bước chính: *Bước 1:* Tách xét quy luật di truyền của từng tính trạng; *Bước 2:* Gộp xét quy luật di truyền của hai tính hoặc từng hai tính trạng một (để xác định xem chúng tuân theo quy luật nào, độc lập hay liên kết); và *Bước 3:* Thiết lập (các) sơ đồ lai kiểm chứng (xem "Phép giải toán lai", *Hoa học trò* số 362 tr.40-41).

Như vậy, cái khó của các bài toán lai thực ra là ở dạng lai một tính, và cái khó nhất của dạng toán này là xác định cho được quy luật chi phối tính trạng và các kiểu quan hệ có thể có giữa các gene allele cũng như giữa các gene không allele với nhau. Từ đó mới có thể xác định các kiểu gene, kiểu hình, và cuối cùng là kiểm chứng các kết quả bằng sơ đồ lai thích hợp.

(iv) Các phương pháp biểu diễn kết quả lai: Bởi vì bất cứ một mệnh đề quy luật nào cũng là "mệnh đề kéo theo" và có tính xác suất (nghĩa là chứa đựng khả năng tiên đoán tiềm tàng), cho nên việc tính toán và biểu diễn các kết quả kỳ vọng của bất kỳ một phép lai nào rõ ràng là đều dựa trên cơ sở xác suất mà chủ yếu là quy tắc nhân. Có nhiều cách khác nhau trong biểu diễn kết quả của các phép lai. Chẳng hạn, trình bày dưới dạng khung Punnett, vốn là phương pháp kinh điển của di truyền học. Một phương pháp thông dụng khác là *phương pháp phân nhánh* (forked-line approach), Phương pháp này có thể dùng để tìm xác suất của các kiểu đời con đối với nhiều gene phân ly độc lập; nó cũng rất tiện dụng khi xác định thành phần allele và xác suất của các loại giao tử từ một kiểu gene cụ thể. Ngoài ra, tùy trường hợp có thể sử dụng trực tiếp *quy tắc nhân* (product rule).

#### 2.4. Công thức xác suất nhị thức

Để đơn giản, ta xét phép thử đồng xu gồm hai sự kiện đối lập là mặt sấp (S) và mặt ngửa (N), với các xác suất tương ứng là p và q, trong đó  $q = 1-p$ . Giả sử trong n phép thử độc lập được tiến hành, sự kiện S xuất hiện k lần và sự kiện N là n-k. Để tính các xác suất này ta phải sử dụng công thức *xác suất nhị thức* (binomial probability) sau đây:



$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k}; \quad k = 0, 1, 2, \dots, n.$$

trong đó,  $C_n^k$  là hệ số nhị thức; ở đây  $C_n^k = n! / k!(n-k)!$

*Lưu ý:* Dãy  $n$  phép thử này còn gọi là dãy *phép thử Bernoulli*, vì nó thoả mãn ba điều kiện: (1) Nó là dãy độc lập; (2) Trong mỗi phép thử, không gian biến cố sơ cấp chỉ có hai biến cố sơ cấp đối lập là  $S$  và  $N$ ; và (3) Xác suất của một biến cố  $S$  hoặc  $N$  là không thay đổi trong mọi phép thử,  $P(S) = p$  và  $P(N) = q = 1 - p$ .

*Ví dụ:* Gieo một đồng xu liên tiếp ba lần, có thể xảy ra  $2^3 = 8$  khả năng độc lập về thứ tự trong 4 nhóm sau đây:

3 mặt sấp	2 sấp và 1 ngửa	1 sấp và 2 ngửa	3 mặt ngửa
SSS	SSN	SNN	NNN
	SNS	NSN	
	NSS	NNS	

Với giả thiết  $p = q = 1/2$ , dựa vào công thức xác suất nhị thức ta dễ dàng tính được xác suất của mỗi trường hợp trên như sau:

$$P(3 \text{ mặt sấp}) = 1 \times (1/2)^3 \times (1/2)^0 = 1/8$$

$$P(3 \text{ mặt ngửa}) = 1 \times (1/2)^0 \times (1/2)^3 = 1/8$$

$$P(2 \text{ sấp } 1 \text{ ngửa}) = 3 \times (1/2)^2 \times (1/2)^1 = 3/8$$

$$P(1 \text{ sấp } 2 \text{ ngửa}) = 3 \times (1/2)^1 \times (1/2)^2 = 3/8$$

## 2.5. Công thức xác suất toàn phần

Giả sử dãy  $B_1, B_2, \dots, B_n$  là một nhóm đầy đủ các biến cố, nghĩa là chúng có hợp là một sự kiện tất yếu ( $B_1 \cup B_2 \cup \dots \cup B_n = \Omega$ ) và gồm từng đôi xung khắc ( $B_i \cap B_j = \emptyset$ , với  $i \neq j$ ;  $i, j = 1, 2, \dots, n$ ); và gọi  $A$  là một biến cố bất kỳ. Khi đó:

$$A = \Omega \cap A = (B_1 \cup B_2 \cup \dots \cup B_n) \cap A = (B_1 \cap A) \cup (B_2 \cap A) \cup \dots \cup (B_n \cap A)$$

Áp dụng các định lý cộng và nhân, ta có *công thức xác suất toàn phần*

$$P(A) = P(B_1) \cdot P(A/B_1) + P(B_2) \cdot P(A/B_2) + \dots + P(B_n) \cdot P(A/B_n)$$

$$= \sum P(B_i) \cdot P(A/B_i); \quad \text{với } i = 1, \dots, n.$$

## 2.6. Công thức Bayes

Với giả thiết như trên cộng thêm một điều kiện mới là phép thử được tiến hành và kết quả là sự kiện  $A$  đã xảy ra. Vì các  $B_i$  ( $i = 1, 2, \dots, n$ ) lập thành một nhóm đầy đủ nên cùng với  $A$  phải có chỉ một sự kiện  $B_k$  nào đó xảy ra. Vậy khi  $A$  xuất hiện thì sự kiện nào trong số các  $B_i$  có nhiều khả năng xảy ra hơn cả? Giải đáp câu hỏi này có nghĩa là cần tính xác suất

$P(B_k/A)$ . Vận dụng định lý nhân xác suất, ta có:

$$P(A \cap B_k) = P(A) \cdot P(B_k/A) = P(B_k) \cdot P(A/B_k)$$

Suy ra:  $P(B_k/A) = P(B_k) \cdot P(A/B_k) : P(A)$  ; với  $P(A) > 0$ .

Thay  $P(A)$  từ công thức xác suất toàn phần ở trên, ta được công thức Bayes như sau:  $P(B_k/A) = P(B_k) \cdot P(A/B_k) : \sum P(B_i) \cdot P(A/B_i)$

Trong đó các  $P(B_i)$  được gọi là các *xác suất tiên nghiệm* (vì chúng được xác định trước khi tiến hành phép thử), còn  $P(B_k/A)$  được gọi là *xác suất hậu nghiệm* (vì nó chỉ được xác định sau khi tiến hành phép thử).

Cần lưu ý rằng, các quy luật Mendel về thực chất là các định luật nhân xác suất, cho nên để tiên đoán các nguy cơ di truyền cần sử dụng đến công thức Bayes, tức là định lý xác suất của nguyên nhân. Và bài toán đặt ra chỉ có thể giải quyết trọn vẹn một khi biết được các xác suất tiên nghiệm.

*Ví dụ:* Lấy ngẫu nhiên một cây đậu hạt vàng ở  $F_2$  (trong thí nghiệm của Mendel) cho lai với cây hạt xanh, và ở thế hệ lai nhận được tất cả là 6 cây hạt vàng. Hãy tìm xác suất của cây hạt vàng  $F_2$  đem lai là thể đồng hợp. (Quy ước: Y- hạt vàng, y - hạt xanh).

*Giải:* Ta biết rằng ở  $F_2$  có tỷ lệ kiểu gene là  $1/4 YY : 1/2 Yy : 1/4 yy$  và tỷ lệ kiểu hình là  $3/4$  vàng :  $1/4$  xanh. Vì cây hạt vàng (Y-) được chọn ngẫu nhiên trong số các cây hạt vàng  $F_2$  nên nó có thể là đồng hợp (YY) hoặc dị hợp (Yy), với xác suất tương ứng là  $1/3$  hoặc  $2/3$ .

Gọi  $B_1$  - sự kiện "cây hạt vàng  $F_2$  lấy ra là thể đồng hợp"

$B_2$  - sự kiện "cây hạt vàng  $F_2$  lấy ra là thể dị hợp"; ( $B_1 \cup B_2 = \Omega$ )

và A là sự kiện "6 cây hạt vàng nhận được ở thế hệ lai".

Ta có các xác suất tiên nghiệm:  $P(B_1) = 1/3$  và  $P(B_2) = 2/3$

Và các xác suất điều kiện:  $P(A/B_1) = 1$  và  $P(A/B_2) = (1/2)^6 = 1/64$

Vậy xác suất (hậu nghiệm) cần tìm là:

$$\begin{aligned} P(B_1/A) &= P(B_1) \cdot P(A/B_1) : [P(B_1) \cdot P(A/B_1) + P(B_2) \cdot P(A/B_2)] \\ &= (1/3 \times 1) : [(1/3 \times 1) + (2/3 \times 1/64)] = 32/33 = 0,97. \end{aligned}$$

*Lưu ý:* Kết quả này cho thấy rằng cây hạt vàng  $F_2$  (được lấy ngẫu nhiên để lai phân tích) có kiểu gene đồng hợp là hầu như chắc chắn (đến 97%), với mức  $\alpha = 0,05$ . Nói cách khác, chỉ với 6 cây hạt vàng đó thôi cũng đủ để khẳng định sự đồng tính của đời con. Đến đây hẳn là chúng ta đã thấy rõ công thức Bayes có tầm quan trọng như thế nào trong việc tiên đoán nguyên nhân và nguy cơ di truyền, cũng như cho phép lý giải một số tình huống phức tạp (xem Hoàng Trọng Phán 1995, 2000b).

## VII. Phương pháp Khi-bình phương (Chi-square method) trong đánh giá độ phù hợp giữa các số liệu quan sát và kỳ vọng

Nói chung, các số liệu thống kê thu được từ các thí nghiệm vốn sai khác ít nhiều so với các con số mang tính chất lý thuyết, tùy thuộc chủ yếu vào mẫu thí nghiệm và phương pháp lấy mẫu. Trong trường hợp đó, chúng ta băn khoăn không rõ liệu sự giả định "xấp xỉ" của chúng ta có thật chắc chắn không? Hay nói theo ngôn ngữ thống kê, "giả thuyết tương đồng  $H_0$  được chấp nhận hay bị bác bỏ", nghĩa là kết quả thu được có thật nghiệm đúng với tỷ lệ của một quy luật nào đó hay không?

Để có được câu trả lời rất ráo cho vấn đề này chỉ có cách là sử dụng *trắc nghiệm Khi-bình phương* ( $\chi^2$ -test). Đây là một công cụ toán thống kê thông dụng cho phép kiểm tra *độ phù hợp* giữa các trị số thực tế *quan sát* được (observed, ký hiệu: O) và các trị số lý thuyết được *kỳ vọng* (expected, ký hiệu: E) của một giả thuyết hay phân phối thực nghiệm khoa học nào đó, hoặc để kiểm tra tính độc lập của hai đại lượng ngẫu nhiên. Nhờ đó ta có thể rút ra quy luật, hoặc hiệu quả của hai phương pháp thí nghiệm nào đó. Đúng về phương diện thực hành, phương pháp này được tiến hành đơn giản như sau:

*Bước 1:* Đặt giả thuyết tương đồng  $H_0$  và sau đó tính trị số  $\chi^2$  thực tế dựa theo công thức:  $\chi^2 = \sum [(O_i - E_i)^2 / E_i]$ ;  $i = 1, 2, \dots, n$ .

*Bước 2:* Tìm trị số  $\chi^2$  lý thuyết bằng cách tra *Bảng các giá trị của phân phối  $\chi^2_\alpha$*  với k bậc tự do. Thông thường người ta sử dụng mức xác suất sai lầm P hay mức ý nghĩa  $\alpha = 0,05$  và  $k = n - 1$ ; trong đó n là số lớp kiểu hình và nó còn tùy trường hợp cụ thể. Mức  $\alpha = 0,05$  thường được dùng làm ranh giới phân chia giữa miền ấn định chấp nhận giả thuyết  $H_0$  và miền ấn định bác bỏ giả thuyết  $H_0$ . Để tiện lợi, dưới đây nêu ra một vài trị số  $\chi^2_{\alpha=0,05}$  lý thuyết thông dụng ứng với một số bậc tự do k như sau:

Số bậc tự do (k)	1	2	3	4
Giá trị $\chi^2_{\alpha=0,05}$	3,84	5,99	7,82	9,49

*Bước 3:* So sánh các giá trị  $\chi^2$  thực tế và lý thuyết.

- Nếu như trị số  $\chi^2$  thực tế nhỏ hơn trị số  $\chi^2$  lý thuyết, tức là có mức xác suất P hay  $\alpha > 0,05$ , giả thuyết  $H_0$  được chấp nhận. Nghĩa là kết quả thu được phù hợp với tỷ lệ được giả định.

- Ngược lại, nếu như trị số  $\chi^2$  thực tế lớn hơn hoặc bằng trị số  $\chi^2$  lý thuyết, tức là có mức xác suất P hay  $\alpha \leq 0,05$ , giả thuyết  $H_0$  bị bác bỏ. Nghĩa là kết quả thu được không phù hợp với tỷ lệ được giả định.

*Ví dụ:* Trong thí nghiệm lai một tính của Mendel, ở  $F_2$  thu được 705 hoa tím và 224 hoa trắng. Trên nguyên tắc, với hai kiểu hình ở  $F_2$  khiến ta có thể nghĩ tới một số tỷ lệ lý thuyết gần với nó như 2:1, 3:1 hoặc 9:7. Nhưng ở đây tỷ lệ "tím : trắng" thực tế là 3,15:1 nên tỷ lệ kỳ vọng hợp lý hơn cả là 3:1. Đó chính là giả thuyết  $H_0$  cần kiểm tra.

Bây giờ ta có thể tính toán giá trị  $\chi^2$  thực tế như sau:

Kiểu hình	Số quan sát ( $O_i$ )	Số kỳ vọng ( $E_i$ )	$(O_i - E_i)^2 / E_i$
Hoa tím	705	$3/4 \times 929 = 696,75$	0,098
Hoa trắng	224	$1/4 \times 929 = 232,25$	0,293
Tổng	929	929	$\chi^2 = 0,391$

Bằng cách tra bảng các giá trị của phân phối  $\chi^2_{\alpha = 0,05}$  với  $k = 2 - 1 = 1$  bậc tự do, ta tìm được trị số  $\chi^2$  lý thuyết là 3,84. Vì trị số  $\chi^2$  thực tế (0,391) nhỏ hơn trị số  $\chi^2$  lý thuyết (3,84) rất nhiều, nên giả thuyết  $H_0$  hoàn toàn được chấp nhận. Nghĩa là kết quả thí nghiệm trên phù hợp một cách sít sao với tỷ lệ 3:1. Điều đó có nghĩa rằng sự sai khác giữa các số liệu thực và các con số lý thuyết tương ứng là rất không đáng kể, không có ý nghĩa về phương diện thống kê.

## Câu hỏi và Bài tập

1. Đối tượng và phương pháp thí nghiệm của Mendel có những đặc điểm độc đáo nào mà từ đó Mendel đã khám phá ra các quy luật di truyền cơ sở đầu tiên, đặt nền móng cho sự ra đời của di truyền học?

2. Khi lai giữa chuột đen và chuột nâu nhận được  $F_1$  toàn chuột đen. Sau đó cho các chuột này tạp giao với nhau thu được  $F_2$  gồm 53 đen và 17 nâu. Hãy giải thích kết quả, viết các sơ đồ lai và tính xác suất của mỗi trường hợp phân ly kiểu hình có thể có ở  $F_2$  (biết rằng số chuột  $F_2$  thu được từ mỗi tổ hợp lai của các chuột  $F_1$  trung bình là bốn con).

3. Từ kết quả lai ở cừu sau đây, hãy xác định kiểu gene của mỗi cá thể.

<u>Phép lai</u>	<u>Đời con</u>
Trắng-1 x trắng-2	6 trắng : 1 đen
Trắng-1 x trắng-3	5 trắng
Trắng-1 x đen-1	3 trắng : 3 đen

4. Cho một ruồi giấm mắt nâu cánh dài giao phối với một ruồi giấm mắt đỏ cánh dài. Đời con thu được gồm 51 đỏ dài, 53 nâu dài, 17 đỏ ngắn và 18 nâu ngắn. Hãy xác định kiểu gene của các ruồi giấm bố mẹ.

5. Dựa vào các phép lai và kết quả dưới đây, hãy biện luận để chỉ ra các quy luật di truyền và kiểu gene của các bố mẹ trong mỗi phép lai.

Phép lai	Đời con			
	Vàng, trơn	Vàng, nhăn	Xanh, trơn	Xanh, nhăn
Vàng, trơn × vàng, trơn	45	15	16	5
Vàng, nhăn × vàng, nhăn	0	42	0	15
Xanh, trơn × vàng, nhăn	31	29	32	30

6. Giả sử đã thực hiện phép lai giữa con mèo lông ngắn, thưa, có đốm trắng (IlddSs) và mèo lông dài, rậm, có đốm trắng (LlDdSs). Hãy trình bày các phương pháp nhân xác suất đơn giản và sơ đồ phân nhánh để tính tỷ lệ kỳ vọng của mỗi một mèo con có kiểu hình sau: lông ngắn, thưa, không có đốm trắng; lông ngắn, rậm, có đốm trắng; và lông dài, rậm, có đốm trắng.

7. Ở ruồi giấm, mắt đỏ ( $se^+$ ) là trội so với mắt màu nâu tối ( $se$ ; viết tắt của *sepia*). Cho hai ruồi giấm mắt đỏ dị hợp lai với nhau, và lấy một con mắt đỏ đời con cho lai trở lại với một bố mẹ mắt đỏ. Cơ may để đời con phép lai ngược đó xuất hiện ruồi giấm mắt màu nâu tối là bao nhiêu?

8. Để khẳng định các nguyên lý di truyền của Mendel, các nhà nghiên cứu đã lặp lại các thí nghiệm ở đậu Hà Lan. Riêng các thí nghiệm về màu sắc hạt, ở  $F_2$  Correns thu được 1.394 vàng và 453 xanh, trong khi Tschermak quan sát thấy có 3.580 vàng và 1.190 xanh. Các trị số này có phù hợp với nguyên lý phân ly không? Bạn hãy chứng minh điều đó bằng công cụ toán thống kê thích hợp.

9. Giả sử người chồng bị bạch tạng, người vợ và các con của họ (một trai và hai gái) đều bình thường. Khi lớn lên, chúng lập gia đình với những người cũng không mắc bệnh này và mỗi gia đình này đều có một trai và một gái bình thường. Giả sử xảy ra sự kết hôn đồng huyết giữa hai đứa cháu có quan hệ cậu cô ruột, khi đó xác suất để sinh ra một trai đầu lòng mắc bệnh này là bao nhiêu? Hãy vẽ sơ đồ phả hệ nói trên và giải thích.

10. Giả sử bạn có một cây đậu Hà Lan thân cao chưa rõ kiểu gene. Bạn đem cây này lai phân tích với cây thân thấp hoặc cho nó tự thụ phấn và đã thu được tất cả bảy cây đều có thân cao. Hỏi xác suất để cây thân cao đem lai kiểm tra có kiểu gene đồng hợp là bao nhiêu?

## Tài liệu Tham khảo

### Tiếng Việt

Nguyễn Ngọc Kiêng 1996. *Thống kê học trong Nghiên cứu Khoa học*.

NXB Giáo Dục, Hà Nội.

Phạm Văn Kiều 1998. *Lý thuyết Xác suất và Thống kê Toán học*. NXB Khoa học và Kỹ thuật, Hà Nội.

Hoàng Trọng Phán 1995. Đánh giá hiện tượng đồng tính trong lai phân tích nhờ công thức Bayes. *Thông tin Khoa học và Giáo dục* số 11/1995, tr. 133-138, Trường ĐHSP Huế.

Hoàng Trọng Phán 2000a. Phép giải toán lai. *Hoa học trẻ* số 362, tr.40-41.

Hoàng Trọng Phán 2000b. Công thức Bayes và phép tính gần đúng trong phân tích di truyền học. *Thông báo Khoa học* số 4/2000, tr. 65-76, Trường ĐHSP Hà Nội.

### **Tiếng Anh**

Brooker RJ. 1999. *Genetics - Analysis and Principles*. Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA.

Campbell NA, Reece JB. 2001. *Essential Biology*. Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc, San Francisco, CA.

Hartl DL, Freifelder D, Snyder LA. 1988. *Basic Genetics*. Jones and Bartlett Publishers, Inc, Boston - Portola Valley.

Lewis R. 2003. *Human Genetics: Concepts and Applications*. 5<sup>th</sup> ed, McGraw-Hill, Inc, NY.

Online Mendelian Inheritance in Man (OMIM<sup>TM</sup>):

<<http://www.ncbi.nlm.nih.gov/Omim/mimstats.html>>

Russell PJ. 2000. *Fundamentals of Genetics*. 2<sup>nd</sup> ed, Benjamin/ Cummings Publishing Company, Inc, Menlo Park, CA.

Suzuki DT, Griffiths AJF, Miller JH, Lewontin RC. 1989. *An Introduction to Genetic Analysis*. 4<sup>th</sup> ed, W-H Freeman and Company, New York.

Tamarin RH. 1999. *Principles of Genetics*. 6<sup>th</sup> ed, McGraw-Hill, Inc, NY.

Weaver RF and Hedrick PW. 1997. *Genetics*. 3<sup>rd</sup> ed, McGraw-Hill Companies, Inc. Wm.C.Brown Publishers, Dubuque, IA.

### **Một số trang web**

<http://www.mhhe.com/lewisgenetics5>

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>

<http://fpt.netSPACE.org/MendelWeb/Mendel.htm>

<http://www.nord-rdb.com/~orphan>

## Chương 2

# Mở rộng và Áp dụng của Di truyền học Mendel

Nét đặc sắc của các thí nghiệm Mendel là ở sự nhất quán hoàn toàn giữa các quan sát và giả thuyết của ông. Tuy nhiên, kể từ sau khi các công trình của Mendel được khám phá lại, người ta còn phát hiện ra nhiều ngoại lệ suy rộng từ mô hình Mendel. Trong chương này, chúng ta sẽ tìm hiểu các vấn đề sau: (i) Sự khác nhau giữa các kiểu trội hoàn toàn, trội không hoàn toàn và đồng trội; (ii) Tác động của gene gây chết; (iii) Hiện tượng đa allele; (iv) Tính đa hiệu của gene; (v) Các kiểu di truyền do nhiều gene cùng tác động lên một tính trạng - tương tác gene; (vi) Cơ sở di truyền của các tính trạng số lượng. Các kiểu di truyền liên kết và liên kết với giới tính tuy cũng được xem như là sự mở rộng của các nguyên lý Mendel, nhưng sẽ được trình bày trong một chương riêng, chương 4.

### I. Các kiểu quan hệ giữa các gene allele đối với một tính trạng

#### 1. Các kiểu trội hoàn toàn, không hoàn toàn và đồng trội

Kể từ sau năm 1900, người ta còn phát hiện thêm một số trường hợp trội khác nhau, bổ sung cho tỷ lệ 3 trội : 1 lặn của Mendel.

##### 1.1. Trội hoàn toàn (*complete dominance*)

Đây là trường hợp di truyền trội-lặn Mendel. Trong hầu hết các trường hợp, allele bình thường (hay kiểu dại) trội hoàn toàn so với các allele đột biến. Điều này có thể lý giải dựa trên cơ sở di truyền sinh hóa ở chỗ, allele trội cho sản phẩm protein hoạt động chức năng bình thường trong khi allele đột biến không tạo ra được sản phẩm có hoạt tính. Do đó các cá thể đồng hợp về allele lặn không hoàn thành được con đường chuyển hóa có liên quan đến gene này. Ở người, đó là trường hợp của các allele đột biến lặn gây bạch tạng, bệnh *phenylxêton-niêu* (*phenylketonuria* = PKU)...

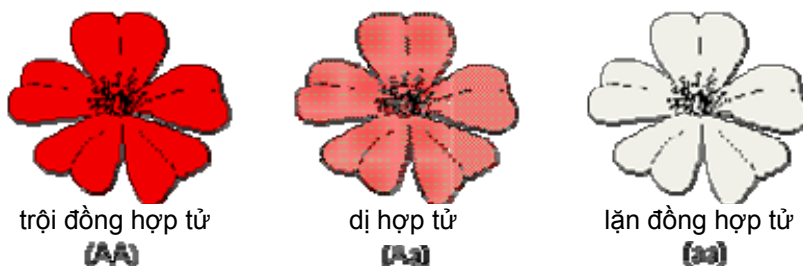
Tuy nhiên, ở một số trường hợp, allele đột biến trội hơn kiểu dại; nghĩa là allele kiểu dại là lặn. Ví dụ: ở người, kiểu lùn phổ biến do không tạo được sụn là trội, cho nên các thể dị hợp biểu hiện kiểu hình đột biến.

##### 1.2. Trội không hoàn toàn (*incomplete dominance*)

Khi lai giữa hai giống hoa bốn giờ (*four-o'clock*; *Mirabilis jalapa*) thuần chủng có hoa màu đỏ và hoa màu trắng, Carl Correns thu được tất cả các cây  $F_1$  có hoa màu hồng, kiểu hình trung gian giữa hai bố mẹ. Sau khi cho các cây  $F_1$  tự thụ phấn, tỷ lệ kiểu hình ở  $F_2$  là 1 đỏ : 2 hồng : 1 trắng. Mặc dù tỷ lệ kiểu hình này có hơi lệch so với của Mendel, nhưng

thực tế nó tương ứng với tỷ lệ kiểu gene 1:2:1 (hình 2.1). Nếu sử dụng quy ước gene A- đỏ là trội không hoàn toàn so với a- trắng, ta có sơ đồ lai sau:

$P_{tc}$	Hoa đỏ (AA) × Hoa trắng (aa)
$F_1$	Aa (Hoa hồng)
$F_2$	Tỷ lệ kiểu gene $\frac{1}{4} AA : \frac{1}{2} Aa : \frac{1}{4} aa$
	Tỷ lệ kiểu hình $\frac{1}{4} \text{đỏ} : \frac{1}{2} \text{hồng} : \frac{1}{4} \text{trắng}$

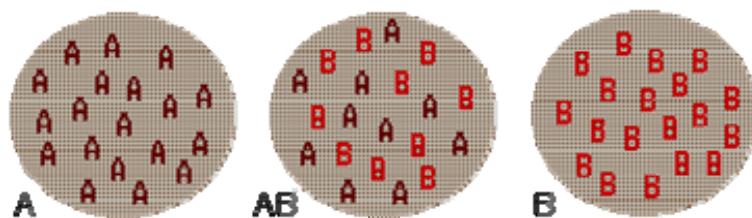


**Hình 2.1** Sự di truyền trung gian đối với màu sắc hoa ở nhiều thực vật.

Bởi kiểu hình của thể dị hợp là trung gian giữa hai thể đồng hợp, vì vậy ta có thể lý giải trên phương diện sinh hóa rằng hàm lượng sản phẩm tích lũy do một allele trội kiểm soát là không đủ để thể hiện kiểu hình màu đỏ như trong trường hợp có mặt cả hai allele trội.

### 1.3. Đồng trội (codominance)

Đồng trội là hiện tượng cả hai allele khác nhau trong một thể dị hợp cùng biểu hiện ra các sản phẩm có hoạt tính khác nhau trong tế bào. Các allele như thế được gọi là các allele đồng trội. Điển hình là trường hợp nhóm máu AB của hệ nhóm máu ABO (hình 2.2; xem giải thích ở mục 3 bên dưới) và nhóm máu MN của hệ nhóm máu M-N ở người.



**Hình 2.2** Kiểu hình các nhóm máu A, AB và B. (Ở đây cho thấy sự đồng trội ở nhóm máu AB. Nhóm máu O không có kháng nguyên nào).

Hệ nhóm máu M-N (do một locus thuộc nhiễm sắc thể thường kiểm soát) có hai allele  $L^M$  và  $L^N$ . Như thế, trong một quần thể sẽ có ba kiểu gene  $L^M L^M$ ,  $L^M L^N$  và  $L^N L^N$  (có thể viết gọn là MM, MN và NN) tương ứng với ba kiểu hình hay nhóm máu là M, MN và N. Nếu cho rằng các phép hôn phối thuận nghịch là tương đương, thì có thể có sáu kiểu hôn phối với các tỷ lệ kiểu gene kỳ vọng ở đời con được cho ở bảng 2.1.



**Bảng 2.1** Các tỷ lệ kỳ vọng ở đời con đối với hệ nhóm máu M-N

Bố mẹ	Đời con		
	$L^M L^M$	$L^M L^N$	$L^N L^N$
$L^M L^M \times L^M L^M$	1	—	—
$L^M L^M \times L^M L^N$	$\frac{1}{2}$	$\frac{1}{2}$	—
$L^M L^M \times L^N L^N$	—	1	—
$L^M L^N \times L^M L^N$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
$L^M L^N \times L^N L^N$	—	$\frac{1}{2}$	$\frac{1}{2}$
$L^N L^N \times L^N L^N$	—	—	1

Một ví dụ khác là allele lặn gây bệnh hồng cầu hình liềm. Ở những người dị hợp tử về allele này ( $Hb^A Hb^S$ ), cả hai allele đều được biểu hiện và các tế bào máu của họ chứa cả hemoglobin bình thường và bất thường.

## 2. Tác động của gene gây chết (lethal)

Các allele gây chết là những đột biến có thể trội hoặc lặn làm giảm sức sống hoặc gây chết đối với các cá thể mang nó và do đó, làm biến đổi tỷ lệ 3:1 của Mendel. Nhiều gene có các allele ảnh hưởng lên tỷ lệ chết chứ không gây chết; các allele này được gọi là các allele *có hại* (deleterious).

**Hình 2.3** Biến đổi màu lông ở chuột.**Hình 2.4** Mèo Manx không đuôi.

Nói chung, các allele gây chết thường là lặn và gây chết ở các thể đồng hợp. Ví dụ, đột biến bạch tạng ở thực vật làm cho cây chết ở giai đoạn non vì không có diệp lục để quang hợp. Bệnh thiếu máu hồng cầu hình liềm ở người (xem mục II) có thể gây chết với tỷ lệ đáng kể ở tuổi trưởng thành khi allele đột biến lặn này ở trạng thái đồng hợp.

Tuy nhiên, một số allele gây chết là những đột biến trội. Điển hình là thí nghiệm lai về màu sắc lông ở chuột của Lucien Cuénot năm 1904. Khi lai giữa hai chuột thân vàng (allele vàng là trội; Hình 2.3), ông thu được tỷ lệ xấp xỉ 2 vàng : 1 kiểu đại. Mặt khác, khi lai giữa các chuột vàng với chuột kiểu đại (màu agouti), ông thấy rằng đời con có tỷ lệ xấp xỉ 1:1.



chúng xác định tên nhóm máu tương ứng. Các cá thể mang allele  $I^A$  và/hoặc allele  $I^B$  có kháng nguyên tương ứng A và/hoặc B; còn allele  $I^O$  không tạo được bất kỳ kháng nguyên nào (nên những người đồng hợp về allele lặn này có kiểu hình O, nghĩa là "không" có kháng nguyên A hoặc B nào cả trên bề mặt các tế bào hồng cầu).

Các *kháng thể* (antibody) là những protein do hệ thống miễn dịch tạo ra với một số lượng lớn nhằm đáp ứng với các kháng nguyên đặc thù từ bên ngoài. Cụ thể, trong huyết thanh ở những người mang nhóm máu O thấy có cả hai loại kháng thể kháng-A và kháng-B gọi là  $\alpha$  và  $\beta$ ; ở những người nhóm máu AB không có bất kỳ loại kháng thể nào; còn ở những người nhóm máu A và B chỉ có một loại kháng thể tương ứng là  $\alpha$  và  $\beta$ .

Ý nghĩa lâm sàng của nguyên tắc truyền máu là ở chỗ khi truyền sai nhóm máu sẽ xảy ra phản ứng ngưng kết chết người do phản ứng giữa kháng nguyên bề mặt của hồng cầu người *cho* (donor), mà lượng huyết thanh là không đáng kể, với kháng thể có trong huyết thanh của người *nhận* (recipient). Vì thế, những người có cùng nhóm máu thì có thể cho và nhận của nhau. Đặc biệt, nhóm máu O do thiếu cả hai loại kháng nguyên A và B nên có thể truyền cho một người mang bất kỳ nhóm máu nào; ngược lại người mang nhóm máu AB có thể nhận bất kỳ nhóm máu nào.

Hiện tượng đa allele tồn tại phổ biến trong các quần thể sinh vật khác nhau. Ví dụ, gene kiểm soát màu mắt đỏ-trắng ở ruồi giấm gồm một chuỗi 12 allele, với tính trội giảm dần từ đỏ kiểu dại cho đến trắng đột biến lặn (w) theo thứ tự từ trái sang phải và trên xuống như sau:

<b>Allele</b>	$W^+$	$W^{sat}$	$W^{co}$	$W^w$	$W^{ap3}$	$W^{ch}$
Màu	đỏ dại	satsuma	coral	wine	apricot3	cherry
<b>Allele</b>	$W^e$	$W^{bl}$	$W^{ap}$	$W^i$	$W^t$	w
Màu	eosin	blood	apricot	ivory	tinged	white

Một số gene ở người, chẳng hạn như các gene đối với kháng nguyên bạch cầu người *HLA* (human leukocyte antigen) xác định các kháng nguyên trên bề mặt của hầu như tất cả các tế bào có thể có nhiều allele. Ví dụ, gene *HLA-B* có nhiều hơn 30 allele được xác định khác nhau về mặt kháng nguyên trong một số quần thể. Kết quả của sự đa dạng này là, trong một quần thể có rất nhiều kiểu gene ở gene *HLA-B* (465 kiểu gene khác nhau, với 30 kiểu đồng hợp và 435 kiểu dị hợp), thực sự tạo nên một dãy biến dị kiểu gene rộng.

## II. Tính đa hiệu của gene (*pleiotropy*)

Hiện tượng một gene ảnh hưởng đến hai hoặc nhiều tính trạng được gọi là *tính đa hiệu* (*pleiotropy*). Ví dụ, trong các thí nghiệm ở đậu Hà Lan,

Mendel đã lưu ý rằng gene kiểm soát màu hoa tím và trắng cũng ảnh hưởng lên màu sắc hạt (vỏ xám hoặc nâu) và gây ra sự có mặt hoặc không có mặt của các vết màu tím ở bẹ lá. Trong ví dụ allele kiểm soát lông vàng ở chuột nói trên, ta thấy rằng nó còn ảnh hưởng lên sức sống ở các thể dị hợp và gây chết ở các thể đồng hợp.

Nhiều bệnh di truyền ở người gây ra bởi các gen có tác dụng đa hiệu. Chẳng hạn, bệnh *phenylxê-tôn-niêu* (phenylketonuria = PKU) xảy ra ở các cá thể đồng hợp về allele lặn đó. Những người mắc bệnh này thiếu hẳn enzyme cần thiết cho sự chuyển hóa bình thường của amino acid phenylalanine thành sản phẩm sinh hóa kế tiếp. Khi so sánh các cá thể bình thường và PKU với nhau cho thấy mức phenylalanine ở nhóm bệnh cao hơn nhiều. Ngoài ra, ở những người bệnh còn có một số biến đổi khác như: chỉ số IQ thấp hơn, đầu bé hơn, và tóc hơi nhạt hơn. Tất cả các hiệu quả đa hiệu này có thể hiểu như là hậu quả của sự rối loạn sinh hóa cơ sở. Chẳng hạn, ở các bệnh nhân PKU có sự tích lũy một độc tố trong đầu khiến cho bộ não bị tổn thương và dẫn tới IQ thấp hơn, đầu bé hơn.

Một ví dụ khác về tính đa hiệu ở người là *bệnh hồng cầu hình liềm* (sickle-cell disease). Vậy tính đa hiệu giải thích trường hợp này như thế nào? Những người đồng hợp về allele đột biến lặn này ( $Hb^S Hb^S$ ) chỉ tạo ra các phân tử hemoglobin bất thường, khiến cho tất cả các tế bào hồng cầu có dạng hình liềm, kích thước bé, màu đỏ nhạt (Hình 2.5). Các tế bào hình liềm nhanh chóng bị cơ thể phá hủy và gây ra sự thiếu máu và suy yếu cơ thể nói chung. Ngoài ra,



**Hình 2.5** Các tế bào hồng cầu bình thường (trên) và dạng hình liềm (dưới), với độ phóng đại gấp đôi của hình trên.

do hình dạng góc cạnh mà các tế bào hình liềm không thể vận chuyển trơn tru trong máu và có xu hướng tích tụ và gây tắc nghẽn các mao mạch. Dòng máu đi đến các bộ phận cơ thể bị giảm bớt, gây sốt định kỳ, đau đớn, và tổn thương các cơ quan khác nhau như não bộ, lách, tim, thận...

### III. Các kiểu tương tác giữa các gene không allele

Kể từ sau khi phát hiện lại các nguyên lý di truyền của Mendel, đã xuất hiện nhiều công trình nghiên cứu trên các đối tượng khác nhau, như: bí ngô, đậu ngọt, lúa mì, ngô, ruồi giấm, gà, chuột, chó, ngựa... Một số kết quả thu được từ các thí nghiệm lai một tính đó cho thấy tỷ lệ phân ly khác với các tỷ lệ 3:1 hoặc 9:3:3:1. Các trường hợp cho tỷ lệ phân ly 9:3:3:1

hoặc một dạng biến đổi của nó được lý giải là do sự tương tác qua lại giữa các gene không allele phân ly độc lập.

Tự trung có hai kiểu tương tác gene, đó là: tương tác giữa các gen allele (đã xét ở trên) và tương tác giữa các gen không allele (thường được gọi tắt là *tương tác gene*). Do các gene không allele có thể nằm trên cùng một cặp nhiễm sắc thể hoặc trên các cặp nhiễm sắc thể tương đồng khác nhau, cho nên về nguyên tắc sẽ có hai dạng tương tác gene không allele tương ứng là tương tác độc lập và tương tác liên kết; và dạng sau phải là chủ yếu bởi vì số lượng gene của một sinh vật nhiều hơn số lượng nhiễm sắc thể của nó rất nhiều. Tuy nhiên, trên thực tế, khi nói đến sự tương tác giữa các gene không allele là ta muốn nói tới sự tương tác giữa các gene không allele phân ly độc lập.

*Lưu ý:* (1) Số lượng các tính trạng hình thành nên một cơ thể sinh vật nói chung là nhiều hơn số lượng gene mã hóa protein có mặt trong bộ gene của nó. Chẳng hạn, theo ước tính gần đây cho thấy trong bộ gene người có khoảng 25.000 gene khác nhau trong khi có tới 50.000 loại protein thực hiện tất cả mọi hoạt động sống của các tế bào. (2) Như thế, sự biểu hiện của đa số tính trạng là kết quả của sự tương tác giữa nhiều gene khác nhau trong quá trình phát triển cá thể; và thực chất của sự tương tác đó là do sự tương tác giữa các sản phẩm hoạt động của các gen có bản chất protein hoặc giữa các sản phẩm sinh ra từ sự xúc tác của các enzyme. (3) Con đường chuyển hóa một chất cụ thể trong tế bào là một chuỗi gồm nhiều phản ứng sinh hóa nối tiếp nhau mà mỗi khâu phản ứng được xúc tác bởi một enzyme (chương 7). Ngay cả một enzyme hay protein có cấu trúc bậc bốn cũng gồm một số chuỗi polypeptide thuộc các gene khác nhau tạo nên (chương 6). (4) Để cho đơn giản, thông thường ta giải thích các kết quả tương tác bằng hai gene nhưng trên thực tế, có thể có nhiều hơn hai gene cùng quy định một tính trạng.

Dưới đây chúng ta xét ba kiểu tương tác gen chính: Tương tác bổ trợ, tương tác át chế và tương tác đa phân cộng gộp. Các ví dụ nêu ra dưới đây là những thí nghiệm kinh điển của di truyền học. Để hiểu được bản chất di truyền của các tính trạng đó, đặc biệt là các tính trạng màu sắc, chúng ta cố gắng làm sáng tỏ cơ chế sinh hóa của chúng trong khả năng có thể.

### 1. Tương tác bổ trợ (*complementary*)

Tương tác bổ trợ là trường hợp tương tác gene làm xuất hiện kiểu hình mới khi có mặt đồng thời các gene không allele trong một kiểu gene. Các gene bổ trợ có thể là gene trội hoặc gene lặn (ở trạng thái đồng hợp). Tương tác kiểu bổ trợ biểu hiện dưới nhiều dạng với các tỷ lệ kiểu hình F<sub>2</sub> khác nhau, như 9:3:3:1; 9:6:1; 9:7.

### 1.1. Tương tác bổ trợ với tỷ lệ 9:3:3:1

*Vi dụ* kinh điển cho trường hợp này là các thí nghiệm của W. Bateson và R.C. Punnett về sự di truyền hình dạng mỏ ở gà. Khi lai giữa các giống gà thuần chủng mỏ hình hoa hồng với mỏ đơn (còn gọi là mỏ hình lá) thu được  $F_1$  toàn mỏ hoa hồng, và sau khi cho tạp giao  $F_1$  thì ở  $F_2$  có tỷ lệ phân ly 3 mỏ hoa hồng : 1 mỏ đơn. Tương tự, khi lai giữa các giống gà thuần chủng mỏ hình hạt đậu với mỏ đơn,  $F_1$  gồm tất cả mỏ hạt đậu và  $F_2$  phân ly 3 mỏ hạt đậu : 1 mỏ đơn. Nhưng khi lai giữa hai giống gà thuần chủng mỏ hoa hồng và mỏ hạt đậu, thì ở  $F_1$  lại thu được tất cả có mỏ hình quả óc chó (hay hạt hồ đào) và tỷ lệ phân ly ở  $F_2$  xấp xỉ 9 quả óc chó : 3 hình hoa hồng : 3 hình hạt đậu : 1 mỏ đơn (hình 2.6).

*Giải thích:* Các kết quả trong hai thí nghiệm đầu cho thấy các dạng mỏ hoa hồng và hạt đậu là trội so với dạng mỏ đơn. Kết quả sau cùng cho thấy  $F_2$  có 16 kiểu tổ hợp với tỷ lệ ngang nhau, trong khi  $F_1$  đồng nhất kiểu gen (vì bố mẹ thuần chủng); điều đó chứng tỏ  $F_1$  đã cho 4 loại giao tử với tỷ lệ tương đương, nghĩa là dị hợp tử về hai cặp gene phân ly độc lập. Suy ra tính trạng này do hai gene khác nhau chi phối, nghĩa là tuân theo *quy luật tương tác gene*. Mặt khác, kiểu hình mới biểu hiện ở  $F_1$  và khoảng 9/16 ở  $F_2$  (ứng với sự có mặt của cả hai gene trội không allele) phải là kết quả của sự tương tác giữa các gene trội không allele theo *kiểu bổ trợ*.

*Quy ước:* Dựa vào tỷ lệ phân ly kiểu hình ở  $F_2$  trong trường hợp hai gene phân ly độc lập và tên gọi các dạng mỏ theo tiếng Anh (rose: hoa hồng, pea: hạt đậu), ta có



óc chó: R-P-



hoa hồng: R-pp



hạt đậu: rrP-



mào đơn: rrpp

**Hình 2.6** Các kiểu mỏ đặc trưng của các giống gà khác nhau và các kiểu gene tương ứng.

thể quy ước và kiểm chứng đơn giản như sau:

R-P-: mỏ hình quả óc chó (do bổ trợ giữa các gene trội R và P)

R-pp: mỏ hình hoa hồng (do biểu hiện của gene trội R)

rrP-: mỏ hình hạt đậu (do biểu hiện của gene trội P)

rrpp: mỏ đơn (do khuyết cả hai gene trội; kiểu đại)

*Kiểm chứng:*

$P_{tc}$             mỏ hoa hồng (RRpp) × mỏ hạt đậu (rrPP)

$F_1$                 mỏ quả óc chó (RrPp)

$$F_1 \times F_1 = RrPp \times RrPp = (Rr \times Rr)(Pp \times Pp)$$

$$\rightarrow F_2 = (3R-:1rr)(3P-:1pp) = 9 R-P- : 3 R-pp : 3 rrP- : 1 rrrp$$

$$= 9 \text{ óc chó} : 3 \text{ hoa hồng} : 3 \text{ hạt đậu} : 1 \text{ đơn}$$

Một ví dụ khác, đó là lai giữa hai giống chuột lông màu đen với màu vỏ quế (cinnamon) được  $F_1$  toàn chuột lông màu xám (có dạng "muối tiêu" khi nhìn gần, còn gọi là *agouti*) và  $F_2$  cho tỷ lệ 9 xám : 3 màu vỏ quế : 3 đen : 1 nâu. Bạn có thể giải thích cơ sở di truyền của trường hợp này?

### 1.2. Tương tác bổ trợ với tỷ lệ 9:7

*Ví dụ:* Thí nghiệm của Bateson và Punnett về sự di truyền màu sắc hoa ở cây đậu ngọt (*Lathyrus odoratus*). Từ phép lai giữa hai giống hoa trắng thuần chủng khác nhau, hai ông thu được  $F_1$  gồm tất cả cây lai có hoa màu đỏ tía (purple). Khi cho các cây  $F_1$  tự thụ phấn ở  $F_2$  nhận được 382 hoa đỏ tía và 269 hoa trắng; kết quả này gần với tỷ lệ 9:7.

*Giải thích:* Với cách lập luận như trên, ta dễ dàng thấy rằng kiểu hình hoa đỏ tía là kết quả của sự tương tác bổ trợ giữa hai gene trội không allele phân ly độc lập. Trước khi đi vào giải thích cơ sở di truyền sinh hóa của các kiểu hình, ta hãy quy ước và kiểm chứng kết quả thí nghiệm trên.

*Quy ước:*

C-P- : hoa đỏ tía (do tác động bổ trợ giữa các gene trội C và P)

C-pp, ccP-, ccpp : hoa trắng (do không có mặt đầy đủ cả hai gene trội)

*Kiểm chứng:*

$P_{tc}$  giống hoa trắng 1(CCpp) × giống hoa trắng 2(ccPP)

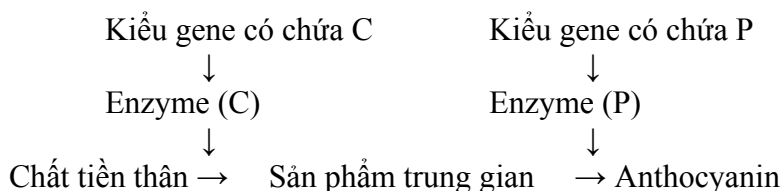
$F_1$  hoa đỏ tía (CcPp)

$F_1 \times F_1 = CcPp \times CcPp = (Cc \times Cc)(Pp \times Pp)$

$\rightarrow F_2 = (3C-:1cc)(3P-:1pp) = 9 C-P- : (3 C-pp + 3 ccP- + 1 ccpp)$

$= 9 \text{ đỏ tía} : 7 \text{ trắng}$

Bây giờ ta thử giải thích cơ sở sinh hóa của các kiểu hình. Sự hình thành màu hoa ở cây đậu ngọt là kết quả của sự tổng hợp một hợp chất gọi là *anthocyanin*; nó xảy ra thông qua một chuỗi các khâu chuyển hóa được xúc tác bởi các enzyme vốn là sản phẩm của các gene. Nếu như bất kỳ khâu nào trong quá trình tổng hợp này bị gián đoạn do vắng mặt của một enzyme hoạt động thì sự hình thành màu sắc không xảy ra. Mặc dù chi tiết chính xác của sự tổng hợp màu sắc ở cây đậu ngọt còn chưa rõ, nhưng có thể minh họa bằng mô hình tổng quát ở hình 2.7.



**Hình 2.7** Sơ đồ minh họa mối quan hệ giữa các gene trội C và P trong quá trình hình thành sắc tố anthocyanin ở cây đậu ngọt.

Đối với các kiểu gene có chứa cả hai gene trội C và P (C-P-), có đầy đủ các enzyme cần thiết cho việc tạo ra anthocyanin và do vậy hoa có màu đỏ tía. Trong khi đó, kiểu gene chứa cc (ccP- hay ccpp), enzyme thứ nhất không được tạo ra hay không có hoạt tính; hậu quả là, phản ứng tạo sản phẩm trung gian không thực hiện được. Tương tự, kiểu gene chứa pp (C-pp hoặc ccpp) thì phản ứng thứ hai biến đổi chất trung gian thành anthocyanin bị dừng lại, vì thiếu hẳn một enzyme tương ứng. Nói cách khác, nếu như kiểu gen có mang cặp allele hoặc là cc hoặc là pp thì con đường tổng hợp bị gián đoạn và các sắc tố không được tạo ra, và kết quả là hoa màu trắng.

### 1.2. Tương tác bổ trợ với tỷ lệ 9:6:1

*Ví dụ:* Sự di truyền hình dạng quả ở bí ngô. Khi lai hai giống bí ngô thuần chủng quả tròn khác nguồn gốc với nhau, ở F<sub>1</sub> xuất hiện toàn dạng quả dẹt, và ở F<sub>2</sub> có sự phân ly kiểu hình xấp xỉ 9 dẹt : 6 tròn : 1 dài.

*Giải thích:* Kết quả này có thể được giải thích theo một trong hai cách sau: bổ trợ giữa các gene trội hoặc bổ trợ giữa các gene lặn. Để đơn giản, dưới đây ta giải thích theo cách đầu, dựa trên quy ước sau đây:

*Quy ước:* D-F- : quả dẹt (do tương tác bổ trợ giữa các gene D và F)

D-ff và ddF- : quả tròn (chỉ có một trong hai gene trội D, F)

ddff : quả dài (do khuyết đồng thời cả hai gene trội)

*Kiểm chứng:*

P<sub>tc</sub> quả tròn-1 (DDff) × quả tròn-2 (ddFF)

F<sub>1</sub> quả dẹt (DdFf)

F<sub>1</sub> × F<sub>1</sub> = DdFf × DdFf = (Dd × Dd)(Ff × Ff)

→ F<sub>2</sub> = (3D-:1dd)(3F-:1ff) = 9 D-F- : (3 D-ff + 3 ddF-) : 1 ddff  
= 9 quả dẹt : 6 quả tròn : 1 quả dài

### 2. Tương tác át chế (epistasis)

Tương tác át chế là hiện tượng một gene này kìm hãm sự biểu hiện của một gene khác không allele với nó. Gene át chế có thể là trội hoặc lặn.

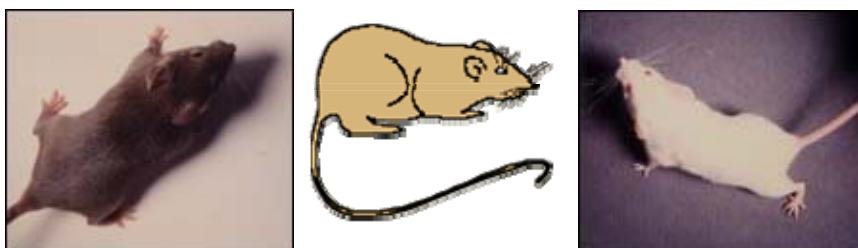


### 2.1. Át chế do gene lặn với tỷ lệ 9:3:4

*Ví dụ:* Sự di truyền màu sắc lông ở chuột (hình 2.8). Khi lai giữa hai dòng chuột thuần chủng lông nâu và bạch tạng, ở  $F_1$  xuất hiện toàn chuột lông đen; và khi cho các chuột  $F_1$  tạp giao với nhau thì ở  $F_2$  có sự phân ly gần với tỷ lệ 9 đen : 3 nâu : 4 bạch tạng. Để giải thích kết quả này, ta dựa vào biện luận như ở ví dụ đầu tiên (mục III-1.1) và quy ước sau đây:

*Quy ước:* B-C- : đen; bbC- : nâu; B-cc và bbcc : bạch tạng

*Giải thích:* Từ quy ước này, ta thấy rằng allele c khi ở trạng thái đồng hợp (cc) ức chế sự biểu hiện của B- và bb khiến cho các kiểu B-cc và bbcc không có sắc tố (bạch tạng, *albino*). Đối với hai kiểu hình còn lại có thể giải thích theo một trong hai cách sau: (1) Allele C là đột biến trội, nên mất khả năng át chế và bản thân nó không tạo màu; khi đó allele trội B quy định màu đen là trội so với allele b quy định màu nâu khi nó ở trạng thái đồng hợp; kết quả là B-C- có kiểu hình lông đen và bbC- cho kiểu hình lông nâu. Như thế ở đây không xảy ra sự tương tác hỗ trợ giữa các gene trội B và C. Cách lý giải này rõ ràng là hợp lý. (2) Allele C quy định màu nâu và B là gene tạo màu, trong khi bb không có khả năng đó. Do đó khi gene trội C (không có khả năng át chế) đứng riêng sẽ cho màu nâu; còn đứng chung với gene trội B sẽ cho hiệu quả hỗ trợ trội với kiểu hình màu đen. Rõ ràng cách giải thích này chỉ hợp lý trên hình thức khi cho rằng allele trội C xác định màu nâu.



**Hình 2.8** Từ trái sang là các chuột agouti đen, agouti nâu và bạch tạng.

*Kiểm chứng:*

$P_{tc}$  Chuột nâu (bbCC) × Chuột bạch tạng (BBcc)

$F_1$  Tất cả chuột đen (BbCc)

$F_1 \times F_1 = BbCc \times BbCc = (Bb \times Bb)(Cc \times Cc)$

→  $F_2 = (3B-:1bb)(3C-:1cc) = 9 B-C-: 3 bbC- : (3 B-cc + 1 bbcc)$   
= 9 đen : 3 nâu : 4 bạch tạng

### 2.2. Át chế do gene trội với tỷ lệ 12:3:1

*Ví dụ:* Sự di truyền màu sắc lông ở ngựa. Khi lai giữa hai giống ngựa

thuần chủng lông xám và hung đỏ với nhau, ở  $F_1$  thu được toàn ngựa lông xám. Sau khi cho tạp giao các ngựa  $F_1$  thì ở  $F_2$  có tỷ lệ kiểu hình xấp xỉ 12 xám : 3 đen : 1 hung đỏ.

*Giải thích:* Để giải thích kết quả này, ta quy ước: A-B- và A-bb : xám; aaB-: đen; và aabb: hung đỏ. Theo đó, allele trội B quy định màu đen là trội so với b- hung đỏ. Và sở dĩ A-B- và A-bb cho kiểu hình lông xám trong khi chúng có chứa B và bb, bởi vì gene trội A át chế các gene không allele và đồng thời nó còn có khả năng tạo màu xám, còn allele a không có cả hai khả năng đó. Điều này được thể hiện qua sơ đồ lai sau đây:

$$\begin{aligned}
 P_{tc} & \quad \text{Ngựa xám (AABB)} \times \text{Ngựa hung đỏ (aabb)} \\
 F_1 & \quad \text{Tất cả ngựa xám (AaBb)} \\
 F_1 \times F_1 & = \text{AaBb} \times \text{AaBb} = (\text{Aa} \times \text{Aa})(\text{Bb} \times \text{Bb}) \\
 \rightarrow F_2 & = (3A-: 1aa)(3B-:1bb) = (9 \text{ A-B-} + 3 \text{ A-bb}) : 3 \text{ aaB-} : 1 \text{ aabb} \\
 & = 12 \text{ xám} : 3 \text{ đen} : 1 \text{ hung đỏ}
 \end{aligned}$$

### 2.3. Át chế do gene trội với tỷ lệ 13:3

*Ví dụ:* Sự di truyền màu sắc lông ở gà. Khi lai giữa giống gà thuần chủng, gà Leghorn trắng với gà Wyandotte trắng, ở  $F_1$  thu được toàn gà lông trắng. Sau khi cho tạp giao các cá thể  $F_1$ , ở  $F_2$  có tỷ lệ phân ly kiểu hình xấp xỉ 13 lông trắng (white): 3 lông có màu (colored).

*Giải thích:* Để giải thích kết quả này, ta quy ước: I-C-, I-cc và iicc : trắng; và iiC-: có màu. Theo đó, allele trội C - gene sản xuất chất tạo màu (*chromogen*) là trội so với allele c- không có khả năng tạo màu; và allele trội I (*inhibitor*) át chế gene không allele với nó và nó cũng không có khả năng tạo màu, còn allele i không có khả năng át chế lẫn tạo màu.

*Sơ đồ lai:*

$$\begin{aligned}
 P_{tc} & \quad \text{Gà Leghorn trắng (IICC)} \times \text{Gà Wyandotte trắng (iicc)} \\
 F_1 & \quad \text{Tất cả gà trắng (IiCc)} \\
 F_1 \times F_1 & = \text{IiCc} \times \text{IiCc} = (\text{Ii} \times \text{Ii})(\text{Cc} \times \text{Cc}) \\
 \rightarrow F_2 & = (3I-:1ii)(3C-:1cc) = (9 \text{ I-C-} + 3 \text{ I-cc} + 1 \text{ iicc}) : 3 \text{ iiC-} \\
 & = 13 \text{ trắng} : 3 \text{ có màu}
 \end{aligned}$$

## 3. Tương tác cộng gộp - sự di truyền đa gene và các tính trạng số lượng

### 3.1. Tương tác cộng gộp (*additive*)

Tương tác cộng gộp hay sự di truyền *đa gene* (polygenic) là hiện tượng di truyền đặc trưng của một số *tính trạng số lượng* (quantitative trait), trong đó các gene không allele tác động cùng hướng lên sự biểu hiện của một tính trạng. Mỗi allele (thường là trội) của các gene đa phân như

thể đóng góp một phần ngang nhau trong sự biểu hiện ra kiểu hình ở một mức độ nhất định. Như vậy, liều lượng các allele tăng dần trong các kiểu gene sẽ tạo ra một dãy biến dị kiểu hình liên tục trong quần thể.

*Vi dụ:* Các thí nghiệm nổi tiếng năm 1909 của nhà di truyền học Thụy Điển (Sweden) Herman Nilsson-Ehle về sự di truyền màu sắc hạt lúa mì (hạt ở đây tức là phôi nhũ - kernel). Khi lai giữa các giống lúa mì thuần chủng hạt đỏ với hạt trắng, ở  $F_1$  ông thu được toàn dạng trung gian có màu hồng; và tùy theo dạng hạt đỏ được sử dụng trong các thí nghiệm mà ở  $F_2$  sẽ có các tỷ lệ phân ly giữa hạt có màu với hạt không màu (trắng) là 3:1, 15:1 hay 63:1. Kết quả phân tích cho thấy chúng do 2-3 gene đa phân chi phối. Sau đây ta hãy xét trường hợp  $F_2$  với tỷ lệ 15 có màu : 1 không màu, hay cụ thể là 1 đỏ: 4 đỏ nhạt: 6 hồng: 4 hồng nhạt: 1 trắng.

*Giải thích:* Do  $F_2$  có 16 kiểu tổ hợp với tỷ lệ tương đương trong khi  $F_1$  đồng nhất kiểu gene, chứng tỏ  $F_1$  cho 4 loại giao tử với tỷ lệ ngang nhau nghĩa là dị hợp tử về hai cặp gene phân ly độc lập. Ở đây,  $F_1$  biểu hiện kiểu hình trung gian của hai bố mẹ và  $F_2$  xuất hiện một dãy biến dị liên tục cùng hướng. Điều đó chứng tỏ tính trạng này tuân theo quy luật tác động cộng gộp hay đa phân tích lũy.

*Quy ước:* Vì allele cho màu đỏ là trội hơn allele cho màu trắng và mức độ biểu hiện của các hạt có màu ở  $F_2$  tùy thuộc vào liều lượng các allele đỏ trong kiểu gene, nên người ta thường ký hiệu các gene không allele bằng các chỉ số 1,2...kèm theo các chữ cái in hoa (A) cho allele trội và chữ cái in thường (a) cho allele lặn, như sau:  $A_1, A_2$  - đỏ;  $a_1, a_2$  - trắng. Từ đây ta có thể dễ dàng xác định kiểu gen của  $F_1$  ( $A_1a_1A_2a_2$ ) và của bố mẹ P: đỏ ( $A_1A_1A_2A_2$ ) và trắng ( $a_1a_1a_2a_2$ ), và thiết lập sơ đồ lai ở hình 2.9.

$P_{tc}$	$A_1A_1A_2A_2$ (đỏ) $\times$ $a_1a_1a_2a_2$ (trắng)				
$F_1$	$A_1a_1A_2a_2$ (hồng)				
$F_2$					
Allele trội	4	3	2	1	0
Kiểu gen	$1A_1A_1A_2A_2$	$2A_1A_1A_2a_2$ $2A_1a_1A_2A_2$	$4A_1a_1A_2a_2$ $1A_1A_1a_2a_2$ $1a_1a_1A_2A_2$	$2A_1a_1a_2a_2$ $2a_1a_1A_2a_2$	$1a_1a_1a_2a_2$
Kiểu hình	đỏ	đỏ nhạt	hồng	hồng nhạt	trắng
Tỷ lệ	1/16	4/16	6/16	4/16	1/16

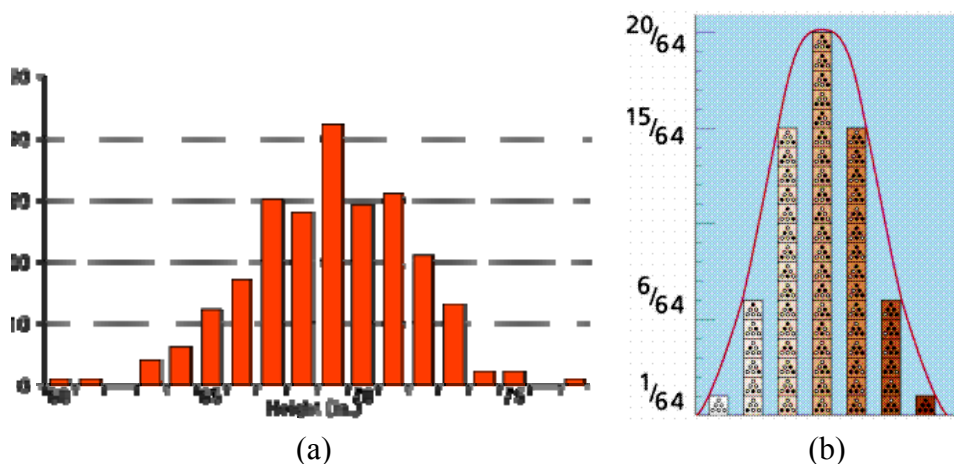
**Hình 2.9** Một phép lai của lúa mì đỏ và trắng do hai gene chi phối, cho thấy mối tương quan giữa tỷ lệ của các kiểu hình  $F_2$  và số allele trội.

Một ví dụ độc đáo khác là trường hợp di truyền số dãy hạt trên bắp ngô (xem trong *Di truyền học đại cương* - Dubinin 1981, tr.135-145).

*Nhận xét:* (1) Bằng cách vẽ một đồ thị biểu diễn mối quan hệ giữa số allele trội có mặt trong kiểu gen (trên trục hoành) và các tần số kiểu hình (trên trục tung) ở  $F_2$ , ta sẽ thu được một đường cong phân bố chuẩn có dạng hình chuông, gọi là phân phối Gauss. Trong đó kiểu hình trung gian hay các kiểu gene chứa hai allele trội (tương ứng với trị trung bình) chiếm tỷ lệ cao nhất, còn các kiểu hình hoặc kiểu gene ở hai đầu mút tương ứng với các ngưỡng "cực đoan" bao giờ cũng chiếm tần số thấp nhất (xem Hình 2.10). (2) Đối với các tính trạng di truyền theo kiểu đa phân công gộp, các hệ số của tỷ lệ kiểu hình có thể xác định bằng cách dựa vào tam giác Pascal.

### 3.2. Tính trạng số lượng (*quantitative trait*)

Thông thường, các tính trạng có trị số kiểu hình liên quan đến kích thước, trọng lượng hay hình dạng... được xác định dựa trên *thang định lượng* (*quantitative scale*), được gọi là các tính trạng số lượng.



**Hình 2.10** Mô hình đơn giản về (a) sự phân bố chiều cao ở người (đơn vị tính ở đây là inch, 1 in = 2,54 cm; trục tung chỉ số lượng); và (b) tương quan giữa các mức độ màu da (trục hoành) và các tỷ lệ trong quần thể (trục tung).

Nói chung, các tính trạng số lượng có các đặc điểm sau: Do nhiều gene quy định; chịu ảnh hưởng lớn của các điều kiện môi trường; và có sự phân bố kiểu hình liên tục trong một quần thể (hình 2.10), nhưng chúng cũng có thể xảy ra dưới dạng các lớp kiểu hình khác nhau, chẳng hạn như trong các ví dụ về dãy màu sắc hạt ở lúa mì hoặc số dãy hạt trên bắp ngô nói trên. Vì vậy, đối với các tính trạng này, không có một mối quan hệ chính xác giữa trị số kiểu hình và một kiểu gene cụ thể. Chẳng hạn, ở người, đó là các tính trạng về chiều cao, trọng lượng, hay *chỉ số thông minh* (*intelligence quotient - IQ*); ở cây lúa, lúa mì đó là số hạt trên mỗi bông,

số bông trên mỗi khóm...Tuy nhiên, trong những năm gần đây nhờ sử dụng các *chỉ thị phân tử* (molecular marker), người ta đã tiến hành lập bản đồ các gene có hiệu quả lớn lên các tính trạng đặc biệt (như các bệnh phức tạp ở người, năng suất cây trồng...) gọi là QTLs, tức các *locus tính trạng số lượng* (quantitative trait loci).

Để nghiên cứu sự biến đổi kiểu hình của các tính trạng số lượng, nhất thiết phải sử dụng các phương pháp của thống kê toán học. Chẳng hạn, phương pháp lấy mẫu sao cho hợp lý, nghĩa là mẫu phải đủ lớn, mang tính ngẫu nhiên và đại diện; việc xử lý số liệu đòi hỏi phải sử dụng một số đại lượng hay tham số thống kê cơ bản sau đây:

- *Trung bình cộng* (mean):  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

- *Biến lượng* (variance):  $v_x = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$

- *Độ lệch chuẩn* (standard deviation) bằng căn bậc hai của biến lượng (hay phương sai mẫu).

Một số đại lượng khác, như *hệ số biến thiên* (coefficient of variation, Cv%), *hệ số tương quan* (r) v.v... có thể xem trong các giáo trình *Di truyền học số lượng* và *Thống kê liên quan* (Trần Văn Diễn và Tô Cẩm Tú 1995; Falconer 1983; Phan Hiếu Hiền 2001).

### 3.3. Mô hình các tính trạng số lượng

Để hiểu và xác định được tầm quan trọng của các tính trạng số lượng, ta cần phải xây dựng một mô hình cho phép chia cắt các trị số kiểu hình thành ra các thành phần di truyền và môi trường. Điều này có thể thực hiện theo cách đơn giản bằng cách biểu thị trị số kiểu hình (P: phenotype) cho một kiểu gene (i) trong một môi trường cụ thể (j) như sau:

$$P_{ij} = G_i + E_j$$

trong đó  $G_i$  là phần đóng góp về mặt di truyền của kiểu gene (genotype) i vào kiểu hình, và  $E_j$  là độ sai lệch do môi trường (environment) j.  $E_j$  có thể âm hoặc dương tùy thuộc vào sự tác động của môi trường j. Cần lưu ý rằng, trong nhiều trường hợp, các quần thể khác nhau có các trị số trung bình khác nhau; và thật khó mà xác định sự khác nhau đó là do các nhân tố di truyền, nhân tố môi trường, hay là sự kết hợp của cả hai gây ra.

Các kiểu gene khác nhau có thể tương tác một cách khác nhau với môi trường của chúng để tạo ra kiểu hình. Nếu như các mối tương tác đặc thù như thế xảy ra giữa các kiểu gene và các môi trường, khi đó ta có thể mở rộng mô hình cơ sở nói trên thành mô hình *tương tác kiểu gene-môi*

*trường* (genotype-environment interaction), với trị số kiểu hình là:

$$P_{ij} = G_i + E_j + GE_{ij}$$

trong đó  $GE_{ij}$  là số đo sự tương tác giữa kiểu gene  $i$  và môi trường  $j$ . Vì môi trường biến động nên sự tương tác đó có thể âm hoặc dương.

Nói chung, mô hình di truyền có thể được định nghĩa một cách chính xác hơn bằng cách phân chia trị số kiểu gene ra hai thành phần như sau:

$$G_i = A_i + D_i$$

trong đó  $A_i$  và  $D_i$  là các thành phần cộng gộp (additive) và trội (dominant). Nếu như thể dị hợp đúng là trung gian giữa các thành phần đồng hợp (như ở ví dụ về sự di truyền màu sắc hạt lúa mì nói trên), thì sự đóng góp của thành phần trội là zero; nghĩa là, toàn bộ thành phần di truyền là cộng gộp. Nói cách khác, nếu cho rằng sự tương tác kiểu gene-môi trường là không đáng kể, thì trị số kiểu hình có thể được biểu thị bằng:

$$P_{ij} = A_i + D_i + E_j$$

Như đã nói từ đầu, không có sự tương ứng chính xác giữa một kiểu gene với một kiểu hình đối với các tính trạng số lượng. Vì vậy, cách thực tế hơn để kiểm tra các tính trạng này là tách sự biến đổi trong một quần thể đối với một kiểu hình cụ thể thành ra hai phần, phần biến đổi do nhân tố di truyền gây ra và phần biến đổi do nhân tố môi trường gây ra. Khi đó phương sai hay *biến lượng kiểu hình* (phenotypic variance) sẽ là:

$$V_P = V_G + V_E$$

trong đó  $V_P$ ,  $V_G$  và  $V_E$  tương ứng là các biến lượng của kiểu hình, di truyền và môi trường. Và theo cách làm ở trên, nếu ta tách các thành phần cộng gộp và trội để kiểm tra biến lượng kiểu hình, ta có:

$$V_P = (V_A + V_D) + V_E$$

trong đó  $V_A$  và  $V_D$  tương ứng là các biến lượng di truyền do các hiệu quả cộng gộp và trội. Nếu ta chia biểu thức này cho  $V_P$ , lúc đó ta thu được tỷ lệ của biến lượng kiểu hình do các thành phần khác nhau gây ra. Cụ thể là, tỷ số giữa biến lượng di truyền cộng gộp và biến lượng kiểu hình được coi là *hệ số di truyền* ( $h^2$ ; heritability):

$$h^2 = V_A/V_P$$

Đại lượng hệ số di truyền này rất quan trọng trong việc xác định tốc độ và hàm lượng đáp ứng đối với sự *chọn lọc định hướng* (directional selection), tức chọn lọc đào thải các kiểu hình cực đoan, và nó thường được các nhà chọn giống động vật và thực vật ước tính trước khi bắt đầu một chương trình chọn lọc nhằm cải thiện năng suất hoặc các tính trạng khác.

### 3.3. Ước lượng hệ số di truyền (*heritability*)

Bây giờ ta thử tìm hiểu vấn đề ước lượng hệ số di truyền dựa trên hệ số tương quan của một số tính trạng ở các *cặp sinh đôi cùng trứng* (monozygotic twins) và sinh đôi *khác trứng* (dizygotic). Ta biết rằng những cặp sinh đôi giống hệt nhau (cùng trứng) thì có cùng kiểu gene; như vậy bất kỳ sự sai khác nào giữa chúng phải là kết quả của các nhân tố môi trường. Trên thực tế, các cặp sinh đôi giống hệt nhau là do sự chia xẻ cùng kiểu gene cũng như các điều kiện môi trường như nhau. Một cách để xác định mức độ tương đồng hay giống nhau là tính *hệ số tương quan r* (correlation) giữa các cá thể về các tính trạng khác nhau. Trị số cao nhất của hệ số  $r$  là 1, trong đó tất cả các cặp cá thể đều có cùng trị số kiểu hình như nhau. Nếu như các tính trạng là không có tương quan giữa các cặp cá thể thì  $r = 0$ . Một kết quả nghiên cứu về các hệ số tương quan của bốn tính trạng khác nhau ở các cặp sinh đôi cùng trứng ( $r_M$ ) và các cặp sinh đôi khác trứng ( $r_D$ ) nhưng cùng giới tính được cho ở bảng 2.3. Những sai khác giữa các trị số này phải là kết quả của sự tương đồng di truyền thấp hơn ở các cặp sinh đôi khác trứng vốn chia xẻ (tính bình quân) chỉ một nửa các gen của chúng nói chung. Công thức để tính hệ số di truyền như sau:

$$h^2 = 2(r_M - r_D)$$

trong đó  $r_M$  và  $r_D$  lần lượt là các hệ số tương quan giữa các cặp sinh đôi cùng trứng và khác trứng. Bằng cách sử dụng công thức này, ta tính được các hệ số di truyền cho bốn tính trạng ở bảng 2.3 biến thiên từ 0,98 đến 0,16. Qua đó cho thấy, hầu hết sự biến đổi các chỉ số IQ và trưởng thành về mặt xã hội dường như là do môi trường, trong khi hầu như sự biến đổi trong số lượng nếp vân tay và chiều cao dường như là do di truyền.

**Bảng 2.3 Sự tương quan của bốn tính trạng giữa các cặp sinh đôi ở người**

Tính trạng	Hệ số tương quan		Hệ số di truyền
	$r_M$	$r_D$	
Số lượng nếp vân tay	0,96	0,47	0,98
Chiều cao	0,90	0,57	0,66
Điểm IQ	0,83	0,66	0,34
Điểm trưởng thành xã hội	0,97	0,89	0,16

## IV. Các mối quan hệ kiểu gene-kiểu hình

### 1. Thường biến và mức phản ứng

Đối với các tính trạng số lượng nói trên, sự biểu hiện ra kiểu hình của một kiểu gene nào đó rõ ràng là tùy thuộc của các nhân tố môi trường, tức là tùy thuộc tùy thuộc vào độ thâm nhập và độ biểu hiện của kiểu gene đó (như chúng ta sẽ xét dưới đây). Hay nói cách khác, kiểu hình là kết quả của sự tương tác giữa kiểu gene và môi trường.

Trong một tình huống lý tưởng, phạm vi tác dụng của môi trường lên một kiểu gene nào đó có thể xác định được bằng cách quan sát kiểu hình của các cá thể giống nhau về mặt di truyền được nuôi dưỡng chăm sóc trong các môi trường khác nhau. Ở người, các cặp sinh đôi cùng trứng được tách nuôi riêng lúc sơ sinh đã chứng minh cho một tình huống như vậy. Tuy nhiên, những chỉ dẫn tốt nhất về các tác dụng của môi trường bắt nguồn từ các nghiên cứu ở loài mà có thể nuôi trồng với số lượng lớn trên các môi trường được kiểm soát.

Thí dụ kinh điển cho nghiên cứu như vậy được J.Clausen và các cộng sự tiến hành trong thập niên 1940. Để xác định hiệu quả của môi trường lên các kiểu hình, các nhà nghiên cứu này đã trồng các *dòng* (clones), tức gồm các cá thể giống nhau về mặt di truyền, cắt theo các vĩ độ khác nhau. Chẳng hạn, họ đã thu thập loài thuộc họ hoa hồng là *Potentilla glandulosa* từ ba vùng: phạm vi ven bờ biển Thái Bình Dương ở vĩ độ 100 mét; vùng lưng chừng dãy núi Sierras ở vĩ độ 1.400 mét (ngọn Mather); và ở độ cao 3.040 mét của dãy Sierras (ngọn Timberline). Từ các mẫu này, họ tạo ra được các dòng bằng cách nhân giống sinh dưỡng và đem trồng ở ba vùng khác nhau đó. Nói khác đi, mỗi kiểu gene thu được từ một trong ba vùng đó được đưa trồng ở cả ba điều kiện môi trường (ứng với các vĩ độ) khác nhau. Từ thí nghiệm này cho thấy rằng dường như cả hai nhân tố di truyền và môi trường đều quan trọng trong việc xác định sự sinh trưởng của cây. Thí dụ, các kiểu gene vùng ven bờ sinh trưởng tốt ở cả vĩ độ thấp và trung bình nhưng không thể sống được ở vĩ độ cao. Điều đó chỉ ra tầm quan trọng của các nhân tố di truyền.

Cách thức mà kiểu hình thuộc một kiểu gene nào đó thay đổi theo môi trường của nó được gọi là *mức phản ứng* (norm of reaction). Các mức phản ứng có thể sai khác rất lớn trong số các kiểu gene đối với các tính trạng liên quan đến tập tính hay hành vi (behaviour). Dù cho đến nay ta không hề xem tập tính hay hành vi như là một kiểu hình, thế nhưng một số các gene đột biến có ảnh hưởng lên tập tính hay hành vi. Chẳng hạn, các cá thể bị bệnh PKU không được điều trị thì có các chỉ số IQ thấp hơn bình thường; IQ là một tính trạng có thể đo được bằng trắc nghiệm IQ.

Tóm lại, mối quan hệ giữa kiểu gene và kiểu hình là phức tạp, không dễ gì căn cứ vào kiểu hình để xác định kiểu gene. Mức phản ứng trên các môi trường khác nhau là một cách để định lượng mối quan hệ đó.

## 2. Độ thâm nhập (penetrance) và độ biểu hiện (expressivity)

Để đánh giá mức độ thể hiện ra kiểu hình của các kiểu gene như thế, năm 1925 Timofeev-Rissofsky đã nêu lên hai khái niệm: độ thâm nhập và độ biểu hiện của gene.



## 2.1. Độ thâm nhập (*penetrance*)

Trong các thí dụ đã xét trước đây, như bảy gene xác định bảy tính trạng ở đậu Hà Lan, các kiểu nhóm máu và nhiều bệnh di truyền ở người đều cho thấy một kiểu tuyệt đối chắc chắn. Tuy nhiên, đối với một số gene, thì một kiểu gene nào đó có thể hoặc không thể cho thấy một kiểu hình nào đó. Hiện tượng này được gọi là *độ thâm nhập* (*penetrance*) của một gene. *Mức độ thâm nhập* (level of penetrance) này có thể tính bằng tỷ lệ của các cá thể mang một kiểu gene nào đó bộc lộ ra một kiểu hình cụ thể. Khi tất cả các cá thể của một kiểu gene cụ thể có cùng kiểu hình như nhau, thì gene đó cho thấy độ thâm nhập hoàn toàn và mức độ thâm nhập này được coi là bằng 1. Mặt khác, gene được coi là thâm nhập không hoàn toàn và mức thâm nhập cụ thể có thể tính được. Chẳng hạn, trong số 160 cá thể thuộc một kiểu gene nào đó có 30 cá thể biểu hiện một kiểu hình bị bệnh, thì mức độ thâm nhập là  $30/160 = 0,1875$ ; nghĩa là gene đó chỉ thể hiện ra được kiểu hình có 18,75%.

Như vậy sự có mặt của độ thâm nhập không hoàn toàn tại một gene có thể khiến cho một kiểu hình tính trạng trội (giả sử kiểu gene là Aa) bị ngắt quãng một thế hệ trong một phả hệ, rồi lại xuất hiện trở lại ở thế hệ tiếp theo. Đó là do allele trội thâm nhập không hoàn toàn. Chẳng hạn, dạng trội của nguyên bào lưới (*retinoblastoma*), một bệnh gây các u ác tính ở mắt trẻ em, chỉ thâm nhập khoảng 90%.

## 2.2. Độ biểu hiện (*expressivity*)

Ngoài ra, một kiểu gen cụ thể nào đó bộc lộ ra kiểu hình kỳ vọng, tức *mức độ biểu hiện* (level of expression) hay *độ biểu hiện* (*expressivity*), có thể sai khác đi. Chẳng hạn, mặc dù một gene gây ra một bệnh có thể phát hiện được ở hầu hết các cá thể mang một kiểu gene nào đó, nhưng một số cá thể có thể bị ảnh hưởng nặng hơn các cá thể khác.

Bệnh Huntington ở người là một tính trạng cho thấy độ biểu hiện biến thiên sai khác theo độ tuổi phát bệnh. Bệnh này nói chung xảy ra ở độ tuổi bắt đầu trung niên, với tuổi trung bình bắt đầu phát bệnh là trên 30 tuổi, nhưng đôi khi có thể xảy ra rất sớm ở tuổi chưa lên 10 hoặc ở một người già hơn. Một số người mang kiểu gene bệnh này chết vì những nguyên nhân khác trước khi bệnh biểu hiện, do độ thâm nhập không hoàn toàn cũng như độ biểu hiện biến thiên đối với gene đó gây ra.

## 2.3. Các nhân tố môi trường và các gene khác

### 2.3.1. Môi trường và các gene gây chết có điều kiện

Vấn đề đặt ra là cái gì là nguyên nhân của độ thâm nhập không hoàn toàn và độ biểu hiện sai khác? Trong một số trường hợp cụ thể, thường thì

không thể xác định được nguyên nhân, nhưng cả môi trường và các gen khác cũng được coi là có ảnh hưởng lên độ thâm nhập và độ biểu hiện của một gene. Ví dụ, ở ruồi giấm, các allele của một số gene tác động lên sức sống và có thể gây chết ở nhiệt độ tăng cao trên 28°C nhưng lại ảnh hưởng chút ít hoặc không ảnh hưởng ở nhiệt độ thấp hơn. Các gene như thế gọi là các *gene gây chết có điều kiện* (conditional lethals); nói chung chúng cho thấy các hiệu quả gây chết trong các tình huống môi trường cực đoan.

Nhiệt độ cũng có thể ảnh hưởng mạnh mẽ lên kiểu hình của các thực vật. Chẳng hạn, cây anh thảo có hoa đỏ khi trồng ở 24°C nhưng lại có hoa trắng khi trồng ở nhiệt độ trên 32°C. Ở mèo Xiêm (Thái Lan) và chó Himalaya, màu sắc bộ lông của chúng tối sẫm hơn ở chân, tai và mũi tại vì thân nhiệt ở các bộ phận xa nhất của cơ thể này tỏ ra thấp hơn. Nếu như để chúng ở nhiệt



**Hình 2.11** Màu lông của mèo Xiêm tối sẫm ở các vùng chân, tai và mũi. độ ấm áp hơn, thì sự sinh trưởng lớp lông mới tại các điểm này sẽ cho màu sáng nhạt (hình 2.11).

### 2.3.2. Các gene sửa đổi (*modifier or modifying genes*)

Nói chung, mức độ biểu hiện của một tính trạng giữa các cá thể có quan hệ họ hàng thường giống nhau hơn là giữa các cá thể không có quan hệ họ hàng, khi cả hai nhóm cá thể đó được nuôi dưỡng chăm sóc trong những môi trường gần giống như nhau. Các gen có ảnh hưởng thứ cấp lên một tính trạng như thế gọi là các *gene sửa đổi* (modifier genes) và chúng có thể gây ảnh hưởng lên kiểu hình một cách đáng kể. Nói cách khác, các gene sửa đổi là những gene gây biến đổi các gene khác trong khi chúng biểu hiện ra kiểu hình.



**Hình 2.12** Thủy tinh thể bình thường (trái) và bệnh đục thủy tinh thể (phải) do một gen trội gây ra có tác động của gene sửa đổi.

Ở nhiều động vật, allele pha loãng (dilute allele) vốn làm giảm cường độ hình thành sắc tố (như từ màu đen sang màu xám ở chuột nhà) khi ở

trạng thái đồng hợp là một ví dụ về gene sửa đổi. Mức độ của hiện tượng không đuôi ở các mèo Manx biến thiên từ không có các đốt sống đuôi cho đến có một ít đốt sống đuôi dính liền nhau, hình như là do ảnh hưởng của các gene sửa đổi lên sự biểu hiện của thể đột biến trội này. Ở người, đó là trường hợp một gene trội gây đục thủy tinh thể (hình 2.12). Gene này gây ra những mức độ suy yếu thị lực khác nhau tùy thuộc vào sự có mặt của một allele đặc thù về một gene sửa đổi đi cùng.

### 2.3.3. Hiện tượng sao hình (*phenocopy*)

Một hiện tượng liên quan gọi là *sao hình* (*phenocopy*), xảy ra khi các nhân tố môi trường gây ra một kiểu hình cụ thể mà thông thường được xác định về mặt di truyền. Trong trường hợp này, một môi trường cực đoan có thể làm gián đoạn sự phát triển bình thường bằng một kiểu tựa như của một gene đột biến. Một trường hợp thảm thương xảy ra ở các nước Tây Âu vào cuối thập niên 1950 và đầu thập niên 1960 khi một số lượng các cháu bé sinh ra có tứ chi bị ngắn đi một cách kinh khủng. Hiện tượng bất thường này giống với một dạng rối loạn di truyền lặn. Chỉ sau khi có một cuộc nghiên cứu rộng rãi người ta mới hay rằng dạng khuyết tật đó được gây ra về mặt môi trường bởi sự bổ sung thalidomide vào bào chế viên thuốc ngủ mà các bà mẹ của các cháu bé bị bệnh đã uống trong suốt thời kỳ đầu mang thai.

## Câu hỏi và Bài tập

- Hãy chỉ ra các mối quan hệ giữa gene và tính trạng. Phân biệt các kiểu tương tác giữa các gene không allele lên sự hình thành một tính trạng.
- Đặc trưng của các tính trạng số lượng là gì? Phân biệt các cặp khái niệm sau: tính trạng số lượng và tính trạng chất lượng, độ thâm nhập và độ biểu hiện, tương tác giữa các gene allele và các gene không allele.
- Một phụ nữ nhóm máu M có một đứa con mang nhóm máu MN. Bà ta nói rằng một người đàn ông nào đó có nhóm máu MN là bố của nó, và ông ấy tuyên bố không phải như vậy. (a) Với thông tin đó, bạn có thể loại trừ người đàn ông này với tư cách người bố không? (b) Việc nghiên cứu xa hơn cho thấy người đàn bà này mang nhóm máu B, đứa con máu O, và người đàn ông kia nhóm máu AB. Với thông tin này, có thể gạt người đàn ông kia ra khỏi tư cách người bố hay không? Tại sao?
- Ở cú mèo, các phép lai giữa các cá thể lông đỏ và bạc đôi khi sinh ra tất cả lông đỏ; đôi khi  $\frac{1}{2}$  đỏ:  $\frac{1}{2}$  bạc; và đôi lúc là  $\frac{1}{2}$  đỏ:  $\frac{1}{4}$  trắng:  $\frac{1}{4}$  bạc. Các phép lai giữa hai con lông đỏ sinh ra hoặc là toàn bộ đỏ; hoặc  $\frac{3}{4}$  đỏ:  $\frac{1}{4}$  trắng; hoặc  $\frac{3}{4}$  đỏ:  $\frac{1}{4}$  bạc. Hãy xác định phương thức di truyền của tính

trạng này và kiểu gene của các cá thể khác nhau tham gia vào các phép lai.

5. Ở thỏ, màu sắc lông do bốn allele của một gene đơn kiểm soát. Thứ tự trội của các allele như sau: C- xám > c<sup>a</sup> - sọc len > c<sup>b</sup> - Himalaya > c - bạch tạng. Ngoài ra, các kiểu gene c<sup>a</sup>c<sup>b</sup> và c<sup>a</sup>c đều cho màu xám nhạt.

Kết quả lai giữa một thỏ xám với ba con thỏ khác như sau:

<u>Phép lai</u>	<u>Đời con</u>
Xám × sọc len	6 xám, 5 xám nhạt
Xám × xám nhạt	8 xám, 3 xám nhạt, 4 himalaya
Xám × bạch tạng	9 xám, 8 himalaya

Hãy chỉ ra phương thức di truyền và kiểu gene của các thỏ bố mẹ.

6. Các màu sắc của một loài thực vật có thể là đỏ, hồng hoặc trắng. Từ kết quả dưới đây, hãy biện luận để xác định phương thức di truyền của tính trạng màu hoa và chỉ ra các kiểu gene của các giống đem lai.

<u>Phép lai</u>	<u>Đời con</u>
Đỏ -1 × Hồng	$\frac{2}{3}$ đỏ : $\frac{1}{3}$ hồng
Đỏ -1 × Trắng	$\frac{1}{2}$ đỏ : $\frac{1}{2}$ hồng
Đỏ -2 × Hồng	$\frac{1}{2}$ đỏ : $\frac{1}{4}$ hồng : $\frac{1}{4}$ trắng
Đỏ -3 × Hồng	Tất cả đỏ
Đỏ -3 × Trắng	Tất cả đỏ

7. Màu tóc người gồm có năm sắc độ: vàng, nâu nhạt, nâu vừa, nâu thẫm và đen. Các phép lai giữa các màu tóc khác nhau cho kết quả sau:

(1) vàng × vàng → tất cả vàng;

(2) đen × đen → tất cả đen;

(3) vàng × nâu vừa → tất cả nâu nhạt; hoặc  $\frac{1}{2}$  vàng :  $\frac{1}{2}$  nâu vừa;

(4) nâu vừa × nâu vừa → tất cả nâu vừa; hoặc  $\frac{1}{2}$  nâu thẫm :  $\frac{1}{2}$  nâu nhạt; hoặc  $\frac{1}{2}$  nâu vừa :  $\frac{1}{4}$  đen :  $\frac{1}{4}$  vàng.

Hãy xác định kiểu gene của năm màu tóc, và cho biết: (a) Nếu lai giữa nâu nhạt và nâu thẫm, tỷ lệ kỳ vọng kiểu hình sinh ra như thế nào?

8. Ở chuột, bất kỳ cá thể nào có cc là bạch tạng (albino). Hai cặp allele khác, A/a và B/b, khi có mặt allele C sẽ cho màu xám, đen, nâu vàng và nâu sẫm. Bốn dòng thuần chủng bạch tạng khác nhau về kiểu gene được cho giao phối với các con lông xám thuần chủng (AABBCC). Tất cả chuột các đời F<sub>1</sub> đều màu xám. Sau đó các chuột F<sub>1</sub> từ mỗi phép lai được cho giao phối giữa chúng với nhau, và kết quả các đời F<sub>2</sub> như sau đây. Hãy xác định kiểu gene mỗi dòng bạch tạng.

Dòng albino	Xám	Đen	Nâu vàng	Nâu sẫm	Bạch tạng
1	44	0	16	0	20
2	31	0	0	0	9
3	48	15	0	0	21
4	145	43	46	15	82

9. Ở một loài thực vật, khi lai giữa một dòng hoa đỏ tía thuần chủng và ba dòng hoa trắng thuần chủng khác nhau với kết quả cho như sau:

	Phép lai 1	Phép lai 2	Phép lai 3
P	Đỏ tía × trắng-1	Đỏ tía × trắng-2	Đỏ tía × trắng-3
F <sub>1</sub>	18 đỏ tía	17 đỏ tía	19 đỏ tía
F <sub>2</sub>	95 đỏ tía 31 trắng	80 đỏ tía 36 trắng 28 đỏ	54 đỏ tía 32 trắng 19 đỏ 17 xanh da trời 6 nâu

Hãy biện luận và xác định: (a) kiểu gene của bốn giống hoa bố mẹ; (b) tỷ lệ kiểu hình của đời con nếu F<sub>1</sub> của phép lai 2 được lai với trắng-2; (c) tỷ lệ kiểu hình của đời con nếu F<sub>1</sub> từ phép lai 3 được cho lai với trắng-3.

10. Khi lai hai giống ngô thuần chủng có 20 dãy hạt và 8 dãy hạt, tất cả bắp ngô F<sub>1</sub> đều có 14 dãy hạt, và ở F<sub>2</sub> có bảy kiểu hình khác nhau: 20, 18, 16, 14, 12, 10 và 8 (dãy hạt) với số lượng thống kê tương ứng là 40 : 245 : 594 : 793 : 603 : 236 : 38 (bắp ngô). Phương thức di truyền của tính trạng nói trên là gì? Giải thích, trình bày sơ đồ lai và minh họa kết quả F<sub>2</sub> bằng một đồ thị. Bạn thử rút ra quy luật chung về sự di truyền của loại tính trạng này và chỉ ra cách xác định tỷ lệ kiểu hình kỳ vọng ở F<sub>2</sub> đối với trường hợp một tính trạng được kiểm soát bởi n locus độc lập.

## Tài liệu Tham khảo

### Tiếng Việt

Trần Văn Diễn và Tô Cẩm Tú. 1995. *Di truyền số lượng* (Giáo trình cao học Nông nghiệp). NXB Nông Nghiệp, Hà Nội.

Dubin NP. 1981. *Di truyền học đại cương* (bản dịch của Trần đình Miên và Phan Cự Nhân). NXB Nông Nghiệp, Hà Nội.

Hutt FB. 1964. *Di truyền học động vật*. (Bản dịch của Phan Cự Nhân). NXB Khoa học và Kỹ thuật, Hà Nội, 1978.

Phan Hiếu Hiền. 2001. *Phương pháp bố trí thí nghiệm và xử lý số liệu (Thống kê thực nghiệm)*. NXB Nông Nghiệp Tp.Hồ Chí Minh.

Phan Cự Nhân (chủ biên), Nguyễn Minh Công, Đặng Hữu Lanh. 1999. *Di truyền học*. NXB Giáo Dục, Hà Nội.

### **Tiếng Anh**

Boon WH. 1984. *Genes and the IQ*. PG Publishing Pte Ltd, Singapore-Hong Kong-New Dehli.

Campbell NA, Reece JB. 2001. *Essential Biology*. Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc, San Francisco, CA.

Clegg CJ, Mackean DG. 2000. *Advanced Biology: Principles and Applications*. 2<sup>nd</sup> ed, John Murray Published Ltd, London.

Falconer DS. 1983. *Introduction to Quantitative Genetics*. 2<sup>nd</sup> ed., Longman, London and New York.

Hartl DL, Freifelder D, Snyder LA. 1988. *Basic Genetics*. Jones and Bartlett Publishers, Inc, Boston - Portola Valley.

Suzuki DT, Griffiths AJF, Miller JH, Lewontin RC. 1989. *An Introduction to Genetic Analysis*. 4<sup>th</sup> ed, W-H Freeman and Company, New York.

Weaver RF, Hedrick PW. 1997. *Genetics*. 3<sup>rd</sup> ed, McGraw-Hill Companies, Inc. Wm.C.Brown Publishers, Dubuque, IA.

Wellnitz WR. 1995. *Genetics: Problem Solving Guide*. 2<sup>nd</sup> ed., Wm.C. Brown Publishers. Dubuque, Iowa.

### **Một số trang web**

<http://www.biology.about.com/library/weekly/aa122602a.htm>

<http://gslc.genetics.utah.edu/units/basis/>

<http://mhhe.com/lewisgenetics5>

<http://www.goodnet.com/nee72478/enable/hotline.htm>

<http://www.cff.org>

<http://consensus.nih.gov>

### Chương 3

## Cơ sở Tế bào của sự Sinh sản, Di truyền và Biến dị

Như chúng ta đều biết, các khám phá của Mendel và những nghiên cứu về cấu trúc tế bào trong nửa cuối thế kỷ XIX vẫn còn chưa có sự gắn kết với nhau. Vào năm 1902, Walter Sutton (USA) và Theodor Boveri (Germany), dựa trên tập tính tương tự của các gene và *nhễm sắc thể* (chromosome) trong quá trình giảm phân đã gợi ý rằng các gene có lẽ nằm trên các nhiễm sắc thể. Sutton nhận định: "Cuối cùng, tôi xin lưu ý rằng xác suất kết hợp của các nhiễm sắc thể bố và mẹ theo các cặp và phân ly sau đó trong sự phân chia giảm nhiễm... tạo thành cơ sở vật lý cho các quy luật di truyền Mendel". Ngày nay, điều ấy quá hiển nhiên. Nhưng vào thời đó nhận định này hết sức quan trọng; nó tạo ra sự kết nối chặt chẽ giữa di truyền học và tế bào học, hình thành nên lĩnh vực mới gọi là di truyền học tế bào với sự ra đời của thuyết di truyền nhiễm sắc thể.

Trong chương này, chúng ta sẽ tìm hiểu các nhiễm sắc thể có hình thái đặc trưng như thế nào; bằng cách nào số lượng nhiễm sắc thể được duy trì ổn định trong quá trình nguyên phân, nhưng lại giảm đi một nửa trong quá trình giảm phân; và các kiểu biến đổi về cấu trúc và số lượng nhiễm sắc thể (đột biến nhiễm sắc thể).

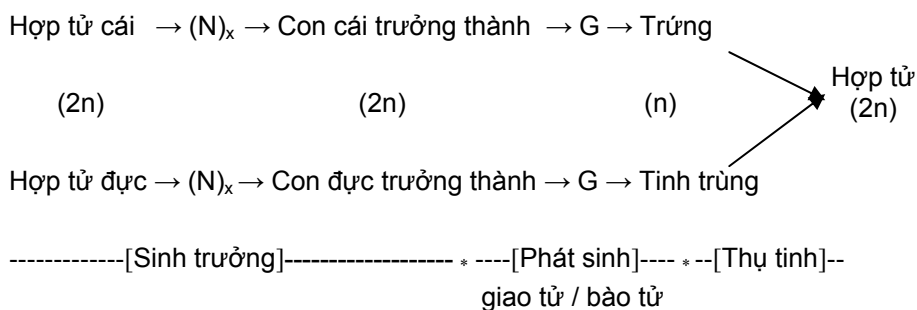
### I. Sinh sản hữu tính và tính ổn định của bộ nhiễm sắc thể

Trước tiên, ta đề cập chủ yếu phương thức sinh sản hữu tính ở eukaryote và mối liên quan giữa nó với sự ổn định về số lượng nhiễm sắc thể đặc trưng của mỗi loài. Thực ra, các eukaryote có hai kiểu sinh sản chính, vô tính và hữu tính.

Sự *sinh sản vô tính* (asexual reproduction) xảy ra khi một cá thể đơn độc tạo ra một cá thể mới giống nó; đây là phương thức sinh sản phổ biến ở thực vật và các động vật đơn giản. Sự *trinh sinh* (parthenogenesis) ở rệp cái chẳng hạn là một trường hợp đặc biệt, cũng sinh con nhưng không qua thụ tinh. Nói chung, con cái sinh ra bằng cách này thì giống với bố mẹ về mặt di truyền.

Sự *sinh sản hữu tính* (sexual reproduction) xảy ra khi các cá thể tạo ra các tế bào sinh dục đực và cái, hay các *giao tử* (gametes), đến lượt chúng kết hợp với nhau tạo thành một tế bào trứng được thụ tinh gọi là *hợp tử* (zygote), tức một tế bào hoàn chỉnh mà từ đó phát triển thành một cá thể mới. Hình thức sinh sản này xảy ra ở hầu như toàn bộ các kiểu sinh vật, kể

cả các động vật đơn giản nhất như con sum (*Balanus*) chẳng hạn, các thực vật, và thậm chí cả vi khuẩn. Ở vi khuẩn, có các kiểu trao đổi thông tin di truyền như tiếp hợp, biến nạp và tải nạp được gọi là *sinh sản cận tính* (parasexual; vấn đề này được trình bày riêng trong giáo trình *Di truyền Vi sinh vật và Ứng dụng*). Thông thường thì các giao tử đực và cái bắt nguồn từ các cá thể khác nhau, cho nên đời con sinh ra khác với bố mẹ chúng về nhiều chi tiết. Đây là nội dung chính mà chương này sẽ tập trung thảo luận. Còn sự tự thụ tinh được xem là trường hợp ngoại lệ quan trọng của sinh sản hữu tính (xem chương 12).



**Hình 3.1** Sơ đồ tổng quát về sinh trưởng và sinh sản ở sinh vật hữu tính.

Ở đây cho thấy số lượng nhiễm sắc thể lưỡng bội ( $2n$ ) và đơn bội ( $n$ ) tương ứng với các giai đoạn khác nhau (hàng dưới cùng). Ký hiệu  $(N)_x$  biểu thị nhiều lần nguyên phân, và  $G$  - giảm phân.

Sơ đồ tổng quát về sự sinh trưởng và sinh sản của sinh vật sinh sản hữu tính được trình bày ở hình 3.1. Trên nguyên tắc, mỗi hợp tử nhân được hai bộ nhiễm sắc thể *đơn bội* (haploid) ký hiệu là  $n$ , một từ giao tử đực và một từ giao tử cái; nên số lượng nhiễm sắc thể trong hợp tử là *lưỡng bội* (diploid), tức  $2n$  đặc trưng và ổn định cho loài. Mỗi bộ đơn bội chứa  $n$  nhiễm sắc thể khác nhau, mỗi chiếc hay kiểu nhiễm sắc thể chỉ có mặt một lần và chứa các gene khác nhau. Tập hợp toàn bộ các gene trong một bộ nhiễm sắc thể đơn bội như thế được gọi là *bộ gene* (genome). Như vậy, trong bộ lưỡng bội đặc trưng của các tế bào soma, các nhiễm sắc thể tồn tại theo từng cặp gồm hai chiếc giống nhau về hình dạng, kích thước và trật tự phân bố các gene - một có nguồn gốc từ bố và một từ mẹ - gọi là các *nhiễm sắc thể tương đồng* (homologous chromosomes).

Ở sinh vật đa bào, hợp tử tăng số lượng tế bào nhờ quá trình *nguyên phân* (mitosis), là kiểu phân chia tế bào tạo ra các tế bào con có số lượng nhiễm sắc thể  $2n$  được giữ nguyên. Khi cơ thể đạt tới độ thành thực sinh dục, một số tế bào của cơ quan sinh sản trải qua *giảm phân* (meiosis), tức kiểu phân chia tế bào tạo ra các giao tử có số lượng nhiễm sắc thể giảm đi



một nửa ( $n$ ). Quá trình này được gọi là *phát sinh giao tử* (gametogenesis) ở động vật và *phát sinh bào tử* (sporogenesis) ở thực vật. Sau đó, trong quá trình *thụ tinh* (fertilization), các giao tử cái hay *trứng* (egg) và giao tử đực hay *tinh trùng* (sperm) hợp nhất với nhau tạo thành các hợp tử đời con. Các hợp tử mới này lại bắt đầu đi vào một chu kỳ sinh trưởng và sinh sản y hệt như vậy.

**Bảng 3.1 Số lượng nhiễm sắc thể  $2n$  của các tế bào soma ở một số loài**

Loài động vật	$2n$	Loài thực vật	$2n$
Người ( <i>Homo sapiens</i> )	46	Đậu Hà Lan ( <i>Pisum sativum</i> )	14
Chimpanzee ( <i>Pan troglodites</i> )	48	Ngô ( <i>Zea mays</i> )	20
Bò ( <i>Bos taurus</i> )	60	Lúa gạo ( <i>Oryza sativa</i> )	24
Ngựa ( <i>Equus caballus</i> )	64	Lúa mạch ( <i>Secal cereale</i> )	14
Lừa ( <i>Equus asinus</i> )	62	Lúa mì ( <i>Triticum durum</i> )	28
Chó ( <i>Canis familiaris</i> )	78	Lúa mì ( <i>Triticum vulgare</i> )	42
Mèo ( <i>Felis catus</i> )	38	Khoai tây ( <i>Solanum tuberosum</i> )	48
Chuột nhắt ( <i>Mus domesticus</i> )	40	Thuốc lá ( <i>Nicotiana tabacum</i> )	48
Gà ( <i>Gallus domesticus</i> )	~78	Loa kèn ( <i>Lilium longiflorum</i> )	12
Ếch cựa ( <i>Rana pipiens</i> )	26	Mận ( <i>Prunus domestica</i> )	48
Cá chép ( <i>Cyprinus carpio</i> )	104	Bông ( <i>Gossypium hirsutum</i> )	52
Ruồi giấm ( <i>D. melanogaster</i> )	8	Hướng dương ( <i>Helianthus annuus</i> )	34

Mỗi loài có một số lượng nhiễm sắc thể đặc trưng; nghĩa là tất cả các cá thể của cùng một loài thì có số lượng nhiễm sắc thể như nhau. Số lượng nhiễm sắc thể này ở hầu hết các loài động-thực vật là lưỡng bội ( $2n$ ) được giới thiệu ở bảng 3.1, và ở một số loài là đơn bội ( $n$ ) - theo nghĩa có pha đơn bội là chính - như các nấm mốc bánh mì hồng (*Neurospora crassa*,  $n = 7$ ), mốc bánh mì đen (*Aspergillus nidulans*,  $n = 8$ ), nấm men bia (*Saccharomyces cerevisiae*,  $n = 17$ )...Rõ ràng là số lượng nhiễm sắc thể không tương ứng với nấc thang tiến hóa của các loài. Tuy nhiên, ta có thể thấy các loài có quan hệ họ hàng gần nhau như người và chimpanzee hay ngựa và lừa thì có số lượng nhiễm sắc thể gần bằng nhau. Cũng cần lưu ý rằng, ở một số loài thuộc bộ cánh màng Hymenoptera như ong mật (*Apis mellifera*) chẳng hạn, con cái là  $2n = 32$  và con đực là  $n = 16$ . Đối với các sinh vật này người ta dùng thuật ngữ là *đơn-lưỡng bội* (haplodiploid). Ở một số loài động-thực vật như ngô, lúa mạch đen... bên cạnh các nhiễm sắc thể chuẩn của bộ nhiễm sắc thể bình thường (gọi là các nhiễm sắc thể A), còn có thêm một vài nhiễm sắc thể rất bé vốn sai khác giữa các cá thể, gọi là các *nhiễm sắc thể thừa số* (supernumerary chromosomes) hay nhiễm

sắc thể *B*, để phân biệt với các nhiễm sắc thể *A*. Các nhiễm sắc thể này được coi là không chứa các gen quan trọng, mặc dù trong một vài trường hợp chúng ảnh hưởng lên độ hữu thụ. Ở đa số loài thuộc lớp chim (*Aves*), bộ nhiễm sắc thể chứa các nhiễm sắc thể rất nhỏ gọi là các *vi nhiễm sắc thể* (microchromosomes) và thường có số lượng lớn.

## II. Hình thái học nhiễm sắc thể eukaryote

Để tìm hiểu sâu hơn về cấu trúc và chức năng của các nhiễm sắc thể, ta cần phải phân biệt giữa các nhiễm sắc thể khác nhau về các điểm sau đây.

### 1. Kích thước nhiễm sắc thể

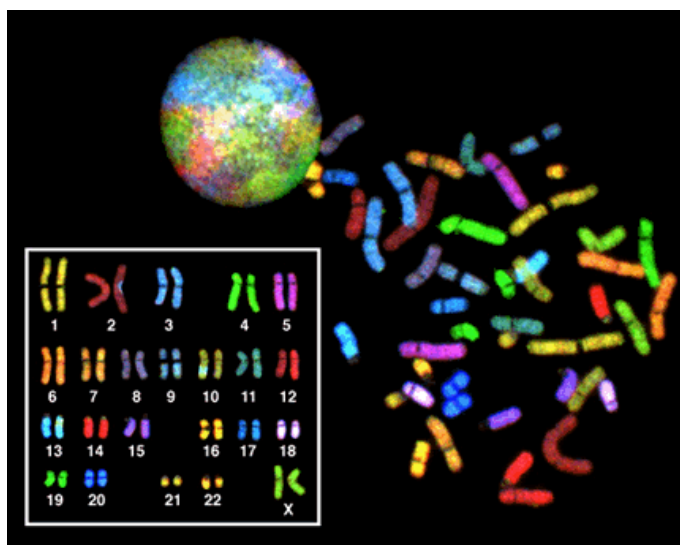
Kích thước các nhiễm sắc thể sai khác nhau rất lớn. Giữa các loài khác nhau sự sai khác này có thể lên tới hơn 100 lần, trong khi một số nhiễm sắc thể trong một loài có kích thước hơn kém nhau khoảng 10 lần.



**Hình 3.2** (a) Bộ nhiễm sắc thể được xử lý bằng thuốc nhuộm phổ biến Giemsa; và (b) xác định các đặc điểm của một nhiễm sắc thể.

Để mô tả bộ nhiễm sắc thể của một loài và xác định các nhiễm sắc thể của nó, cần phải thiết lập *kiểu nhân* (karyotype) và sử dụng một hệ thống các quy ước dựa trên kích thước của chúng, vị trí tâm động, và các kiểu băng (các hình 3.2 và 3.5). Theo truyền thống, người ta sử dụng kỹ thuật hiện băng (ví dụ băng G ở hình 3.2; xem thêm mục 3) và chụp ảnh các nhiễm sắc thể ở kỳ giữa của nguyên phân, gọi là phương pháp kỳ giữa. Sau đó phóng đại ảnh, cắt rời từng chiếc (gồm hai chromatid chị em) và sắp thành từng cặp tương đồng (vai ngắn quay lên trên và vai dài hướng xuống dưới), rồi xếp theo thứ tự nhỏ dần về kích thước và đánh số từ lớn nhất đến nhỏ nhất. Chẳng hạn, đối với người (xem các hình 3.3 và 3.5): (i) các nhiễm sắc thể không phải giới tính gọi là *nhiễm sắc thể thường* (autosome) được đánh số từ 1 đến 22 trên cơ sở chiều dài, còn các nhiễm sắc thể giới tính X và Y được tách riêng. (ii) Vai ngắn của nhiễm sắc thể được biểu thị bằng p và vai dài bằng q dựa trên vị trí của tâm động. (iii) mỗi vai trước tiên được chia thành các vùng và sau đó được xác định bằng các băng. Ví dụ, ký hiệu 9q34 có nghĩa là vai dài của nhiễm sắc thể số 9,

vùng 3, và băng 4. Đây là nơi khu trú của gene xác định hệ nhóm máu ABO.



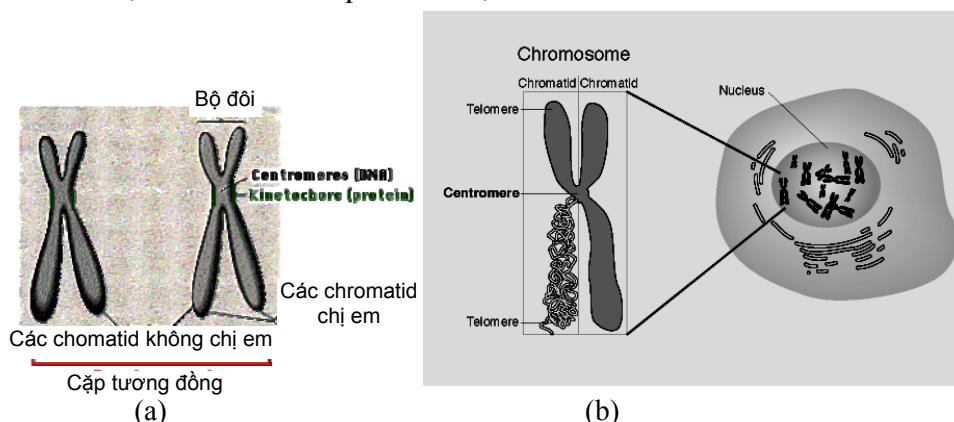
**Hình 3.3** Bộ nhiễm sắc thể người được thiết lập bằng phương pháp mới - kiểu nhân phổ (spectral karyotype). *Nguồn:* Schrock và cs (1996).

Gần đây, Schrock và cs (1996) đã giới thiệu các phương pháp thiết lập kiểu nhân mới sử dụng các thuốc nhuộm huỳnh quang có khả năng bám vào các vùng đặc thù của các nhiễm sắc thể, gọi là *thiết lập kiểu nhân phổ* (spectral karyotyping; Hình 3.3). Bằng cách sử dụng một loạt các vật dò (*probes*) đặc thù có hàm lượng các thuốc nhuộm khác biệt, các cặp nhiễm sắc thể khác nhau sẽ có các đặc trưng phổ riêng. Một đặc điểm độc đáo của công nghệ này là sự sử dụng một nhiễu kế (interferometer) giống như các nhà thiên văn dùng để đo quang phổ phát ra từ các vì sao. Công nghệ này cho phép phát hiện các biến đổi nhẹ về màu sắc (mà mắt thường chúng ta không thể phát hiện) bằng chương trình máy tính rồi sau đó phân loại từng cặp nhiễm sắc thể theo các màu sắc khác nhau. Việc sắp xếp các nhiễm sắc thể theo từng cặp rất đơn giản vì các cặp tương đồng thì có cùng màu như nhau, và các sai hình hoặc các trao đổi chéo có thể nhận biết được một cách dễ dàng hơn.

## 2. Tâm động và các kiểu nhiễm sắc thể

Các nhiễm sắc thể thường khác nhau về vị trí *tâm động* (centromere), là nơi mà các sợi thoi bám vào trong quá trình phân bào (hình 3.4). Thường thì mỗi nhiễm sắc thể chỉ có một eo thắt lớn chứa một tâm động gọi là *eo sơ cấp* (primary constriction). (Thuật ngữ *kinetochore* dùng để chỉ cấu trúc protein bao quanh vùng tâm động). Nói chung, có ba kiểu

nhiễm sắc thể chính: nếu tâm động nằm gần chính giữa, gọi là nhiễm sắc thể *tâm giữa* (metacentric); nếu tâm động nằm lệch về một đầu, gọi là nhiễm sắc thể *tâm đầu* (acrocentric); và nếu tâm động ở gần sát đầu mút, gọi là nhiễm sắc thể *tâm mút* (telocentric). Nói cách khác, tâm động chia nhiễm sắc thể thành hai vai hay cánh (arms) dài ngắn khác nhau; chẳng hạn, trong khi nhiễm sắc thể tâm giữa có hai vai gần như bằng nhau, thì nhiễm sắc thể tâm lệch có một vai ngắn và một vai dài. Các kết quả nghiên cứu cho thấy ở chuột nhà tất cả 40 cặp nhiễm sắc thể đều thuộc kiểu tâm giữa, trong khi ở bộ nhiễm sắc thể người gồm hai kiểu tâm giữa và tâm đầu, không có kiểu tâm mút (hình 3.3). Ngoài eo sơ cấp, trên một số nhiễm sắc thể cụ thể còn có thể có *eo thứ cấp* (secondary constriction), và nếu eo này ở gần đầu mút và xảy ra sự thắt sâu sẽ tạo nên một vệ tinh (satellite) của nhiễm sắc thể (ở người, đó là các nhiễm sắc thể 13, 14, 15, 21 và 22; xem hình 3.5b). Eo thứ cấp thường chứa các gene tổng hợp các RNA ribosome, tích tụ tạm thời và hình thành nên tổ chức gọi là hạch nhân (nucleolus); nó có mặt trong *nhân* (nucleus) ở một số giai đoạn khác nhau của sự phân bào. Hai đầu mút của mỗi nhiễm sắc thể eukaryote có cấu trúc đặc biệt gọi là *telomere* (sẽ được thảo luận ở chương 5) và thường có trình tự DNA khác với phần còn lại của nhiễm sắc thể.

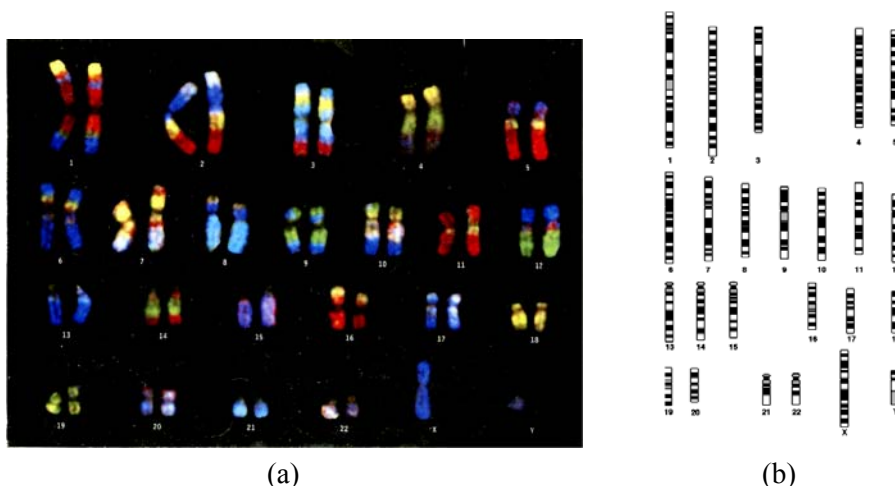


**Hình 3.4 (a)** Một cặp nhiễm sắc thể tương đồng kiểu tâm đầu ở kỳ giữa. **(b)** Mỗi chiếc lấy ra từ nhân kỳ giữa gồm hai chromatid chị em dính nhau ở tâm động, gọi là bộ đôi (dyad) và có các đầu mút đặc trưng gọi là telomere.

### 3. Các kiểu băng nhiễm sắc thể (chromosomal bands)

Người ta có thể xác định các vùng khác nhau trên các nhiễm sắc thể bằng các kỹ thuật hiện băng (band techniques) có xử lý các hóa chất khác nhau, như các băng G, Q và R (tương ứng là Giemsa, quinacrine và reverse bands- băng đảo ngược) (xem Verma và Babu 1995). Thực ra, thuật ngữ *chromosome* (= *chromo*: màu + *soma*: thể) theo nghĩa đen có

nghĩa là "thể bắt màu". Các kỹ thuật sử dụng các thuốc nhuộm khác nhau để xác định các vùng của nhiễm sắc thể nào đó làm hiện ra các *vạch, dải* hay *băng* (bands) có độ đậm nhạt, sáng tối hoặc hiển thị bằng các màu khác nhau. Chẳng hạn, các băng sẫm hơn thường xuất hiện ở gần tâm động hoặc ở các *đầu mút* (telomeres) của các nhiễm sắc thể, trong khi các vùng khác bắt màu yếu hơn nên nhạt hơn. Các vùng bắt màu đậm gọi là *chất dị nhiễm sắc* (heterochromatin) và các vùng bắt màu nhạt gọi là *chất đồng nhiễm sắc* (euchromatin). Nói chung, các vùng đặc trưng này được duy trì ổn định trong các tế bào hay các cá thể khác nhau của một loài và do đó chúng tỏ ra hữu ích cho việc xác định các nhiễm sắc thể cụ thể. Hình 3.5 trình bày kiểu nhân và *biểu đồ nhiễm sắc thể* (idogram) người.



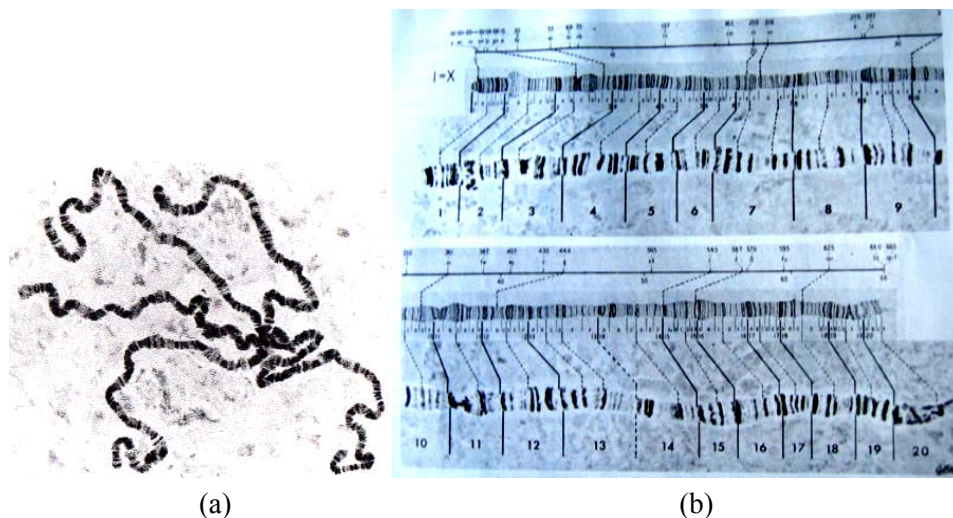
**Hình 3.5 (a) Kiểu nhân, và (b) biểu đồ nhiễm sắc thể người.**

*Nguồn:* © Current Opinion in Genetics & Development (1998)

Các vùng euchromatin thường trải qua một chu kỳ đóng và mở xoắn đều đặn, tương ứng với sự hoạt động của hầu hết các gene khu trú trên đó. Các vùng heterochromatin gồm hai kiểu: *ổn định* (constitutive) và *không ổn định* (facultative). Heterochromatin ổn định, tức là phần cố định của bộ gene và không biến đổi được thành euchromatin; chúng được coi là bất hoạt về mặt di truyền. Ngược lại, heterochromatin không ổn định bao gồm cả euchromatin là những vùng bắt giữ màu và cô đặc (theo nghĩa đóng xoắn chặt) đặc trưng của euchromatin trong một số giai đoạn phát triển; hay nói cách khác, đó thực chất là các vùng euchromatin nhưng chỉ biến đổi thành heterochromatin trong kỳ trung gian (interphase) ở một số mô.

Trên thực tế, người ta còn phát hiện một số dạng nhiễm sắc thể đặc biệt có kích thước khổng lồ, như: (i) các nhiễm sắc thể kiểu *chổi lông* hay *bàn chải đèn* (lampbrush chromosomes) bắt gặp trong noãn bào sơ cấp của

động vật có xương sống (có thể dài tới 1mm như ở lưỡng thê có đuôi); và (ii) các *nhiễm sắc thể đa sợi* (polytene chromosomes) có mặt trong nhân tế bào tuyến nước bọt của ấu trùng bộ hai cánh (*Diptera*).



**Hình 3.6** (a) Nhiễm sắc thể khổng lồ trong một nhân tế bào tuyến nước bọt của ấu trùng *D. melanogaster*, và (b) cấu trúc chi tiết của nhiễm sắc thể X được tách làm hai nửa, mỗi nửa có 10 vùng và mỗi vùng bắt đầu bằng một băng màu đậm.

Chẳng hạn, ở ruồi giấm *D. melanogaster*, các nhiễm sắc thể đa sợi này sinh ra do hiện tượng *nội nguyên phân* (endomitosis), tức các nhiễm sắc thể tự nhân đôi liên tiếp cả chục lần nhưng tế bào không phân chia và hầu như ở trạng thái mở xoắn. Vì vậy chúng có thể chứa tới cả ngàn sợi chromatid dính nhau trong một bó, với chiều dài hơn 100 lần so với các nhiễm sắc thể bình thường (hình 3.6). Và đã xác định được tất cả là 102 vùng, với hơn 5.000 băng. Riêng một mình nhiễm sắc thể X (tức nhiễm sắc thể số 1) có 20 vùng được đánh số từ 1 đến 20, với hơn 1.000 băng. Ở hình 3.7b cho thấy nó là nhiễm sắc thể tâm mút, có tâm động ở vùng 20 nằm sát đầu mút bên phải của nửa dưới, và gene gây ra mắt trắng nằm gần đầu mút bên trái ở vùng 3 thuộc nửa trên.

### III. Chu kỳ tế bào và nguyên phân

*Nguyên phân* (mitosis) là quá trình phân chia tế bào trong đó các tế bào con được tạo ra có số lượng nhiễm sắc thể giống với các tế bào bố mẹ. Kiểu phân bào này đặc trưng cho các tế bào soma, kể cả các tế bào sinh dục ( $2n$ ) ở pha sinh sản của sự phát sinh giao tử ở các động-thực vật (mục IV.2), và xảy ra theo cấp số nhân với công bội bằng 2, nghĩa là: từ một tế bào ban đầu trải qua  $k$  lần nguyên phân liên tiếp sẽ cho ra  $2^k$  tế bào giống

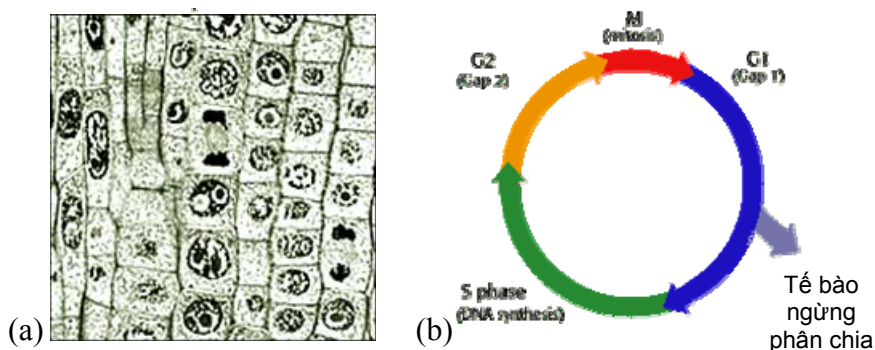


nó. Nhờ vậy mà cơ thể lớn lên và các tế bào trong cơ thể thường xuyên được đổi mới. Quy luật phân bào này được minh họa đơn giản như sau:

$$1 \rightarrow 2 \rightarrow 4 \rightarrow 8 \rightarrow 16 \rightarrow 32 \rightarrow 64 \rightarrow 128 \rightarrow 256 \rightarrow 512 \rightarrow \dots$$

### 1. Chu kỳ tế bào (cell cycle)

Quá trình nguyên phân lặp lại theo chu kỳ như vậy được gọi là chu kỳ nguyên phân hay *chu kỳ tế bào* (cell cycle). Nguyên phân là một phần của toàn bộ chu kỳ tế bào đối với các tế bào trải qua nguyên phân (như hợp tử, các tế bào phôi, các tế bào thuộc các mô sinh trưởng hay mô phân chia; hình 3.7a). Nói chung, một chu kỳ tế bào bao gồm hai giai đoạn chính là nguyên phân (ký hiệu: M), là một phần tương đối nhỏ của toàn bộ chu kỳ



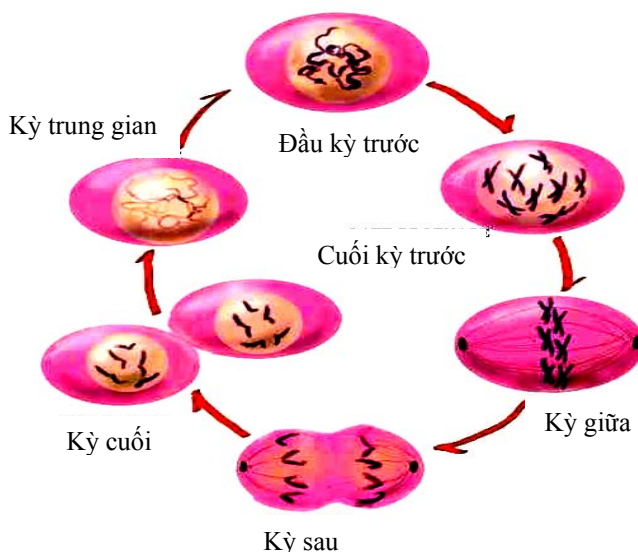
**Hình 3.7 (a) Sự phân chia của các tế bào chóp rễ hành tây (*Allium cepa*).**

**(b) Sơ đồ tổng quát của một chu kỳ tế bào.**

tế bào, và phần còn lại của chu kỳ tế bào gọi là *kỳ trung gian* (interphase). Gọi là kỳ trung gian bởi vì nó nằm giữa hai lần phân chia liên tiếp. Đây là giai đoạn tế bào diễn ra các hoạt động chuyển hóa cao độ, tổng hợp và tái bản vật chất di truyền - DNA, chuẩn bị tích cực cho tế bào bước vào nguyên phân. Nó được chia thành ba phần, gọi là G<sub>1</sub>, S và G<sub>2</sub>. Như vậy, theo nguyên tắc, một chu kỳ tế bào bao gồm bốn giai đoạn theo thứ tự sau đây (hình 3.7b): (1) G<sub>1</sub> (first gap) = giai đoạn khởi đầu trong đó tế bào sinh trưởng, chuyển hóa và chuẩn bị cho sự tái bản bộ gene; (2) S (DNA synthesis) = tổng hợp DNA (về chi tiết, xem chương 5); (3) G<sub>2</sub> (second gap) = chuẩn bị cho quá trình nguyên phân; và (4) M = *nguyên phân* (mitosis). Thời gian của các giai đoạn khác nhau trong chu kỳ tế bào khác nhau một cách đáng kể, tùy thuộc vào từng loài, từng kiểu tế bào, nhiệt độ và các nhân tố khác. Chẳng hạn, thời lượng tương ứng với bốn giai đoạn G<sub>1</sub>, S, G<sub>2</sub> và M đối với các tế bào máu trắng của người đang phân chia là 11, 7, 4 và 2 giờ (thời gian toàn bộ là 24 giờ).

Khi một hợp tử vừa được hình thành hay một cơ thể đang sinh trưởng,

chu kỳ này được lặp lại nhiều lần để hình thành nên một cá thể với hàng tỷ tế bào. Một số kiểu tế bào trưởng thành, như các tế bào thần kinh và tế bào cơ vẫn giữ nguyên ở kỳ trung gian, thực hiện các chức năng đã được biệt hóa trong cơ thể cho đến lúc chết và không bao giờ phân chia nữa; giai đoạn đó được gọi là pha  $G_0$ . Tuy nhiên, một số tế bào có thể từ pha  $G_0$  quay lại đi vào chu kỳ tế bào. Mặc dù hầu hết các tế bào lympho trong máu người ở pha  $G_0$ , nhưng nếu có sự kích thích thích hợp như khi bắt gặp kháng nguyên phù hợp chẳng hạn, chúng có thể bị kích thích để quay lại chu kỳ tế bào. Có thể nói,  $G_0$  không đơn thuần chỉ ra sự vắng mặt của các tín hiệu cho nguyên phân mà là một sự ức chế hoạt tính của các gene cần thiết cho nguyên phân. Các tế bào ung thư thì không thể đi vào pha  $G_0$  và được định trước để lặp lại chu kỳ tế bào một cách vô hạn (xem chương 5).



**Hình 3.8** Sơ đồ biểu diễn các kỳ của nguyên phân và chu kỳ của nó.

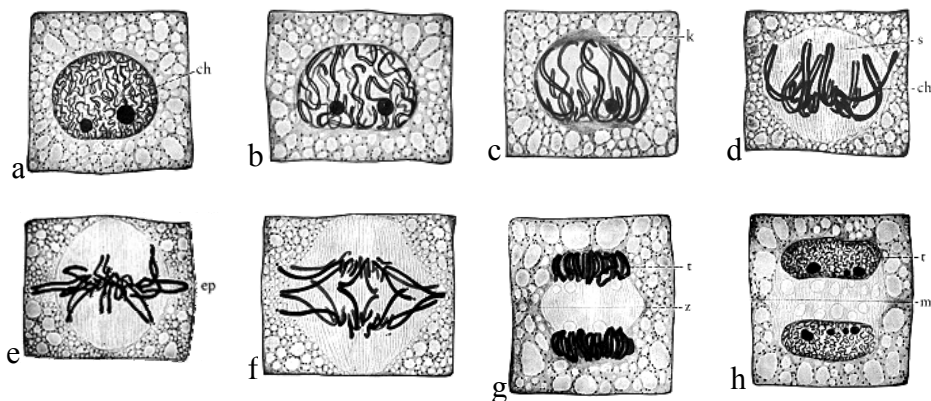
## 2. Nguyên phân (mitosis)

Nguyên phân tự nó có thể chia làm bốn giai đoạn khác nhau, diễn tiến theo một trình tự như sau: *kỳ trước* (prophase), *kỳ giữa* (metaphase), *kỳ sau* (anaphase) và *kỳ cuối* (telophase). Mỗi giai đoạn có một nét đặc trưng riêng, đặc biệt là mối liên quan với tập tính của nhiễm sắc thể, nhờ đó mà ta có thể xác định chúng (hình 3.8 và 3.9).

Sau khi tự nhân đôi ở kỳ trung gian (cụ thể là pha S) và hoàn tất việc chuẩn bị bước vào nguyên phân (pha  $G_2$ ), lúc này các nhiễm sắc thể tiếp tục đóng xoắn và kết đặc, nhờ vậy chúng hiện rõ dần dưới kính hiển vi quang học. Mỗi nhiễm sắc thể bây giờ gồm hai *chromatid chị em* (sister



chromatids) dính nhau ở tâm động. Theo nguyên tắc, các chromatid này hoàn toàn giống nhau do kết quả của sự tái bản bán bảo toàn DNA ở pha S (chương 5). Hơn nữa, vì hai chromatid chị em dính nhau tại vùng tâm động, nên chúng được xem là một nhiễm sắc thể.



**Hình 3.9** Các giai đoạn của quá trình nguyên phân ở một tế bào chóp rễ hành tây (*Allium cepa*).

### 2.1. Kỳ trước (*prophase*)

Cũng trong giai đoạn này, *hạch nhân* (nucleolus) thường biến mất và màng nhân bắt đầu tan vỡ. *Trung thể* (centriole) phân chia và hình thành xung quanh nó một cấu trúc mới gồm rất nhiều *sợi thoi* (spindle fiber) trải dài tới các cực của tế bào. Một số sợi thoi dính trực tiếp vào tâm động (cụ thể là *kinetochore*) của nhiễm sắc thể.

### 2.2. Kỳ giữa (*metaphase*)

Vào kỳ giữa, màng nhân tan biến hoàn toàn, các sợi thoi dính vào tâm động của các nhiễm sắc thể và đẩy chúng về *mặt phẳng xích đạo* (equatorial plane) của tế bào và xếp thành một vòng. Nói chung, lúc này các nhiễm sắc thể đóng xoắn cực đại (nghĩa là chiều dài rút ngắn tối đa và do đó đường kính cũng nở ra tối đa), với cấu trúc điển hình đặc trưng cho từng loài. Do đó, kỳ giữa là thời điểm thuận lợi nhất cho việc thiết lập các kiểu nhân và nghiên cứu hình thái học các nhiễm sắc thể như đã nói ở trên.

### 2.3. Kỳ sau (*anaphase*)

Vào đầu kỳ sau, tại mỗi nhiễm sắc thể xảy ra sự phân tách tâm động, các chromatid chị em bây giờ rời nhau và được gọi là *các nhiễm sắc thể con* (daughter chromosomes). Kế đó, các sợi thoi co rút và gây ra sự chuyển động của các nhiễm sắc thể con giống nhau về hai cực đối diện. Nếu nhìn dưới kính hiển vi lúc này, ta thấy nhiễm sắc thể xuất hiện dưới

dạng chữ V, J hoặc I, tùy theo kiểu tâm giữa, tâm đầu hay tâm mút.

Như vậy, chính sự sắp xếp thành một vòng của các nhiễm sắc thể ở kỳ giữa và sự phân ly đồng đều của chúng về hai cực ở kỳ sau làm thành bản chất hay là quy luật đặc trưng cho quá trình nguyên phân.

#### 2.4. Kỳ cuối (*telophase*)

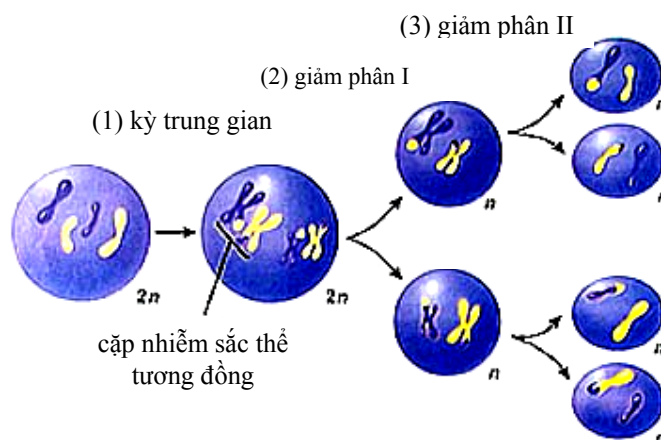
Vào kỳ cuối, hai bộ nhiễm sắc thể con đã về tới các cực đối diện và bắt đầu mở xoắn. Lúc này màng nhân xuất hiện trở lại và bao bọc các bộ nhiễm sắc thể; các sợi thoi tan biến, hạch nhân và các nhân được hình thành trở lại.

Kế đó, ở động vật, màng tế bào hình thành một *eo thắt* (*furrow*) từ ngoài vào trong; ở thực vật, một *phiến tế bào* (*cell plate*) phát triển từ trung tâm ra ngoài. Điều này làm phân cách hai bộ nhiễm sắc thể con và tế bào chất giữa hai tế bào con. Các tế bào con có số lượng nhiễm sắc thể giống với tế bào ban đầu. Như vậy, thực ra, nguyên phân gồm hai quá trình phân chia: *phân chia nhân* (*karyokinesis*) và *phân chia tế bào chất* (*cytokinesis*); nhưng thực chất của nguyên phân là sự phân chia nhân.

### IV. Giảm phân, sự phát sinh giao tử và thụ tinh

#### 1. Giảm phân (*meiosis*)

Giảm phân là kiểu phân bào đặc trưng cho các tế bào sinh dục, trong đó các tế bào con sinh ra (gọi chung là các giao tử) có số lượng nhiễm sắc thể giảm đi một nửa (hình 3.10).



**Hình 3.10** Kết quả của giảm phân với hai lần phân chia. Ở đây cho thấy hậu quả của sự tái tổ hợp và phân chia giảm nhiễm trong giảm phân I.

Giảm phân là một giai đoạn trong quá trình phát sinh giao tử, xảy ra ở pha trưởng thành (*mature*) sau khi bộ nhiễm sắc thể đã được nhân đôi ở kỳ trung gian thuộc pha sinh trưởng. Quá trình giảm phân gồm hai lần phân

chia nối tiếp nhau, giảm phân I và giảm phân II (hình 3.10). Mỗi lần phân chia này cũng được chia làm bốn kỳ. *Giảm phân I* (meiosis I) còn gọi là *phân chia giảm nhiễm* (reductional division), vì số lượng nhiễm sắc thể ( $2n$ ) giảm xuống còn đơn bội ( $n$ ). Trong giảm phân I, các chromatid chị em vẫn còn dính nhau trong khi các nhiễm sắc thể tương đồng phân ly. *Giảm phân II* (meiosis II) còn gọi là *phân chia đồng đều* (equational division) và rất giống với nguyên phân ở chỗ phân tách các chromatid chị em và số lượng nhiễm sắc thể giữ nguyên không đổi.

Bởi vì giảm phân là quá trình di truyền quan trọng và căn bản nhất ở cấp độ tế bào, cơ sở cho việc lý giải các quy luật di truyền và biến dị, cho nên trước khi đi vào mô tả giảm phân, cần nắm ba điểm chính sau đây: (1) Giảm phân cùng với sự thụ tinh sau đó cho phép duy trì số lượng nhiễm sắc thể ở các loài sinh sản hữu tính. (2) Giảm phân I cho phép các nhiễm sắc thể bố và mẹ khác nhau phân ly ngẫu nhiên về mỗi giao tử. (3) Sự trao đổi chéo giữa các chromatid trên các cặp nhiễm sắc thể tương đồng ở kỳ trước trong giảm phân I tạo ra các tổ hợp allele mới ở các gene khác nhau.

### 1.1. Giảm phân I (*meiosis I*)

#### 1.1.1. Kỳ trước I (*prophase I*)

Đây là pha phức tạp nhất của toàn bộ quá trình giảm phân, được chia thành năm giai đoạn khác nhau (hình 3.11a-e). Giai đoạn thứ nhất gọi là giai đoạn *leptotene* (sợi mảnh), đặc trưng bằng sự xuất hiện của các nhiễm sắc thể ở dạng các sợi mảnh khi nhìn dưới kính hiển vi quang học.

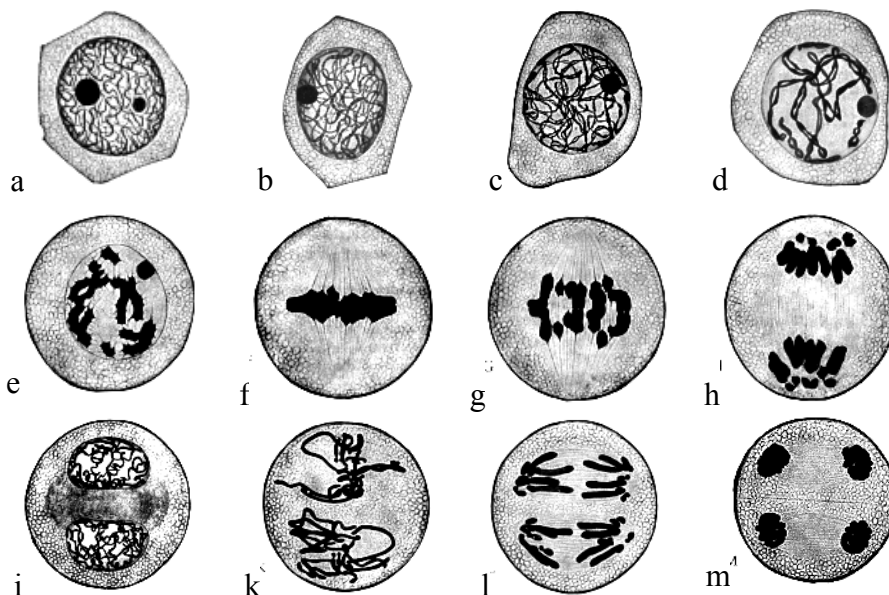
Kế tiếp là giai đoạn *zygotene* (sợi kết hợp), các nhiễm sắc thể tương đồng tiến lại gần nhau và các gene tương ứng đối diện nhau. Quá trình kết hợp của các nhiễm sắc thể tương đồng sau đó gọi là *tiếp hợp* (synapsis), là nét đặc trưng cơ bản phân biệt giữa giảm phân và nguyên phân. Cấu trúc gồm hai nhiễm sắc thể tương đồng chứa bốn chromatid kết cặp như vậy gọi là *thể lưỡng trị* (bivalent).

Giai đoạn thứ ba có tên là *pachytene* (sợi dày), vì lúc này các thể lưỡng trị ngắn lại và dày lên và sự tiếp hợp hoàn tất. Trong giai đoạn này có thể xảy ra hiện tượng *trao đổi chéo* (crossing over) hay *tái tổ hợp* (recombination) ở từng phần của mỗi cặp nhiễm sắc thể tương đồng.

Giai đoạn thứ tư gọi là *diploctene* (sợi kép), các nhiễm sắc thể tương đồng bắt đầu tách ra, đặc biệt là ở các vùng nằm hai bên tâm động. (Đĩ nhiên, các chromatid chị em còn dính nhau ở tâm động cho tới đầu kỳ sau II). Thông thường mỗi cặp nhiễm sắc thể tương đồng có thể có một hoặc hai vùng trong đó chúng vẫn còn kề sát nhau hoặc tiếp xúc với nhau, gọi là các *hình chéo* (chiasmata). Các hình chéo này (hình 3.12) là bằng chứng

vật lý cho sự tái tổ hợp xảy ra sau khi các nhiễm sắc thể tương đồng đã tiếp hợp. Nói chung, có ít nhất một hình chéo trên một vai nhiễm sắc thể, nhưng dọc theo các nhiễm sắc thể thì có thể có một số hình chéo.

Giai đoạn cuối của kỳ trước I, *diakinesis*, được đặc trưng bằng sự ngăn lại của các nhiễm sắc thể và chấm dứt các hình chéo. Điều đó có nghĩa là, các hình chéo bị đẩy về các đầu mút của các nhiễm sắc thể. Lúc này, hạch nhân và màng nhân cũng biến mất.



**Hình 3.11** Các giai đoạn của quá trình giảm phân ở hoa hành tây (*Allium cepa*). Kỳ trước I gồm 5 giai đoạn: (a) *leptotene*, (b) *zygotene*, (c) *pachytene*, (d) *diplotene*, và (e) *diakinesis*. Tiếp theo là (f) kỳ giữa I, (g) kỳ sau I, (h) kỳ cuối I, (i) kỳ nghỉ ngắn giữa hai lần phân chia (*interkinesis*), (k) kỳ giữa II, (l) kỳ sau II, và (m) kỳ cuối II.

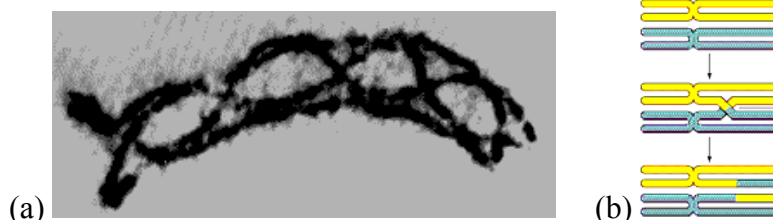
### 1.1.2. Kỳ giữa I (*metaphase I*)

Ở kỳ giữa I (hình 3.11f), các nhiễm sắc thể xếp trên mặt phẳng xích đạo thành hai vòng và các tâm động được đính vào các sợi thoi, sao cho cứ hai nhiễm sắc thể của mỗi cặp tương đồng (tức thể lưỡng trị) nằm đối diện nhau qua mặt phẳng kỳ giữa, với các hình chéo xếp thẳng hàng dọc theo nó. Sự kiện này khác biệt với kỳ giữa nguyên phân và là cơ sở cho sự phân chia giảm nhiễm và phân ly ngẫu nhiên của các nhiễm sắc thể ở kỳ sau I.

### 1.1.3. Kỳ sau I (*anaphase I*)

Vào lúc này, hai tâm động trong các nhiễm sắc thể tương đồng của

mỗi thể lưỡng trị đẩy nhau ra xa, sao cho một nhiễm sắc thể của mỗi cặp tương đồng đi về một cực (hình 3.11g). Khi các nhiễm sắc thể tương đồng đẩy nhau ra, các hình chéo hoàn toàn chấm dứt. Khác với kỳ sau nguyên phân, các chromatid chị em ở kỳ sau I vẫn còn dính nhau ở tâm động.



**Hình 3.12 (a)** Các hình chéo trên một cặp nhiễm sắc thể. **(b)** Sự hình thành các chromatid tái tổ hợp bởi một hình chéo từ hai nhiễm sắc thể tương đồng.

#### 1.1.4. Kỳ cuối I (*telophase I*)

Ở hầu hết sinh vật, sau khi các nhiễm sắc thể di chuyển tới các cực thì màng nhân hình thành xung quanh chúng và tế bào này phân chia thành hai tế bào con (hình 3.11h). Tuy nhiên, các chi tiết chính xác về mặt tế bào học của kỳ cuối I còn nhiều sai biệt, đặc biệt là ở các thực vật.

#### 1.2. Giảm phân II (*meiosis II*)

Giữa giảm phân I và giảm phân II thường chỉ có một kỳ trung gian ngắn ngủi gọi là *interkinesis* (hình 3.11i) Trong thời gian này không xảy ra sự tổng hợp DNA; mỗi tế bào chứa một bộ nhiễm sắc thể đơn bội ( $n$ ), trong đó mỗi nhiễm sắc thể chứa hai chromatid chị em. *Kỳ trước II* (*prophase II*) thường xảy ra rất nhanh và không rõ nét; các kỳ còn lại tương tự như trong nguyên phân. Thật vậy, ở *kỳ giữa II* (*metaphase II*), các tâm động dính vào các sợi thoi và di chuyển về mặt phẳng xích đạo (hình 3.11k). Vào đầu *kỳ sau II* (*anaphase II*), các nhiễm sắc thể tách nhau ở tâm động và sau đó các nhiễm sắc thể con phân ly về các cực đối diện (hình 3.11l). *Kỳ cuối II* (*telophase II*) bắt đầu khi tại mỗi cực có một bộ nhiễm sắc thể đơn bội đơn ( $n$ ) và màng nhân hình thành xung quanh chúng (hình 3.11m).

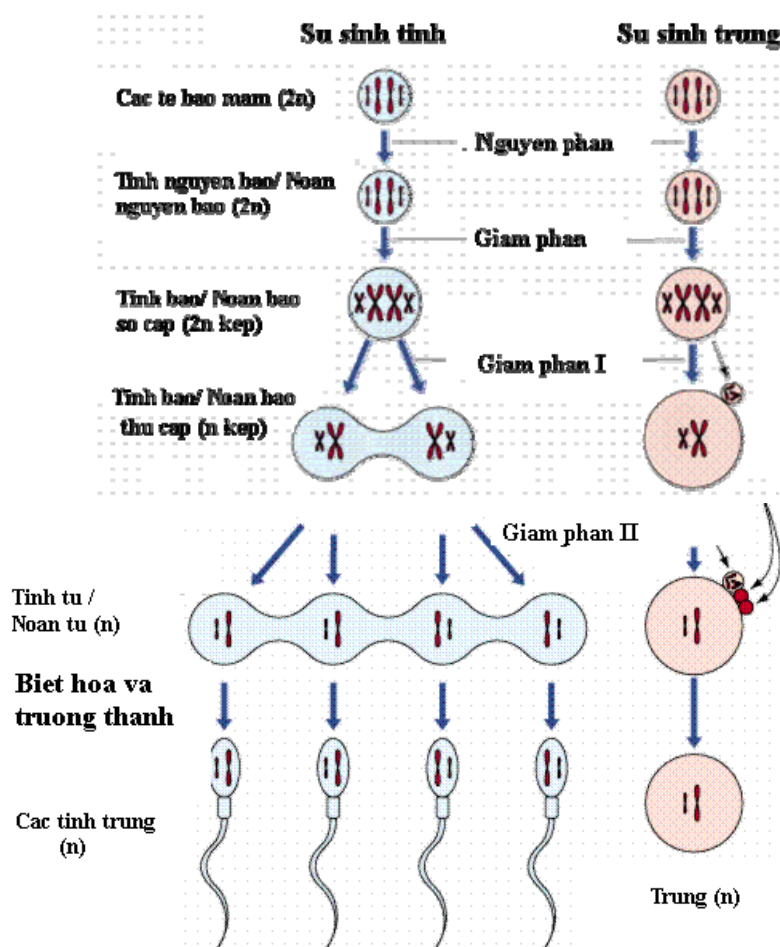
Như vậy, theo nguyên tắc, từ một tế bào sinh dục  $2n$  trải qua giảm phân cho ra bốn tế bào sinh dục đơn bội, gọi là các giao tử (hình 3.10).

#### 2. Sự phát sinh giao tử (*gametogenesis*)

Như đã biết, giảm phân là thời kỳ hay pha (phase) quan trọng nhất của quá trình phát sinh giao tử ở động vật và phát sinh bào tử ở thực vật. Dưới đây chỉ đề cập chủ yếu quá trình phát sinh giao tử ở động vật (hình 3.13).

Ở động vật, sự tạo thành các giao tử đực hay  *tinh trùng* (sperm), gọi là sự *sinh tinh* (spermatogenesis), xảy ra trong  *tinh hoàn* (teste) - cơ quan

sinh sản đực. Quá trình này bắt đầu với sự sinh trưởng của một tế bào lưỡng bội chưa biệt hóa gọi là *tế bào mẹ tinh trùng* (spermatogonium). Tại pha sinh sản, tế bào này thực hiện nhiều lần quá trình nguyên phân để gia tăng số lượng tế bào mẹ tinh trùng ( $2n$ ). Sau đó, mỗi tế bào này chuyển qua pha sinh trưởng, biệt hóa thành một *tinh bào sơ cấp* (primary spermatocyte;  $2n$  kép). Kế đó, tế bào này trải qua pha trưởng thành với hai lần phân chia liên tiếp của giảm phân. Sau giảm phân I, một tinh bào sơ cấp cho ra hai *tinh bào thứ cấp* đơn bội kép (secondary spermatocyte;  $n$  kép). Sau giảm phân II, mỗi tế bào này phân chia thành hai *tinh tử* đơn bội (spermatid;  $n$ ). Bước cuối cùng là sự biệt hóa của các tinh tử thành các tế bào *tinh trùng* (sperm), có cấu tạo đầy đủ các bộ phận (như đầu, cổ và đuôi dài; hình 3.13 và 3.14).



**Hình 3.13** Sơ đồ minh họa quá trình sinh tinh và sinh trứng ở động vật.

Sự *sinh trứng* (oogenesis) hay tạo các giao tử cái ở động vật xảy ra

trong các *buồng trứng* (ovary), cơ quan sinh sản cái. Quá trình này cũng bắt đầu bằng một *tế bào mẹ của trứng* (oogonium;  $2n$ ). Sau khi trải qua pha sinh sản, mỗi tế bào sinh noãn này biệt hóa thành *noãn bào sơ cấp* (primary oocyte;  $2n$  kép) ở pha sinh trưởng, với kích thước tăng trưởng một cách đặc biệt. Tại pha trưởng thành, sau giảm phân I, từ noãn bào sơ cấp tạo ra một *noãn bào thứ cấp* (secondary oocyte;  $n$  kép) và *thể cực* (polar body) thứ nhất. Sau lần giảm phân II, từ noãn bào thứ cấp cho ra một *noãn tử* (ootid;  $n$ ) và một thể cực, còn thể cực kia tạo ra hai thể cực thứ hai. Chung cuộc, từ một tế bào sinh noãn ( $2n$ ) cho ra chỉ một noãn ( $n$ ) kích thước rất lớn và ba thể cực ( $n$ ) kích thước rất nhỏ. Nguyên nhân là do giảm phân xảy ra gần màng tế bào, nên các thể cực chỉ nhận được một ít tế bào chất và chúng vẫn còn bám trên bề mặt của trứng cho đến lúc tiêu biến, tách ra. Sự tập trung tế bào chất trong một noãn tử để rồi sau đó biệt hóa thành *trứng* (ovum, egg) chính là nguồn cung cấp chất dinh dưỡng cho sự phát triển phôi sau khi thụ tinh.

Sự sinh tinh diễn ra liên tục ở các động vật trưởng thành sinh sản quanh năm, chẳng hạn như người và tùy theo mùa ở các động vật sinh sản theo mùa. Khả năng cho tinh trùng là cực kỳ cao ở hầu hết các động vật; một con đực (male) trưởng thành sản xuất hàng tỷ tinh trùng trong suốt quãng đời của chúng. Ngược lại, việc sinh trứng ở các con cái (female) nói chung là thấp hơn rất nhiều; chẳng hạn, một người nữ suốt đời chỉ có thể sản sinh chừng 500 trứng chín. Tất cả số trứng này phát triển thành các noãn bào sơ cấp trước khi con cái (female) được sinh ra và chúng dừng lại ở kỳ trước I cho tới tuổi dậy thì. Lúc này chúng tiếp tục giảm phân và bắt đầu chín lần lượt thành các trứng và mỗi tháng rụng một trứng. Tế bào trứng người có đường kính chừng  $0,1\text{ mm}$  ( $100\ \mu\text{m}$ ) trong khi đó phần rộng nhất của tinh trùng có đường kính là  $2,5\ \mu\text{m}$ .

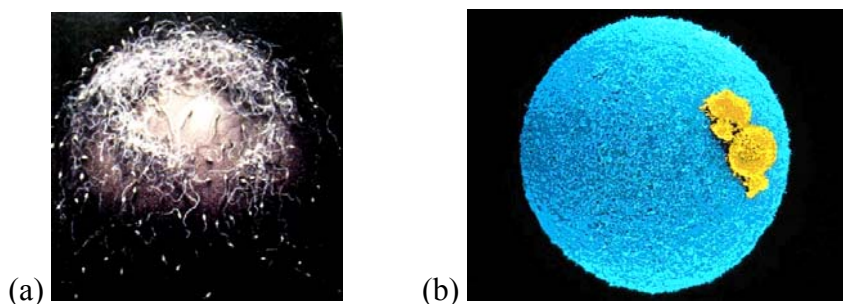
Ở thực vật bậc cao, sự hình thành các giao tử đực được gọi là sự hình thành hạt phấn hay *phát sinh tiểu bào tử* (microsporogenesis). Quá trình này giống như sự sinh tinh ở động vật ở chỗ giảm phân cho ra bốn tế bào đơn bội có cùng kích thước và tiềm năng chức năng. Sự hình thành noãn ở thực vật còn gọi là sự hình thành túi phôi hay *phát sinh đại bào tử* (megasporogenesis). Nó cũng tương tự như quá trình giảm phân ở động vật. Tuy nhiên, các sản phẩm sai khác nhau đáng kể.

### 3. Sự thụ tinh (fertilization)

Nói chung, thụ tinh (hình 3.14) là hiện tượng kết hợp ngẫu nhiên giữa các giao tử đực và giao tử cái mang bộ nhiễm sắc thể đơn bội ( $n$ ) tạo thành các *hợp tử* (zygotes) có bộ nhiễm sắc thể lưỡng bội ( $2n$ ) đặc trưng của loài. Thông thường, một trứng chỉ được thụ tinh bởi một tinh trùng và tạo

ra một hợp tử (hình 3.14b).

Ở các thực vật bậc cao có quá trình *thụ tinh kép* (double fertilization). Khi ống phấn xâm nhập *túi phôi* (embryo sac) và đưa hai tinh tử vào, sẽ xảy ra hai quá trình thụ tinh: một tinh tử sẽ kết hợp với *noãn* (egg) tạo thành hợp tử ( $2n$ ) và tinh tử kia sẽ dung hợp với *nhân nội nhũ* (endosperm nucleus;  $2n$ ) tạo thành một nhân nội nhũ ( $3n$ ). Sau đó, trong quá trình hình thành hạt, hợp tử sẽ phát triển thành phôi hạt và nhân nội nhũ sẽ phát triển thành nội nhũ nuôi phôi, gọi là phôi nhũ.



**Hình 3.14** (a) Các tinh trùng nhím biển (sea urchin) vây quanh một trứng nhím biển. (b) Một tế bào trứng người vừa được thụ tinh, các nhân của tinh trùng và trứng thể hiện bằng hai hình dạng không đều nhau, có màu vàng-nâu. Khi hai nhân này dung hợp thì sự thụ tinh coi như hoàn tất.

*Nguồn:* (a) Francis Leroy, Bicosmos/SPL/Photo Researchers, Inc.

(b) Phototake NYC/Dr. Nikas/Jason Burns © 1993-2003 Microsoft Corporation.

## V. Các biến đổi của nhiễm sắc thể

Trên đây chúng ta mới chỉ tìm hiểu các đặc tính di truyền ổn định của những cá thể mang bộ nhiễm sắc thể bình thường mà chưa đề cập đến mặt biến đổi của nó. Trên thực tế, các nhiễm sắc thể có thể bị biến đổi (đột biến) theo hai cách căn bản, đó là các biến đổi về cấu trúc và về số lượng (bảng 3.2). Các biến đổi này có thể gây những hậu quả khác nhau tùy loại, nhưng thường là nghiêm trọng. Chúng có thể gây hiệu quả tức thời lên các nhiễm sắc thể, gây rối loạn trong hoạt động của bộ gene và giảm phân; và từ đó ảnh hưởng tới sức sống, độ hữu thụ... của các cá thể mang chúng.

### 1. Các biến đổi về cấu trúc nhiễm sắc thể

Có bốn kiểu biến đổi cấu trúc nhiễm sắc thể, đó là: lặp đoạn (hay nhân đoạn), mất đoạn (hay khuyết đoạn), đảo đoạn và chuyển đoạn như đã giới thiệu ở bảng 3.2 (xem Hình 3.15). Các kiểu đột biến cấu trúc nhiễm sắc thể xảy ra là do sự đứt gãy và nối lại bất thường của các đoạn trên cùng một cặp tương đồng hoặc thậm chí giữa các cặp khác nhau như trong trường hợp chuyển đoạn. Nói chung, các đột biến này thường ít gặp, trung



binh cứ 1.000 giao tử mới hình thành có một vài đột biến liên quan kiểu nào đó. Các hậu quả về sau đều là kết quả của sự tiếp hợp gene-đôi-gene giữa các nhiễm sắc thể bình thường và bất thường (kiểu kết cặp không hoàn toàn tương xứng này được gọi là kiểu *dị nhân*, heterokaryotypic) cũng như các vấn đề sau này xảy ra ở các kỳ sau I và II trong giảm phân.

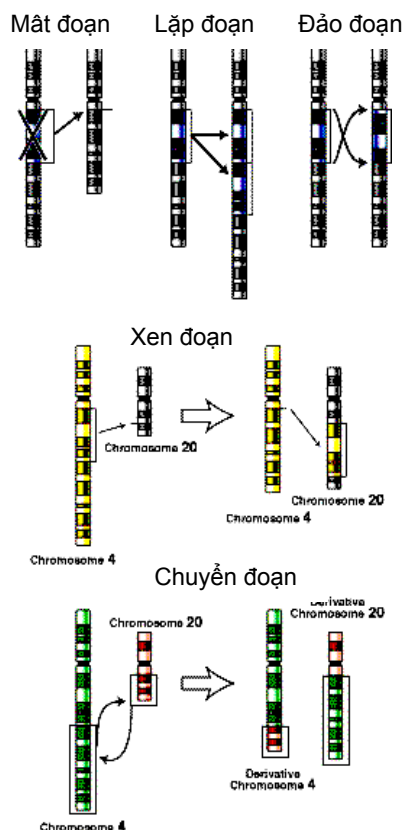
**Bảng 3.2 Các kiểu biến đổi (đột biến) nhiễm sắc thể**

Các kiểu biến đổi về cấu trúc	Các kiểu biến đổi về số lượng
Lặp đoạn (duplication)	Mức nguyên bội (Euploidy)
Mất đoạn (deletion)	Các thể tự đa bội (autopolyploids)
Đảo đoạn (inversion)	Các thể dị đa bội (allopolyploids)
Chuyển đoạn (translocation)	Lệch bội (Aneuploidy)
	Các thể một (monosomies; $2n - 1$ )
	Các thể ba (trisomies; $2n + 1$ )

### 1.1. Lặp đoạn (*duplication*)

Lặp đoạn hay nhân đoạn là trường hợp một đoạn nhiễm sắc thể nào đó có mặt hai lần trên một nhiễm sắc thể. Tùy theo vị trí và trật tự của các vùng lặp đoạn, ta có thể phân ra hai kiểu chính sau: (i) Lặp đoạn có thể nằm kề sát vùng nhiễm sắc thể ban đầu. Trong trường hợp trật tự các băng (hoặc các gene) vẫn giữ nguyên như ban đầu thì gọi là lặp đoạn *nối tiếp* (tandem), hoặc có trật tự ngược lại, gọi là lặp đoạn *đảo ngược* (reverse). (ii) Nếu như vùng lặp lại nằm cách xa đoạn gốc ban đầu thì gọi là lặp đoạn *chuyển chỗ* (displaced).

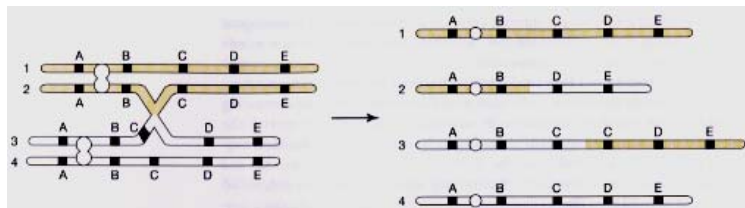
Có thể hình dung cách lặp đoạn nối tiếp xảy ra là do các nhiễm sắc thể tương đồng gối nhau và có các chỗ đứt đồng thời trên hai nhiễm sắc thể tại các điểm khác nhau. Nếu các nhiễm sắc thể tương đồng khác nhau này nối lại, thì một chiếc sẽ có một đoạn lặp nối tiếp và chiếc kia sẽ mất đi một đoạn tương ứng.



**Hình 3.15 Sơ đồ tổng quát các kiểu**

Kiểu *xen đoạn* ở đây là sự *chuyển đoạn giữa* (interstitial translocation).

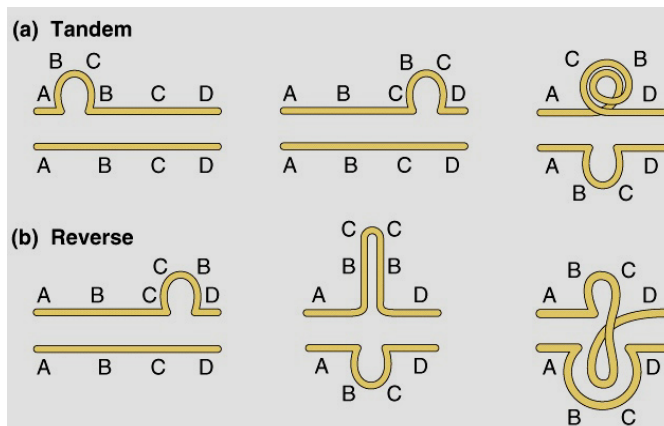
Trong trường hợp đó, lặp đoạn và mất đoạn là hai kiểu thuận nghịch của nhau. Điều đó cũng có thể do cơ chế *trao đổi chéo không đều* (unequal crossing over) gây ra (Hình 3.16).



**Hình 3.16** Lặp đoạn và mất đoạn do cơ chế trao đổi chéo không đều gây ra.

*Hậu quả là:*

(1) Về phương diện tế bào học, khi một cá thể là dị hợp kiểu nhân về một nhiễm sắc thể lặp đoạn và một nhiễm sắc thể bình thường thì vùng lặp đoạn đó không có đoạn tương đồng để kết cặp trong giảm phân. Lúc đó đoạn lặp sẽ phình ra dưới dạng một cấu trúc *hình vòng* (loop) sao cho các phần còn lại của cả hai nhiễm sắc thể có thể bắt cặp với nhau, tức tiếp hợp gene-đối-gene (Hình 3.17). Ở một số trường hợp, một vai nhiễm sắc thể có chứa đoạn lặp có thể uốn vòng lại sao cho hai đầu của đoạn lặp giao nhau.



**Hình 3.17** Sự kết cặp của các nhiễm sắc thể ở các cá thể có kiểu nhân dị hợp về (a) lặp đoạn nối tiếp - *tandem* và (b) lặp đoạn đảo ngược - *reverse*.

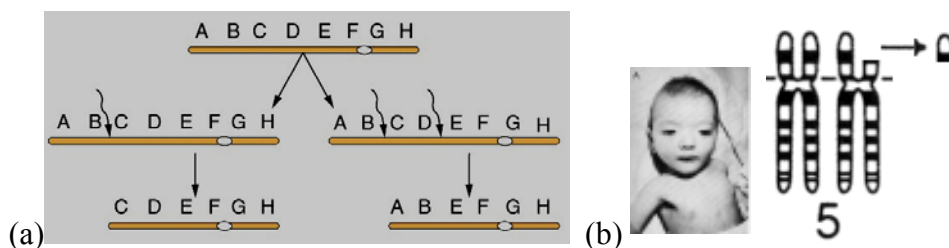
(2) Về phương diện kiểu hình, các cá thể dị hợp hoặc đồng hợp về các đoạn lặp bé vẫn có thể sống, mặc dù chúng thường bộc lộ ra một vài hiệu quả kiểu hình (như trong trường hợp các đột biến mất *Bar* ở ruồi giấm).

(3) Về phương diện tiến hóa, nếu như các cá thể mang các đột biến này vẫn sống bình thường, thì chúng sẽ có tiềm năng cho sự biến đổi tiến hóa xa hơn nữa ở các gene phụ thêm này, gọi là tiến hóa bằng cơ chế nhân đôi gene. Thực ra, có rất nhiều bằng chứng chứng tỏ sự nhân đôi gene là một

cơ chế quan trọng nhất để sinh ra các gen mới và các quá trình tiến hóa mới, tạo thuận lợi cho sự tiến hóa của các sinh vật phức tạp từ các sinh vật nguyên sơ. Chẳng hạn, các gene khác nhau mã hóa cho hemoglobin ở các động vật có xương sống có thể bắt nguồn từ một gene tổ tiên đã được nhân đôi, và sau đó các bản sao nhân đôi này phân ly về mặt chức năng của chúng (có thể xem thêm trong: Kimura 1983; Li 1983).

## 1.2. Mất đoạn (*deletion*)

Một nhiễm sắc thể bị khuyết mất đi một đoạn được gọi là đột biến *mất đoạn* (*deletion*) hay *khuyết đoạn* (*deficiency*) (hình 3.18). Mất đoạn có thể xảy ra tại một phần bên trong nhiễm sắc thể, gọi là mất đoạn giữa hay *mất đoạn khe* (*interstitial deletion*). Còn nếu như một đoạn nhiễm sắc thể bị đứt và không nối lại được, gọi là *mất đoạn mút* (*terminal deletion*). Trong trường hợp đoạn đứt gãy có chứa tâm động thì cả đoạn này lẫn nhiễm sắc thể mất đoạn thường sẽ bị dung giải trong quá trình phân bào. Cơ chế mất đoạn do sự trao đổi chéo không đều đã nói tới ở phần trên.



**Hình 3.18 (a) Cơ chế gây mất đoạn mút (trái) và đoạn khe; (b) Hội chứng *cri du chat* do mất đoạn mút trên vai ngắn nhiễm sắc thể số 5.**

*Hậu quả:* Khi các thể mất đoạn ở trạng thái đồng hợp, chúng thường bị chết do khuyết hẳn các gene thiết yếu. Thậm chí khi ở trạng thái dị hợp, các gene gây chết có thể gây ra sự phát triển bất thường. Một ví dụ nổi tiếng ở người đó là đột biến mất đoạn ở một phần đáng kể của vai ngắn trên nhiễm sắc thể số 5 (5p), khi ở trạng thái dị hợp gây ra hội chứng "mèo kêu" (tiếng Pháp: *cri du chat*; tiếng Anh: *cry-of-the-cat*). Những đứa trẻ mắc hội chứng này nói chung có giọng cao the the đặc trưng giống như tiếng mèo kêu, có đầu nhỏ và trí tuệ; và chúng thường bị chết sớm ở độ tuổi sơ sinh hoặc thơ ấu (Hình 3.18).

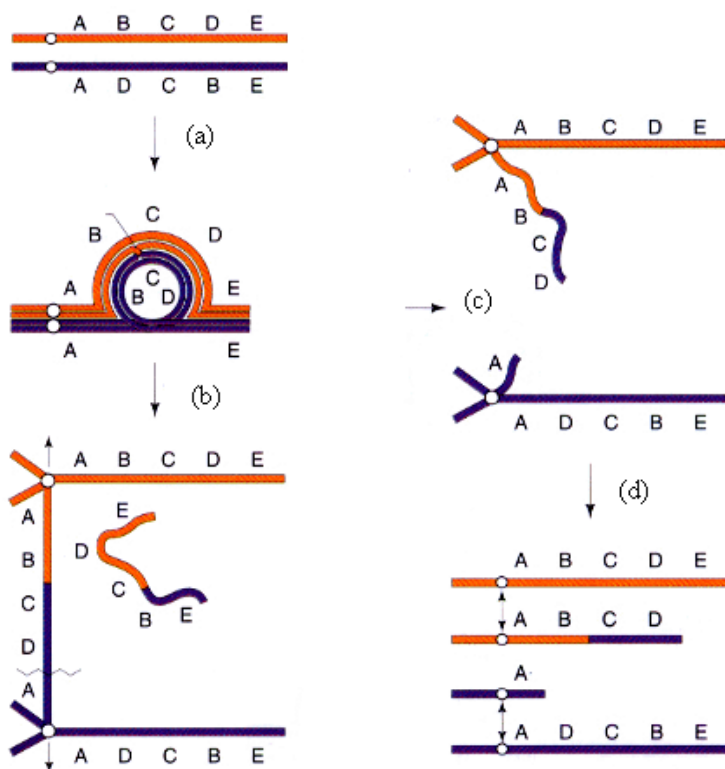
Ngoài ra, các thể dị hợp về mất đoạn thường cho thấy sự kết cặp nhiễm sắc thể bất thường trong giảm phân. Do nó không có vùng tương đồng để mà kết cặp với, nên trên nhiễm sắc thể bình thường sẽ hình thành một cái vòng gọi là *vòng mất đoạn* (*deletion loop*).

Tuy nhiên, một số đặc điểm khác lại tỏ ra hữu ích cho việc xác định các mất đoạn, chẳng hạn: (1) Không như các đột biến khác, các đột biến

mắt đoạn là không đảo ngược trở thành nhiễm sắc thể kiểu đại đực (tức dạng bình thường vốn có của tự nhiên). (2) Ở các thể dị hợp mắt đoạn, các allele lặn trên nhiễm sắc thể bình thường sẽ biểu hiện ra kiểu hình bởi vì nhiễm sắc thể mắt đoạn bị khuyết vùng tương ứng. Sự biểu hiện của các allele lặn trong trường hợp như vậy, gọi là hiện tượng *trội giả* (pseudo-dominance); nó có ích trong việc xác định chiều dài của đoạn bị mất.

### 1.3. Đảo đoạn (*inversion*)

Đảo đoạn là hiện tượng xảy ra do sự đứt đồng thời tại hai điểm trên một nhiễm sắc thể và sau đó đoạn bị đứt xoay  $180^\circ$  rồi nối lại. Hậu quả là, trật tự các gene trong đoạn đảo ngược lại với trật tự bình thường (hình 3.15). Tùy theo sự tương quan của đoạn đảo với vị trí tâm động, có thể chia làm hai kiểu đảo đoạn. Nếu đoạn đảo không chứa tâm động, gọi là *đảo đoạn cận tâm* (paracentric inversion); ngược lại, nếu đoạn đảo băng qua cả tâm động thì gọi là *đảo đoạn quanh tâm* (pericentric inversion).



**Hình 3.19** Sự trao đổi chéo xảy ra bên trong vòng của thể dị hợp đảo đoạn cận tâm tạo ra các giao tử chứa các khuyết đoạn lớn. (a) Kết cặp và xuất hiện vòng; (b) Phân ly làm xuất hiện cầu nối và đoạn không tâm; (c) Cầu nối hai tâm đứt gãy ngẫu nhiên; và (d) Các sản phẩm được tạo ra.

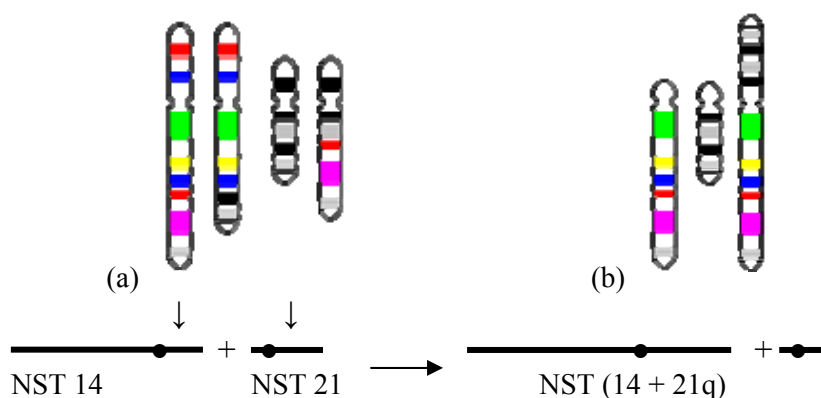
Về cơ chế, có thể hình dung như sau: vì lý do nào đó, một nhiễm sắc thể xoắn lại ở một vai (không liên quan tâm động) hoặc băng qua cả hai vai (tâm động nằm trên vòng cuộn), sau đó xảy ra sự đứt gãy tại vùng giao nhau rồi nối lại không đúng. Các cá thể dị hợp về một đoạn đảo có thể nhận diện bởi sự có mặt của các vòng đảo đảo (inversion loops) trên các nhiễm sắc thể thuộc giai đoạn sợi dày (pachytene) của giảm phân. Đó là do ái lực của hai nhiễm sắc thể tương đồng. Cách duy nhất mà hai nhiễm sắc thể tương đồng này cặp được với nhau là một chiếc tự nó xoắn lại thành một cấu trúc hình vòng trong khi chiếc kia làm thành một nửa vòng bao quanh nó mà không xoắn lại. Nếu như sự trao đổi chéo xảy ra trong vùng có đoạn đảo (Hình 3.19), có thể dẫn tới hình thành các giao tử chứa các đoạn lặp và khuyết đoạn lớn và thường gây chết ở giai đoạn hợp tử.

#### 1.4. Chuyển đoạn (*translocation*)

Chuyển đoạn là di chuyển của một đoạn nhiễm sắc thể từ một nhiễm sắc thể này sang một nhiễm sắc thể khác, không tương đồng với nó và gây ra sự cấu trúc lại nhiễm sắc thể đáng kể (hình 3.15, 3.20). Có hai kiểu chuyển đoạn: (i) *chuyển đoạn giữa* (interstitial translocation) là sự di chuyển một đoạn cho một nhiễm sắc thể khác mà không nhận lại, tức chuyển đoạn không thuận nghịch; và (ii) *chuyển đoạn thuận nghịch* (reciprocal translocation), tức có sự trao đổi qua lại hai chiều. Kiểu thứ hai phổ biến hơn. Nếu như hai đoạn được nối lại trong chuyển đoạn thuận nghịch tạo ra hai chiếc có kích thước lớn và hai chiếc kia bé, thì các nhiễm sắc thể được chuyển đoạn có kích thước bé hơn thường bị biến mất. Khi đó, số lượng nhiễm sắc thể bị giảm đi do sự trao đổi nhiễm sắc thể. Như thế, các chuyển đoạn có thể làm thay đổi cả kích thước lẫn vị trí tâm động.

Những cá thể dị hợp về chuyển đoạn chỉ sinh ra khoảng một nửa số con bình thường; hiện tượng đó được gọi là *bán bất dục* (semi-sterility). Khi các thể dị hợp chuyển đoạn trải qua giảm phân sẽ cho ra khoảng một nửa số giao tử chứa cả hai phần của chuyển đoạn thuận nghịch và khoảng một nửa kia chứa cả hai nhiễm sắc thể bình thường. Số giao tử còn lại chỉ chứa một phần của chuyển đoạn và một nhiễm sắc thể bình thường; các giao tử loại này sẽ gây chết hợp tử do mất cân bằng về vật chất di truyền.

Diễn hình là trường hợp *chuyển đoạn Robertson* (Robertsonian translocation), trong đó các vai dài của hai nhiễm sắc thể tâm đầu không tương đồng gắn với nhau và chỉ có một tâm động (hình 3.20b). Còn nhiễm sắc thể tâm đầu bị đứt gãy ở vai dài và đoạn đứt từ nhiễm sắc thể kia được nối lại thường biến mất trong quá trình phân bào hoặc sau vài thế hệ. Các chuyển đoạn Robertson có ý nghĩa quan trọng trong di truyền học người.



**Hình 3.20 (a) Chuyển đoạn thuận nghịch; (b) Chuyển đoạn Robertson, và sơ đồ bên dưới cho thấy sự đứt - nối của hai nhiễm sắc thể không tương đồng tạo ra kiểu chuyển đoạn này. Mũi tên (↓) chỉ các điểm đứt.**

Mặc dù các chuyển đoạn có thể tạo ra các nhiễm sắc thể bình thường, song chúng cũng có thể gây ra nhiều bệnh di truyền ở người. Chẳng hạn, khoảng 5% số người mắc *hội chứng Down* (Down syndrome) có một bố mẹ là dị hợp về một chuyển đoạn. Cụ thể trong trường hợp này, vai dài nhiễm sắc thể 21 (đoạn 21q) bị đứt ra và gắn vào chỗ đứt trên vai ngắn nhiễm sắc thể 14; người bố hoặc mẹ này ở trạng thái dị hợp với kiểu hình bình thường, nhưng có nguy cơ gây hội chứng Down ở đời con do sự phân ly sai hình xảy ra trong giảm phân. Theo nguyên tắc, khi giảm phân, thể dị hợp chuyển đoạn có thể tạo ra bốn loại giao tử, với xác suất ngang nhau, thuộc hai nhóm có bộ nhiễm sắc thể cân bằng và không cân bằng sau đây:

$$[14 ; 21] : [14 + 21q] : [14] : [(14 + 21q) ; 21]$$

Khi các giao tử này thụ tinh với các giao tử bình thường  $[14 ; 21]$  sẽ cho ra bốn kiểu hợp tử tương ứng là:

$$[(14)_2 ; (21)_2] : [(14;14 + 21q) + 21] : [(14)_2 + 21] : [(14;14 + 21q) + (21)_2]$$

$$\begin{array}{cccc} \downarrow & \downarrow & \downarrow & \downarrow \\ \text{Bình thường} & \text{Thể dị hợp chuyển đoạn} & \text{Thể một 21} & \text{Thể ba 21} \\ (46) & (45) & (45; \text{chết}) & (46) \end{array}$$

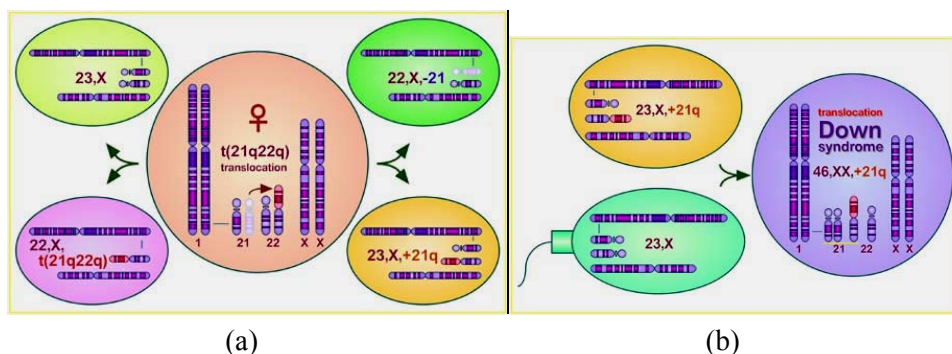
Bởi vì các cá thể nhận được chỉ một nhiễm sắc thể 21 sẽ bị chết, nên chung cuộc có khoảng 1/3 số trẻ sống (từ một thể dị hợp chuyển đoạn như vậy) được kỳ vọng là mắc hội chứng Down. Trên thực tế, tỷ lệ này thấp hơn 1/3, chủ yếu là do một số cá thể không sống sót được ở thời kỳ thai.

Cần lưu ý rằng, nguyên nhân này của hội chứng Down đưa ra các gợi ý cho *tư vấn di truyền* (genetic counseling) ở chỗ: (1) Hội chứng Down có thể tái diễn ở các con cái của một kiểu nhân dị hợp về chuyển đoạn, trong

khi đó hội chứng Down bình thường thì không tái diễn trong các anh chị em ruột vì vô sinh (xem mục 2.2 dưới đây). (2) Một nửa số anh chị em ruột bình thường về kiểu hình của các cá thể Down chính họ lại là những thể dị hợp về chuyển đoạn, và do đó có thể sinh ra con cái bị Down.

Ở nước ta, theo Nguyễn Văn Rục (2004), có khoảng 6% trường hợp hội chứng Down do các thể chuyển đoạn gây ra, bao gồm các dạng  $t(14q; 21q)$ ,  $t(21q; 21q)$  và  $t(21q; 22q)$ .

Trên hình 3.21 trình bày cơ chế hình thành các loại giao tử từ người nữ thuộc dạng chuyển đoạn  $t(21q; 22q)$  và sự thụ tinh giữa một trứng bất thường có nhiễm sắc thể 22 được chuyển đoạn 21q (ký hiệu là:  $23,X,+21q$ ) với tinh trùng bình thường ( $23,X$ ) tạo ra hội chứng Down thuộc dạng chuyển đoạn ( $46,XX,+21q$ ).



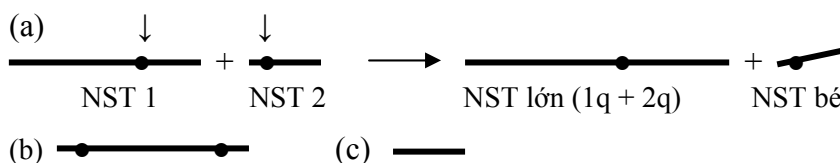
**Hình 3.21 Hội chứng Down do thể chuyển đoạn dạng  $t(21q; 22q)$  gây ra:** (a) Sự hình thành các loại giao tử khác nhau từ người nữ  $t(21q; 22q)$ ; và (b) Sự thụ tinh giữa một trứng bất thường ( $23,X,+21q$ ) với tinh trùng bình thường ( $23,X$ ) tạo ra hội chứng Down chuyển đoạn dạng ( $46,XX,+21q$ ).

## 2. Các biến đổi về số lượng nhiễm sắc thể

Từ bảng 3.2 cho thấy sự biến đổi số lượng nhiễm sắc thể sai khác nhau ở hai cách cơ bản: (1) Các thể biến dị nguyên bội hay thể đa bội (euploid variants) có số lượng các bộ nhiễm sắc thể khác nhau; và (2) các thể biến dị lệch bội hay thể dị bội (aneuploid variants) có số lượng nhiễm sắc thể cụ thể không phải là lưỡng bội. Nói chung, các kiểu biến đổi số lượng nhiễm sắc thể ảnh hưởng lên sự sống sót lớn hơn các kiểu biến đổi cấu trúc nhiễm sắc thể. Trên thực tế, ở người, hơn một nửa các trường hợp sảy thai tự nhiên xảy ra trong ba tháng mang thai đầu có liên quan tới các thai bị lệch bội, đa bội, hoặc các sai hình nhiễm sắc thể lớn khác.

Trước khi thảo luận các thể đột biến số lượng nhiễm sắc thể liên quan tới sự tăng hoặc giảm hàm lượng vật chất di truyền, ta hãy xét qua các cơ chế khác mà qua đó số lượng nhiễm sắc thể có thể thay đổi (hình 3.22).



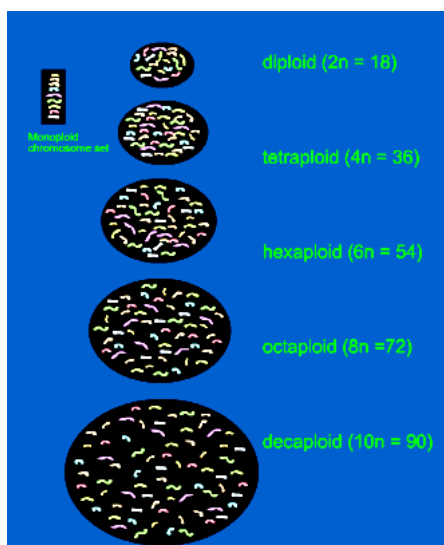


**Hình 3.22** Sự hình thành nhiễm sắc thể tâm giữa do dung hợp tâm động của hai nhiễm sắc thể tâm đầu không tương đồng (a), và các dạng nhiễm sắc thể hai tâm và không tâm (b và c).

Như đã đề cập, nếu như các chuyển đoạn thuận nghịch xảy ra giữa hai nhiễm sắc thể tâm đầu mà các đoạn lớn (tức vai dài) được nối lại với nhau, kết quả là tạo ra một nhiễm sắc thể lớn có tâm giữa và một nhiễm sắc thể bé mà nó có thể biến mất trong quá trình phân bào. Cơ chế hình thành một nhiễm sắc thể tâm giữa lớn do sự đứt-nối xảy ra tại các tâm động của hai nhiễm sắc thể tâm đầu không tương đồng như thế gọi là sự *dung hợp tâm động* (centric fusion; hình 3.22a). Ngoài ra, thỉnh thoảng cũng bắt gặp các dạng nhiễm sắc thể bất thường có hai tâm động và không tâm động (hình 3.22b-c). Tuy nhiên, các dạng bất thường như thế có tính không ổn định bởi vì chúng thường bị loại bỏ khỏi tế bào khi phân ly về hai cực trong quá trình phân bào; chẳng hạn, dạng hai tâm động thường xuất hiện ở dạng cầu nối và sẽ bị đứt gãy do các sợi thoi kéo căng về hai cực.

### 2.1. Hiện tượng đa bội (*polyploidy*)

Các *thể đa bội* (polyploids) là trường hợp trong đó các sinh vật có ba, bộ nhiễm sắc thể trở lên (hình 3.23). Nếu ta gọi  $x$  là số nhiễm sắc thể đơn bội cơ bản, khi đó các sinh vật có ba, bốn bộ nhiễm sắc thể... sẽ có số nhiễm sắc thể và tên gọi tương ứng là  $3x$  (*thể tam bội*: triploid),  $4x$  (*thể tứ bội*: tetraploid)... Lưu ý: thay vì dùng ký hiệu  $n$  để chỉ số nhiễm sắc thể đơn bội như trước đây, ở đây  $x$  chỉ số nhiễm sắc thể trong một bộ và  $n$  biểu thị cho số nhiễm sắc thể trong một giao tử. Ví dụ: một sinh vật lục bội với 60 nhiễm sắc thể ( $6x = 2n = 60$ ), thì  $x = 10$  và  $n = 30$ .



**Hình 3.23** Hiện tượng đa bội hóa điển hình ở chi *Chrysanthemum*, với rất nhiều loài đa bội thể khác nhau sinh ra từ loài lưỡng bội ( $2n = 18$ ).



Nói chung, hiện tượng đa bội thể tương đối phổ biến ở các thực vật nhưng hiếm gặp ở hầu hết các động vật. Gần như một nửa số thực vật có hoa đều là các thể đa bội, kể cả các loài cây trồng quan trọng. Chẳng hạn, khoai tây tứ bội ( $4x = 48$ ), lúa mì mềm lục bội ( $6x = 42$ ), và cây dâu tây bát bội ( $8x = 56$ ). Ở thực vật bậc cao, *Chrysanthemum* là một chi điển hình về hiện tượng đa bội hóa (hình 3.23). Trong quá trình giảm phân ở các loài thuộc chi này, các nhiễm sắc thể kết đôi tạo thành các thể lưỡng trị, loài 118 nhiễm sắc thể tạo thành 9 thể lưỡng trị, loài 36 nhiễm sắc thể tạo thành 18 thể lưỡng trị v.v... Mỗi giao tử nhận một nhiễm sắc thể từ mỗi thể lưỡng trị, vì vậy số lượng nhiễm sắc thể trong mỗi giao tử của bất kỳ loài nào cũng chính bằng một nửa số lượng nhiễm sắc thể được tìm thấy trong mỗi tế bào soma của nó. Ví dụ, loài thập bội có 90 nhiễm sắc thể thì tạo thành 45 thể lưỡng trị, do đó mỗi giao tử sẽ mang 45 nhiễm sắc thể. Nhờ vậy qua thụ tinh, bộ đầy đủ 90 nhiễm sắc thể của loài này được phục hồi. Như vậy, các giao tử của cơ thể đa bội rõ ràng là không phải đơn bội như ở cơ thể lưỡng bội.

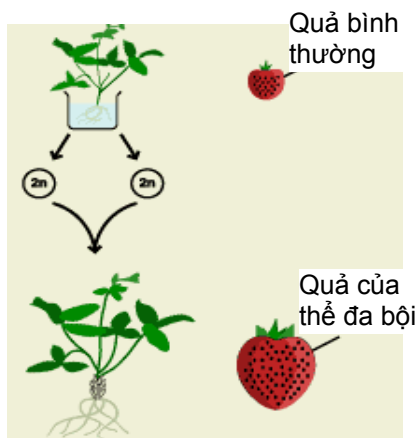
Thông thường, người ta phân biệt hai kiểu thể đa bội: (1) các *thể đa bội cùng nguồn* hay *thể tự đa bội* (autopolyploids) là các thể đa bội nhận được tất cả các bộ nhiễm sắc thể của chúng từ cùng một loài; và (2) các *thể đa bội khác nguồn* hay *thể dị đa bội* (allopolyploids) là các thể đa bội nhận được các bộ nhiễm sắc thể từ các loài khác nhau. Chẳng hạn, nếu như một hạt phấn lưỡng bội không giảm nhiễm từ một loài lưỡng bội thụ tinh cho một trứng lưỡng bội cũng của loài đó, đời con sinh ra là các *thể tự tứ bội* (autotetraploids), hay AAAA, trong đó A biểu thị một bộ nhiễm sắc thể hoàn chỉnh hay *bộ gene* (genome) của kiểu A. Mặt khác, nếu như hạt phấn lưỡng bội của một loài thụ tinh cho một trứng lưỡng bội của một loài khác có quan hệ họ hàng với loài này, đời con sinh ra sẽ là các *thể dị tứ bội* (allotetraploids), hay AABB, trong đó B chỉ bộ gene của loài thứ hai. Tất cả các bộ nhiễm sắc thể trong một thể tự đa bội đều là tương đồng, giống như khi chúng ở trong một thể lưỡng bội. Nhưng trong các thể dị đa bội, các bộ nhiễm sắc thể khác nhau nói chung sai khác nhau ở một mức độ nào đó, và được gọi là *tương đồng một phần* (homeologous), hay *tương đồng từng phần* (partially homologous).

Trong tự nhiên, các thể đa bội xảy ra với tần số rất thấp, khi một tế bào trải qua sự nguyên phân hoặc giảm phân bất thường. Chẳng hạn, nếu trong nguyên phân tất cả các nhiễm sắc thể đi về một cực, thì tế bào đó sẽ có số nhiễm sắc thể là tứ bội. Nếu như xảy ra giảm phân bất thường, có thể tạo ra một giao tử không giảm nhiễm có  $2n$  nhiễm sắc thể. Tuy nhiên, trong hầu hết các tình huống, giao tử lưỡng bội này sẽ kết hợp với một giao tử đơn bội bình thường và sinh ra một thể tam bội. Người ta cũng có

thể tạo ra các thể đa bội bằng cách xử lý colchicine, một loại hóa chất gây rối loạn sự hình thành thoi vô sắc. Kết quả là, các nhiễm sắc thể không phân ly về các cực được, và thường thì xuất hiện các thể tự tứ bội.

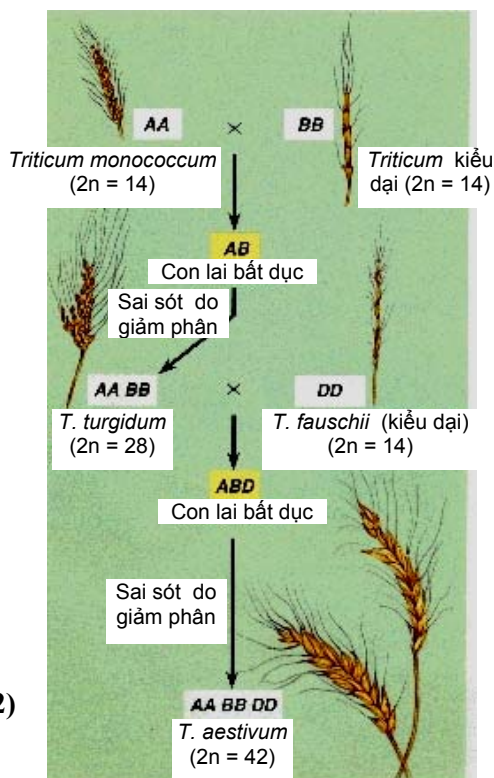
2.1.1. Các thể tự đa bội (*autopolyploids*)

Các cơ thể tam bội (AAA) thường là các thể tự đa bội sinh ra do thụ tinh giữa các giao tử đơn bội và lưỡng bội. Chúng thường bất dục bởi vì xác suất sinh ra các giao tử có được cân bằng là rất thấp. Trong giảm phân, ba cái tương đồng có thể kết cặp và hình thành một thể lưỡng tri, hoặc hai cái tương đồng kết cặp như là một thể lưỡng tri, để lại nhiễm sắc thể thứ ba không kết cặp. Tuy nhiên, do tập tính của các nhiễm sắc thể không tương đồng là độc lập, nên xác suất để một giao tử có chính xác  $n$  nhiễm sắc thể là  $(\frac{1}{2})^n$  (sử dụng quy tắc nhân), và xác suất để một giao tử có được chính xác  $2n$  nhiễm sắc thể cũng là  $(\frac{1}{2})^n$ . Tất cả các giao tử khác còn lại sẽ là không có sự cân bằng và nói chung là không hoạt động chức năng trong các hợp tử có chứa chúng. Chẳng hạn, hầu hết các cây chuối là các thể tam bội; chúng sinh ra các giao tử không cân bằng, và kết quả là không có hạt.



**Hình 3.24** Sự hình thành thể đa bội với hoa trái lớn hơn thể  $2n$  bình thường.

**Hình 3.25** Sự hình thành lúa mì *Triticum aestivum* dị lục bội ( $2n = 42$ ) bằng con đường dị đa bội. ➔



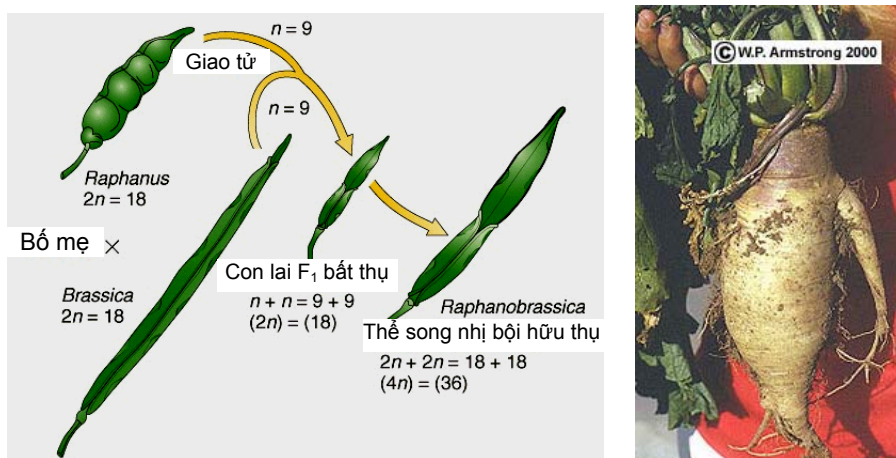
Các thể đa bội thường lớn hơn các thể lưỡng bội họ hàng (ví dụ, cho hoa trái lớn hơn; Hình 3.24). Các thể tự đa bội có thể giảm phân bình

thường nếu như chúng chỉ tạo thành các thể lưỡng trị hoặc các thể tứ trị. Còn nếu như bốn nhiễm sắc thể tương đồng tạo thành một thể tam trị và một thể đơn trị, thì các giao tử nói chung sẽ có quá nhiều hoặc quá ít các nhiễm sắc thể.

### 2.1.2. Các thể dị đa bội (*allopolyploids*)

Hầu hết các thể đa bội trong tự nhiên đều là các thể dị đa bội, và chúng có thể cho ra một loài mới. Chẳng hạn, lúa mì *Triticum aestivum* là một dạng dị lục bội với 42 nhiễm sắc thể. Qua kiểm tra các loài hoang dại họ hàng cho thấy lúa mì này bắt nguồn từ ba dạng tổ tiên lưỡng bội khác nhau, mỗi dạng đóng góp hai bộ nhiễm sắc thể (ở đây ta ký hiệu AABBDD). Sự kết cặp chỉ xảy ra giữa các bộ nhiễm sắc thể tương đồng, vì vậy giảm phân là bình thường và cho ra các giao tử cân bằng có  $n = 21$ . Rõ ràng, hiện tượng dị đa bội đóng vai trò quan trọng trong sự tiến hoá của lúa mì (xem Hình 3.25).

Năm 1928, nhà khoa học người Nga, G. Karpechenko đã tạo ra một thể dị tứ bội rất là đặc biệt; khi ông lai giữa cải bắp *Brassica* và cải củ *Raphanus sativus*, cả hai đều có số nhiễm sắc thể lưỡng bội là 18. Ông muốn tạo ra con lai có lá của cây cải bắp và củ của cây cải củ. Sau khi thu được các hạt lai từ một cây lai nhân tạo, đem gieo trồng và phát hiện rằng chúng có 36 nhiễm sắc thể. Tuy nhiên, thay vì thu được các tính trạng như ông mong đợi, cây lai này có lá của cây cải củ và củ của cây cải bắp! Hình 3.26 cho thấy sự hình thành con lai giữa hai loài cải củ và cải bắp nói trên, được gọi là *Raphanobrassica*.

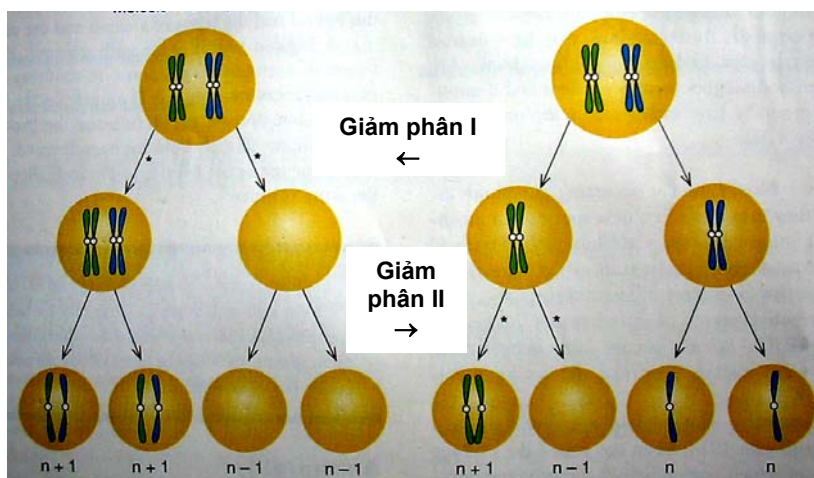


**Hình 3.26** Sự tạo thành thể song nhị bội hữu thụ *Raphanobrassica* từ hai loài cải bắp *Brassica* và cải củ *Raphanus* đều có  $2n = 18$  (trái); và một kết quả cụ thể của con lai  $F_1$  *Raphanobrassica* (theo W.P.Armstrong 2000).

Trong trường hợp nếu một hạt phấn đơn bội có bộ gene A thụ phấn cho hoa của loài có bộ gene B, sẽ cho ra một con lai bất thụ có thành phần bộ gene AB. Nếu như sau đó nguyên phân không xảy ra được trên một nhánh, có thể sinh ra các tế bào AABB. Nếu các tế bào này tự thụ phấn thì sẽ tạo ra một thể dị đa bội. Lợi dụng đặc điểm này, các nhà chọn giống sử dụng colchicine tác động lên con lai bất thụ để tạo ra các thể dị đa bội.

## 2.2. Hiện tượng lệch bội (*aneuploidy*)

Trong tự nhiên, thỉnh thoảng ta bắt gặp các cá thể có số lượng nhiễm sắc thể không phải là bội số của số nhiễm sắc thể đơn bội, do chúng bị thừa hoặc thiếu một hoặc vài nhiễm sắc thể cụ thể nào đó. Đó là các thể lệch bội hay *thể dị bội* (*aneuploids*).

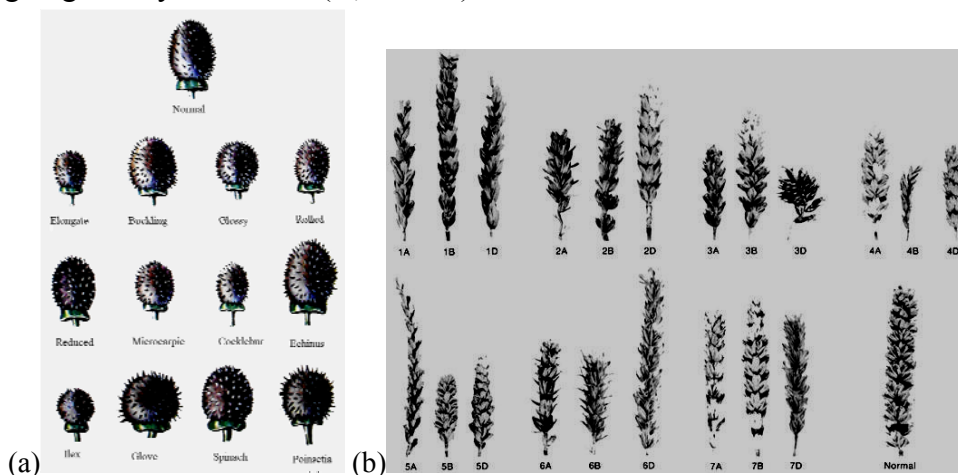


**Hình 3.27** Sự không phân tách xảy ra trong giảm phân I (bên trái) và giảm phân II (bên phải) với các giao tử được tạo ra.

Nguyên nhân của hiện tượng lệch bội là sự *không phân tách* (*nondisjunction*) của hai nhiễm sắc thể tương đồng trong quá trình giảm phân hoặc nguyên phân. Sự không phân tách trong giảm phân tự nó được coi là kết quả của sự kết cặp không đúng cách của các nhiễm sắc thể tương đồng trong giảm phân sớm đến nỗi các tâm động không đối diện nhau trên mặt phẳng kỳ giữa, hoặc là không hình thành được hình chéo. Kết quả là cả hai nhiễm sắc thể cùng đi về một cực, làm cho một tế bào con thừa một nhiễm sắc thể và tế bào con kia không có nhiễm sắc thể đó. Khi các giao tử bất thường ( $n + 1$ ) và ( $n - 1$ ) này thụ tinh với các giao tử bình thường ( $n$ ) sẽ sinh ra các hợp tử có bộ nhiễm sắc thể bất thường tương ứng: thừa một chiếc ( $2n + 1$ ), gọi là *thể ba* (*trisomy*) và thiếu một chiếc ( $2n - 1$ ), gọi là *thể một* (*monosomy*). Sự không phân tách phổ biến nhất là trong giảm phân I, nhưng cũng có thể xảy ra cả trong giảm phân II (hình 3.27). Các

kiểu tổ hợp nhiễm sắc thể khác như  $2n + 2$  (*thể bốn*: tetrasomy) hoặc  $2n - 2$  (*thể không*: nullisomy) cũng có thể xảy ra, nhưng ở đây ta không quan tâm. Sự không phân tách cũng có thể xảy ra trong nguyên phân gây ra các thể khảm về các tế bào bình thường và lệch bội.

Các thể ba được biết đến ở nhiều loài. Ở thực vật, ví dụ điển hình đó là một loạt các thể ba với những đặc tính kỳ lạ ở loài cà độc dược *Datura stramonium* được Alfred Blakeslee nghiên cứu vào khoảng năm 1920. Thực ra, người ta đã phát hiện được tất cả các thể ba về từng nhiễm sắc thể trong số 12 nhiễm sắc thể khác nhau ở loài này; và mỗi thể ba có một kiểu hình đặc trưng, thể hiện rõ nhất là ở vỏ quả. Điều đó chứng tỏ các nhiễm sắc thể khác nhau có các hiệu quả di truyền khác nhau lên tính trạng này (hình 3.28a). Các thể ba cũng được nghiên cứu ở nhiều loài cây trồng như ngô, lúa gạo và lúa mì nhằm xác định các nhiễm sắc thể mang các gene khác nhau. Ở hình 3.28b còn cho thấy hiệu quả di truyền của các *thể khuyết nhiễm* liên quan 7 nhiễm sắc thể khác nhau đối với kiểu bông ở ba giống lúa mì khác nhau (A, B và D).



**Hình 3.28 (a)** Quả bình thường của cà độc dược *Datura* (trên cùng) và 12 kiểu *thể ba* khác nhau, mỗi kiểu có một vẻ ngoài và tên gọi khác nhau.

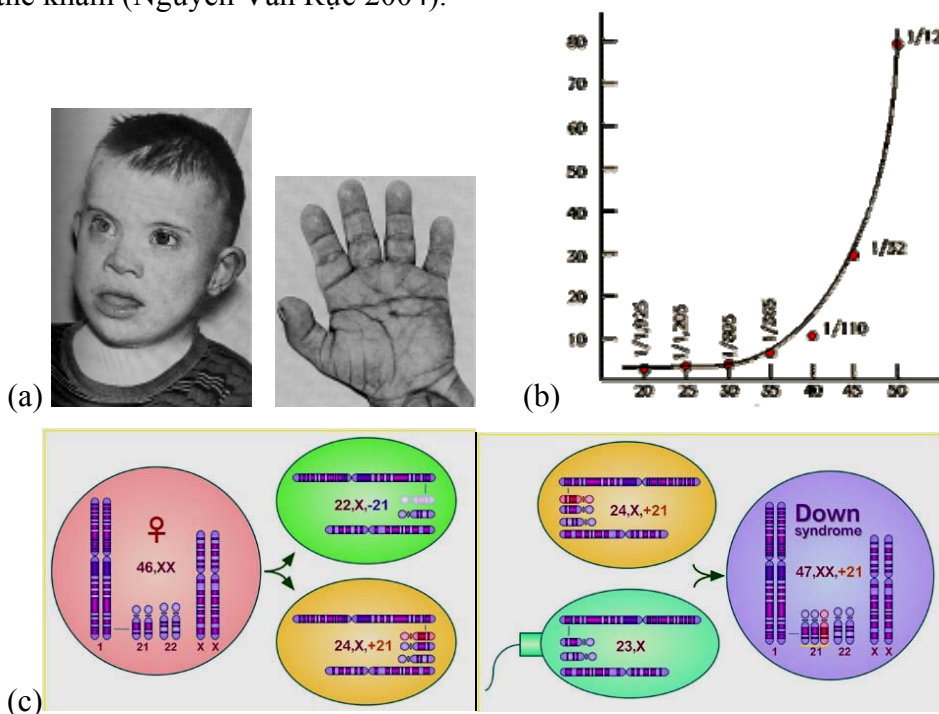
**(b)** Các *thể khuyết nhiễm* liên quan 7 nhiễm sắc thể khác nhau ở ba bộ gen lúa mì (A, B và D) cho các hiệu quả di truyền khác nhau đối với kiểu hình bông so với dạng bình thường (hình cuối).

Ở người, việc phân tích các thành phần nhiễm sắc thể của các trường hợp sảy thai tự phát cho thấy hầu như tất cả các thể một và nhiều thể ba đều là các dạng gây chết thai. Tuy nhiên, một số trường hợp vẫn được sinh ra với các hội chứng khác nhau. Phổ biến nhất là hội chứng Down, thể ba nhiễm sắc thể 21, với tần số 1/700 số trẻ sơ sinh và thường tỷ lệ với tuổi



tác người mẹ, đặc biệt là từ độ tuổi 35 trở đi (Hình 3.29). Hội chứng Down được mô tả cách đây chừng 160 năm, nói chung có các đặc trưng là trí tuệ và vẻ ngoài chung dễ thấy như: đầu to, trán vát, khe mắt xếch, lưỡi hay thè ra ..., thường tử vong ở độ tuổi 10-40 và hiếm khi sinh sản.

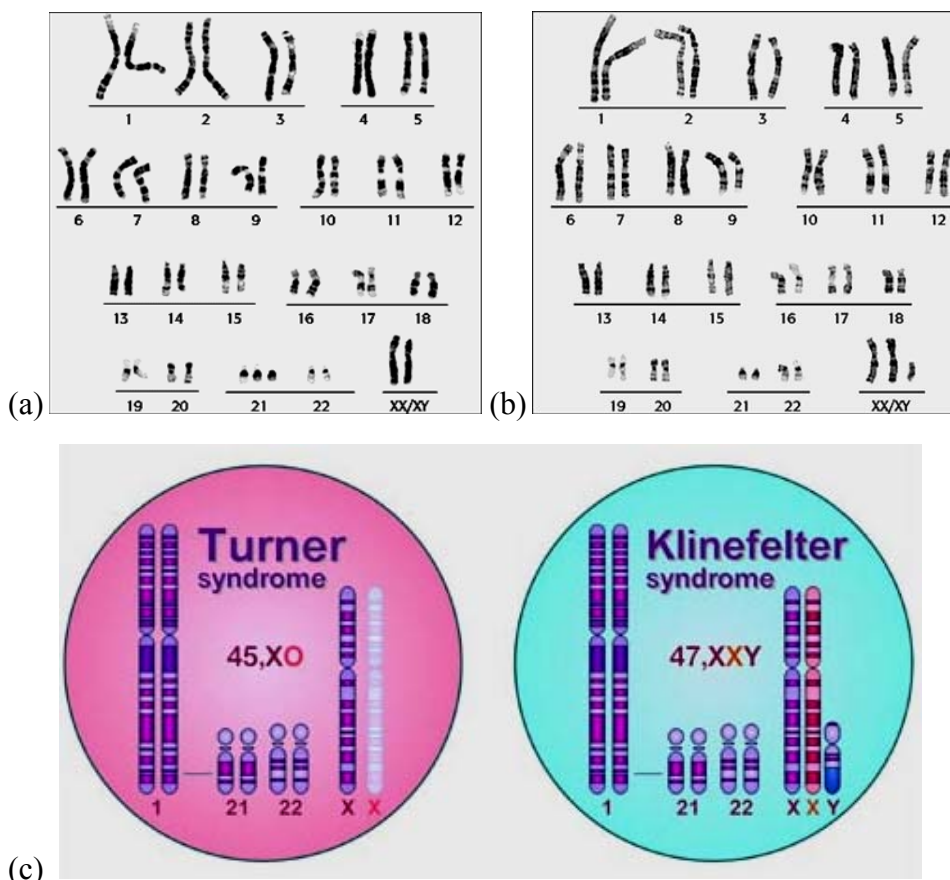
Cơ sở nhiễm sắc thể của hội chứng Down được khám phá đầu tiên vào năm 1959. Kiểu nhân của những người này được cho thấy ở các hình 3.29c và 3.30a. Danh pháp hiện hành để chỉ cá thể có thể ba 21 là (47,+21), trong đó con số 47 chỉ toàn bộ số nhiễm sắc thể và +21 chỉ ra rằng có ba bản sao của nhiễm sắc thể 21. Trong số những người mắc hội chứng Down, chỉ có chừng 5% là kiểu nhân dị hợp về chuyển đoạn thuận nghịch như đã nói trước đây, và hầu hết (khoảng 95% trường hợp) là kết quả của sự không phân tách trong giảm phân. Một kết quả nghiên cứu gần đây cho thấy đặc điểm karyotype của các trường hợp Down ở nước ta, như sau: 91% thể trisomy 21 thuần, 6% thể chuyển đoạn như đã nói ở trên và 3% là thể khảm (Nguyễn Văn Rục 2004).



**Hình 3.29 Hội chứng Down**

(a) Một cháu bé mắc hội chứng Down (trái) với đặc điểm cấu tạo bàn tay. (b) Một đồ thị về nguy cơ mắc hội chứng Down tính trên 1.000 trẻ sơ sinh (trục tung) liên quan với các độ tuổi người mẹ (trục hoành). (c) Sự hình thành các trường thừa và thiếu NST 21 (trái) và sự thụ tinh giữa trường thừa NST 21 với tinh trùng bình thường tạo ra thể ba 21.

Các thể ba nhiễm sắc thể thường khác là rất hiếm, chủ yếu là bởi vì chúng không sống được khi còn là thai nhi. Ví dụ, *hội chứng Patau*, thể ba 13 (47, +13), có tần số sảy thai là 1/33, tần số bắt gặp là 1/15.000 trẻ sơ sinh, và sống chưa đầy sáu tháng. Trường hợp khác là *hội chứng Edward*, thể ba 18 (47, +18), tần số bắt gặp ở trẻ sơ sinh là 1/5.000 và sống không đến một năm.



**Hình 3.30 Kiểu nhân của một số thể đột biến lệch bội thường gặp.**

(a) Kiểu nhân người nữ mắc hội chứng Down (47, XX, +21); (b) Kiểu nhân người mắc hội chứng Klinefelter (47, XXY). (c) Một trình bày đơn giản kiểu nhân của những người mắc hội chứng Turner và Klinefelter.

Ngoài ra, sự không phân tách của các nhiễm sắc thể giới tính ở người gây ra một số kiểu nhiễm sắc thể bất thường phổ biến như: XO, XXX, XXY và XYY. Chẳng hạn, *hội chứng Klinefelter* XXY (hay 47, XXY) bắt gặp ở trẻ sơ sinh với tần số khoảng 1/2.000; đây là trường hợp của những người nam có các đặc điểm như vô sinh, thân hình cao không cân đối...

Kiểu nhân của cá thể mắc hội chứng Klinefelter được mô tả ở hình 3.30b.

Những người mắc *hội chứng XYY* (có thể gọi là "siêu nam"), xuất hiện với tần số khoảng 1/1.000 trong số các trẻ nam sơ sinh; những cá thể này không có vẻ gì là bệnh tật, sức sống và sinh sản bình thường. Các nghiên cứu xa hơn cho thấy dường như những người này có khuynh hướng phạm tội, có thể có tương quan tối thiểu với hành vi. Tần số các cá thể XYY bị ngời tù cao đáng kể so với cộng đồng dân cư nói chung; tuy nhiên, chưa tới 5% tổng số các cá thể XYY là được giáo dục dạy dỗ đàng hoàng. *Hội chứng Turner XO* (hay 45, X) xuất hiện ở trẻ sơ sinh với tần số chừng 1/5.000 và tần số sảy thai khoảng 1/18; đây là trường hợp của những người nữ vô sinh, lùn không cân đối, cổ ngắn... Còn *hội chứng "siêu nữ" XXX* (hay 47, XXX) bắt gặp ở trẻ sơ sinh với tần số khoảng 1/700; hầu như tất cả những người nữ mắc hội chứng này đều vô sinh.

## Câu hỏi và Bài tập

1. Thế nào là chu kỳ tế bào? Bằng cách nào số lượng nhiễm sắc thể đặc trưng của mỗi loài được duy trì ổn định trong quá trình nguyên phân?
2. Bằng cách nào số lượng nhiễm sắc thể từ trạng thái lưỡng bội giảm xuống còn đơn bội trong quá trình giảm phân? Tại sao nói các sự kiện ở kỳ trước và kỳ sau của giảm phân I góp phần quan trọng trong việc tạo ra nguồn biến dị tổ hợp phong phú đa dạng ở các loài sinh sản hữu tính?
3. So sánh nguyên phân và giảm phân, và cho biết ý nghĩa của chúng.
4. Tại sao trong mỗi tế bào số lượng nhiễm sắc thể (mức  $n$ ) và hàm lượng ADN (mức  $c$ ) là không đồng bộ trong mỗi giai đoạn của phân bào?
5. Hai điểm sai khác chính giữa sự sinh tinh và sinh trứng ở người và hầu hết các động vật có vú khác là gì? Các quá trình phát sinh giao tử ở động vật và thực vật bậc cao có những điểm giống và khác nhau nào?
6. Một sinh vật có số lượng nhiễm sắc thể lưỡng bội bằng 12, được ký hiệu: Aa, Bb, Cc, Dd, Ee và Ff. (a) Có bao nhiêu tổ hợp nhiễm sắc thể khác nhau có thể xuất hiện trong các giao tử? (b) Xác suất để một giao tử nhận được tất cả các nhiễm sắc thể viết hoa là gì?
7. (a) Có thể có những kiểu biến đổi nào xảy ra trong cấu trúc nhiễm sắc thể? Mô tả những nét chính xảy ra trong mỗi kiểu cùng với hậu quả và ứng dụng của chúng. (b) Gọi tên các trường hợp lệch bội phổ biến ở người và mô tả các thành phần nhiễm sắc thể của chúng.
8. Một ruồi giấm cái thân đen mun ee (e = ebony) đồng hợp được cho



lai với một con đực hoang dại đồng hợp ( $e^+e^+$ ) đã được chiếu tia X. Ở đời con có xuất hiện một ruồi cái thân mun. Khi cho ruồi giấm này lai với ruồi giấm  $F_1$  thu được  $\frac{1}{2}$  thân mun:  $\frac{1}{2}$  hoang dại. Hãy giải thích kết quả này.

9. Giả sử rằng ở  $F_1$  của phép lai ở bài tập 8 ta phát hiện được hai con ruồi giấm đực và cái đều có thân đen mun. Có thể nói gì về số lượng tương đối của đời con nếu cho hai con ruồi giấm  $F_1$  đó lai với nhau? (Cho biết các mắt đoạn thường gây chết ở trạng thái đồng hợp).

10. Cho một cây dị hợp ABCDE/abcde lai với một cây abcde/abcde, và đời con xuất hiện sáu dạng sau: ABCDE; abcde; Abcde; aBCDE; ABCDe; abcDE. Có điều gì bất thường trong kết quả này và bạn có thể giải thích điều đó ra sao? (Gợi ý: sự thiếu vắng của hai hoặc nhiều kiểu hình được kỳ vọng gợi ra một thể dị hợp đảo đoạn có liên quan đến các gene liên kết).

## Tài liệu Tham khảo

### Tiếng Việt

Dubinín NP. 1981. *Di truyền học đại cương* (bản dịch của Trần Đình Miên và Phan Cự Nhân). NXB Nông Nghiệp, Hà Nội.

Kimura M. 1983. *Thuyết tiến hóa phân tử trung tính*. (Bản dịch của Hoàng Trọng Phán). NXB Thuận Hóa - Huế, 1993 (tr.17-33).

Lê Đình Lương, Phan Cự Nhân. 1997. *Cơ sở di truyền học*. NXB Giáo Dục, Hà Nội.

Phan Cự Nhân (chủ biên), Nguyễn Minh Công, Đặng Hữu Lan. 1999. *Di truyền học*. NXB Giáo Dục, Hà Nội.

Nguyễn Văn Rực. 2004. *Nghiên cứu đặc điểm karyotype, kiểu hình của trẻ Down và karyotype của bố mẹ*. Luận án Tiến sĩ Y học, Trường Đại học Y Hà Nội, Hà Nội.

Phạm Quang Vinh. 2003. *Nghiên cứu bất thường nhiễm sắc thể trong các thể bệnh Leucémie cấp ở người lớn tại Viện Huyết học Truyền máu*. Luận án Tiến sĩ Y học, Trường Đại học Y Hà Nội, Hà Nội.

### Tiếng Anh

Campbell NA, Reece JB. 2001. *Essential Biology*. Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc, San Francisco, CA.

Clarke L. 1998. Centromeres: proteins, protein complexes, and repeated domains at centromere of simple eukaryotes. In: *Current Opinion in Genetics & Development*, Vol 8, No2 (Allis CD and Gasser SM, eds.), pp 212-218.

Clegg CJ, Mackean DG. 2000. *Advanced Biology: Principles and Applications*. 2<sup>nd</sup> ed, John Murray Published Ltd, London.

Dobie KW, Hari KL, Maggert KA, Karpen GH. 1999. Centromere proteins and chromosome inheritance: a complex affair. In: *Current Opinion in Genetics & Development*, Vol 9, No2 (Kadonaga JT and Grunstein M, eds.), pp 206-217.

Hartl DL, Freifelder D, Snyder LA. 1988. *Basic Genetics*. Jones and Bartlett Publishers, Inc, Boston - Portola Valley.

Lewis R. 2003. *Human Genetics: Concepts and Applications*. 5<sup>th</sup> ed, McGraw-Hill, Inc, NY.

Li W-H. 1983. Evolution of duplication genes and pseudogenes. In: *Evolution of Genes and Proteins* (Nei M, Koehn RK, eds.), pp. 14-37. Sinauer Associates Inc.-Publishers, Sunderland, Massachusetts, USA.

Biggins S. and Murray AW. 1999. Sister chromatid cohesion in mitosis. In: *Current Opinion in Genetics & Development*, Vol 9, No2 (Kadonaga JT and Grunstein M, eds.), pp 230-236.

Russell PJ. 2003. *Essential Genetics*. Benjamin/Cummings Publishing Company, Inc, Menlo Park, CA.

Schrock E, du Manoir S, Veldman T, Schoell B, Weinberg J, Ferguson-Smith MA, Ning Y, Ledbetter DH, Bar-Am I, Soenksen D, Garini Y, Ried T. 1996. Multicolor spectral karyotyping of human chromosomes. *Science* 273: 495.

Verma RS and Babu A. 1995. *Human chromosomes: Principles and Techniques*. 2nd ed., McGraw-Hill, Inc, NY. (Ch2-6, pp.6-231).

Weaver RF, Hedrick PW. 1997. *Genetics*. 3<sup>rd</sup> ed, McGraw-Hill Companies, Inc. Wm.C.Brown Publishers, Dubuque, IA.

### **Một số trang web**

<http://www.mhle.com/lewisgenetics5>

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>

<http://www.genetic.org>

<http://www.fivepminus.org>

<http://www.ndss.org>

<http://www.members.aol.com/cdousa/cdo.htm>

<http://www.turner-syndrom-us.org>

<http://www.trisomy.org>

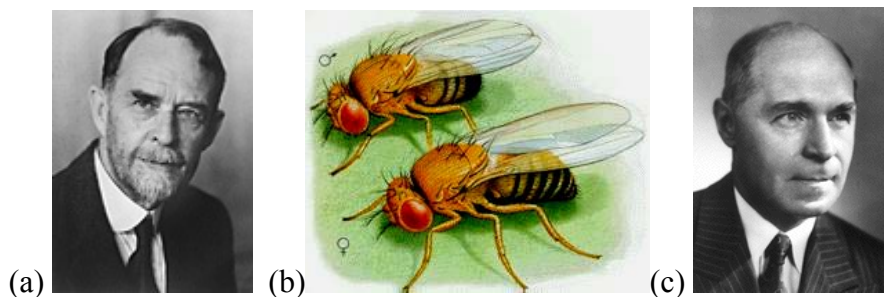
## Chương 4

# Di truyền học Nhiễm sắc thể

Như đã thảo luận ở các chương trước, nguyên lý phân ly độc lập của Mendel chỉ nghiệm đúng trong trường hợp các gene nằm trên các nhiễm sắc thể khác nhau và được lý giải bằng sự phân ly ngẫu nhiên của các nhiễm sắc thể trong giảm phân. Tuy nhiên, vào đầu thập niên 1910, Thomas Hunt Morgan (hình 4.1a) dựa vào các kết quả nghiên cứu ở ruồi giấm *Drosophila melanogaster* (hình 4.1b) đã nhận ra rằng có nhiều gene cùng nằm trên một nhiễm sắc thể. Điều khẳng định đúng đắn này đã sớm bổ sung và làm sáng tỏ cho các nguyên lý di truyền Mendel, đồng thời đặt nền tảng vững chắc cho sự phát triển của di truyền học trong suốt nửa đầu thế kỷ XX. Trong chương này, chúng ta lần lượt tìm hiểu thuyết di truyền nhiễm sắc thể với các vấn đề xác định giới tính, các kiểu di truyền liên kết đối với hai hoặc nhiều gene trên cùng một nhiễm sắc thể, cũng như các phương pháp kinh điển dùng để lập bản đồ di truyền ở các sinh vật.

### I. Trường phái Morgan với thuyết di truyền nhiễm sắc thể

Từ 1910, Morgan cùng với ba cộng sự là Alfred H. Sturtevant, Calvin Bridges và Herman J. Muller (hình 4.1c) đã xây dựng thành công *thuyết di truyền nhiễm sắc thể* (chromosome theory of inheritance) dựa trên đối tượng nghiên cứu nổi tiếng là ruồi giấm *D. melanogaster*. Trước tiên, chúng ta hãy tìm hiểu đặc điểm của đối tượng, phương pháp nghiên cứu và nội dung của học thuyết này.



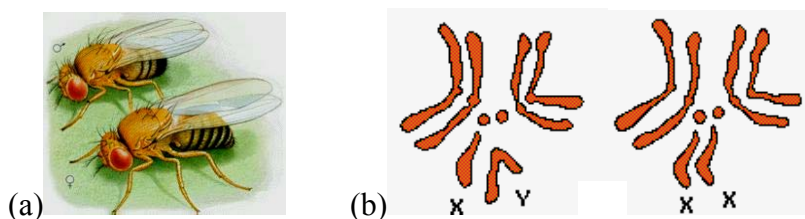
**Hình 4.1** (a) T.H.Morgan; (b) Ruồi giấm *D. melanogaster*, đối tượng nghiên cứu nổi tiếng của Morgan; và (c) Herman Muller, một trong ba môn đệ xuất sắc của Morgan với phương pháp gây đột biến bằng tia X.

#### 1. Tầm quan trọng của ruồi giấm *Drosophila*

##### 1.1. Các đặc điểm và giá trị của ruồi giấm trong nghiên cứu di truyền học

*Drosophila melanogaster* (hình 4.1 và 4.2) có lẽ là sinh vật nổi tiếng

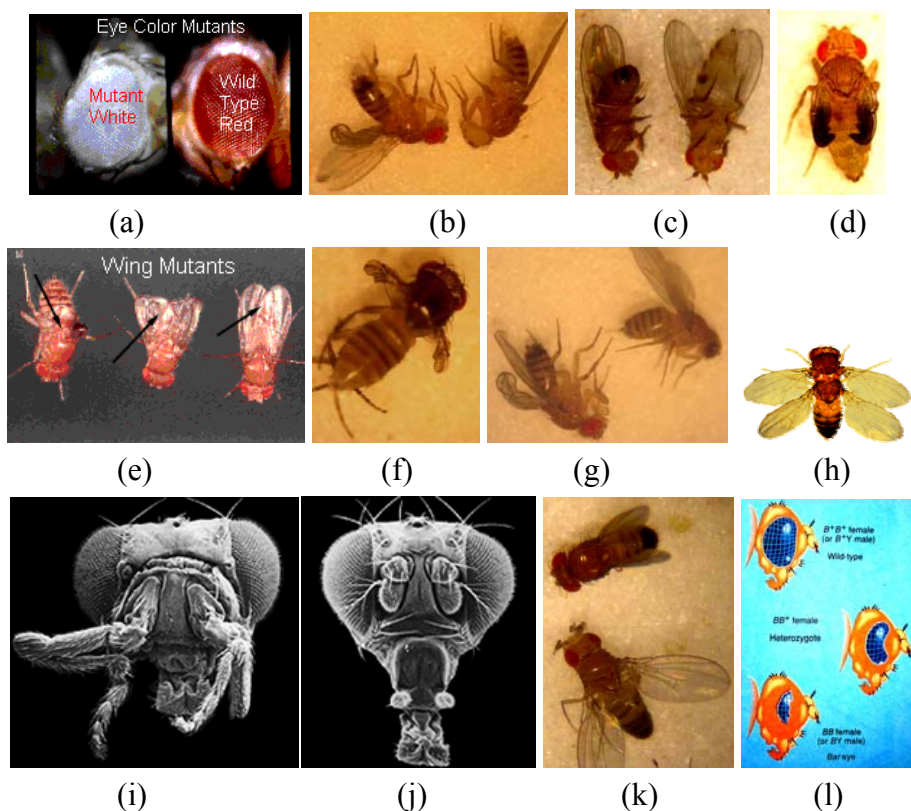
nhất được dùng làm *sinh vật mô hình* (model organism) cho các nhà di truyền học. Ruồi giấm thuộc lớp côn trùng (Insecta), bộ hai cánh (Diptera). Chúng rất thích mùi lên men của các hũ dưa vại cả và đặc biệt là những trái cây chín muồi như chuối, mít hay cam, chanh..., vì vậy chúng được biết dưới cái tên thông dụng là ruồi giấm hay "ruồi trái cây" (fruit-flies). Ruồi giấm phân bố rộng khắp các vùng ôn đới và nhiệt đới trên hành tinh chúng ta.



**Hình 4.2 (a)** Sự khác nhau về hình thái ngoài giữa ruồi giấm đực (trên) và ruồi giấm cái (dưới); và **(b)** bộ nhiễm sắc thể lưỡng bội  $2n = 8$  của chúng, với cặp nhiễm sắc thể giới tính XY- con đực (trái) và XX- con cái.

Giá trị của ruồi giấm *Drosophila* trong các thí nghiệm di truyền nằm trong các đặc điểm sau (các hình 4.2, 4.3 và 4.4):

- Mỗi cặp ruồi giấm sinh được hàng trăm con trong một lứa;
- Vòng đời ngắn, chỉ có hai tuần lễ là chúng có thể nhanh chóng đạt tới tuổi trưởng thành để tham gia sinh sản; và chu kỳ sống có thể rút xuống còn 10 ngày, nếu ở nhiệt độ  $25^{\circ}\text{C}$ . Các ruồi cái trưởng thành về mặt sinh dục nội trong 12 giờ, và chúng lại đẻ trứng hóa nhộng trong hai ngày..
- Sau khi giao phối, các con cái có thể bảo quản các tinh trùng, vì vậy cần thiết phải tiến hành các phép lai với các con cái đang còn trinh (virgin). Ruồi cái còn trinh có thể dễ dàng nhận ra qua màu cứt su (meconium) hay xám nhọt của cơ thể chúng (hình 4.3d) và phân nhộng dưới dạng chấm đen có thể nhìn thấy xuyên qua vùng bụng.
- Ruồi giấm *Drosophila* tương đối dễ nuôi, và dễ dàng phân biệt đực-cái ở các giai đoạn non và trưởng thành để cách ly và tiến hành lai. Bên cạnh sự phân biệt ngoại hình các ruồi non còn trinh như đã nói trên, ở giai đoạn trưởng thành ruồi đực thường khác với ruồi cái ở các điểm sau: cơ thể bé hơn; vùng bụng dưới có ba vạch đen với vạch dưới cùng rộng, trong khi ruồi cái có năm vạch rời nhau; chòm bụng ở con đực hơi tròn và ở con cái nhọn (hình 4.2a).
- Bộ nhiễm sắc thể đầy đủ của các tế bào soma ruồi giấm chỉ có bốn cặp,  $2n = 8$  (hình 4.2b). Toàn bộ bộ gene của *Drosophila* đã được xác định trình tự trong thời gian gần đây.



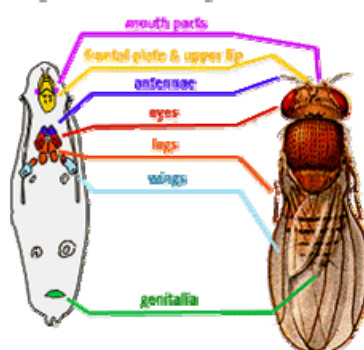
**Hình 4.3 Một số thể đột biến quan trọng của ruồi giấm *Drosophila*.**

Chú thích: (a-b) thể đột biến mắt trắng và mắt đỏ kiểu dại; (c) con đực trưởng thành (trái) và con đực còn trinh có màu cắt su; (d) con cái trinh có màu cắt su; (e) các dạng đột biến khác nhau về cánh; (f) thể đột biến này là một ví dụ về hai kiểu hình cánh ngắn/thân màu đen mun; (g) thể đột biến cánh quăn ở con cái (trái) so với con cái bình thường; (h) thể đột biến hai đốm ngực; (i) thể đột biến antennapedia- kiểu chân râu và (j) anten kiểu dại; (k) so sánh thể đột biến anten (dưới) và dạng bình thường (trên); (l) Mắt kiểu dại với các dạng mắt thỏ (Bar eye) dị hợp và đồng hợp (theo thứ tự từ trên xuống).

- Ruồi giấm có nhiều tính trạng, đặc biệt là các thể đột biến tự phát như mắt trắng hoặc được tạo ra trong phòng thí nghiệm (như thân đen, cánh ngắn, cánh quăn, mắt nâu...); các tính trạng này có thể phân biệt bằng mắt thường, kính lúp hoặc kính hiển vi quang học (hình 4.3 và 4.5).

- Các tế bào tuyến nước bọt của ấu trùng ruồi giấm (hình 4.4) có chứa các nhiễm sắc thể khổng lồ đa sợi (như đã giới thiệu ở chương 3); đây là điểm thuận lợi cho việc xác định các phân cụ thể của các nhiễm sắc thể. Các băng này tự chúng không phải là các gene nhưng rất hữu ích cho việc lập bản đồ các gene trên các nhiễm sắc thể.

**Hình 4.4** Các đĩa mầm (trong ấu trùng) từ đó hình thành nên các cơ quan của ruồi giấm trưởng thành. Từ trên xuống: các phần phụ miệng, mảnh trán và môi trên, anten, mắt, chân, cánh và cơ quan sinh dục.



Ngoài ra, các dụng cụ và hóa chất được sử dụng trong các thí nghiệm ở ruồi giấm là tương đối đơn giản, được giới thiệu ở hình 4.5 dưới đây.



**Hình 4.5** Các dụng cụ thí nghiệm với ruồi giấm. (a) ống nghiệm nuôi và lai ruồi giấm; (b) khay đựng các lọ hóa chất và một số vật dụng khác; và (c) kính hiển vi quang học dùng cho nghiên cứu hình thái và tế bào học.

### 1.2. Các tính trạng được kiểm soát về mặt di truyền của ruồi giấm

Một quần thể bình thường của ruồi giấm *Drosophila* bao gồm các con ruồi điển hình có thân xám, cánh dài và mắt đỏ. Dạng ruồi giấm này là phổ biến nhất, và được các nhà di truyền học xếp vào kiểu dại (wild type).

Nhiều dạng đột biến cũng được phát hiện trong tự nhiên. Một con ruồi đột biến có một xuất phát điểm di truyền ít nhất là một trong số các tính trạng của dạng ruồi giấm bình thường. Các tính trạng đột biến được biết đến bằng một tên gọi. Ví dụ:

đột biến	tên gọi	ký hiệu	đặc tính
mắt trắng	trắng (white)	w	lặn
thân đen mun	mun (ebony)	e	lặn
cánh ngắn	ngắn (vestigial)	vg	lặn
mắt nâu	nâu (brown)	bw	lặn
mắt thoi (Bar)	Bar	B	trội

Các thể đột biến lặn được ký hiệu bằng các chữ cái viết thường, và các tính trạng trội tương ứng được biểu thị bằng các chữ cái viết hoa.

Trong các thí nghiệm lai di truyền, allele bình thường ở một locus cụ thể được các nhà di truyền học trước đây ký hiệu bằng dấu "+". Ký hiệu này thường được bỏ qua ở mức nhập môn, mặc dù nó thường được dùng cho ruồi giấm. Tuy nhiên, đối với bậc Đại học, các allele bình thường được quy cho "kiểu dại" và sử dụng các ký hiệu của các nhà di truyền học.

## 2. Thuyết di truyền nhiễm sắc thể

Ruồi giấm *Drosophila melanogaster* đã được nhà di truyền học người Mỹ (USA), T.H. Morgan (1866-1945), sử dụng trong nghiên cứu di truyền học từ những năm đầu của thế kỷ XX, trong khi đang làm việc tại Học viện Công nghệ California (California Institute of Technology). Nhờ sử dụng ruồi giấm này, Morgan và các cộng sự của mình đã xây dựng thành công học thuyết di truyền nhiễm sắc thể (chromosome theory of heredity).

Trước tiên, ta hãy tóm lược các nội dung chính của thuyết di truyền nhiễm sắc thể và những đóng góp của trường phái Morgan cho sự phát triển của di truyền học.

(1) Học thuyết này xác nhận rằng: gene là đơn vị cơ sở của tính di truyền nằm trên nhiễm sắc thể. Trên mỗi nhiễm sắc thể có nhiều gene phân bố thẳng hàng, mỗi gene chiếm một vị trí xác định gọi là locus; các gene trên mỗi nhiễm sắc thể hợp thành một nhóm liên kết. Số nhóm liên kết gene chính là số nhiễm sắc thể đơn bội, còn gọi là bộ gene (genome).

(2) Trong quá trình giảm phân, các gene trên cùng nhiễm sắc thể có xu hướng phân ly cùng nhau về giao tử. Đây là cơ sở của hiện tượng di truyền liên kết gene hoàn toàn và di truyền liên kết (linkage) nói chung.

(3) Tuy nhiên, trong quá trình giảm phân có thể xảy ra sự trao đổi chéo (crossing over) ở một số đoạn giữa hai nhiễm sắc thể tương đồng, kéo theo sự trao đổi các gen giữa chúng, còn gọi là tái tổ hợp hay hoán vị gene. Đây chính là cơ sở của hiện tượng liên kết gene không hoàn toàn.

(4) Tần số tái tổ hợp (rate of recombination) của các gene là một số hữu tỷ, thỏa mãn giới hạn từ 0,0 đến 0,5 (tức không vượt quá 50%). Đại lượng này tỷ lệ thuận với khoảng cách giữa các gene và phụ thuộc vào một số yếu tố khác như giới tính cũng như mức độ ức chế bởi các trao đổi chéo đồng thời tại nhiều điểm trên một cặp tương đồng. Dựa vào các tần số tái tổ hợp gene này ta có thể thiết lập bản đồ nhiễm sắc thể (chromosome map) của các loài, hay còn gọi là bản đồ liên kết (linkage map).

(5) Vấn đề các nhiễm sắc thể xác định giới tính (sex-determining chromosomes) và hiện tượng di truyền liên kết với giới tính (sex linkage) được làm sáng tỏ lần đầu tiên bởi học thuyết này.

(6) Ngoài ra, trường phái của Morgan đã xác định rằng: gene - đơn vị



di truyền học then chốt này đóng ba vai trò: (i) Gene là *đơn vị chức năng*, nghĩa là gene được xem như một thể thống nhất toàn vẹn kiểm soát một tính trạng cụ thể. (ii) Gene là *đơn vị tái tổ hợp*, nghĩa là gene không bị chia nhỏ bởi sự trao đổi chéo (vì theo quan điểm này, trao đổi chéo không xảy ra bên trong phạm vi một gene mà chỉ xảy ra giữa các gene); như thế gene được coi là *đơn vị cấu trúc cơ sở của vật chất di truyền*, nhiễm sắc thể. (iii) Gene là *đơn vị đột biến*, nghĩa là nếu đột biến xảy ra trong gene dù ở bất kỳ vị trí nào hoặc với phạm vi ra sao, chỉ gây ra một trạng thái cấu trúc mới tương ứng với một kiểu hình mới, kiểu hình đột biến, khác với kiểu hình bình thường. Tuy nhiên, quan niệm này vẫn còn chưa rõ ràng và không thực sự chính xác theo quan điểm của di truyền học hiện đại (sẽ được thảo luận ở chương 6).

Trong việc xây dựng học thuyết di truyền nhiễm sắc thể, phải kể đến những đóng góp to lớn của các môn đệ xuất sắc của Morgan, đó là: Alfred H. Sturtevant với việc đề xuất phương pháp xây dựng bản đồ di truyền năm 1913; Calvin Bridges với việc phát hiện ra cơ chế xác định các kiểu hình giới tính ở ruồi giấm năm 1916; và Herman J. Muller với sự phát triển phương pháp gây đột biến bằng tia X năm 1927. Nhờ đóng góp to lớn đó, Morgan đã được trao giải thưởng Nobel năm 1933 và Muller - năm 1946.

## II. Sự xác định giới tính (sex determination)

Cơ chế tự nhiên mà trong đó một cá thể của một loài phân tính (dioecious species) trở thành con đực hoặc con cái (hay lưỡng tính, hermaphroditic) được gọi là *xác định giới tính* (sex determination). Thực ra, không có một phương thức phổ biến nào cho việc xác định giới tính; nhưng cho đến nay, các nhà sinh học đã khám phá ra nhiều cơ chế xác định giới tính, tập trung vào hai tiêu chí sau (xem các Bảng 4.1- 4.3). Ở một số loài, giới tính được xác định sau khi thụ tinh bởi các nhân tố môi trường như nhiệt độ hoặc sự có mặt của các thể kèm giới tính. Kiểu xác định giới tính này được gọi là *xác định giới tính do môi trường* (environmental sex determination = ESD). Ở các loài khác, giới tính được xác định lúc thụ tinh bằng sự tổ hợp của các gene mà hợp tử nhận được. Kiểu xác định giới tính này được gọi là *xác định giới tính do kiểu gene* (genotypic sex determination = GSD). Dưới đây chúng ta lần lượt xét qua hai hệ thống này (về chi tiết, xem trong: Kalthoff 1997, tr.660-685; và Yablokov 1986, tr.17-30). Riêng vấn đề bất hoạt nhiễm sắc thể X sẽ được thảo luận ở mục II-3-2, vì nó có liên quan đến di truyền liên kết giới tính.

### 1. Sự xác định giới tính do kiểu gen (GSD)

Ở hầu hết các sinh vật, giới tính được xác định bằng sự sai khác nhiễm sắc thể và các gene trên chúng (bảng 4.1). Ví dụ đầu tiên về sự xác định



giới tính do nhiễm sắc thể được ghi nhận ở rệp *Protenor* năm 1905, trong đó các con cái được khám phá có hai nhiễm sắc thể X và con đực chỉ có một X. Các giao tử đực có thể có một X hoặc không có nhiễm sắc thể X được biết là có tỷ lệ ngang nhau, và các giao tử cái thì lúc nào cũng mang một X, vì vậy số lượng đời con có XX (cái) và X (đực) là ngang nhau.

Sau đó không lâu, nhiều nhà nghiên cứu đã xác định sự có mặt của một nhiễm sắc thể giới tính Y ở các sinh vật khác. Ở các loài này, sự có mặt của Y cùng với một X cho ra con đực, trong khi hai bản sao của X sinh ra một con cái. Với lại, các giao tử đực gồm hai kiểu, vì vậy có số lượng tinh trùng hay hạt phấn mang X và mang Y bằng nhau. Trong trường hợp giới tính có cả hai kiểu nhiễm sắc thể giới tính như thế, thì kiểu XY được gọi là *giới tính dị giao tử* (heterogametic sex), còn kiểu XX được gọi là *giới tính đồng giao tử* (homogametic sex). Ở hình 4.6 cho thấy các nhiễm sắc thể X và Y, và cơ chế xác định giới tính ở người.

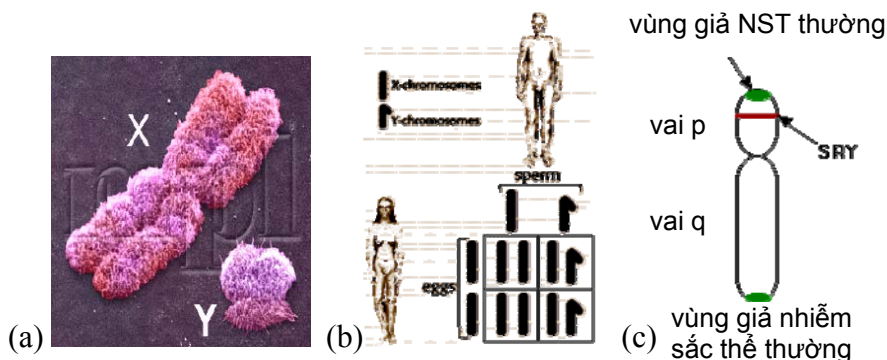
Ở một số sinh vật thuộc các lớp như chim và bò sát thì ngược lại, các con đực (trống) là đồng giao tử và các con cái (mái) là dị giao tử. Để tránh sự nhầm lẫn, người ta thường gọi các nhiễm sắc thể giới tính này là Z và W; như vậy cặp nhiễm sắc thể giới tính con đực là ZZ và con cái là ZW. Kết quả là, các con cái cho hai loại giao tử (một loại mang Z và một loại mang W với tỷ lệ ngang nhau), trong khi các con đực chỉ cho một loại giao tử tất cả đều mang một nhiễm sắc thể Z. Sự hiểu biết về giới tính này cùng với sự di truyền liên kết giới tính có ý nghĩa thực tiễn to lớn trong chăn nuôi gia cầm. Vấn đề này sẽ được thảo luận trở lại ở cuối mục III.

**Bảng 4.1 Các kiểu xác định giới tính khác nhau bởi nhiễm sắc thể**

Sinh vật	Con cái	Con đực
Hầu hết động vật có vú, một số côn trùng (tất cả bọ hai cánh), một số cá, một số thực vật	XX	XY
Rệp <i>Protenor</i> , một số côn trùng khác, kangaroo	XX	X
Lớp chim, hầu hết các bò sát, bướm đêm	ZW	ZZ
Bộ cánh màng (Hymenoptera)	Lưỡng bội	Đơn bội

Hầu như ở tất cả các động vật có vú, sự có mặt của một nhiễm sắc thể Y cần thiết cho sự phát triển của kiểu hình giống đực. Chẳng hạn, những người mắc hội chứng Turner (45, X) đều là nữ. Hơn nữa, những người mắc hội chứng Klinefelter (47, XXY hoặc 48, XXXY) đều là nam mặc dù họ có thể có tới hai hoặc ba nhiễm sắc thể X. Điều đó chứng tỏ nhiễm sắc thể Y có chứa gene xác định tính đực (maleness). Ngày nay ta biết rõ rằng đó chính là *nhân tố xác định tinh hoàn* (TDF = testis-determining factor) nằm trên vai ngắn nhiễm sắc thể Y (hình 4.6c). Sinclair và cs (1990) gọi

nó là gene *SRY* (sex-determining region Y).



**Hình 4.6** (a) Các nhiễm sắc thể X và Y ở động vật có vú, đại diện là người. (b) Cơ chế xác định giới tính ở người. (c) Nhiễm sắc thể Y của người với gene *SRY* nằm kề vùng giả nhiễm sắc thể thường ở đầu mút của vai ngắn.

Cần lưu ý rằng, gần đây người ta cũng đã xác định được gene "chuyển đổi" (switch gene) *SRY* trên nhiễm sắc thể Y bằng cách kiểm tra một vài cá thể hiếm hoi bị *đảo ngược giới tính* (*sex-reversed*), nghĩa là con đực XX và con cái XY. Các con đực XX thực tế có mang một mẫu của Y chứa gene xác định tính đực, còn các con cái XY thì lại thiếu hẳn vùng như vậy ở Y. Chẳng hạn, ở một vài bệnh nhân đảo ngược giới tính vốn là những người nam vô sinh rõ ràng là mang cả hai nhiễm sắc thể X và không có Y, và một số người nữ mắc hội chứng Turner thì lại mang một X và một Y rõ ràng. Các kết quả phân tích kỹ hơn (trên bốn người bệnh đảo ngược giới tính nhờ áp dụng các vật dò DNA được tạo dòng từ nhiễm sắc thể Y vào *Southern blots* toàn bộ DNA bộ gene của họ) cho thấy rằng hầu hết các trường hợp này đều có liên quan đến sự chuyển một đoạn DNA đặc thù trên nhiễm sắc thể Y (nằm sát *vùng giả nhiễm sắc thể thường* - pseudoautosomal region); và chính sự chuyển đoạn này đã gây ra hiện tượng đảo ngược giới tính. Điều này được giải thích là do các sự kiện trao đổi chéo hiếm hoi giữa X và Y trong giảm phân ở người nam hay con đực. Lưu ý rằng gene *SRY*<sup>+</sup> mã hóa một protein bám DNA và nó được bảo tồn cao độ ở các động vật có vú. John Gubbay và cs (1990) phát hiện ra một gene tương đồng ở chuột, ký hiệu là *Sry*<sup>+</sup>, không có mặt trong chuột cái XY bị đảo ngược giới tính. Ở ruồi giấm, có một thể đột biến autosome của gene "chuyển đổi", *tra* (transformation), mà khi ở trạng thái đồng hợp trong các cá thể XX biến đổi chúng thành những con đực, chứ không phải là các con cái như kỳ vọng. Tình huống ngược lại được biết ở người, trong đó các cá thể mang một đột biến trên nhiễm sắc thể X có một trạng thái gọi là *nữ hóa tinh hoàn* (testicular feminization). Trong trường hợp này, các cá thể XY về mặt kiểu hình là nữ, có ngực nở nang và âm đạo, nhưng

họ đều vô sinh. Đột biến này hiển nhiên là có liên quan tới một chất mang hormone mà nó ngăn không cho những người này tiết ra testosterone; và hậu quả là họ không phát triển được các tính trạng nam thứ cấp.

**Bảng 4.2 Tỷ lệ X/A và các kiểu hình giới tính của ruồi giấm *D. melanogaster***  
(theo Bridges 1921, trích từ Yablokov 1986)

Bộ nhiễm sắc thể <sup>(*)</sup>	Tỷ lệ X/A	Kiểu hình giới tính	Ghi chú
3X : 2A	1,50	Siêu cái	Bất thụ, tính trạng thiên về cái
4X : 3A	1,33	Siêu cái	Bất thụ, tính trạng thiên về cái
4X : 4A	1,00	Cái tứ bội	Hữu thụ
3X : 3A	1,00	Cái tam bội	Độ hữu thụ giảm
2X : 2A	1,00	Cái lưỡng bội	Hữu thụ
1X : 1A	1,00	Cái đơn bội	Bất thụ
3X : 4A	0,75	Giới tính trung gian	Bất thụ
2X : 3A	0,67	Giới tính trung gian	Bất thụ
1X : 2A	0,50	Đực lưỡng bội	Hữu thụ
2X : 4A	0,50	Đực tứ bội	Hữu thụ
1X : 3A	0,33	Siêu đực	Bất thụ, tính trạng thiên về đực

(\*) X - số nhiễm sắc thể X, A - số bộ đơn bội nhiễm sắc thể thường.

Mặc dù ở ruồi giấm *Drosophila* những con cái là XX và con đực là XY, song sự có mặt của Y không nhất thiết cho kiểu hình đực (như ở động vật có vú). Chẳng hạn, các kiểu hình ngoại lệ này đã được Calvin Bridges, một học trò xuất sắc của Morgan, phát hiện từ năm 1916 đó là các con đực XO có mắt đỏ và các con cái XXY mắt trắng. Sau đó, Bridges phát triển các nòi ruồi giấm tam bội. Khi kiểm tra các cá thể có số lượng nhiễm sắc thể X khác nhau, ông phát hiện ra rằng kiểu hình giới tính ở ruồi giấm là một hàm số của tỷ lệ giữa số lượng các nhiễm sắc thể X (X) và số lượng các bộ nhiễm sắc thể thường (A). Ví dụ, ở bảng 4.2 cho thấy các con cái bình thường có hai nhiễm sắc thể X và hai bộ nhiễm sắc thể thường, nghĩa là tỷ lệ  $X/A = 2/2 = 1$ . Các con đực bình thường có một X và hai bộ nhiễm sắc thể thường, do đó  $X/A = 1/2 = 0,5$ . Cũng vậy, các con đực và cái ngoại lệ vẫn có các tỷ lệ tương ứng là 0,5 và 1,0. Khi Bridges kiểm tra các con ruồi có ba nhiễm sắc thể X và hai bộ nhiễm sắc thể thường, tức  $3/2 = 1,5$ , ông thấy rằng chúng đều là các con cái bất dục (đôi khi gọi là "siêu cái", metafemales), trong khi các con ruồi có một X và ba bộ nhiễm sắc thể thường,  $X/A = 1/3 = 0,33$ , đều là các con đực bất dục hay gọi là dạng siêu đực (metamales). Cuối cùng, các con ruồi có hai X (XX) và ba bộ nhiễm sắc thể thường,  $X/A = 2/3 = 0,67$  và chúng có các đặc điểm kiểu hình của cả hai giới và được gọi là giới tính trung gian (intersexes). Từ các phát hiện này, Bridges đề nghị rằng, nhìn chung, các nhiễm sắc thể X

xác định giới cái và các *nhiễm sắc thể thường* (autosomes) xác định giới đực. Nói cách khác, khi tỷ lệ X/A là 1 hoặc lớn hơn, các ruồi giấm sẽ có kiểu hình cái, và khi tỷ lệ X/A là 0,5 hoặc nhỏ hơn thì chúng sẽ có kiểu hình đực. Tuy nhiên, ta cần lưu ý rằng mặc dù thậm chí nhiễm sắc thể Y không nhất thiết cho kiểu hình đực ở ruồi giấm, nhưng nó nhất thiết cần cho độ hữu thụ (các con đực XO và các con ruồi siêu đực đều bất thụ).

Ngoài ra, các sinh vật thuộc bộ cánh màng (Hymenoptera) như ong mật và ong bắp cày, chúng không có nhiễm sắc thể giới tính; trong trường hợp này, như đã nói ở chương 3, con cái là lưỡng bội và con đực là đơn bội. Ví dụ, ở ong mật (*Apis mellifera*), các ong cái bao gồm cả ong chúa lẫn ong thợ là lưỡng bội ( $2n = 32$ ) và các ong đực - đơn bội ( $n = 16$ ). Khi giảm phân, ong chúa cho các giao tử cái ( $n = 16$ ), và ong đực với hiện tượng lưỡng bội giả giảm phân cho các giao tử đực ( $n = 16$ ). Các trứng được thụ tinh ( $2n = 32$ ) phát triển thành các ong cái, chỉ một vài con đực ưu tiên mớm sữa chúa liên tục trong 3-5 ngày đầu sẽ phát triển thành ong chúa; còn các trứng không được thụ tinh sẽ phát triển thành các ong đực.

## 2. Sự xác định giới tính do môi trường (ESD)

Một trường hợp nổi bật nhất về xác định giới tính do môi trường là *Bonellia viridis*, một loại ấu trùng giun biển (Leutert 1975). Các con cái trưởng thành của loài này bám vào các mỏm đá ở đại dương. Cơ thể con cái dài hơn 10 cm, con đực rất bé (1-3 mm) sống ký sinh bên trong con cái. Các ấu trùng của *Bonellia* sống như là thành phần của *sinh vật trôi nổi* (plankton), nghĩa là, các sinh vật nhỏ bé di chuyển thụ động trong nước. Sự xác định giới tính xảy ra khi các ấu trùng được thụ tinh trên một giá thể và biến thái. Những ấu trùng được thụ tinh trong sự cách ly (sống tự do) thì phát triển thành các con cái, trong khi đó các ấu trùng chui vào trong một con cái trưởng thành thì trở thành các con đực. Một số trường hợp khác, chẳng hạn như các trứng cá sấu châu Mỹ (*Alligator*) được đưa vào nhiệt độ cao cho ra hầu hết là con đực, còn nếu ở nhiệt độ thấp cho ra hầu hết là con cái. Ở rùa thì ngược lại (xem bảng 4.3).

**Bảng 4.3 Các kiểu xác định giới tính khác nhau do môi trường (ESD)**

Loài	Cơ chế	Các giới tính	
Rùa (Turtles)	Nhiệt độ	Ấm: cái	Lạnh: đực
Cá sấu (Alligators)	Nhiệt độ	Lạnh: cái	Ấm: đực
<i>Meloidogyne incognita</i>	Mật độ quần thể	Phân tán: cái	Tập trung: đực
<i>Bonellia viridis</i>	Sự có mặt con cái	Có: đực	Không: cái

Nguồn: Hodgkin (1992), trích chọn theo sửa đổi của Kalthoff (1997, tr.662).

Ngoài ra, giới tính của một số loài cá chịu ảnh hưởng bởi ưu thế sống thành đàn, còn các loài thực vật nào đó lại cho các giới tính khác nhau tùy thuộc vào độ dài ngày hoặc các nhân tố khác ảnh hưởng lên tốc độ sinh trưởng của chúng.

Tóm lại, ở các loài sinh sản hữu tính giao phối, nói chung tỷ lệ đực/cái trên quy mô quần thể-loài là xấp xỉ 1:1, và có thể giải thích bằng cơ chế phân ly của các nhiễm sắc thể giới tính về các giao tử trong quá trình giảm phân và sự kết hợp ngẫu nhiên của chúng trong thụ tinh.

### III. Sự di truyền liên kết với giới tính (sex-linked inheritance)

Trong các chương 1 và 2 chúng ta mới chỉ đề cập đến các gene trên *nhiễm sắc thể thường* (autosomes), tức các nhiễm sắc thể tồn tại trong các tế bào soma đúng nghĩa là các cặp tương đồng; vì vậy mà các kết quả lai thuận nghịch là tương đương nhau. Tuy nhiên, một số thí nghiệm của Morgan ở ruồi giấm vào năm 1910 cho thấy các kết quả *lai thuận nghịch* (reciprocal crosses) là khác nhau. Dưới đây ta xét một số trường hợp di truyền liên kết với giới tính chủ yếu ở ruồi giấm và ở người, mối liên quan giữa di truyền liên kết giới tính và sự bất hoạt của nhiễm sắc thể X, cũng như sự di truyền của một số tính trạng bị giới hạn bởi giới tính và chịu ảnh hưởng của giới tính.

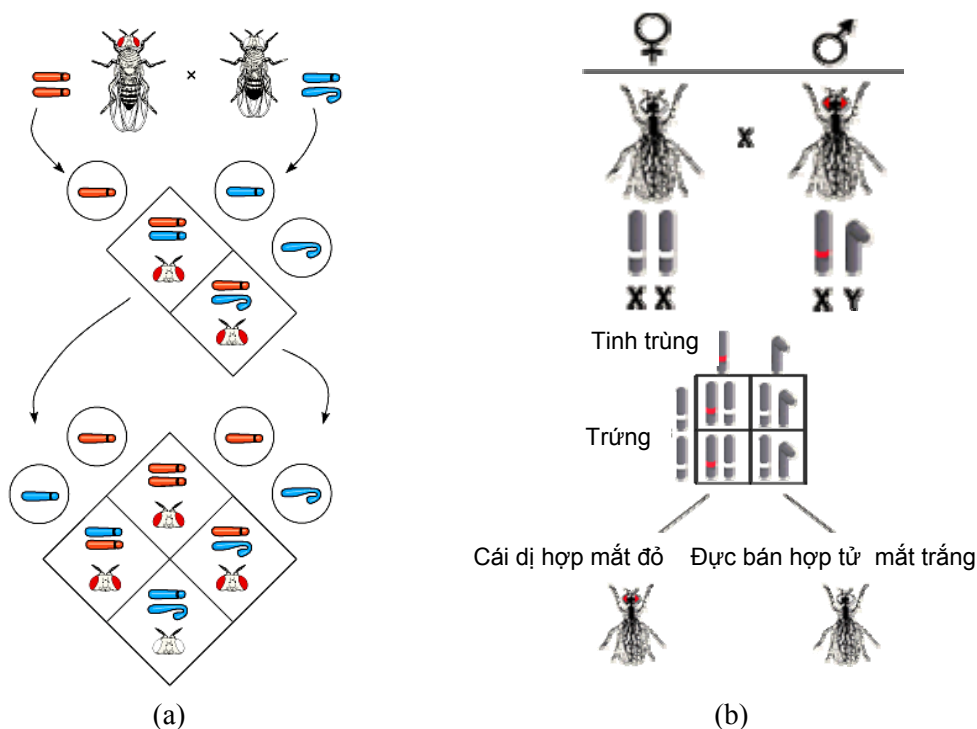
#### 1. Đặc điểm di truyền của các gen trên nhiễm sắc thể X và Y

Bản chất của các nhiễm sắc thể giới tính X và Y là khác nhau (hình 4.6). Trong khi nhiễm sắc thể X thường rất lớn và mang nhiều gene kiểm soát chủ yếu các tính trạng thường, thì nhiễm sắc thể Y lại rất bé và chứa ít gene. Vì vậy sự *di truyền liên kết với giới tính* (sex-linked inheritance), tức sự di truyền của các gen trên các nhiễm sắc thể giới tính cũng khác nhau.

##### 1.1. Sự di truyền liên kết - X

Hiện tượng di truyền liên kết giới tính được Morgan khám phá lần đầu tiên ở ruồi giấm *Drosophila*. Bình thường màu mắt của ruồi giấm có màu đỏ. Trong một quần thể *Drosophila* nuôi qua nhiều thế hệ, Morgan đã phát hiện một thể đột biến mắt trắng ở ruồi đực (xem hình 4.3a-b). Khi cho lai giữa ruồi đực mắt trắng này với ruồi cái mắt đỏ thuần chủng, ở đời con F<sub>1</sub> nhận được tất cả mắt đỏ (đúng như kỳ vọng allele mắt trắng là lặn). Sau khi cho tạp giao các ruồi F<sub>1</sub> với nhau, ở F<sub>2</sub> thu được 2.459 ruồi cái mắt đỏ, 1.011 ruồi đực mắt đỏ và 782 ruồi đực mắt trắng. Nhìn chung, tỷ lệ đỏ : trắng không gần với tỷ lệ 3:1. Tuy nhiên, kết quả bất ngờ là tất cả các con ruồi mắt trắng đều là đực (hình 4.7a)! Hơn nữa, khi cho ruồi đực mắt trắng lai với ruồi cái mắt đỏ F<sub>1</sub>, đời con sinh ra gồm cả hai loại mắt đỏ và trắng ở mỗi giới tính với tỷ lệ tương đương. Điều đó chứng tỏ tính trạng mắt

trắng phân bố ở cả hai giới, chứ không phải chỉ ở giới đực. Mặt khác, khi tiến hành phép lai nghịch giữa ruồi cái mắt trắng ở  $F_2$  trong thí nghiệm nói trên với ruồi đực mắt đỏ, kết quả thu được khác với phép lai thuận trước đây, tất cả ruồi cái mắt đỏ và tất cả ruồi đực mắt trắng (hình 4.7b).



**Hình 4.7** Các phép lai thuận nghịch về mắt trắng và mắt đỏ ở ruồi giấm. (a) Phép lai thuận:  $P = \text{cái mắt đỏ} \times \text{đực mắt trắng} \rightarrow F_1 \text{ tất cả đỏ} \rightarrow F_2 \text{ 3 đỏ : 1 trắng}$  (mắt trắng chỉ có ở giới đực); và (b) Phép lai nghịch:  $P = \text{cái mắt trắng} \times \text{đực mắt đỏ} \rightarrow F_1 \text{ gồm tất cả ruồi cái mắt đỏ và tất cả ruồi đực mắt trắng}$ .

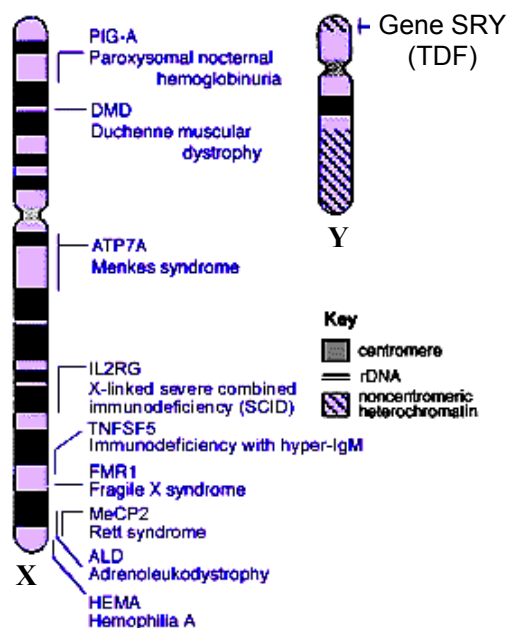
Với các kết quả đó, Morgan cho rằng các tính trạng mắt đỏ và mắt trắng này di truyền liên kết giới tính, do các allele trội ( $w^+$ ) và ( $w$ ) lặn tương ứng nằm trên nhiễm sắc thể X kiểm soát; ở con cái là XX và ở con đực là XY. Và thuật ngữ *bán hợp tử* (hemi-zygote) được ông dùng để mô tả trường hợp này.

Ta có thể tóm tắt các đặc điểm của sự di truyền liên kết X như sau: (1) Một cơ thể cái dị hợp tử sẽ truyền allele lặn cho một nửa con cái và một nửa con đực của nó. (2) Allele lặn ở con đực tồn tại ở trạng thái bán hợp tử sẽ biểu hiện ra kiểu hình lặn (nên kiểu hình lặn này phổ biến ở giới đực), và con đực sẽ truyền allele lặn này cho các con cái (female) của nó. (3) Hiện tượng di truyền allele lặn từ bố cho con gái và biểu hiện ở cháu ngoại

traï như thế rõ ràng là có sự cách quãng thể hệ, và nó được gọi là *di truyền chéo*. (4) Nói chung, việc nhận biết một tính trạng nào đó tuân theo quy luật di truyền liên kết-X có thể dựa vào các tỷ lệ kiểu hình khác nhau từ các phép lai thuận nghịch, hoặc sự phân bố không đều của các kiểu hình ở hai giới, hoặc bằng cách sử dụng "phép thử độc lập" (quy tắc nhân) như đã nói ở chương 1.

Bây giờ ta xét sơ qua sự di truyền liên kết giới tính (các gene trên X) ở người. Năm 1994, McKusick liệt kê 161 locus đã được xác định trên nhiễm sắc thể X người. Tuy nhiên, theo thống kê của OMIM, tính đến ngày 8/2/2005 con số này là 539 (OMIM 2005) trong tổng số hơn 1.400 gene chứa trong 150 triệu cặp base của nhiễm sắc thể này (NCBI 2005). Dù nhiễm sắc thể X mang hàng trăm gene như vậy nhưng rất ít gene này, nếu có, trực tiếp tác động lên giới tính. Tuy nhiên, như đã nói, sự di truyền của các gene này tuân theo các quy tắc đặc biệt, vì: (i) người nam chỉ có một nhiễm sắc thể X đơn độc; (ii) hầu như tất cả các gene trên X không có các gene tương ứng trên Y; do vậy (iii) bất kỳ gene nào trên X sẽ được biểu hiện ở nam giới, dù nó là lặn ở nữ giới. Những người nữ dị hợp tử được gọi là "*thể mang*" (carrier) vì mặc dù họ không cho thấy các triệu chứng, nhưng lại truyền gene này cho khoảng một nửa số con trai (sẽ phát triển bệnh) và một nửa cho con gái của họ (cũng là các thể mang).

Một ví dụ quen thuộc về sự di truyền liên kết-X là bệnh *máu khó đông* (hemophilia) do một allele lặn của gene hemophilia A nằm gần đầu mút vai dài gây ra (hình 4.8). Các cá thể bị bệnh này mất khả năng tổng hợp một protein (nhân tố VIII) cần thiết cho sự đông máu bình thường. Một phả hệ nổi tiếng của bệnh này bắt đầu từ gia đình nữ hoàng Victoria nước Anh (England). Một trong số các con trai của nữ hoàng bị bệnh máu khó đông, và hai trong số các cô con gái của bà là các thể dị hợp tử về gene này ( đã sinh ra



**Hình 4.8** Bản đồ nhiễm sắc thể X của người (bên trái) cho thấy vị trí một số gene bệnh, và sự không tương đồng giữa nó với nhiễm sắc thể Y.

ba cháu ngoại trai của bà đều mắc bệnh máu khó đông và bốn cháu ngoại gái tất cả đều dị hợp tử). Các allele lặn từ hai trong số bốn cô gái (cháu ngoại) dị hợp tử này đã được nhập vào các gia đình hoàng tộc của Nga và Tây Ban Nha. Còn gia đình hoàng tộc nước Anh hiện giờ do phát xuất từ một hoàng tử bình thường của nữ hoàng nên không mắc bệnh này. Hình 4.9 cho thấy gia đình của Czar Nicholas II nước Nga (Russia), trong đó người vợ Alexandra chính là cháu ngoại gái của nữ hoàng Victoria. Cậu con trai Alexis của họ (phía trước hình) mắc bệnh hemophilia và đã bị hành hình ở tuổi 14. Trong khi bốn cô con gái của họ đều bình thường. Đối với bệnh hemophilia, các kiểu gene và kiểu hình của những người nữ và nam thường được biểu diễn như sau:

Nữ	Kiểu gene	Kiểu hình	Nam	Kiểu gene	Kiểu hình
	$X^H X^H$	bình thường		$X^H Y$	bình thường
	$X^H X^h$	thể mang		$X^h Y$	bệnh
	$X^h X^h$	bệnh			

Bệnh hemophilia chỉ ảnh hưởng đến 0,004% (hay 1/25.000) trong số nam giới của quần thể, nghĩa là tần số allele này bằng 0,00004. Vậy xác suất để cho một người nữ mắc bệnh này là khoảng  $(0,00004)^2$ .

**Hình 4.9** Một bức ảnh của gia đình Hoàng đế Czar Nicholas II. Vợ ông, Czarina Alexandra, ngồi bên trái, với bốn cô con gái tất cả đều có máu đông bình thường, và riêng mỗi cậu con trai Alexis (phía trước) bị bệnh hemophilia.



Một trường hợp khác khá phổ biến là bệnh mù màu đỏ-lục (red-green colour blindness). Những người mắc bệnh này không thể phân biệt các màu mà những người khác nhìn thấy như màu lục, vàng, cam và đỏ (chúng được nhìn thấy như là một màu). Để phát hiện dạng bệnh này người ta sử dụng các sơ đồ màu chẳng hạn như *Phiếu trắc nghiệm mù màu của Ishihara* (Ishihara's tests for colour blindness). Bệnh này do một allele lặn trên X (nằm giữa hai gene hemophilia A và B) gây ra; ở nam giới có khoảng 8% mắc bệnh này trong khi ở nữ chỉ khoảng 0,4%.

Ngoài ra, *rối loạn dưỡng cơ Duchenne* (Duchenne muscular dystrophy = DMD) là một căn bệnh quái ác thảm thương mà hầu hết xảy ra ở trẻ em, cuộc đời chúng gắn chặt với chiếc xe lăn, hít thở cũng khó khăn. Hiện giờ



một người bệnh này cũng hiếm khi sống sót quá 20 tuổi. Bệnh DMD ảnh hưởng tới 1 trên 4.000 trẻ em nam, và khoảng 1/3 trong số đó là hậu quả của các đột biến tự phát trong gene này. Gene DMD nằm trên vai ngắn của nhiễm sắc thể X (xem hình 4.8). Theo số liệu hiện giờ, gene DMD là gene có kích thước lớn nhất, khoảng 2,4 triệu cặp base, tương ứng khoảng 1,5% chiều dài của nhiễm sắc thể X ( $2,4 \times 10^6 : 150 \times 10^6 = 0,016$ ).

Trước khi kết thúc phần này, ta hãy đề cập một vài ứng dụng thực tiễn của sự hiểu biết về di truyền giới tính và liên kết với giới tính. Trong chăn nuôi gia cầm, việc xác định giới tính lúc chúng còn non hoặc thậm chí ở giai đoạn trứng thật là quan trọng. Chẳng hạn, ở gà, một gene trên nhiễm sắc thể Z có một allele lặn xác định màu lông vàng ( $Z^S Z^S$  hoặc  $Z^S W$ ), và một allele trội xác định lông bạc ( $Z^S Z^S$ ,  $Z^S Z^s$  hoặc  $Z^S W$ ). Khi cho lai giữa con mái lông bạc ( $Z^S W$ ) và con trống lông vàng ( $Z^S Z^s$ ) sẽ cho đời con tất cả con trống lông bạc ( $Z^S Z^s$ ) và tất cả con mái lông vàng ( $Z^S W$ ). Rõ ràng là, với hiểu biết quy luật di truyền chéo, cho phép ta dễ dàng xác định giới tính khi gà con mới một ngày tuổi. Bạn hãy kiểm tra phép lai ngược lại (chẳng hạn  $Z^S Z^S \times Z^S W$ ) xem thử liệu phép lai này có thể dùng để xác định giới tính của các gà con mới nở hay không?

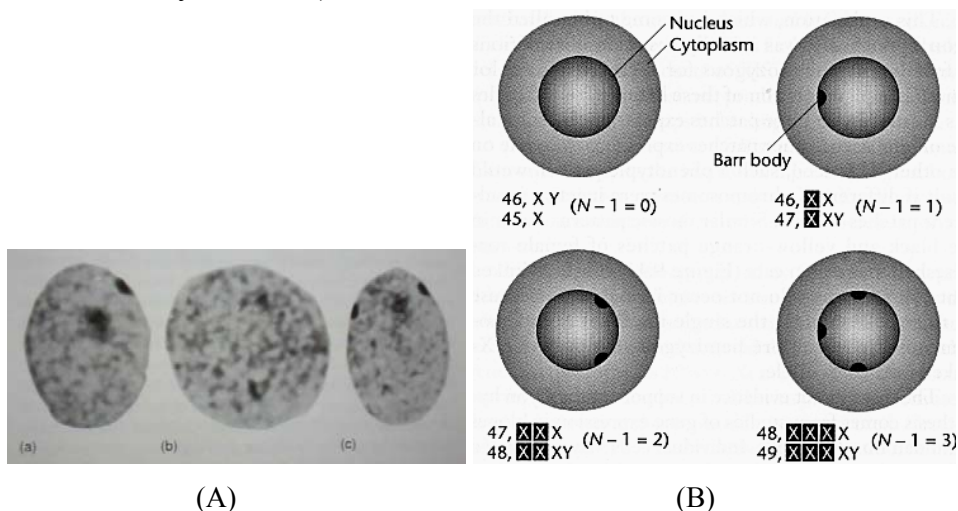
## 1.2. Sự di truyền liên kết-Y

Nói chung, nhiễm sắc thể Y rất bé, chứa ít gene. Nhiễm sắc thể Y ở ruồi giấm hầu như không mang gene, trong khi đó ở người có 48 gene đã biết trình tự trong tổng số hơn 200 gene được ước tính là chứa trong khoảng 50 triệu cặp base (OMIM-NCBI 2005). Ở hai đầu mút của nhiễm sắc thể Y có hai vùng được gọi là các *vùng giả nhiễm sắc thể thường* (pseudautosomal regions; hình 4.6c) bởi vì các gene định khu bên trong chúng (cho đến nay chỉ phát hiện được 9 gene) đều được di truyền giống như bất kỳ các gene nào thuộc nhiễm sắc thể thường. Và sự trao đổi chéo giữa X và Y chỉ có thể xảy ra ở hai vùng tương đồng rất nhỏ này của Y.

Mặc dù 95% của nhiễm sắc thể Y nằm giữa các vùng giả tương đồng, nhưng số gene được phát hiện ở đây là chưa tới 80. Một số gene này mã hóa các protein dùng chung cho tất cả các tế bào (và cả hai giới tính). Những gen còn lại mã hóa các protein hình như chỉ hoạt động trong các tinh hoàn. Gene chủ chốt nhất ở nhóm sau là gene xác định tinh hoàn *TDF*, hay gene *SRY*, định khu trên vai ngắn ngay bên ngoài vùng giả nhiễm sắc thể thường (hình 4.6c). Các tính trạng được quy định bởi các gene nằm ở vùng không tương đồng này của Y được di truyền theo *đường thẳng* (straightforward), nghĩa là: (i) Chúng chỉ biểu hiện ở nam giới; và (ii) chúng luôn luôn được truyền từ bố cho con trai.

## 2. Sự bất hoạt của nhiễm sắc thể X và một số vấn đề liên quan

Ở tất cả các động vật có vú kể cả người, nhiễm sắc thể X khác biệt với các nhiễm sắc thể khác ở chỗ chỉ có một chiếc hoạt động trong một tế bào. Các con cái (gái) có hai nhiễm sắc thể X, bình thường chỉ một chiếc là hoạt động và chiếc kia bất hoạt. Nhiễm sắc thể X bất hoạt kết vón tạo thành một cấu trúc bắt màu tối sẫm, thường có dạng thấu kính lồi bám vào mặt trong màng nhân, gọi là *thể Barr* (Barr body) (hình 4.10). Nó mang tên của người khám phá đầu tiên là Murray Barr. Các con đực bình thường chỉ có một X và ở trạng thái hoạt động trong tất cả các tế bào, do vậy không có thể Barr. Về nguyên tắc, số thể Barr bằng số nhiễm sắc thể X (ở hình 4.10B, ký hiệu là N) trừ đi một.



**Hình 4.10** (A) Ảnh chụp các nhân từ các tế bào của (a) một con cái (XX) với một thể Barr, (b) một con đực (XY) với không thể Barr nào, và (c) một con cái XXX với hai thể Barr. (B) Sơ đồ minh họa sự hình thành thể Barr trong các trường hợp khác nhau.

• **Cơ chế bất hoạt-X:** Sự bất hoạt của một nhiễm sắc thể X cần đến một gene trên nhiễm sắc thể này gọi là *XIST*. Gene *XIST* mã hóa một phân tử RNA lớn (một kiểu khác với các RNA, như mRNA, được sử dụng trong tổng hợp protein). Phân tử RNA của gene *XIST* tích tụ dọc theo nhiễm sắc thể X có chứa gene *XIST* hoạt động và gây bất hoạt tất cả (hoặc hầu như tất cả) hàng trăm gene khác trên nhiễm sắc thể đó. Một điểm cần lưu ý là, RNA của gene *XIST* không di chuyển trên nhiễm sắc thể X nào khác ở trong nhân. Các thể Barr là những nhiễm sắc thể X bất hoạt được "son phết" (painted) bởi RNA *XIST*. (Rastan 1994; Lee và Jaenisch 1997).

Trong những giai đoạn đầu của sự phát triển phôi cái, locus *XIST* trên mỗi trong số hai nhiễm sắc thể X của nó được biểu hiện nhưng RNA *XIST* nhanh chóng bị bẻ gãy. Rồi một điều gì đó xảy ra để thông báo trước sự

cân bằng đối với một chiếc này hoặc chiếc kia của các nhiễm sắc thể X. Sự phiên mã tiếp tục trên một trong số các nhiễm sắc thể X, dẫn tới việc tích tụ của RNA *XIST* và làm biến đổi nhiễm sắc thể đó thành ra một thể Barr bất hoạt. Sự phiên mã của *XIST* dừng lại trên nhiễm sắc thể X khác cho phép tất cả hàng trăm gene khác của nó được biểu hiện. Sự tắt ngấm của locus *XIST* trên nhiễm sắc thể X hoạt động được tiến hành bằng việc methyl hóa các *trình tự điều hòa XIST* (*XIST* regulator sequences). Sự *methyl hóa DNA* (DNA methylation) thường xảy ra trong biểu hiện gene cho nên sự methyl hóa khóa chặt hẳn sự biểu hiện của gene *XIST* và nó cho phép biểu hiện liên tục của tất cả các gene liên kết-X còn lại. Ngoài ra, sự bất hoạt-X ở phôi cái là hoàn toàn ngẫu nhiên đối với một trong hai X (hình 4.10B), và nó xảy ra sớm trong quá trình phát triển phôi (ở giai đoạn khoảng 2.000 tế bào). Không thể dự đoán liệu X từ bố hay X từ mẹ bị bất hoạt trong một tế bào nào đó. Nhưng đây không phải là trường hợp cho các *màng ngoài phôi* (extraembryonic membranes); ở tất cả các tế bào của các màng này thì chỉ có nhiễm sắc thể X của người bố là bị bất hoạt (Rastan 1994; Lee và Jaenisch 1997; Kimball 2004).

- **Một số gene trên X thoát khỏi sự bất hoạt:** Vấn đề đặt ra là, với 18 gene được phát hiện trên Y cũng như X thì sao? Sẽ không có nhu cầu nào cho các con cái (females) làm bất hoạt một bản sao của các gene đó để giữ cân bằng với tình huống ở các con đực. Thế thì chúng quản lý điều tiết việc này ra sao vẫn còn chưa được khám phá (Kimball 2004).

- **Chẩn đoán các bất thường nhiễm sắc thể X:** Rõ ràng là, sự hiểu biết về hiện tượng này có ứng dụng thực tế to lớn trong chẩn đoán trước sinh hoặc kiểm tra các bệnh tật ở những người trưởng thành có thể có liên quan đến các đột biến lệch bội nhiễm sắc thể X. Chẳng hạn, để chẩn đoán trước sinh, người ta sử dụng phương pháp *chọc ối* (amniocentesis), rồi làm tiêu bản kiểm tra sự bất thường nhiễm sắc thể và thiết lập kiểu nhân. Công thức tổng quát cho trường hợp này là:

$$\text{Số nhiễm sắc thể X} = \text{số thể Barr} + 1$$

Như vậy, những người nam bình thường và các cá thể mắc hội chứng Turner (XO) không có thể Barr nào; những người nữ bình thường và các cá thể mắc hội chứng Klinefelter (XXY) có một thể Barr; những cá thể XXX có hai thể Barr v.v. (xem Hình 4.10B).

- **Giải thích sự hình thành các thể khảm di truyền:** Với cơ chế bất hoạt ngẫu nhiên của các nhiễm sắc thể X, các tế bào của con cái sẽ là *thể khảm di truyền* (genetic mosaics) về hai nhiễm sắc thể X của nó. Đây là nguyên nhân dẫn tới sự hình thành màu lông "sọc vằn" hay tam thể ở mèo cái (*Felix catus*). Ta có thể hình dung mối quan hệ giữa kiểu gene và kiểu

hình ở mèo liên quan tới một gene xác định màu lông trên nhiễm sắc thể X như sau: Nếu quy ước allele B - màu đen và allele O- màu da cam, ta có:

Cái	Kiểu gene	Kiểu hình	Đực	Kiểu gene	Kiểu hình
	$X^B X^B$	đen		$X^B Y$	đen
	$X^O X^O$	da cam		$X^O Y$	da cam
	$X^B X^O$	tam thể			

Như vậy, để cho mèo đực là tam thể rõ ràng là nó phải có hai nhiễm sắc thể X,  $X^B X^O Y$ , nghĩa là thể dị bội về nhiễm sắc thể giới tính. Điều này rất hiếm khi xảy ra trên thực tế. Đến đây ta thấy rằng việc giải thích sự xuất hiện mèo tam thể bằng cơ chế tương tác giữa các gene allele theo kiểu đồng trội trong nhiều tài liệu trước đây và cả hiện nay là không đúng! Đó cũng là lý do tại sao chúng tôi đưa mục này vào đây mà không đặt nó trong chương 3 (khi xem xét các thể đột biến lệch bội) hoặc sau mục II-1 của chương này như cách làm truyền thống.



**Hình 4.11** Sự hình thành màu sắc bộ lông ở mèo tam thể được giải thích bằng sự bất hoạt ngẫu nhiên của các nhiễm sắc thể X.

### 3. Các tính trạng giới hạn bởi giới tính và chịu ảnh hưởng của giới tính

Như đã nói ở trên, thông thường các tính trạng có sự phân bố không đồng đều giữa hai giới tính hoặc tập trung ở một giới làm ta liên tưởng tới quy luật di truyền liên kết với giới tính do gene hoặc trên X (di truyền chéo) hoặc trên Y (di truyền thẳng). Tuy nhiên, một số tính trạng được biểu hiện một cách khác biệt ở cả hai giới nhưng không phải là liên kết với giới tính. Đó là những tính trạng bị giới hạn bởi giới tính hoặc là chịu ảnh hưởng của giới tính.

(1) *Các tính trạng bị giới hạn bởi giới tính* (sex-limited traits): đó là những tính trạng chỉ biểu hiện ở một giới. Có nhiều tính trạng như vậy liên quan với các đặc điểm sinh sản ở các động vật cái. Chẳng hạn, khả năng cho sữa ở các gia súc như trâu-bò, khả năng đẻ trứng ở gia cầm và lớp chim nói chung, hoặc tập tính đẻ trứng ở các côn trùng.

**Bảng 4.4** Ví dụ về tính trạng chịu ảnh hưởng của giới tính ở Cừu

	<i>HH</i> (Suffolk)	<i>Hh</i> (Con lai)	<i>hh</i> (Dorset có sừng)
Con đực	Không sừng	Có sừng	Có sừng
Con cái	Không sừng	Không sừng	Có sừng



Rõ ràng là tỷ lệ  $F_2$  được quan sát và kỳ vọng theo tỷ lệ 9:3:3:1 là hoàn toàn không khớp nhau; nếu kiểm tra bằng phương pháp  $\chi^2$ , sự sai khác này cao một cách đáng kể và không phù hợp với giả thuyết phân ly độc lập. Hai ông đã cố gắng đưa nó về tỷ lệ 7:1:7:1 để giải thích bằng kiểu tương tác gene, nhưng không thành công. Và cuối cùng, họ đã đưa ra gợi ý rằng bởi vì hai kiểu hình bố mẹ vượt quá mức cho phép ở  $F_2$ , có lẽ có một sự nối kết giữa các allele dạng bố mẹ. Đối với hiện tượng gây nhiễu lên sự phân ly độc lập này, họ gọi là kiểu *kết nối* (coupling). Mặt khác, họ cho rằng số lượng thấp của các kiểu  $F_2$  vốn có một kiểu hình trội ở gene này và một kiểu hình lặn ở gene kia, là do ái lực âm tính tự nhiên của các allele trội và lặn, và họ gọi là kiểu *đẩy nhau* (repulsion).

### 1. Khám phá về sự trao đổi chéo ở ruồi giấm

Để giải thích về mặt vật lý các quan sát của Bateson và Punnett, vài năm sau đó Morgan đã tiến hành hàng loạt thí nghiệm sử dụng đồng thời hai gene ở *Drosophila*, mà mỗi gene đều có hai allele lặn và trội (ký hiệu allele dựa theo tên của thể đột biến lặn trong tiếng Anh như đã đề cập). Có thể kể một số thí nghiệm mà ông đã làm sau đây: (i) một gene kiểm soát màu mắt ( $pr$ : tím và  $pr^+$ : đỏ đại) và gene kia kiểm soát cánh ( $vg$ : ngắn và  $vg^+$ : dài bình thường); (ii) màu mắt ( $w$ : trắng và  $w^+$ : đỏ) và dạng cánh ( $m$ : cánh bé và  $m^+$ : cánh bình thường); hoặc (iii) màu thân ( $e$ : mun và  $e^+$ : xám) và dạng cánh ( $c$ : cánh cong và  $c^+$ : cánh thẳng) v.v... Bây giờ ta hãy xét một trường hợp quen thuộc, đó là: gene quy định màu sắc thân (allele lặn  $b$ : đen và allele trội  $B$ : xám) và gen xác định chiều dài cánh ( $vg$ : cánh ngắn và  $Vg$ : cánh dài bình thường).

• Morgan cho lai giữa hai dòng ruồi thuần chủng thân xám, cánh dài ( $BBVgVg$ ) và thân đen, cánh ngắn ( $bbvvgv$ ), ở đời con  $F_1$  ông thu được tất cả các con đực và cái đều có thân xám, cánh dài ( $BbVgvg$ ). Nhưng khi đem lai phân tích các con đực và cái  $F_1$ , kết quả hoàn toàn khác nhau.

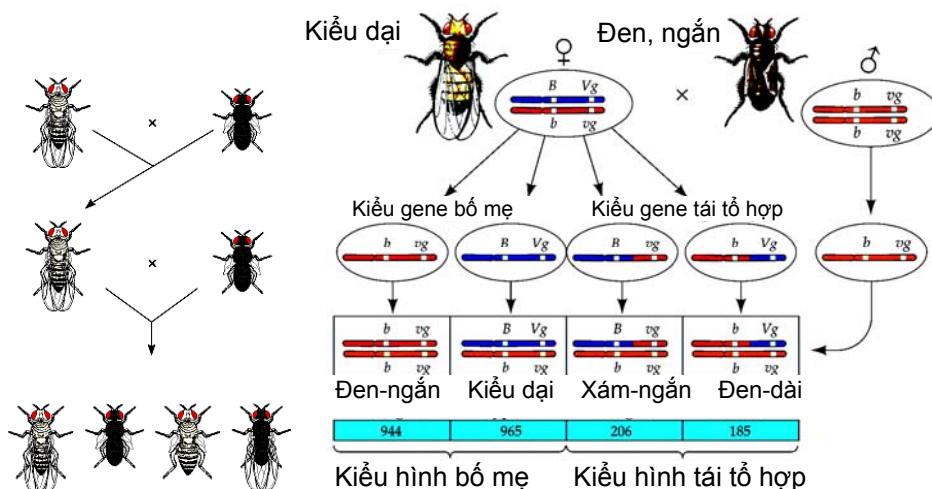
Trước tiên, ta xét phép lai phân tích giữa các con cái xám-dài  $F_1$  với con đực đen-ngắn. Kết quả các phép lai này cho đời con có đủ cả bốn kiểu hình (xem hình 4.13a), nhưng với tỷ lệ quan sát khác xa với tỷ lệ kỳ vọng 1:1:1:1 (xem hình 4.13b), như sau:

Kiểu hình	xám-dài	đen-ngắn	xám-ngắn	đen-dài	Tổng
Số liệu quan sát	965	944	206	185	2300
Số liệu kỳ vọng	575	575	575	575	2300
	Các kiểu hình bố mẹ		Các kiểu hình tái tổ hợp		

Từ kết quả đời con của lai phân tích ruồi cái  $F_1$  còn cho thấy các kiểu hình bố mẹ ban đầu là xấp xỉ nhau, chiếm 83% ( $[965 + 944]: 2300 =$

0,83); các kiểu hình tái tổ hợp cũng có số lượng gần như tương đương, chiếm 17% ([206 + 185]: 2300 = 100% - 83% = 0,17). Điều này chỉ có thể lý giải nếu như các gene nằm trên cùng nhiễm sắc thể, hay gọi là liên kết (linkage), và trong quá trình giảm phân ở ruồi cái có xảy ra sự *trao đổi chéo* (crossing over) hay *tái tổ hợp* (recombination) giữa hai gene này.

Thật vậy, trong phép lai phân tích này, ruồi đực đen-ngắn ( $bvg/bvg$ ) chỉ cho một loại giao tử chứa cả hai allele lặn,  $bvg$ . Từ đây suy ra ruồi cái  $F_1$  đã xảy ra trao đổi chéo và cho bốn loại giao tử tương ứng với tỷ lệ các kiểu hình ở đời con: (i) Các kiểu bố mẹ (*parental*) mỗi kiểu có tỷ lệ trung bình là 41,5% ( $BVg = bvg = 0,415$ ); và (ii) Các kiểu tái tổ hợp (*recombinant*) mỗi kiểu có tỷ lệ trung bình là 8,5% ( $Bvg = bVg = 0,085$ ). Như thế, dựa vào thành phần gene của P hoặc thành phần allele của các giao tử chiếm tỷ lệ cao (tức các kiểu bố mẹ) ta dễ dàng xác định được kiểu gene của  $F_1$  là  $BVg/bvg$  (kiểu *cis*).



**Hình 4.13** Kết quả của các phép lai P và lai phân tích ruồi cái  $F_1$  (trái) và giải thích bằng cơ sở tế bào học của kết quả lai phân tích ruồi cái  $F_1$ .

• Trong thí nghiệm thứ hai, Morgan lai hai kiểu gene, mỗi kiểu là đồng hợp về allele trội ở một gene này và đồng hợp về allele lặn ở một gene kia ( $BBvgvg \times bbVgVg$ ). Các ruồi  $F_1$  dị hợp tử kép ( $BbVgvg$ ) sau đó được lai với dòng kiểm tra  $bbvgvg$ . Kết quả của phép lai này, so với thí nghiệm trước, cho thấy ở đời con có sự đảo ngược giữa nhóm các kiểu bố mẹ và nhóm các kiểu tái tổ hợp; nghĩa là, số lượng đời con mang các tính trạng trội đơn ( $Bvg/bvg$  và  $bVg/bvg$ ) cao hơn nhiều so với kỳ vọng, trong khi số lượng đời con mang các tính trạng trội kép ( $BVg/bvg$ ) và các tính trạng lặn kép ( $bvg/bvg$ ) thấp hơn nhiều so với kỳ vọng. Như vậy,  $F_1$  có kiểu gene  $Bvg/bVg$  (kiểu *trans*).

• Ghi nhớ:

(1) Sự kiện trao đổi chéo thường được chỉ ra trên các sơ đồ bằng một 'dấu chéo' (×) nối giữa các nhiễm sắc thể tương đồng (hình 4.14).

(2) Đối với trường hợp dị hợp tử kép về hai cặp gen liên kết, có hai cách sắp xếp các allele trên một cặp nhiễm sắc thể:  $\frac{AB}{ab}$  và  $\frac{Ab}{aB}$ , mà có thể viết đơn giản là:  $AB/ab$  và  $Ab/aB$ .

- Cách sắp xếp các allele trong đó hai allele trội trên một nhiễm sắc thể và hai allele lặn trên chiếc kia,  $AB/ab$ , được gọi là kiểu kết nối (*coupling*) hay kiểu đều (*cis*). Với kiểu gene này,  $\underline{AB}$  và  $\underline{ab}$  là các giao tử thuộc kiểu cha mẹ, còn  $\underline{Ab}$  và  $\underline{aB}$  là các giao tử thuộc kiểu tái tổ hợp.

- Cách sắp xếp các allele trong đó mỗi nhiễm sắc thể mang một allele trội và một allele lặn sao cho trong thể dị hợp tử kép chúng ở các vị trí chéo nhau,  $Ab/aB$ , được gọi là kiểu đẩy nhau (*repulsion*) hay kiểu lệch (*trans*). Với kiểu gene này,  $\underline{Ab}$  và  $\underline{aB}$  là các giao tử thuộc kiểu cha mẹ, còn  $\underline{AB}$  và  $\underline{ab}$  là các giao tử thuộc kiểu tái tổ hợp.



**Hình 4.14** Mô hình một trao đổi chéo đơn (trái) và cách biểu diễn nó.

(3) Sự tái tổ hợp của các gene liên kết ở kiểu đều (*cis*) cũng như kiểu chéo (*trans*) từ cả hai phép lai phân tích nói trên và ở các phép lai thuận nghịch nói chung đều xảy ra với tần số gần như nhau;

(4) Tổng tỷ lệ của các kiểu tái tổ hợp ở đời con của phép lai phân tích được gọi là tần số tái tổ hợp, và tần số này đặc trưng cho mỗi cặp gene riêng biệt (như sẽ thảo luận dưới đây).

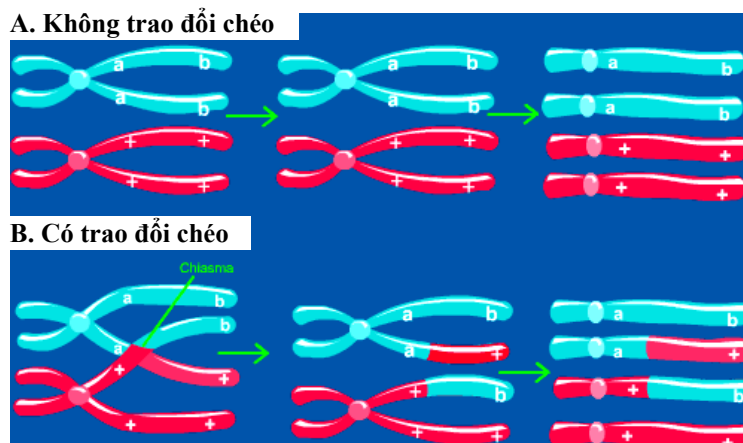
## 2. Liên kết gene hoàn toàn (hay giảm phân không có trao đổi chéo)

Bây giờ ta trở lại thí nghiệm đầu tiên của Morgan. Khi thực hiện phép lai phân tích giữa các con đực xám-dài  $F_1$  với con cái đen-ngắn, đời con chỉ có hai kiểu hình xám-dài và đen-ngắn với tỷ lệ xấp xỉ 1:1, chứ không phải bốn kiểu với tỷ lệ đều nhau như dự đoán. Điều này chứng tỏ ruồi đực  $F_1$  chỉ cho hai loại giao tử với tỷ lệ tương đương ( $\underline{BVg} = \underline{bv}g = 50\%$ ), tương ứng với hai kiểu hình đời con (vì ruồi cái đen-ngắn chỉ cho một loại giao tử chứa hai allele lặn,  $\underline{bv}g$ ); nghĩa là các gene này cùng nằm trên một nhiễm sắc thể và giữa chúng có sự liên kết hoàn toàn (*complete linkage*). Nói cách khác, tái tổ hợp không xảy ra ở ruồi giấm đực.



Giảm phân mà không có bất kỳ sự trao đổi chéo nào chỉ xảy ra ở một số ít loài. Ở các sinh vật này, tại kỳ giữa I các nhiễm sắc thể tương đồng nằm dọc bên nhau ở mặt phẳng của thoi. Chúng phân tách theo cách thông thường ở kỳ sau I, và giảm phân tiến hành một cách bình thường sau đó. Kiểu giảm phân này thấy có ở một vài côn trùng, kể cả các con đực của ruồi giấm, và ở một số thực vật có hoa. Ở tất cả các loài còn lại có ít nhất một trao đổi chéo được hình thành trong mỗi thể lưỡng trị ở kỳ trước I.

Tóm lại, ở tất cả các loài sinh sản hữu tính, trong quá trình giảm phân có thể không xảy ra trao đổi chéo hoặc có ít nhất một trao đổi chéo được hình thành trong mỗi thể lưỡng trị ở kỳ trước I. Đối với hai cặp gene bất kỳ trên một cặp nhiễm sắc thể tương đồng, các sản phẩm tương ứng được tạo thành trong các giao tử được minh họa ở hình 4.15 dưới đây.



**Hình 4.15** Giảm phân không có trao đổi chéo (A) và có một trao đổi chéo đơn (B), và các sản phẩm được tạo thành đối với hai cặp gene bất kỳ trên một cặp nhiễm sắc thể tương đồng.

## V. Trao đổi chéo và lập bản đồ di truyền

### 1. Tần số tái tổ hợp

Sự trao đổi chéo là một quá trình trao đổi giữa các nhiễm sắc thể trong giảm phân (kết hợp với sự hình thành giao tử một cách bình thường) cho ra các tổ hợp tính trạng mới, gọi là các biến dị tổ hợp (xem hình 4.13). *Tần số tái tổ hợp* (frequency of recombination or rate of recombination), còn được gọi là *trị số trao đổi chéo* (cross over value), ký hiệu là  $r$ , được tính bằng tỷ lệ phần trăm của các thể tái tổ hợp sinh ra trong một phép lai phân tích, theo công thức sau:

$$r = \frac{n}{m} \times 100$$

trong đó:  $n$  - số lượng các cá thể tái tổ hợp được sinh ra, và  $m$  - tổng số cá thể của đời con của phép lai phân tích; và  $r$  - tần số tái tổ hợp (thường gọi là *tần số hoán vị gene*), là một số hữu tỷ thỏa mãn miền giới hạn  $[0; 0,5]$ , có thể biểu diễn bằng số thập phân hoặc phần trăm.

Ở các sinh vật mà chỉ số trao đổi chéo đã được nghiên cứu rộng rãi, chẳng hạn như ngô và ruồi giấm, tỷ lệ của các thể tái tổ hợp sinh ra trong một phép lai cụ thể là khá ổn định. Tuy nhiên, các tỷ lệ này là khác nhau đối với các gene khác nhau trên cùng một nhiễm sắc thể. Điều này được giải thích dựa trên cơ sở rằng mỗi một gene có một vị trí cố định (tức locus của nó) trên một nhiễm sắc thể cụ thể, và rằng sự trao đổi chéo có thể xảy ra giữa các gene nằm xa nhau. Với ví dụ về hai gene  $b$  và  $vg$  ở ruồi giấm nói trên, ta có thể nói rằng các locus này nằm khá gần nhau, bởi vì có rất ít các thể tái tổ hợp được tạo ra. Áp dụng công thức tính tần số tái tổ hợp ở trên, ta tính được:  $r = \frac{206 + 185}{2300} \times 100 = 17\%$ . Từ đây suy ra tỷ lệ

kỳ vọng của các loại giao tử tái tổ hợp và không tái tổ hợp (dạng bố mẹ):

giao tử kiểu tái tổ hợp  $Bvg = bVg = 17\%: 2 = 8,5\%$

giao tử kiểu bố mẹ  $BVg = bv g = (100\% - 17\%): 2 = 41,5\%$

Một cách tổng quát, dựa theo tần số tái tổ hợp ( $r$ ), ta có thể biểu diễn tỷ lệ của các loại giao tử tái tổ hợp và không tái tổ hợp đối với hai kiểu gene dị hợp tử đều ( $AB/ab$ ) và dị hợp tử chéo ( $Ab/aB$ ) ở bảng 4.5.

**Bảng 4.5 Tỷ lệ của các loại giao tử**

Giao tử	Kiểu gene bố mẹ	
	$AB/ab$	$Ab/aB$
$AB$	$\frac{1}{2}(1 - r)$	$\frac{1}{2} r$
$ab$	$\frac{1}{2}(1 - r)$	$\frac{1}{2} r$
$Ab$	$\frac{1}{2} r$	$\frac{1}{2}(1 - r)$
$aB$	$\frac{1}{2} r$	$\frac{1}{2}(1 - r)$
Tổng	1	1

Bây giờ ta đề cập một ít về mối quan hệ giữa hai gene không allele trong quá trình giảm phân tạo giao tử, mà ta có thể kiểm tra bằng phương pháp kinh điển được áp dụng trong nghiên cứu di truyền học là lai phân tích. Ngoài ra, cũng có thể sử dụng các phương pháp khác, không được phổ biến lắm hoặc khá phức tạp như là tự thụ phân hoặc tạp giao.

Đối với hai gene ở trạng thái dị hợp tử kép quy định hai cặp tính trạng tương phản khác nhau, trong một phép lai phân tích:

(i) Nếu như hai gene nằm trên hai nhiễm sắc thể khác nhau, sẽ có bốn kiểu giao tử (hay kiểu hình) được tạo thành với tỷ lệ ngang nhau;

(ii) Nếu như hai gene liên kết hoàn toàn trên một nhiễm sắc thể, chỉ có thể tạo ra hai loại giao tử (hay kiểu hình) với tỷ lệ ngang nhau ( $r = 0$ );

(iii) Nếu như hai gene liên kết không hoàn toàn trên một nhiễm sắc thể, có thể tạo ra bốn loại giao tử (hay kiểu hình) thuộc hai nhóm, trong đó các kiểu giao tử tái tổ hợp mỗi kiểu chiếm tỷ lệ bằng  $r/2$ , và các kiểu giao tử bố mẹ mỗi kiểu chiếm tỷ lệ bằng  $(1-r)/2$ ;

(iv) Tần số tái tổ hợp ( $r$ ) biến thiên trong khoảng  $0 \rightarrow 0,5$ ; thông thường  $r$  thấp hơn 50%, phụ thuộc vào khoảng cách vật lý giữa hai gene trên nhiễm sắc thể. Trong các trường hợp cực đoan, trị số  $r$  có thể bằng zero ( $r = 0$ ), nếu như hai gene liên kết hoàn toàn hoặc nằm gần sát nhau; và trị số  $r$  này cũng có thể đạt cực đại, nghĩa là  $r = 0,5$  hay 50%, trong trường hợp các gene nằm cách nhau hơn 50 đơn vị bản đồ (như chúng ta sẽ đề cập ngay dưới đây) hay nói cách khác, tất cả các tế bào sinh giao tử trải qua giảm phân đều xảy ra trao đổi chéo giữa hai gene được xét đến. Khi đó tỷ lệ các giao tử (hay tỷ lệ phân ly kiểu hình) trùng với trường hợp các gene phân ly độc lập. Tuy nhiên, rất hiếm khi xảy ra như vậy.

## 2. Bản đồ di truyền

Việc thiết lập bản đồ gene của các nhiễm sắc thể lần đầu tiên được áp dụng ở ruồi giấm *Drosophila*, được đề xuất vào năm 1913 bởi Alfred Sturtevant, một môn đệ của Morgan. Theo ông, ta có thể dựa vào các tần số tái tổ hợp của các gene thu được trong các phép lai phân tích để mô tả mối quan hệ vật lý của các gene trên một nhiễm sắc thể theo trật tự tuyến tính, gọi là *bản đồ liên kết* (linkage map) hay *bản đồ di truyền* (genetic map). *Khoảng cách bản đồ* (map distance) giữa hai gene, ví dụ  $b$  và  $vg$  là 17%, được coi là cách nhau 17 đơn vị bản đồ (map unit; viết tắt là: m.u.); hay nói cách khác, 1 đơn vị bản đồ là 1% tái tổ hợp. Các nghiên cứu về sau cho thấy rằng khoảng cách di truyền được đo bằng phương pháp thống kê (tức tỷ lệ phần trăm tái tổ hợp, do Morgan và Sturtevant đề xuất) nói chung là giống với các khoảng cách trên nhiễm sắc thể đo được về mặt tế bào học hay hóa sinh học.

Nguyên tắc chung của việc xây dựng bản đồ nhiễm sắc thể ở một loài nào đó, theo Morgan và các đồng sự của ông, có thể tóm tắt thế này:

(1) Xác lập số nhóm liên kết (số nhiễm sắc thể đơn bội) của loài. Điều này có thể tiến hành bằng cách thực hiện hàng loạt các phép lai có thể được để xác định các mối quan hệ (độc lập và liên kết) trong số hàng loạt gene, kết hợp với các nghiên cứu tế bào học (đếm số lượng nhiễm sắc thể).

(2) Xác định thành phần gene của mỗi một nhóm liên kết (dựa vào các kết quả lai khác nhau).

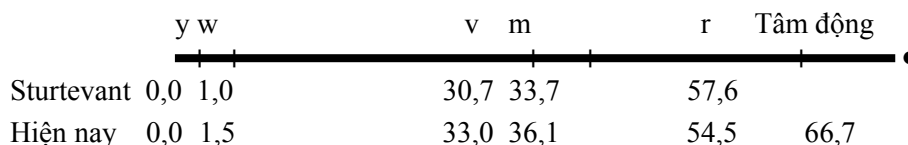
(3) Xác định vị trí, trật tự và khoảng cách giữa các gene trên mỗi nhóm liên kết. Nguyên tắc chung là, giữa các gene liên kết có tần số tái tổ hợp càng thấp chừng nào chúng tỏ chúng nằm càng gần nhau chừng ấy; theo chuỗi suy luận đó ta có thể biết được vị trí của một gene khởi đầu và các gene kế tiếp theo chiều dọc trên một nhiễm sắc thể, và đặt gene khởi đầu nằm ở đầu mút của vai ngắn ứng với vị trí zero (0,0); sau đó dùng phương pháp cộng dồn để biểu diễn vị trí và khoảng cách trên bản đồ của các gene kế tiếp với so với gene đầu mút này. Khi đó ta có được bản đồ của một nhóm liên kết. Ta có thể hình dung nó như một đoạn thẳng nằm ngang có nhiều vạch chỉ ra trật tự thẳng hàng của các gene, mỗi vạch ứng với một locus-gene cụ thể (phía trên vạch là ký hiệu và có thể cả tên đầy đủ của thể đột biến gene đó và phía dưới vạch là vị trí trên bản đồ của gene so với gene khởi đầu, được ghi bằng một con số cụ thể; hình 4.17a). Nếu ta dựng đứng bản đồ lên, theo nguyên tắc, ký hiệu các gen sẽ nằm về phía bên phải và khoảng cách bản đồ của các gene nằm phía bên trái (hình 4.17b).

Ở đây, có một số điểm cần lưu ý: (i) Độ chính xác của bất kỳ bản đồ liên kết nào được thiết lập trước đây đều có thể thay đổi và ngày càng chính xác hơn; đó là do các phát hiện bổ sung những gene mới nằm trung gian giữa các gene đã được định vị cùng với các tần số tái tổ hợp của chúng; (ii) Riêng ruồi giấm, nhiễm sắc thể X được Morgan gán cho số I, kể đến là các số II, III và IV tương ứng với sự nhỏ dần về kích thước. Ngày nay, mặc dù hệ thống đánh số nhiễm sắc thể có thay đổi, song đối với ruồi giấm vẫn được giữ nguyên để biểu thị lòng tôn kính đối với Morgan - người sáng lập thuyết di truyền nhiễm sắc thể; và (3) Morgan đã cải tiến cách lập bản đồ vốn công kênh phức tạp nói trên bằng phương pháp lai phân tích ba điểm (như sẽ được thảo luận dưới đây).

**Bảng 4.6 Tần số tái tổ hợp của năm cặp gene liên kết-X**

Các gen	Tần số tái tổ hợp
thân vàng (y) – mắt trắng (w)	$214/21.736 = 0,010$
thân vàng (y) – mắt đỏ tươi (v)	$1.464/4.551 = 0,322$
mắt trắng (w) – mắt đỏ tươi (v)	$471/1.584 = 0,297$
mắt đỏ tươi (v) – cánh bé (m)	$17/573 = 0,030$
mắt trắng (w) – cánh bé (m)	$2.062/6.116 = 0,337$
mắt trắng (w) – cánh không phát triển (r)	$406/898 = 0,452$
mắt đỏ tươi (v) – cánh không phát triển (r)	$109/405 = 0,269$

Chẳng hạn, bằng cách sử dụng số liệu ở bảng 4.6, Sturtevant đã xác định các mối quan hệ vật lý của năm gene này và gợi ý rằng chúng xếp theo một đường thẳng. Cách làm như sau: (1) Hai gene  $y$  và  $w$  có tần số nhỏ nhất ( $r_{y-w} = 1\%$ ) nên nằm gần nhau. (2) vì  $r_{w-v} = 29,7\% < r_{y-v} = 32,2\% \Rightarrow$  trật tự ba gene này phải là  $y-w-v$ . (3) Gene gần nhất với  $v$  là  $m$  ( $r_{v-m} = 3\%$ ), mà  $r_{w-m} \approx r_{w-v} + r_{v-m} \Rightarrow m$  phải nằm bên phải  $v$ . (4)  $r_{v-r} = 26,9\% < r_{w-r} = 45,2\% \Rightarrow r$  phải nằm bên phải  $v$  và  $m$ . Vậy, trật tự của năm gene liên kết trên X là:  $y-w-v-m-r$ . Từ đây, Sturtevant xây dựng nên một bản đồ bắt đầu bằng  $y$  ở vị trí 0,0 về phía bên trái và sử dụng các tần số tái tổ hợp giữa các gene kề nhau để biểu thị vị trí của chúng trên bản đồ ứng với từng con số cụ thể. Theo đó,  $w$  ở vị trí là 1,0,  $v$  ở vị trí 30,7 ( $= 1,0 + 29,7$ ),  $m$  ở vị trí 33,7 ( $= 30,7 + 3,0$ ), và  $r$  ở vị trí 57,6 ( $= 30,7 + 26,9$ ). So với bản đồ hiện nay, vị trí năm gene trên bản đồ của Sturtevant có hơi khác một chút bởi vì có nhiều gene trung gian được phát hiện bổ sung (hình 4.16).

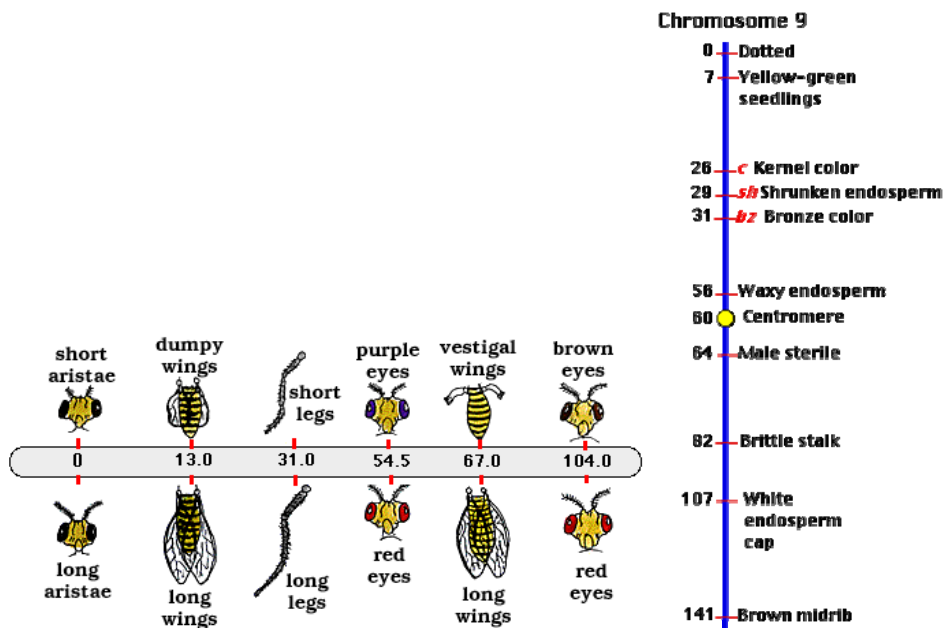


**Hình 4.16** Bản đồ Sturtevant về năm gene liên kết-X và so với hiện nay.

Dưới đây cho thấy bản đồ của nhiễm sắc thể số 2 ở ruồi giấm (hình 4.17a) và nhiễm sắc thể số 9 ở ngô (hình 4.17b) với những gene chính yếu. Chẳng hạn, trên hình 4.17a, hình ống ở giữa tượng trưng cho nhiễm sắc thể số 2 ruồi giấm chỉ ra vị trí của sáu gene; phía trên là các kiểu hình đột biến, và phía dưới là các kiểu hình bình thường. Từ trái sang: vị trí 0 - gene  $al$  (aristaleless antenna) chỉ đột biến râu ngắn đối với râu dài bình thường; vị trí 13 - gene  $dp$  (dumpy wings) chỉ đột biến cánh xén với cánh dài; vị trí 31 - gene  $d$  (dach's) chỉ đột biến chân ngắn với chân dài), vị trí 54,5 - gene  $pr$  (purple eyes) chỉ đột biến mắt đỏ tía với mắt đỏ), vị trí 67,0 - gene  $vg$  (vestigial) chỉ đột biến cánh ngắn với cánh dài), vị trí 104 - gene  $bw$  (brown eyes) chỉ đột biến mắt nâu với mắt đỏ).

Cần lưu ý rằng, ngày nay, sau khi định vị bản đồ các gene xác định bảy tính trạng mà Mendel đã nghiên cứu ở đậu Hà Lan, người ta không khỏi ngạc nhiên tại sao Mendel lại không quan sát hiện tượng liên kết trong các phép lai hai tính của mình. Mặc dù Mendel chọn ra bảy tính trạng và có bảy nhiễm sắc thể ở đậu Hà Lan ( $2n = 14$ ), hai trong số các tính trạng ông nghiên cứu được xác định bởi các gene trên nhiễm sắc thể số 1 và ba tính trạng được xác định bởi các gene trên nhiễm sắc thể số 4 (bảng 4.7). Thực ra, Mendel chỉ chọn nghiên cứu một vài tính trạng và có lẽ, thật may mắn là ông đã không quan sát thấy các kiểu mà khác với kỳ

vọng về sự phân ly độc lập. Chính điều này đã làm nên sự nhất quán hoàn toàn và hết sức rõ ràng, đơn giản trong các kết quả mà ông đã công bố.



**Hình 4.17** Bản đồ nhiễm sắc thể số 2 của ruồi giấm (bên trái) và nhiễm sắc thể số 9 của ngô (bên phải).

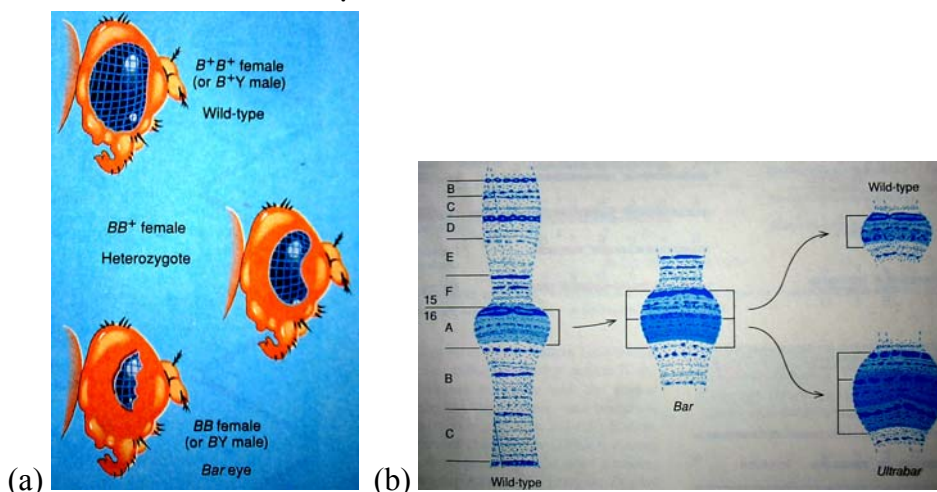
Quả đúng như Watson đã viết trong *Lời nói đầu* cuốn "*Chuỗi xoắn kép - Hồi ký về việc phát minh ra cấu trúc của DNA*" của mình, rằng: "Cách giải quyết đúng không những phải đẹp mà còn phải đơn giản" (Watson 1968). Chân lý bao giờ cũng giản dị nhưng giản dị chưa hẳn là chân lý; chân lý bao giờ cũng đẹp nhưng đẹp thì chưa đủ để là chân lý. Tuy nhiên, để nhận chân được cái giản dị, cái đẹp đích thực trong các khám phá của Mendel, như chúng ta đều biết, nhân loại đã phải mất chừng 35 năm!

**Bảng 4.7** Các vị trí nhiễm sắc thể đối với bảy tính trạng ở đậu Hà Lan

Tính trạng	Kiểu hình	Allele	Nhiễm sắc thể
Dạng hạt	tròn-nhăn	<i>R-r</i>	7
Màu sắc hạt	vàng-xanh	<i>I-i</i>	1
Màu sắc quả	xanh-vàng	<i>Gp-gp</i>	5
Kết cấu quả	trơn-nhăn	<i>V-v</i>	4
Màu sắc hoa	tím-trắng	<i>A-a</i>	1
Vị trí hoa	trục-đỉnh	<i>Fa-fa</i>	4
Chiều cao cây	cao-thấp	<i>Le-le</i>	4



X-dính ở trạng thái dị hợp và do đó có kiểu hình mắt Bar dị hợp. (ii) Nếu trao đổi chéo xảy ra ở giai đoạn hai sợi thì sẽ có dính lúu tới cả hai sợi. Vì vậy, khi tái bản xảy ra sau đó, tất cả các giao tử sẽ lại là dị hợp. Nói cách khác, các kết quả quan sát được trong cả hai trường hợp không có trao đổi chéo và trao đổi chéo hai sợi là như nhau.



**Hình 4.19** (a) Các kiểu hình đại, mắt Bar dị hợp và đồng hợp (ultraBar); và (b) Các kiểu băng tương ứng có một, hai và ba bản sao của băng 16A.

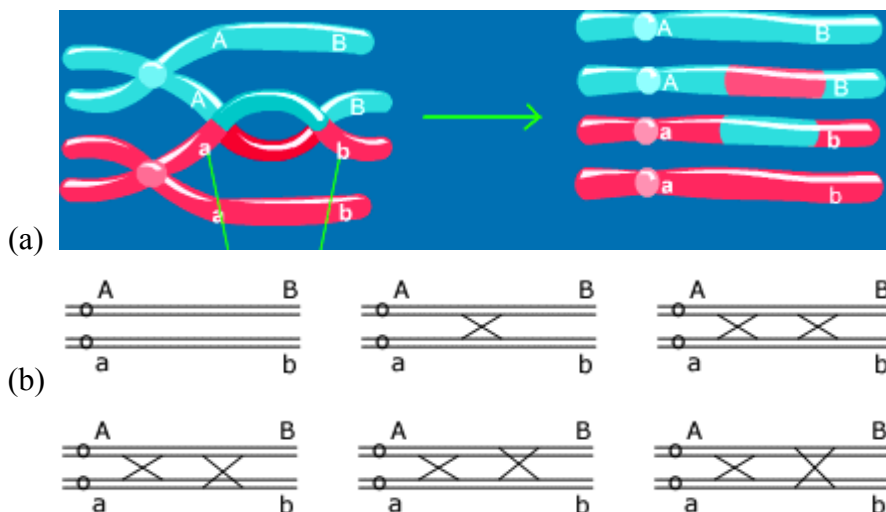
Tái tổ hợp ở giai đoạn bốn sợi sinh ra các kết quả khác nhau ở đời con. Trong trường hợp này, chỉ có hai trong bốn sợi tham gia vào tái tổ hợp, vì vậy sẽ sinh ra các kiểu mới. Với tái tổ hợp ở giai đoạn bốn sợi, cả kiểu đại lẫn kiểu đồng hợp BB sẽ được tạo ra. Trên thực tế, điều này được quan sát trong các thí nghiệm, chỉ ra rằng sự tái tổ hợp xảy ra sau khi tái bản - nghĩa là, trao đổi chéo phải xảy ra ở giai đoạn bốn sợi (xem hình 4.15). Chúng ta sẽ đề cập trở lại bằng chứng bổ sung về trao đổi chéo bốn sợi ở vi nấm *Neurospora* trong phần sau (mục VII).

• **Sơ lược về trao đổi chéo phức (multiple crossing over)**

Trên thực tế, khi hai gene nằm xa nhau trên một nhiễm sắc thể, giữa chúng có thể xảy ra nhiều trao đổi chéo trong một lần giảm phân và điều này gây ra những rắc rối trong việc tính toán, giải thích các số liệu tái tổ hợp và lập bản đồ dựa trên các số liệu này. Trên hình 4.20 cho thấy trao đổi chéo kép ở giai đoạn bốn sợi có thể xảy ra theo nhiều cách khác nhau. Chẳng hạn, hình 4.20a cho thấy trao đổi chéo kép xảy ra giữa hai cặp gene Aa và Bb nhưng các nhiễm sắc thể này không tái tổ hợp đối với các allele của hai gene, và do đó không thể phân biệt được với các nhiễm sắc thể không trao đổi chéo về mặt di truyền. Kết quả là cho ra chỉ hai loại giao tử với tỷ lệ ngang nhau ( $\underline{AB} = \underline{ab}$ ) giống như trường hợp không trao đổi chéo



(hình 4.15A), thay vì cho bốn loại giao tử thuộc hai kiểu (hình 4.15B). Hiện tượng này làm cho giá trị tái tổ hợp nhận được sẽ thấp hơn tần số trao đổi chéo thực và ảnh hưởng tới khoảng cách bản đồ giữa các gene.



**Hình 4.20** (a) Hai trao đổi chéo đơn xảy ra ở vùng giữa hai gene A và B, và kết quả là sinh ra cả bốn loại giao tử không tái tổ hợp đối với hai gene này, nghĩa là không phát hiện được thể trao đổi chéo kép nào. (b) Trao đổi chéo đơn và các dạng trao đổi chéo kép liên quan đến hai, ba và cả bốn sợi có thể xảy ra trong một cặp nhiễm sắc thể tương đồng.

Như đã nói, tần số tái tổ hợp cực đại giữa hai gene bất kỳ trên một cặp nhiễm sắc thể tương đồng là 50%. Điều này chỉ có thể xảy ra khi hai gene nằm cách xa nhau tới mức ít nhất có một đoạn trao đổi chéo hầu như lúc nào cũng xảy ra giữa chúng. Trên hình 4.15B ta thấy rõ ràng, cứ một trao đổi chéo đơn trong mỗi lần giảm phân sẽ cho ra một nửa số sản phẩm có các tổ hợp giống bố mẹ và nửa kia là các sản phẩm có tái tổ hợp gene.

Trên hình 4.20b cũng cho thấy khi xảy ra hai trao đổi chéo đơn mà có một chromatid chung hay trao đổi chéo kép liên quan ba sợi (hai hình đầu của hàng dưới), thì kết quả sẽ không khác với kết quả của trao đổi chéo đơn như ở hình 4.15B. Một khả năng khác, đó là trao đổi chéo kép liên quan cả bốn sợi (hình 4.20b, xem hình cuối của hàng dưới), trong đó sự trao đổi chéo thứ hai xảy ra giữa hai chromatid (ở ngoài) khác với hai chromatid tham gia vào trao đổi chéo thứ nhất (ở trong). Trong trường hợp này, tất cả các sản phẩm đều được tái tổ hợp. Nếu các chromatid tham gia vào hai trao đổi chéo đơn một cách ngẫu nhiên thì tỷ lệ kỳ vọng của ba kiểu trao đổi chéo kép sẽ tương ứng với tỷ lệ  $\frac{1}{4} : \frac{1}{2} : \frac{1}{4}$ . Điều đó có nghĩa là trung bình sẽ có hai trong bốn sản phẩm của giảm phân là có tái tổ hợp:

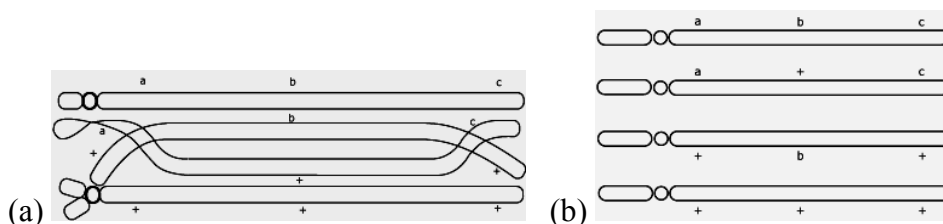
$(\frac{1}{4})(0) + (\frac{1}{2})(2) + (\frac{1}{4})(4) = 2$ . Tỷ lệ này rõ ràng là giống như khi một trao đổi chéo đơn mà lúc nào cũng xảy ra giữa hai gene.

Thực ra, như sẽ đề cập dưới đây, trao đổi chéo kép có thể phát hiện được trong các thí nghiệm tái tổ hợp sử dụng ba gene, gọi là lai ba điểm.

## VI. Lập bản đồ gene từ các phép lai phân tích ba điểm

### 1. Trao đổi chéo kép với việc xác định trật tự và khoảng cách các gene

Giả sử ta có ba gene liên kết A, B và C cần lập bản đồ. Theo lý thuyết, có thể có ba cách sắp xếp khác nhau tùy vào vị trí gene ở giữa, đó là: (i) A-B-C; (ii) A-C-B; và (iii) B-A-C. Tuy nhiên, trong trường hợp chưa biết chắc về sự liên kết của ba gene, và giả sử ta đã xác định được tần số tái tổ hợp giữa hai gene A và B là 15% và giữa hai gene B và C là 20%. Theo suy luận, ta biết rằng ba gene này cùng nhóm liên kết; và theo nguyên tắc, chúng có thể liên kết với nhau theo một trong hai cách: (i) A-B-C; hoặc (ii) B-A-C. Trong trường hợp này, ta cần phải tiến hành một phép lai phân tích giữa A và C để xem liệu tần số tái tổ hợp là 35% hay 5%. Như vậy, để xác định được vị trí tương đối của ba gene ta phải tiến hành ba phép lai phân tích riêng biệt. Trong khi đó, nếu dựa vào khả năng xảy ra một trao đổi chéo kép hay là hai trao đổi chéo đơn đồng thời (trong trường hợp ba gene ở khoảng cách đủ lớn), ta chỉ cần thực hiện một *phép lai phân tích ba điểm* (three-point testcross) là đủ. Điều đó có nghĩa là, việc lập bản đồ ba gene trở nên gọn nhẹ, không còn công kênh phức tạp như trước nữa.



**Hình 4.21 (a) Một trao đổi chéo kép, và (b) các sản phẩm tạo thành.**

Thật vậy, theo lý thuyết xác suất của các sự kiện ngẫu nhiên, một khi trao đổi chéo kép xảy ra (ở một số tế bào này), thì các trao đổi chéo đơn riêng rẽ (dĩ nhiên là xảy ra ở các nhóm tế bào khác) tất yếu phải xảy ra và có các tần số cao hơn hẳn tần số trao đổi chéo kép. Trong trường hợp đó, đời con của phép lai phân tích thu được (nếu đầy đủ) sẽ có tám kiểu hình thuộc bốn nhóm khác nhau: Trước hết là các kiểu cha mẹ với tần số cao nhất; kế đến là các kiểu trao đổi chéo đơn khác nhau; và cuối cùng là các kiểu trao đổi chéo kép với tần số thấp nhất. Và theo nguyên tắc, cứ hai kiểu hình của mỗi nhóm sẽ có số lượng tương đương nhau. Trên hình 4.21 cho thấy một trao đổi chéo kép xảy ra ở đoạn chứa gene b (nằm giữa a và

c), và kết quả là cho ra bốn sản phẩm gồm hai kiểu cha mẹ (abc và +++) và hai kiểu tái tổ hợp đối với gene b (a+c và +b+). Với kết quả này ta dễ dàng xác định gene ở giữa vì nó là sản phẩm của trao đổi chéo kép.

Bây giờ ta hãy tìm hiểu hai ví dụ kinh điển về lai phân tích ba điểm, một thí nghiệm ở ruồi giấm và một thí nghiệm chọn giống ở ngô.

• **Ví dụ 1:** Trong thí nghiệm sử dụng lại ba gene liên kết-X ở ruồi giấm (y, w và m), sau khi thu được các con cái F<sub>1</sub> dị hợp tử về ba cặp gen, Morgan cho chúng lai trở lại với con đực dòng kiểm tra (hình 4.22). Kết quả phân tích cho thấy các con cái F<sub>1</sub> cho tám loại giao tử (tức tám kiểu hình ở đời con của nó), hai trong số đó là các kiểu giao tử bố mẹ và sáu kiểu còn lại là các giao tử tái tổ hợp.

Theo sự trình bày ở trên, ta có thể tiến hành phân loại và tính tần số của mỗi kiểu giao tử dựa vào số lượng các kiểu hình thống kê được (hình 4.22) rồi vẽ bản đồ cho ba gene này, theo các bước chung nhất như sau:

P	♀ <u>+++ / +++</u>	×	♂ <u>y w m / Y</u>
	↓		
	<u>+++ / y w m</u>	×	<u>y w m / Y</u>
	(♀ F <sub>1</sub> )		(♂ dòng kiểm tra)
		↓	
	<u>Giao tử</u>	<u>Số lượng</u>	<u>Tần số</u>
Kiểu bố mẹ (không tái tổ hợp)	+++	3.501	0,664
	y w m	3.471	
Kiểu tái tổ hợp đơn m với w	++m	1.754	0,329
	y w +	1.700	
Kiểu tái tổ hợp đơn y với w	y ++	28	0,0057
	+ w m	32	
Kiểu tái tổ hợp kép w với y và m	+ w +	6	0,00086
	y + m	3	
Tổng		10.495	1,0

**Hình 4.22** Một phép lai phân tích các ruồi giấm cái dị hợp tử về ba gene liên kết-X (có cho sẵn số lượng và tần số của các kiểu khác nhau ở đời con).

**Bước 1:** *Xác định thành phần gene trội lặn trong nhóm liên kết* (dựa vào các kiểu giao tử bố mẹ, tức hai kiểu hình có số lượng nhiều nhất). Ở đây là y w m và +++ (đã biết trước).

**Bước 2:** *Xác định vị trí gene ở giữa.* Căn cứ vào hai kiểu hình có số

lượng ít nhất để suy ra đây là sản phẩm của trao đổi chéo kép (Lưu ý: trên thực tế, có thể chỉ xuất hiện một kiểu hình có thể là do số lượng chưa đủ lớn hoặc do các gene nằm khá gần nhau nên gây nhiễu, như sẽ thảo luận ở phần sau). Qua so sánh thành phần gene của các kiểu tái tổ hợp kép với các kiểu bố mẹ cho phép tìm ra gene thay đổi vị trí chính là gene ở giữa. Trong ví dụ này, gene ở giữa là *w*. Vậy trật tự ba gene là *y-w-m*, và kiểu gene của con cái  $F_1$  là:  $+++ / ywm$ . Khi biết được gene ở giữa, ta dễ dàng xác định được các kiểu trao đổi chéo đơn bằng cách sắp xếp lại trật tự gene ở các kiểu giao tử rồi so sánh với các kiểu bố mẹ để tìm ra gene có sự thay đổi vị trí. Ví dụ, với hai kiểu giao tử  $++m$  và  $yw+$ , vì *m* đổi chỗ nên đây là kết quả của trao đổi chéo đơn giữa *w* và *m*. Vậy hai kiểu còn lại là do trao đổi chéo đơn giữa *y* và *w*.

**Bước 3:** *Tính khoảng cách giữa các gene.* Để tìm khoảng cách giữa các gene có thể tiến hành theo một trong hai cách sau:

(i) Tính trực tiếp dựa vào số lượng của các kiểu tái tổ hợp. Theo nguyên tắc, khoảng cách của hai gene nằm kề nhau thì bằng số lượng cá thể của các kiểu tái tổ hợp giữa hai gene đó cộng với số lượng cá thể của các kiểu tái tổ hợp kép rồi chia cho tổng số cá thể ở đời con. Và khoảng cách giữa hai gene đầu mút thì bằng tổng các khoảng cách của các gene nằm trong nó. Cụ thể, khoảng cách *y* và *w* =  $(28 + 32 + 6 = 3) / 10.495 = 0,0066$ . Tương tự, ta tính được khoảng cách hai gene *w* và *m* =  $(1.754 + 1.700 + 6 = 3) / 10.495 = 0,330$ . Khi đó khoảng cách bản đồ giữa *y* và *m* =  $0,0066 + 0,330 = 0,3366$ .

(ii) Tính gián tiếp thông qua các tần số kiểu hình thuộc các nhóm khác nhau (xem ở hình 4.22). Khi có được các tần số này rồi, ta tính khoảng cách giữa hai gene kề nhau bằng cách cộng tần số trao đổi chéo đơn thực tế của hai gene đó với tần số trao đổi chéo kép thực tế. Ví dụ, khoảng cách hai gene *y* và *w* =  $0,0057 + 0,00086 = 0,0066$ ; và khoảng cách hai gene *w* và *m* =  $0,329 + 0,00086 = 0,330$ . Khi đó khoảng cách *y* và *m* là 0,3366 (Cũng có thể tính khoảng cách *y* và *m* bằng cách cộng hai tần số trao đổi chéo đơn thực tế; nghĩa là bằng  $0,0057 + 0,329 = 0,335$ . Con số này có hơi khác một chút so với con số tính được ở trên. Tuy nhiên, trên thực tế, việc tính khoảng cách hai gene đầu mút là không cần thiết).

**Bước 4:** *Vẽ bản đồ gene.* Bạn hãy tự vẽ bản đồ cho ba gene này.

• **Ví dụ 2:** Tóm tắt kết quả thí nghiệm lai phân tích ba gene trên trên một nhiễm sắc thể ở ngô (*Zea mays*). Các allele lần được sử dụng trong phép lai một cá thể  $F_1$  dị hợp tử ba cặp gene (*Lz Gl Su / lz gl su*) này là: *lz* (tập tính mọc bò lan), *gl* (lá bóng), và *su* (nội nhũ đường). Phép lai và kết quả được cho ở bảng 4.8.



quan sát được là 0,00086; rõ ràng là thấp hơn tỷ lệ được kỳ vọng. Đúng ra, theo lý thuyết, số cá thể trao đổi chéo kép xuất hiện phải bằng  $0,00218 \times 10495 \approx 23$  (con số thực tế xuất hiện ít hơn  $23 - 9 = 14$  cá thể).

Sự chênh lệch như thế nói chung rất phổ biến ở các sinh vật, và nó được coi là hiện tượng *nhiều* (interference); nghĩa là sự xuất hiện một trao đổi chéo ở một vùng nhiễm sắc thể làm giảm xác suất xảy ra trao đổi chéo ở vùng thứ hai lân cận nó.

Để mô tả sự khác nhau giữa các số lượng trao đổi chéo quan sát và kỳ vọng đó, lần đầu tiên H.J. Muller đưa ra một phương pháp chuẩn cho sự tính toán này. Theo đó, *hệ số trùng hợp* (coefficient of coincidence) được đo bằng số lượng hay tần số các thể tái tổ hợp kép quan sát (O: observed) chia cho số lượng hay tần số kỳ vọng tương ứng (E: expected); và nếu ta ký hiệu hệ số trùng hợp là C, thì  $C = O/E$ ; trong đó:  $0 \leq C \leq 1$ . Trị số này là một tiêu chuẩn cho phép đo mức độ nhiễu I (*interference*) của một trao đổi chéo này lên một trao đổi chéo khác. Công thức dùng để tính độ nhiễu này như sau:

$$I = 1 - C$$

trong đó  $0 \leq I \leq 1$ . Từ ví dụ 1 ở trên, ta tính được  $C = 0,00086/0,00218 = 0,3945$ ; điều này có nghĩa là số lượng trao đổi chéo kép thực tế đã xảy ra chỉ bằng 39,45% con số lý thuyết. Vậy  $I = 1 - 0,3945 = 0,6055$ . Trị số này cho thấy mức độ nhiễu cao đáng kể.

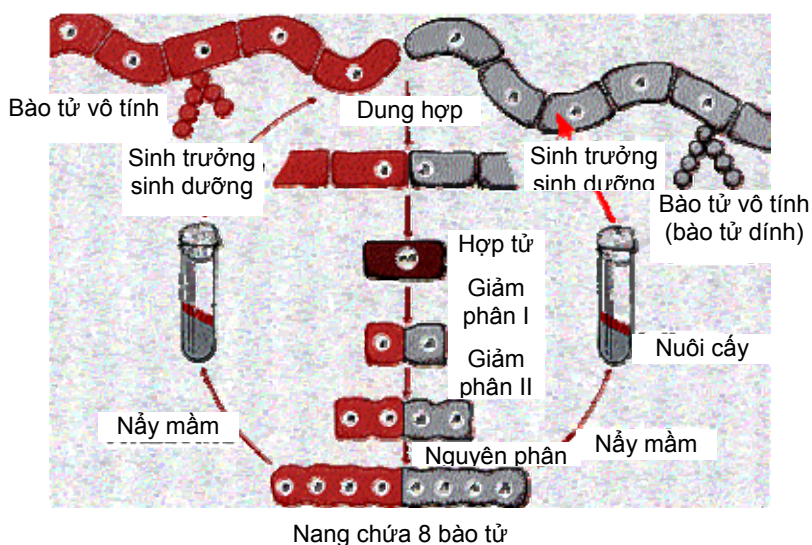
Rõ ràng hai đại lượng này là nghịch nhau và tổng của chúng bằng đơn vị ( $I + C = 1$ ). Nếu  $I = 1$ , lúc đó  $C = 0$ ; điều đó có nghĩa là, hiện tượng nhiễu tăng dần lên khi mà khoảng cách hai gene phía bên ngoài càng ngắn lại, cho tới khi trao đổi chéo kép không thể xảy ra tại một điểm giới hạn (nhiều hoàn toàn) và hệ số phù hợp bằng zero. Ở nhiều sinh vật khoảng cách này là khoảng 10 đơn vị bản đồ. Ngược lại, nếu  $I = 0$ , thì  $C = 1$ ; nghĩa là khi các gene nằm rất xa nhau đến nỗi các trao đổi chéo của chúng không ảnh hưởng gì lên nhau (hoàn toàn không nhiễu), thì hiện tượng nhiễu biến mất và hệ số phù hợp bằng 1. Các thực nghiệm cho thấy  $I = 0$  khi tổng khoảng cách giữa các gene lớn hơn chừng 45 đơn vị bản đồ.

Một lý thuyết để giải thích tần số trao đổi chéo kép thấp hiếm là ở chỗ trị số nhiễu cao đối với các gene liên kết chặt rõ ràng là có liên quan với hiện tượng "xơ cứng", không thể vặn xoắn về mặt vật lý của các chromatid trong một vùng rất ngắn. Tuy nhiên, hiện tượng nhiễu có thể xảy ra ở các gene cách xa nhau tới 30 đơn vị bản đồ; điều này chỉ ra rằng các nhân tố khác cũng có tầm quan trọng không kém. Thật là thú vị, các tâm động của các nhiễm sắc thể tâm giữa ở ruồi giấm (và có lẽ cả ở nhiều sinh vật khác)

hoạt động như là các vật cản đối với hiện tượng nhiễu. Nói cách khác, sự trao đổi chéo ở một vai của nhiễm sắc thể tâm giữa không gây hiệu quả ức chế lên sự trao đổi chéo ở vai kia.

## VII. Lập bản đồ bằng phân tích bộ bốn (tetrad analysis)

Việc khẳng định trao đổi chéo bốn sợi cũng thu được ở *Neurospora crassa*, một loại vi nấm được nghiên cứu nhiều trong di truyền học, có các sản phẩm của giảm phân chứa đựng trong một dãy các bào tử có trật tự được sắp xếp theo kiểu thẳng hàng (hình 4.23). Một nhân hợp tử chứa trong một cấu trúc dạng túi gọi là nang (ascus), trải qua giảm phân hầu như ngay sau khi nó vừa được tạo thành. Bốn nhân sinh ra từ giảm phân nằm trong một cái nang theo một trật tự thẳng hàng, và mỗi nhân này lại trải qua một lần nguyên phân nữa tạo thành hai bào tử nang (ascospores) nằm kế nhau và giống hệt nhau về mặt di truyền. Do đó, mỗi nang trưởng thành có chứa tám bào tử nang theo bốn cặp, mỗi cặp biểu thị một sản phẩm của giảm phân. Lúc này, ta có thể tách các bào tử nang theo trình tự của mỗi nang và cho nảy mầm trong các ống nuôi cấy riêng để xác định các kiểu gene của chúng.



**Hình 4.23** Vòng đời của *Neurospora crassa*.

Chú thích: Các bào tử sắp xếp thành một dãy thẳng hàng trong nang sao cho mỗi nửa bắt nguồn từ các thể phân ly ở giảm phân I. Mỗi một phần tư bắt nguồn từ các chromatid kết hợp với một tâm động cụ thể.

Sự sắp xếp có trật tự của các sản phẩm giảm phân cho phép xác định tần số tái tổ hợp giữa một gene và tâm động của nó - ấy là cách lập bản đồ

*tâm động* (map the centromere). Kỹ thuật lập bản đồ này dựa trên một đặc điểm của giảm phân, đó là: *các tâm động tương đồng của các nhiễm sắc thể bố mẹ tách ra ở giảm phân I, và các tâm động của các chromatid chị em tách nhau ở giảm phân II*. Như vậy, trong trường hợp không có trao đổi chéo giữa một gene và tâm động của nó, các allele của gene (ví dụ A và a) tách nhau ở giảm phân I, sự phân ly này được gọi là *phân ly giảm phân I* (first-division segregation). Thay vì, nếu như một trao đổi chéo xảy ra giữa gene và tâm động của nó, các allele A và a của nó sẽ không được phân ly cho đến lúc xảy ra giảm phân II; sự phân ly này được gọi là *phân ly giảm phân II* (second-division segregation). Như thế, khi phân ly giảm phân I xảy ra, chỉ có thể có hai kiểu trật tự sắp xếp của các sản phẩm giảm phân, đó là AAaa và aaAA. Tuy nhiên, có thể có bốn kiểu của phân ly giảm phân II, bởi vì sự sắp xếp ngẫu nhiên của các nhiễm sắc thể tương đồng ở kỳ giữa I và của các chromatid ở kỳ giữa II. Bốn sự sắp xếp này được hình dung như thế này: AaAa, aAaA, AaaA, và aAAa.

Tỷ lệ phân trăm của các nang có kiểu phân ly giảm phân II của một gene có thể được sử dụng để lập bản đồ cho gene gắn liền với tâm động của nó. Chẳng hạn, ta giả sử rằng có 30% số nang từ một phép lai có kiểu phân ly giảm phân II đối với các allele A và a. Điều này có nghĩa là 30% số tế bào trải qua giảm phân đã có một thể trao đổi chéo giữa A và tâm động của nó. Hơn nữa, trong mỗi tế bào đã xảy ra một trao đổi chéo, hai trong số các chromatid là các thể tái tổ hợp và hai cái còn lại là những thể không tái tổ hợp. Nói cách khác, một trao đổi chéo đơn cho ra bốn sản phẩm giảm phân, mà một nửa trong số đó là các thể tái tổ hợp và một nửa là các thể không tái tổ hợp. Như vậy, tần số trao đổi chéo bằng 30% tương ứng với tần số tái tổ hợp là 15%. Theo quy ước, khoảng cách bản đồ phản ánh tần số của các sản phẩm giảm phân có tái tổ hợp hơn là phản ánh tần số của các tế bào mà trong đó trao đổi chéo đã xảy ra. Vì vậy, khoảng cách bản đồ giữa một gene và tâm động của nó được cho ở biểu thức sau:

$$\frac{1/2 (\text{Số nang có các kiểu phân ly giảm phân II})}{\text{Tổng số các nang}} \times 100$$

Công thức này có giá trị chừng nào gene nằm đủ gần tâm động sao cho một vài trao đổi chéo phức có thể xảy ra. Như thế, các trị số liên kết đáng tin cậy được xác định tốt nhất cho các gene mà chúng nằm khá gần tâm động. Vị trí của các gene ở xa hơn lúc đó được thực hiện bằng cách lập bản đồ của các gene này gắn liền với các gene nằm gần tâm động hơn.

• **Tái tổ hợp nguyên phân** (*mitotic recombination*)

Sự trao đổi chéo cũng có thể xảy ra trong nguyên phân, mặc dù với tần

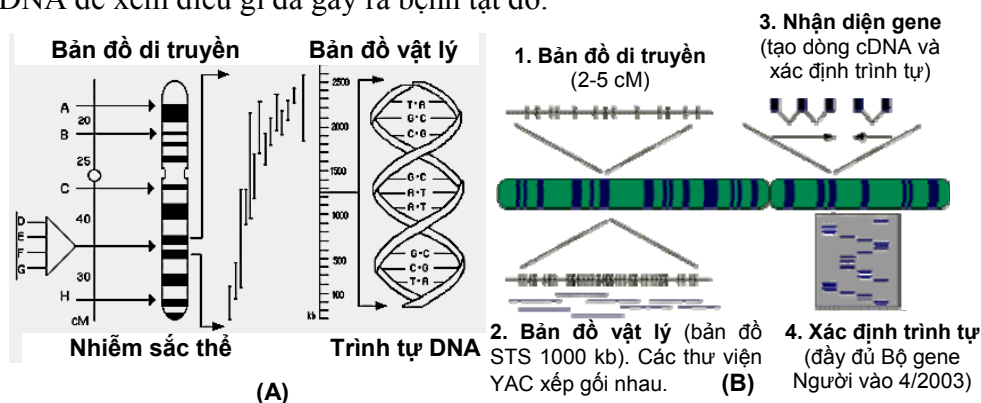


số thấp hơn khoảng 1000 lần so với giảm phân. Bằng chứng đầu tiên về *tái tổ hợp nguyên phân* (mitotic recombination) thu được ở ruồi giấm, nhưng lại được nghiên cứu kỹ nhất ở các vi nấm, đặc biệt là nấm men bia và *Aspergillus*. Và các bản đồ di truyền có thể thiết lập từ các tần số tái tổ hợp trong nguyên phân. Khoảng cách bản đồ tương đối giữa các gene cụ thể đôi khi tương xứng với khoảng cách bản đồ dựa trên các tần số tái tổ hợp trong giảm phân, nhưng chưa rõ lý do vì sao các khoảng cách này lại thường khác nhau một cách đáng kể. Những sự trái ngược, không nhất quán này có thể là kết quả hoặc là của các cơ chế tiếp hợp khác nhau hoặc là của sự phân bố không ngẫu nhiên của các vị trí trao đổi tiềm năng.

### • Khoảng cách bản đồ và khoảng cách vật lý

Như đã đề cập trước đây, khoảng cách bản đồ giữa các gene được ước lượng bằng các phép lai di truyền nói chung là tương ứng với khoảng cách vật lý, hay độ dài của DNA, mà nó tách bạch các gene. Một trường hợp ngoại lệ chính là vùng tâm động, không có tái tổ hợp nhưng có thể chứa một độ dài DNA khá lớn. Hơn nữa, rõ ràng là tái tổ hợp có thể thường được khu trú trong các vùng đặc biệt, và không phải hoàn toàn ngẫu nhiên trên suốt chiều dài của nhiễm sắc thể.

Bằng cách so sánh các *bản đồ di truyền* (genetic map) và *bản đồ vật lý* (physical map), hình 4.24 cho thấy cả khoảng cách bản đồ ước tính được (bằng centiMorgan, viết tắt là cM) lẫn số lượng kilobase (1kbp = 1000 bp) trong DNA (bản đồ vật lý). Kế bên bản đồ di truyền của mỗi nhiễm sắc thể là bản đồ vật lý dựa trên các nghiên cứu xác định bằng nhiễm sắc thể (hình 4.24). Bản đồ mới này có tiềm năng hết sức to lớn trong việc giúp các nhà khoa học xác định vị trí các gene gây ra các bệnh cụ thể và sau đó kiểm tra DNA để xem điều gì đã gây ra bệnh tật đó.



**Hình 4.24** Hai cách so sánh khoảng cách bản đồ (A và B); khoảng cách bản đồ di truyền được cho bằng đơn vị cM trên một nhiễm sắc thể và khoảng cách vật lý được cho bằng số kilobase trong trình tự DNA.

Gần đây, một nỗ lực to lớn đã được tập trung hướng tới thu nhận cho được một bản đồ di truyền của bộ gen người (xem hình 4.24, và các trang web <http://www.genome.gov/> và <http://www.ncbi.nlm.nih.gov/Omim/>). Bằng cách sử dụng các *chỉ thị phân tử* (molecular markers) khác nhau (hiện giờ được biết có tới hàng ngàn marker trải khắp bộ gene người) như: (i) *các đa hình độ dài phân đoạn DNA* (restriction fragment length polymorphism = RFLPs), nghĩa là các biến đổi trong các trình tự DNA của các sinh vật thuộc một loài nào đó mà có thể xác định được bằng sự tách đoạn các trình tự này nhờ sử dụng các enzyme giới hạn (chương 10), vì sự biến đổi nằm bên trong vị trí giới hạn; và RFLPs có thể được sử dụng để đo độ đa dạng của một gene trong quần thể; hay (ii) *Vị trí trình tự đích* (sequence tagged site = STS), đó là một trình tự độc nhất từ một vị trí nhiễm sắc thể đã biết mà có thể khuếch đại bằng phương pháp PCR; các STS hoạt động như là những marker vật lý cho việc lập *bản đồ bộ gene* (genome mapping) và *tạo dòng* (cloning) v.v...

• *Lưu ý*: Vấn đề lập bản đồ bằng phân tích bộ bốn ở các vi nấm và vi tảo tạm dừng tại đây bởi vì nó sẽ được trình bày đầy đủ trong giáo trình *Di truyền học vi sinh vật và ứng dụng*. Các hiện tượng *tái tổ hợp bên trong gene* (recombination within genes) và *phép phân tích bổ trợ* (complementation) hay là *trắc nghiệm cis-trans* (cis-trans test) sẽ được đề cập trong chương 6 và trình bày chi tiết trong giáo trình vừa nói trên.

## Câu hỏi và Bài tập

1. (a) Trình bày các cơ chế xác định giới tính ở các sinh vật, và nêu ý nghĩa thực tiễn của chúng.

(b) Cơ sở di truyền học của hiện tượng bất hoạt nhiễm sắc thể X là gì?

(c) Phân biệt các tính trạng liên kết với giới tính, tính trạng bị giới hạn bởi giới tính và tính trạng chịu ảnh hưởng của giới tính, và cho các ví dụ.

2. (a) Nội dung cơ bản của thuyết di truyền nhiễm sắc thể là gì? Tần số tần số tái tổ hợp là gì và được ứng dụng trong di truyền học như thế nào? Cho ví dụ.

(b) Hãy chứng minh: Tần số tái tổ hợp ( $r$ ) của hai cặp gene bất kỳ trên một cặp nhiễm sắc thể tương đồng nằm trong giới hạn  $0 \leq r \leq 0,5$ . Có nhận xét gì về mối liên hệ giữa các quy luật di truyền liên kết và phân ly độc lập?

3. Ở ruồi giấm, người ta đã tiến hành một số phép lai và nhận được kết quả dưới đây. Hãy giải thích cơ sở di truyền cho mỗi tính trạng và xác định kiểu gene của các cá thể.

P: Cái thân thâm, mắt trắng  $\times$  đực thân vàng nhạt, mắt đỏ

F<sub>1</sub>: Giới cái = tất cả vàng nhạt, mắt đỏ

Giới đực = tất cả vàng nhạt, mắt trắng

F<sub>2</sub>: 27 vàng nhạt, mắt đỏ: 25 vàng nhạt, mắt trắng: 9 thâm, mắt đỏ: 8 thâm, mắt trắng. (Biết rằng không có sự khác nhau về tỷ lệ phân bố đực - cái ở mỗi kiểu hình của F<sub>2</sub>).

4. Một phép lai giữa ruồi giấm cái cánh khía với ruồi giấm đực cánh bình thường, đời con gồm có 35 con cái cánh khía, 38 con cái cánh bình thường và 33 con đực bình thường. (a) Hãy xác định phương thức di truyền của kiểu hình cánh có khía và sơ đồ lai. (b) Trên thực tế có xuất hiện ruồi giấm cái cánh khía thuần chủng hay không? Tại sao?

5. Khi tiến hành lai giữa một ruồi giấm cái mắt đỏ và cánh ngắn với ruồi giấm đực mắt nâu và cánh dài thu được tất cả đời con có mắt đỏ và cánh dài. Kết quả tạp giao của các ruồi giấm F<sub>1</sub> thu được F<sub>2</sub> như sau:

Ruồi giấm cái	Ruồi giấm đực
76 đỏ, dài	39 đỏ, dài
24 đỏ, ngắn	37 nâu, dài
	13 đỏ, ngắn
	11 nâu, ngắn

Hãy xác định quy luật di truyền của các tính trạng và lập các sơ đồ lai.

6. Một nòi ruồi giấm kiểu dại được cho lai với một nòi khác đồng hợp về các thể đột biến liên kết trên X là m, r và v. Sau đó đem các ruồi giấm cái F<sub>1</sub> lai ngược với các con đực thuộc nòi lặn thì nhận được số lượng các kiểu hình đời con của chúng như dưới đây.

Kiểu hình	Số lượng	Kiểu hình	Số lượng
+ + +	331	m r +	10
m + +	2	m + v	81
+ r +	73	+ r v	0
+ + v	14	m r v	309

Hãy xác định trật tự các gene và lập bản đồ với khoảng cách tương đối giữa các thể đột biến này.

7. Xét các phép lai sau đây ở ruồi giấm về ba gene liên kết với giới tính:

P: + + + / + + +  $\times$  y w m / Y

F<sub>1</sub>: Tất cả đều kiểu dại

Nếu đem các ruồi giấm cái  $F_1$  lai phân tích với các con đực dòng kiểm tra đã tính được tần số của các kiểu hình như sau:

Các lớp kiểu hình	Tần số
Nhóm bố mẹ (không tái tổ hợp)	0,664
Nhóm tái tổ hợp của m với y và w	0,329
Nhóm tái tổ hợp của y với w và m	0,0057
Nhóm tái tổ hợp của w với y và m	0,00086

- Hãy xác định kiểu gene của các ruồi giấm  $F_1$ .
- Biểu diễn vị trí và khoảng cách tương đối của các gene.
- Tính hệ số trùng hợp và độ nhiễu.
- Tính số cá thể kỳ vọng cho mỗi kiểu hình ở đời con (biết rằng tổng số cá thể là 10.495).

8. Một ruồi giấm cái dị hợp về cả ba gene  $+/sc$ ;  $+/ec$ ;  $+/vg$  được cho lai với một con đực đồng hợp lặn về cả ba gene này. Đời con gồm có:

sc ec vg	233	sc + vg	12
+ + +	239	+ ec +	14
sc ec +	241	sc + +	14
+ + vg	231	+ ec vg	16

Hãy giải thích kết quả này và chỉ ra bản đồ liên kết phù hợp.

9. Khi lai hai ruồi giấm *Drosophila* đều dị hợp về các tính trạng lặn liên kết là mắt nâu và thân đen cho kết quả như sau: 52 thân vàng nhạt, mắt đỏ : 9 thân vàng nhạt, mắt nâu : 10 thân đen, mắt đỏ : 9 thân đen, mắt nâu.

Hãy giải thích kết quả và lập sơ đồ lai.

10. Ở ruồi giấm, mắt quả thận (k), mắt dạng máu (cd), và thân mun (e) là ba allele lặn liên kết với nhau. Nếu đem các ruồi cái mắt thận, dạng máu đồng hợp lai với các ruồi đực thân mun đồng hợp, thu được các ruồi  $F_1$  tất cả đều có kiểu dại. Nếu đem các ruồi cái  $F_1$  dị hợp lai với các con đực mắt thận, dạng máu, thân mun thu được đời con gồm 1000 con, trong đó:

Mắt thận, dạng máu	440	Thân mun, mắt máu	23
Thân mun	443	Mắt thận	25
Mắt thận, thân mun	32	Mắt thận, máu, mun	1
Mắt máu	34	Kiểu dại	2

- Xác định thành phần gene trên một nhiễm sắc thể của ruồi cái  $F_1$ .
- Vẽ bản đồ của các gene và tính hệ số trùng hợp và độ nhiễu.

## Tài liệu Tham khảo

### Tiếng Việt

Dubinin NP. 1981. *Di truyền học đại cương* (Bản dịch của Trần Đình Miên và Phan Cự Nhân). NXB Nông Nghiệp, Hà Nội

Hutt FB. 1964. *Di truyền học động vật*. (Bản dịch của Phan Cự Nhân). NXB Khoa học và Kỹ thuật, Hà Nội, 1978.

Phan Cự Nhân. 2001. *Di truyền học động vật*. NXB Khoa học và Kỹ thuật, Hà Nội.

Phan Cự Nhân (chủ biên), Nguyễn Minh Công, Đặng Hữu Lanh. 1999. *Di truyền học* (2 tập). NXB Giáo Dục, Hà Nội.

Watson JD. 1968. *Chuỗi xoắn kép: Hồi ký về việc phát minh ra cấu trúc của DNA*. (Bản dịch của Lê Đình Lương và Thái Doãn Tĩnh). NXB Khoa học và Kỹ thuật, Hà Nội, 1984.

### Tiếng Anh

Ayala FJ, Kiger JA. 1980. *Modern Genetics*. Benjamin/Cummings Publishing Company, Inc, Menlo Park, CA.

Brooker RJ. 1999. *Genetics - Analysis and Principles*. Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA.

Campbell NA, Reece JB. 2001. *Essential Biology*. Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc, San Francisco, CA.

Clegg CJ, Mackean DG. 2000. *Advanced Biology : Principles and Applications*. 2<sup>nd</sup> ed., John Murray Published Ltd, London.

Hartl DL, Freifelder D, Snyder LA. 1988. *Basic Genetics*. Jones and Bartlett Publishers, Inc, Boston - Portola Valley.

Kalthoff K. 1996. *Analysis of Biological Development*. McGraw-Hill, Inc., New York.

Karpen GH. 1994. Position-effect variegation and the new biology of heterochromatin. In: *Current Opinion in Genetics & Development*, Vol 4, No 2 (Stillman B and Green M, eds.), pp 281-291.

Kimball J. 2004. <http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/>

Lewis R. 2003. *Human Genetics: Concepts and Applications*. 5<sup>th</sup> ed, McGraw-Hill, Inc, NY.

Martin E, Ruse M, Holmes E. 1996. *Oxford Dictionary of Biology*. 3rd. ed., Oxford University Press, Oxford, UK.

McKusick V. 1998. *Mendelian Inheritance in Man*. 12th ed., Johns Hopkins University Press, Baltimore.

Rastan S. 1994. X chromosome inactivation and the *Xist* gene. In: *Current Opinion in Genetics & Development*, Vol 4, No2 (Stillman B and Green M, eds.), pp 292-297.

Russell PJ. 2003. *Essential Genetics*. Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA.

Suzuki DT, Griffiths AJF, Miller JH, Lewontin RC. 1989. *An Introduction to Genetic Analysis*. 4<sup>th</sup> ed, W-H Freeman and Company, New York.

Tamarin RH. 1999. *Principles of Genetics*. 6<sup>th</sup> ed, McGraw-Hill, Inc, NY.

Weaver RF, Hedrick PW. 1997. *Genetics*. 3<sup>rd</sup> ed, McGraw-Hill Companies, Inc. Wm.C.Brown Publishers, Dubuque, IA.

Wellnitz WR. 1995. *Genetics: Problem Solving Guide*. 2<sup>nd</sup> ed., Wm.C. Brown Publishers. Dubuque, Iowa.

Yablokov AV. 1986. *Population Biology* (Progress and problems of studies on natural populations). MIR Publishers, Moscow. pp17-30.

### **Một số trang web**

National Human Genome Research Institute (NHGRI)/ National Institute of Health (NIH). 2005: <http://www.genome.gov/>

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

Online Mendelian Inheritance in Man (OMIM<sup>TM</sup>):

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>

## Chương 5

# Bản chất Hóa học và Tái bản của Vật chất Di truyền

Trước tiên, chúng ta cần nắm các đặc tính thiết yếu của vật chất di truyền sau đây: (1) Đặc tính *thông tin sinh học*: chứa đựng thông tin cần thiết cho việc xác định cấu trúc của tất cả các protein đặc thù của loài (các gene cấu trúc) và điều khiển các hoạt động sinh trưởng, phân chia và biệt hoá tế bào; (2) Đặc tính *tái bản*: khả năng tự sao chép chính xác, đảm bảo thông tin di truyền của thế hệ sau giống với thế hệ trước; (3) Đặc tính *hoạt động của các gene*: các gene trong bộ gene có khả năng tổng hợp ra các sản phẩm là những phân tử tham gia vào mọi động sống căn bản của tế bào (*phiên mã và dịch mã*); (3) Đặc tính *biến đổi*: khả năng bị biến đổi từ nhiều quá trình khác nhau (như đột biến, tái tổ hợp, các yếu tố di truyền vận động). Chính sự biến đổi đa dạng này tạo ra các nguồn biến dị di truyền đa dạng và phong phú cho sự chọn lọc và tiến hoá.

Ngày nay chúng ta đều biết rằng vật chất di truyền chính là các nucleic acid mà ở hầu hết sinh vật là loại *deoxyribonucleic acid* (DNA) và ở một số virus là *ribonucleic acid* (RNA). Trong chương này, chúng ta sẽ tìm hiểu thành phần hóa học và cấu trúc cũng như cơ chế tái bản của vật chất di truyền là các đại phân tử DNA và RNA.

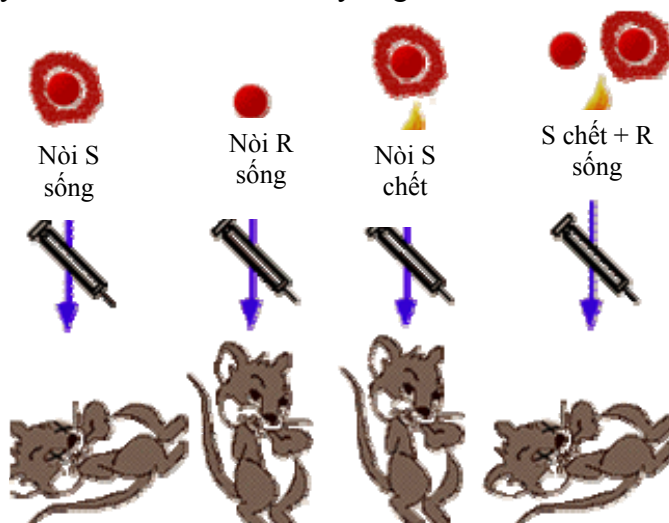
## I. Bằng chứng vật chất di truyền là các nucleic acid

### 1. Các thí nghiệm biến nạp ở vi khuẩn

Năm 1928, Frederick Griffith đặt nền tảng cho việc xác định DNA là vật chất di truyền thông qua thí nghiệm *biến nạp* (transformation) trên phế cầu khuẩn *Streptococcus pneumoniae*. Các vi khuẩn kiểu dại có vỏ bọc bằng polysaccharide mọc thành các khuẩn lạc hay dòng đặc trưng bằng dạng nhẵn (S: smooth); chúng được coi là *độc* (virulent) vì có khả năng gây viêm phổi làm chuột chết. Một nòi đột biến của nó mất khả năng hình thành vỏ bọc và mọc thành các khuẩn lạc bé, xù xì (R: rough); chúng là các vi khuẩn *không độc* (avirulent).

Griffith tiến hành các thí nghiệm khác nhau và nhận thấy rằng: (1) Nếu tiêm các vi khuẩn sống nòi S, chuột chết; (2) Nếu tiêm các vi khuẩn sống nòi R, chuột sống; (3) Nếu tiêm các vi khuẩn sống nòi S đã bị giết chết bằng nhiệt, chuột sống bình thường; (4) Nhưng khi tiêm hỗn hợp gồm các vi khuẩn nòi R còn sống và vi khuẩn nòi S đã bị giết chết bằng nhiệt, kết quả bất ngờ là chuột chết. Từ mẫu máu của các con chuột chết này ông đã phân lập được các vi khuẩn nòi S còn sống. Điều đó chứng tỏ rằng bằng

cách nào đó các mảnh vỡ tế bào vi khuẩn nòi S bị giết bởi nhiệt đã xâm nhập vào tế bào R sống và làm biến đổi nó thành vi khuẩn S sống và gây độc. Quá trình này gọi là biến nạp và chất trong mảnh vỡ tế bào S đã gây sự biến nạp đặc thù đó được gọi là tác nhân biến nạp (*transforming agent*). Vậy bản chất của tác nhân này là gì?



**Hình 5.1** Thí nghiệm của Griffith (từ trái sang là 1-4 như đã mô tả trong bài).

Trong hơn mười năm trời kể từ sau thí nghiệm của Griffith, Oswald Avery cùng với MacLeod và McCarty đã tiến hành trở lại thí nghiệm biến nạp này nhưng trong ống nghiệm (*in vitro*). Họ tinh chiết các thành phần của tác nhân gây biến nạp và cho các enzyme khác nhau tác động thăm dò lên hoạt tính của chất chiết này. Kết quả cho thấy hoạt tính của chất chiết chỉ bị phân hủy bởi deoxyribonuclease (DNase) - enzyme có khả năng phân cắt làm suy thoái chỉ đối với DNA, mà không bị tác động phân hủy của ribonuclease (RNase) - enzyme làm suy thoái chỉ các RNA, cũng như của các protease (như trypsin và chymotrypsin) - enzyme phân giải protein. Vào năm 1944, nhóm nghiên cứu của Avery đã chứng minh được rằng: DNA là vật chất mang thông tin di truyền, chứ không phải protein.

## 2. Tính ổn định của hàm lượng DNA ở các sinh vật bậc cao

Trước khi Avery và cs công bố kết quả này, đa số các nhà hóa sinh học vẫn tin rằng gene chính là protein. Vì vậy sau kết quả hiển nhiên đó, nhiều nhà khoa học vẫn còn nghi ngờ tính phổ quát của nó.

Vào lúc này, các nghiên cứu bản chất hóa học của nhiễm sắc thể đã khẳng định rằng, hầu hết DNA được khu trú trong các nhiễm sắc thể ở trong nhân. Hơn nữa, các phân tích hóa học của A.E.Mirsky và H.Ris

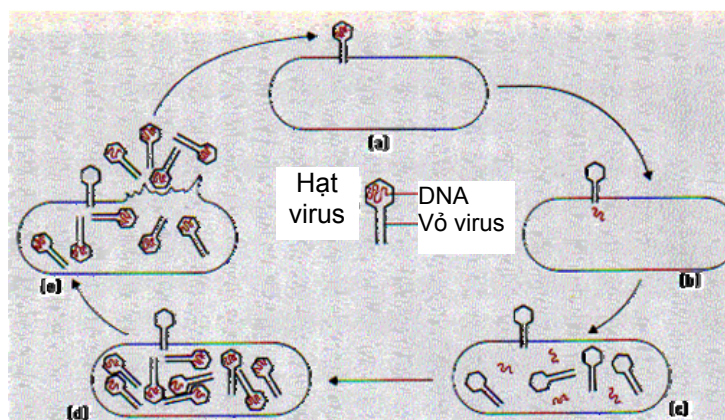


(USA) và của R.Vendrely (France) cho thấy hàm lượng DNA mỗi bộ nhiễm sắc thể lưỡng bội luôn luôn ổn định đối với một cơ thể nhất định và bằng hai lần hàm lượng DNA mỗi tế bào tinh trùng đơn bội. Chẳng hạn, các số liệu tương ứng ở bò là xấp xỉ 6,6 và 3,3 picogam ( $1\text{pg} = 10^{-12}\text{g}$ ).

### 3. Các thí nghiệm ở virus

#### 3.1. Thí nghiệm Hershey-Chase

Năm 1952, Alfred D.Hershey và Martha Chase tiến hành các nghiên cứu đánh dấu đồng vị phóng xạ ở loại *thể thực khuẩn* (bacteriophage, thường nói gọn là *phage*) gọi là T2 vốn ký sinh ở vi khuẩn *Escherichia coli*. Phage này chứa một lõi DNA và lớp vỏ bảo vệ bằng protein. Dựa vào đặc điểm nguyên tử lưu huỳnh (S: sulfur) chỉ có trong protein và nguyên tử phosphor (P) chỉ có trong nucleic acid, họ đã đánh dấu protein bằng chất đồng vị phóng xạ  $S^{35}$  và DNA bằng  $P^{32}$  bằng cách cho các phage này sinh trưởng trên các môi trường có  $S^{35}$  và sinh sản khi có  $S^{32}$ . Các virus đánh dấu được sử dụng để theo dõi thành phần nào từ virus ban đầu đã đi vào tế bào chủ và sau đó xuất hiện trong các phage thế hệ con.



**Hình 5.2** Thí nghiệm Hershey-Chase. (a) Phage bám ở bề mặt ngoài màng tế bào vi khuẩn; (b) bơm lõi DNA vào trong và để lại lớp vỏ protein bám ở ngoài; (c) tái bản, phiên mã và tổng hợp hàng loạt protein của virus; (d) lắp ráp lõi + vỏ để tạo ra các virion; (e) enzym lysozyme phá hủy vách tế bào vi khuẩn và phóng thích các phage thế hệ con.

Kết quả cho thấy trong các phage thế hệ con chỉ có hầu như DNA dạng ban đầu và không hề có protein dạng này (hình 5.2). Hơn nữa, lớp vỏ protein dạng ban đầu bám ở bên ngoài màng tế bào vi khuẩn, không thực hiện chức năng nào sau khi DNA đã được tiêm vào trong vi khuẩn. Như vậy, vật chất di truyền của phage T2 là DNA, chứ không phải protein.

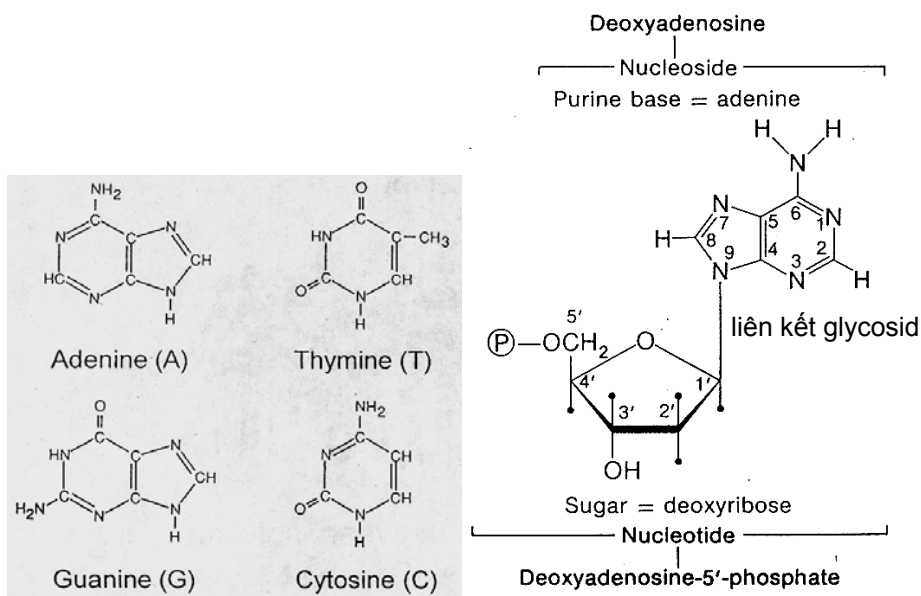
#### 3.2. Thí nghiệm chứng minh vật chất di truyền của một số virus là RNA

Sau này, năm 1956 A.Gierer chứng minh rằng RNA được tinh sạch từ *virus đốm thuốc lá* (tobacco mosaic virus = TMV) có khả năng lây nhiễm khi không có bất kỳ protein nào thuộc virus. Kết quả phân tích các virus thể hệ con cho thấy chúng hoàn toàn giống với các virus sinh ra từ sự lây nhiễm bằng các virus còn nguyên vẹn. Bằng chứng RNA là thành phần di truyền của TMV cũng đã được Fraenkel Conrat và B.Singer tái xác nhận năm 1957 trong thí nghiệm "lắp ráp chéo" giữa lõi RNA và vỏ protein của hao kiểu TMV là S và HR.

## II. Thành phần hóa học và cấu trúc của DNA

Các *nucleic acid* (DNA, RNA) là những polymer sinh học, có trọng lượng phân tử lớn, gồm nhiều đơn phân (monomer) là các *nucleotide* nối với nhau tạo thành các chuỗi hay mạch *polynucleotide* - cấu trúc sơ cấp của các phân tử nucleic acid.

### 1. Cấu trúc của các nucleotide



**Hình 5.3** Bốn loại base của DNA và cấu trúc một nucleotide (dAMP).

Mỗi nucleotide gồm ba thành phần: một nhóm *phosphate*, một gốc đường pentose và một trong bốn loại *base nitơ* (các dẫn xuất của *purine* hoặc *pyrimidine*). Về cấu trúc, nhóm phosphate nối với gốc đường tại nguyên tử carbon số 5 (C<sub>5'</sub>) bằng một *liên kết ester* và base nitơ nối với gốc đường tại nguyên tử carbon số 1 (C<sub>1'</sub>) bằng một *liên kết β-glycosid* (Hình 5.3). *Lưu ý*: (1) Các base là thành phần đặc trưng qui định tên gọi các nucleotide mang chúng. Trong DNA chứa bốn loại base cơ bản:

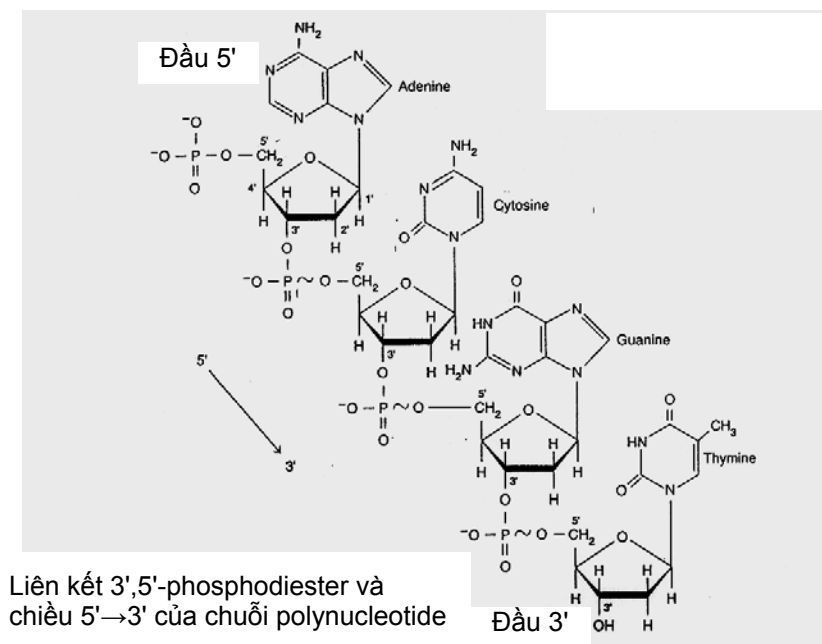
*adenine* (A), *thymine* (T), *guanine* (G) và *cytosine* (C); trong RNA cũng chứa bốn loại nhưng chỉ khác là *thymine* được thay bởi *uracil* (U).

(2) Đường pentose của RNA là *D-ribose* và của DNA là *2-deoxy-D-ribose* (ký hiệu D-: chỉ dạng đường quay phải để phân biệt với dạng L-quay trái không có trong thành phần của các nucleic acid tự nhiên). Hai gốc đường này khác nhau ở C<sub>2</sub>; trong ribose là nhóm -OH và trong deoxyribose là -H. Sự có mặt của thành phần đường trong các nucleotide phân biệt hai loại *ribonucleotide* và *deoxyribonucleotide* cấu tạo nên hai loại nucleic acid tương ứng là RNA và DNA.

(3) Cấu trúc "base + đường" gọi là *nucleoside*; cách đọc và viết tắt tương ứng với các base A, T, G và C trong DNA như sau: deoxyadenosine (dA), deoxythymidine (dT), deoxyguanosine (dG) và deoxycytidine (dC); cách gọi cho các nucleoside trong RNA chỉ cần bỏ tiền ngữ "deoxy" (d), và thay cho dT là uridine (U). Và nếu đọc đầy đủ tên gọi của một nucleotide chỉ cần thêm cái "đuôi" 5'-monophosphate, ví dụ ở hình 5.3.

(4) Tính phân cực trong cấu trúc một nucleotide còn thể hiện ở các nhóm hydroxyl (-OH) ở hai vị trí C<sub>5'</sub> (hình thành liên kết ester với nhóm phosphate để tạo ra nucleotide) và C<sub>3'</sub> (hình thành liên kết phosphodiester với nucleotide khác để tạo chuỗi polynucleotide).

## 2. Cấu trúc chuỗi polynucleotide



**Hình 5.4** Cấu trúc chuỗi polynucleotide của DNA.

Các nucleotide trong DNA hoặc RNA nối với nhau bằng các mối liên kết 3',5'-phosphodiester giữa gốc đường của nucleotide này với nhóm phosphate của nucleotide kế tiếp, tạo thành chuỗi *polynucleotide* (nhờ sự xúc tác của các enzyme tương ứng là DNA- hoặc RNA-polymerase). Vì vậy các chuỗi này bao giờ cũng được kéo dài theo chiều 5'→3' (đầu 5' mang nhóm phosphate tự do và đầu 3' chứa nhóm -OH tự do) và có bộ khung gồm các gốc đường và phosphate luân phiên nhau, còn các base nhờ vậy có trình tự được đọc theo một chiều xác định (hình 5.4).

Thông thường người ta biểu diễn trình tự base 5'→3' theo chiều từ trái sang phải. Ví dụ dưới đây cho thấy các chuỗi RNA và DNA chỉ khác nhau bởi base U và T và gốc đường trong các nucleotide của chúng.

chuỗi RNA 5'- AUAGACUACG-3'

chuỗi DNA 5'- dAdTdAdGdAdCdTdAdCdG-3'

### 3. Thành phần hóa học và cấu trúc của chuỗi xoắn kép DNA

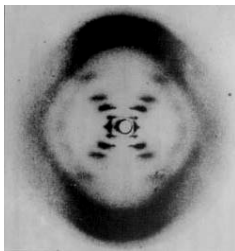
#### 3.1. Thành phần hóa học của DNA

Năm 1949, E.Chargaff áp dụng phương pháp sắc ký giấy vào phân tích thành phần hóa học của DNA các loài khác nhau (Bảng 5.1; có thể xem trong Watson *et al* 1987, tr.73) đã khám phá ra ba điểm quan trọng sau đây: (1) Số lượng bốn loại base trong DNA là không bằng nhau; (2) Tỷ lệ tương đối của các base là không ngẫu nhiên; và trong tất cả các mẫu DNA nghiên cứu tồn tại mối tương quan về hàm lượng (%) giữa các base như sau:  $A \approx T$  và  $G \approx C$ , nghĩa là tỷ số  $(A+G)/(T+C) \approx 1$ ; và (3) Mỗi loài có một tỷ lệ  $(A+T)/(G+C)$  đặc thù.

**Bảng 5.1 Thành phần base của DNA ở một số loài (từ nhiều tác giả)**

Sinh vật	A%	T%	G%	C%	$\frac{A+G}{T+C}$	$\frac{A+T}{G+C}$
Phage lambda	21,3	22,9	28,6	27,2	1,00	0,79
Phage T7	26,0	26,0	24,0	24,0	1,00	1,08
<i>Mycobacterium tuberculosis</i>	15,1	14,6	34,9	35,4	1,00	0,42
<i>Escherichia coli</i>	24,7	23,6	26,0	25,7	1,03	0,93
<i>Aspergillus niger</i> (nấm mốc)	25,0	24,9	25,1	25,0	1,00	1,00
<i>Saccharomyces cerevisiae</i>	31,3	32,9	18,7	17,1	1,00	1,79
<i>Triticum</i> (lúa mì)	27,3	27,1	22,7	22,8	1,00	1,19
<i>Zea mays</i> (ngô)	26,8	27,2	22,8	23,2	0,98	1,17
<i>Salmo salar</i> (cá hồi)	29,7	29,1	20,8	20,4	1,02	1,43
<i>Gallus domestica</i> (gà nhà)	29,5	27,7	22,4	20,4	1,08	1,34
<i>Homo sapiens</i> (người)	30,9	29,4	19,9	19,8	1,01	1,52

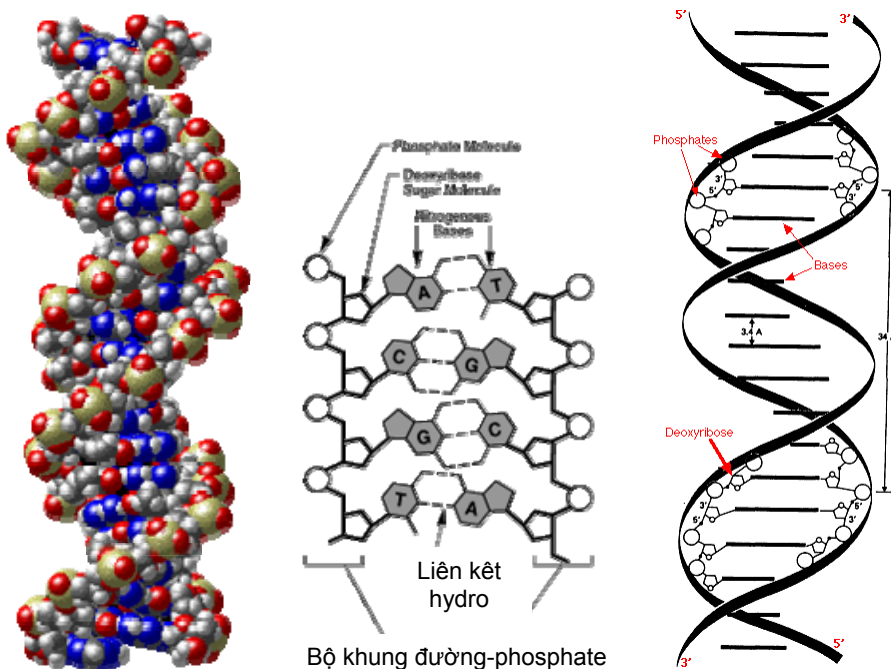
### 3.2. Cấu trúc chuỗi xoắn kép



Vào năm 1951-52, việc nghiên cứu cấu trúc ba chiều của DNA bằng phân tích nhiễu xạ tia X được bắt đầu bởi Maurice Wilkins và Rosalind Franklin. Các bức ảnh chụp được 1952 (hình 5.5) gợi ý rằng DNA có cấu trúc xoắn gồm hai hoặc ba chuỗi. Lúc này ở Anh còn có một số nghiên cứu khác nhằm phát triển lý thuyết nhiễu xạ của Linus Pauling để tìm hiểu cấu

**Hình 5.5** Ảnh chụp DNA tinh thể bằng tia X của Franklin.

trúc DNA. Tuy nhiên, giải pháp đúng đắn nhất là *chuỗi xoắn kép* (double helix) bổ sung do James Watson và Francis Crick đã đưa ra năm 1953. Mô hình này (hình 5.6) phù hợp với các số liệu của Wilkins và Franklin cũng như của Chargaff. Sự kiện này mở ra một bước ngoặt mới cho sự ra đời và phát triển với tốc độ như vũ bão của di truyền học phân tử và sinh học nói chung.



**Hình 5.6** Các mô hình cấu trúc DNA. Bên trái là mô hình không gian lấp đầy. Hình giữa là chuỗi xoắn duỗi thẳng cho thấy các cặp base ở giữa và hai bộ khung đường-phosphate ở hai bên, với các thành phần được chỉ ra bằng các mũi tên. Bên phải là sơ đồ chuỗi xoắn kép với hai sợi đối song song.

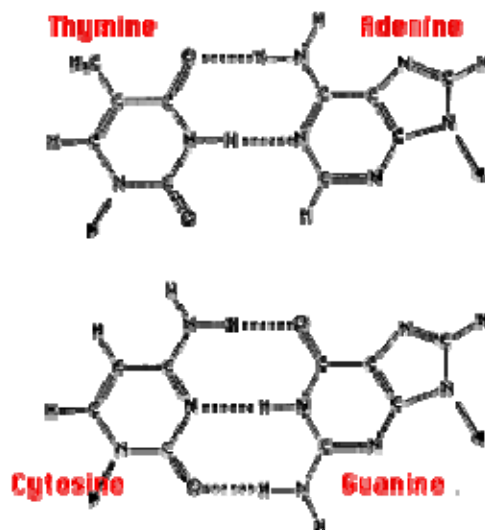
Mô hình Watson-Crick hay DNA dạng B có các đặc điểm sau:

(1) DNA gồm hai chuỗi *đối song song* (antiparallel) cùng uốn quanh một trục trung tâm theo chiều xoắn phải, với đường kính  $20\text{Å}$  ( $1\text{ångstrom} = 10^{-10}\text{m}$ ), gồm nhiều vòng xoắn lặp lại một cách đều đặn và chiều cao mỗi vòng xoắn là  $34\text{Å}$ , ứng với 10 *cặp base* (base pair, viết tắt là bp).

(2) Các bộ khung đường-phosphate phân bố ở mặt ngoài chuỗi xoắn và các base nằm ở bên trong; chúng xếp trên những mặt phẳng song song với nhau và thẳng góc với trục phân tử, với khoảng cách trung bình  $3,4\text{Å}$ .

(3) Hai sợi đơn gắn bó với nhau bằng các mối *liên kết hydro* (vốn là lực hóa học yếu) được hình thành giữa các cặp base đối diện theo *nguyên tắc bổ sung* "một purine - một pyrimidine". Cụ thể là, trong DNA chỉ tồn tại hai kiểu kết cặp base đặc thù là A-T (với hai liên kết hydro) và G-C (với ba liên kết hydro) (xem các hình 5.6 và 5.7).

(4) Tính chất bổ sung theo cặp base dẫn đến sự bổ sung về trình tự các base giữa hai sợi đơn của mỗi chuỗi xoắn kép. Vì vậy, trong DNA sợi kép bao giờ cũng có:  $A = T$  và  $G = C$  (quy luật Chargaff); nghĩa là  $(A+G) = (T+C)$  hay tỷ số  $\frac{A+G}{T+C} = 1$ , còn tỷ lệ  $\frac{A+T}{G+C}$  đặc thù cho từng loài.



**Hình 5.7** Hai kiểu kết cặp base của DNA. Cặp G-C nối với nhau bằng ba liên kết hydro (biểu thị bằng các đường chấm), có cùng hình dạng chính xác như cặp A-T nối với nhau bằng hai liên kết hydro. Các mũi tên chỉ vị trí liên kết giữa base với gốc đường.

Cần lưu ý rằng, hai đặc điểm quan trọng nhất trong cấu trúc DNA là sự *phân cực ngược chiều* của hai sợi đơn ( $5' \rightarrow 3'$  và  $3' \rightarrow 5'$ ) và *nguyên tắc bổ sung* giữa các cặp base. Hai đặc điểm này cho phép giải thích một cách cơ

bản các cơ chế di truyền ở cấp độ phân tử (tái bản, phiên mã và dịch mã).

#### 4. Sơ lược về các đặc tính hóa lý của các nucleic acid

##### 4.1. Các dạng của DNA

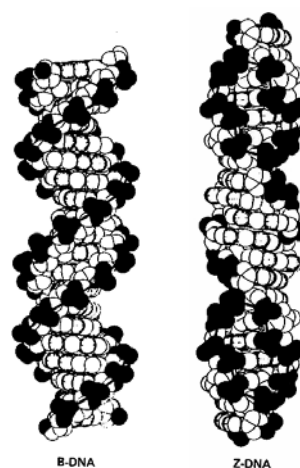
Mô hình Watson-Crick hay DNA *dạng B* là cấu trúc phổ biến. Tuy nhiên, sau này người ta còn phát hiện ra nhiều dạng xoắn phải khác (A,C,D...); chúng có một số biến đổi so với DNA-B (xem bảng 5.2).

Bên cạnh các dạng DNA xoắn phải, từ năm 1979 Alexander Rich và đồng sự còn phát hiện thêm một dạng DNA xoắn trái duy nhất cho đến nay, gọi là *DNA-Z*. Dạng DNA này có các bộ khung zigzag uốn gập khúc theo chiều trái, mỗi vòng xoắn dài  $45,6\text{\AA}$  chứa 12 cặp base (hình 5.8 và bảng 5.2). DNA dạng Z chứa các sợi gồm các purine và pyrimidine xếp xen kẽ nhau, thí dụ:

--GCGCGCGC--

--CGCGCGCG--

Mặc dù Rich khám phá DNA-Z khi nghiên cứu về các hợp chất mô hình, song cấu trúc này dường như cũng có mặt trong các tế bào sống ở một tỷ lệ nhỏ và vẫn còn chưa thật sự hiểu rõ chức năng của nó.



**Hình 5.8** DNA-B và DNA-Z

**Bảng 5.2** Đặc điểm của các dạng DNA - A, B, C và Z

Dạng	Chiều xoắn	Số bp/vòng xoắn	Đường kính chuỗi xoắn
A	Phải	11,0	$23\text{\AA}$
B	Phải	10,0	$19\text{\AA}$
C	Phải	9,3	$19\text{\AA}$
Z	Trái	12,0	$18\text{\AA}$

\*Nguồn: J.Kimball (từ internet). Về chi tiết, xem trong Watson *et al* 1987, tr.249.

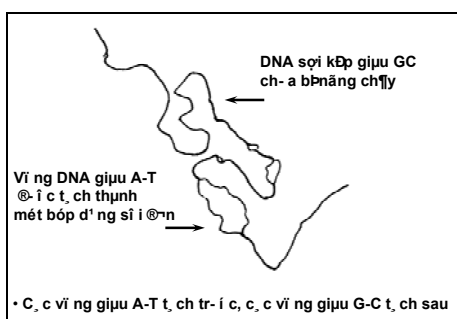
##### 4.2. Biến tính và hồi tính của DNA

###### 4.2.1. Khái niệm biến tính và hồi tính

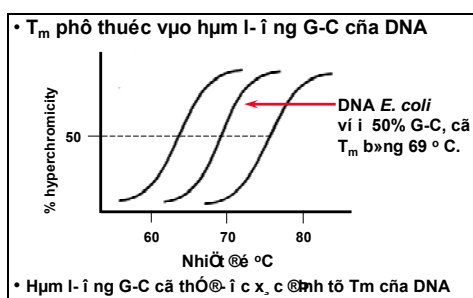
Một trong những đặc điểm quan trọng nhất của DNA là hai mạch đơn bổ sung của nó gắn với nhau bằng các mối liên kết hydro, vốn là lực liên kết hóa học yếu nên chúng có thể bị phân hủy dưới tác dụng của các enzyme, năng lượng... làm cho hai mạch đơn của chuỗi xoắn kép tách rời nhau. Nhờ đó DNA mới có thể tái bản, các gene mới có thể biểu hiện ra

các sản phẩm của mình. Mặt khác, DNA có thể phục hồi trở lại trạng thái ban đầu theo một quá trình ngược lại, nghĩa là có tính thuận-nghịch.

Bằng thực nghiệm, người ta đã chứng minh điều đó bằng cách sử dụng các tác nhân vật lý và hóa học khác nhau. Chẳng hạn, khi đun nóng từ từ các phân tử DNA lên tới nhiệt độ gần  $100^{\circ}\text{C}$  (thường là  $90-95^{\circ}\text{C}$ ), thì các liên kết hydro của chúng bị phá hủy hoàn toàn và hai sợi bổ sung tách ra. Hiện tượng này được gọi là *biến tính* (denaturation). Ngược lại, khi làm nguội từ từ dung dịch đột nóng chứa DNA bị biến tính hoàn toàn, các sợi đơn thường cặp đôi với sợi bổ sung của chúng và làm phục hồi chuỗi xoắn kép như lúc đầu. Hiện tượng đó gọi là *hồi tính* (renaturation).



**Hình 5.9** DNA biến tính từng phần.



**Hình 5.10** Sự phụ thuộc của  $T_m$  vào hàm lượng GC của 3 DNA khác nhau

Ở nhiệt độ vừa phải hoặc khi có mặt các tác nhân gây mất ổn định như alkali hay formamide, thì các phân tử DNA bị biến tính từng phần. Khi đó tại các vùng giàu cặp AT sẽ tách từng phần trước, trong khi các vùng giàu cặp GC vẫn giữ nguyên đặc tính xoắn kép. Điều này có thể lý giải là do mỗi cặp AT chỉ có hai liên kết hydro rõ ràng là kém bền hơn so với mỗi cặp GC có tới ba liên kết như thế.

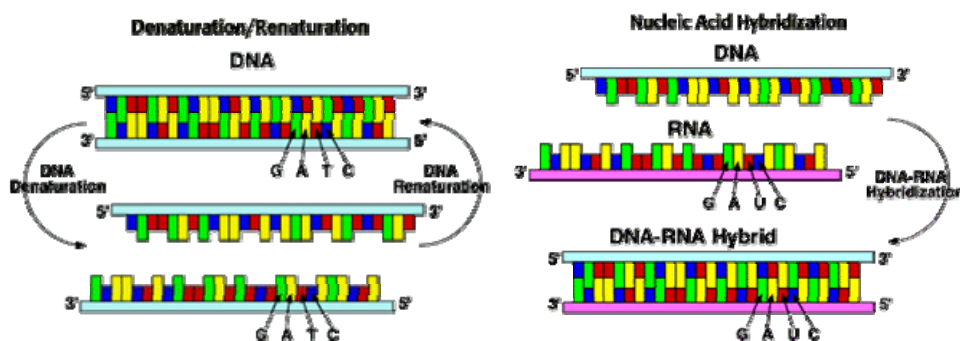
Nhiệt độ mà tại đó các sợi DNA bị biến tính hay tách nhau một nửa được gọi là *nhệt độ nóng chảy* (melting temperature), hay  $T_m$ .  $T_m$  là điểm giữa của pha chuyển tiếp (hình 5.9) và nó tùy thuộc vào hàm lượng G-C của DNA, nghĩa là đặc trưng cho DNA mỗi loài. Ví dụ, DNA của *E. coli* với 50% G-C thì có  $T_m$  là  $69^{\circ}\text{C}$  (hình 5.10).

Nói chung, hàm lượng GC của một DNA có thể biến thiên từ 22% ở nấm mốc nhầy *Dictyostelium* đến 73% ở *Mycobacterium phlei*. Điều này có thể gây một hiệu quả mạnh lên các đặc tính hóa lý của DNA, đặc biệt là lên nhiệt độ nóng chảy, vốn nó tăng tuyến tính với hàm lượng GC. Nhiệt độ nóng chảy ( $T_m$ ) của một DNA là nhiệt độ mà tại đó hai sợi bị biến tính hay tách nhau một nửa. Ngoài ra, nồng độ ion thấp và các dung môi hữu cơ cũng thúc đẩy sự biến tính của DNA.



#### 4.2.2. Ứng dụng trong lai phân tử (hình 5.11)

Người ta lợi dụng khả năng nói trên của DNA để tạo ra các phân tử DNA lai nhân tạo bằng cách làm lạnh từ từ hỗn hợp các DNA biến tính từ hai loài khác nhau. *Kỹ thuật lai phân tử* (molecular hybridization) này đã được ứng dụng rộng rãi để xác định mức độ tương đồng DNA của các nhóm phân loại khác nhau. Ví dụ, các thực nghiệm cho thấy có khoảng 25% tổng số DNA người và chuột có thể lai với nhau (Watson *et al* 1987).



**Hình 5.11** Biến tính và hồi tính của DNA (trái) và lai nucleic acid (phải).

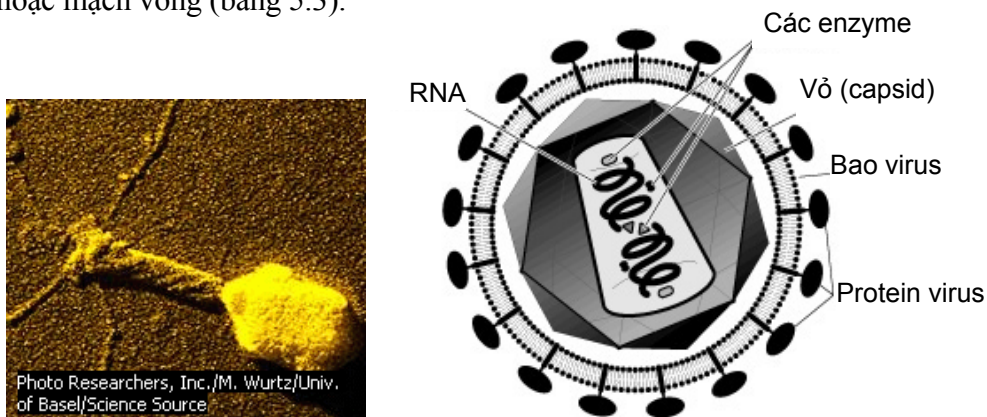
Theo nghĩa rộng, lai phân tử hay lai nucleic acid được hiểu là sự tương tác giữa các sợi nucleic acid bổ sung. Nó có thể xảy ra giữa hai sợi DNA hay giữa các sợi DNA và RNA, và đó chính là cơ sở của nhiều kỹ thuật, chẳng hạn như: lai một *mẫu dò có đánh dấu đồng vị phóng xạ* (radioactive probe) để lọc ra DNA hay RNA bám vào mang lai là một trong những thí nghiệm cho nhiều thông tin bổ ích nhất được tiến hành trong di truyền học phân tử. Hai kiểu cơ bản của lai phân tử là *phương pháp thấm Southern* (Southern blots) và *phương pháp thấm Northern* (Northern blots). Trong đó, kiểu đầu là lai bằng một mẫu dò để lọc DNA, còn kiểu sau dùng để lọc RNA. *Mẫu dò* (probe) là các công cụ sơ cấp được dùng để xác định các trình tự bổ sung cần quan tâm; nó là một nucleic acid sợi đơn được đánh dấu phóng xạ và được dùng để xác định một trình tự nucleic acid bổ sung bám trên màng lọc nitrocellulose hay màng lai bằng nylon. Nói chung, mẫu dò là một *dòng* (clone) được tạo ra bằng cách xen DNA vào một *vector*. Thông thường hầu hết các vật dò này là các dòng thuộc plasmid.

### III. Kích thước bộ gene và tính phức tạp về mặt tiến hóa

#### 1. Sơ lược về bộ gene của các virus, prokaryote và eukaryote

Virus là nhóm "sinh vật" bé nhất chưa có cấu tạo tế bào, không tồn tại đơn độc mà ký sinh bắt buộc ở các tế bào *sinh vật tiền nhân* (prokaryote) hoặc *sinh vật nhân chuẩn* (eukaryote); chỉ trong điều kiện đó chúng mới có khả năng tái bản. Các virus ký sinh hoặc gây nhiễm vi khuẩn gọi là *thể*

*thực khuẩn* (bacteriophage) hay *phage*. Cấu trúc của các virus tương đối đơn giản, gồm hai phần chính là lõi axit nucleic và vỏ protein (hình 5.12). Một số virus ở thực vật hoặc các virus gây ung thư, AIDS, SARS ở người và động vật có bộ gene là RNA. Số còn lại bao gồm nhiều virus ký sinh ở vi khuẩn và động vật có bộ gene là DNA sợi kép hoặc sợi đơn, có mạch thẳng hoặc mạch vòng (bảng 5.3).



**Hình 5.12** Ảnh chụp phage T4 (trái) và sơ đồ cấu trúc của HIV (phải).



**Hình 5.13** Ảnh chụp các tế bào *E. coli* (trái) và vật chất di truyền của nó (phải), và bên dưới là ảnh hiển vi điện tử của plasmid pSC101.

Nhóm prokaryote bao gồm các *vi khuẩn* (bacteria) và *vi khuẩn cổ* (archae) là các sinh vật có cấu tạo tế bào đơn giản nhất. Vi khuẩn *Escherichia coli* (hình 5.13) là đối tượng được sử dụng rộng rãi trong các nghiên cứu di truyền phân tử. Bộ gene chính của nó là một phân tử DNA sợi kép vòng có kích thước lớn (4.639.221 bp, với 4290 gene mã hóa protein và 53 gene RNA) gọi là nhiễm sắc thể chính. Nó thường tập trung ở một "vùng nhân" (nucleoid), không có màng nhân bao bọc, và ở trạng thái *siêu xoắn* (supercoiled DNA) dưới sự kiểm soát của các enzyme

*topoisomerase*. Ngoài ra còn có nhiều phân tử DNA sợi kép trần mạch vòng khác có kính thước bé hơn nhiều, gọi là các *plasmid* (Vấn đề này được trình bày kỹ trong giáo trình *Di truyền học Vi sinh vật và Ứng dụng*).

Nhóm eukaryote là nhóm lớn nhất và tiến hóa đa dạng nhất về trình độ tổ chức cơ thể, bao gồm tất cả các sinh vật có cấu tạo tế bào (trừ vi khuẩn và vi khuẩn lam), có thể là đơn bào hoặc đa bào. Trong tế bào chứa hai hệ thống di truyền, bộ gene nhân (đã trình bày ở chương 3 và 4) và bộ gene tế bào chất (xem chương 9). Kích thước bộ gene của một số sinh vật thường gặp được giới thiệu ở các bảng 5.3 và 5.4.

## 2. Mối quan hệ giữa kích thước bộ gene và tính phức tạp về tiến hóa

Dựa trên các kết quả phân tích bộ gene của các virus và vi khuẩn cũng như của bộ gene ở eukaryote (bộ nhiễm sắc thể đơn bội), cho phép khái quát như sau: *Kích thước của bộ gene tăng lên tương đối cùng với mức độ phức tạp về tiến hóa*. Thật vậy, từ bảng 5.3 ta thấy rằng kích thước bộ gene của các virus nói chung là rất nhỏ so với nhóm prokaryote; và kích thước bộ gene của các prokaryote lại tỏ ra quá đơn giản so với ngay cả một eukaryote đơn bào như nấm men.

*Ví dụ:* Trong khi DNA của một vi khuẩn điển hình như *E. coli* hơn 4,6 triệu cặp base và chứa 4.377 gene mã hóa protein (không kể 53 gene RNA); thì phage MS2 là một trong các virus bé nhất, bộ gene RNA của nó cũng chỉ có 3.569 base với tất cả 4 gene; hoặc có kích thước lớn như virus Epstein-Barr - bộ gene DNA sợi kép mạch vòng - cũng chỉ có 172.282 cặp base với tất cả 80 gene. Nếu xét trên cả hai nhóm prokaryote và eukaryote, ta thấy rằng kích thước các bộ gene biến thiên rất rộng: bộ gene đối với một sinh vật sống tự do (một vi khuẩn) được biết là bé nhất chứa khoảng 600.000 cặp base DNA, trong khi các bộ gen người và chuột là khoảng 3 tỷ. Nếu xét riêng ở nhóm eukaryote, ta đã biết mỗi loài có một số lượng nhiễm sắc thể đặc trưng và nó không phản ánh trình độ tiến hóa của các loài. Tuy nhiên, về mặt nào đó rõ ràng là có sự tương quan thuận giữa hàm lượng DNA của các bộ gene đơn bội và nấc thang tiến hóa của các động-thực vật. Dù vậy vẫn có một số ngoại lệ so với quy tắc này!

## 3. Hàm lượng DNA và nghịch lý giá trị C

Chẳng hạn, mặc dù *Psilotum nudum*, đôi khi gọi là "dương xỉ lông", là một thực vật đơn giản hơn nhiều so với loài *Arabidopsis thaliana* vốn là một thực vật có hoa thuộc họ cải, nhưng nó lại có hàm lượng DNA lớn hơn tới 3.000 lần. Mặc dù ta hiểu tại sao, nhưng có điều thú vị là trên 80% bộ gene của nó là DNA lặp lại không chứa thông tin di truyền nào cả! Hay một số thực vật (như ngô, loa kèn) hay một số động vật (như lưỡng thê, cá) lại có kích thước bộ gene lớn gấp nhiều lần so với lớp thú (bảng 5.3).

Một số lượng thể chứa DNA nhiều gấp bộ gene chúng ta đến 30 lần, nhưng chắc chắn không phải là chúng phức tạp gấp chúng ta 30 lần.

**Bảng 5.3 Kích thước bộ gene (genome) một số sinh vật thường gặp**

Bộ gene sinh vật	Số bp	Số gene	Ghi chú
<b>Virus</b>			
Phage Ø-X174 (ở <i>E. coli</i> )	5.386	10	DNA sợi đơn vòng
Phage lambda (ở <i>E. coli</i> )	48.502	~61	DNA sợi kép thẳng
Phage T2 hoặc T4 (ở <i>E. coli</i> )	$\sim 2 \times 10^5$	150-200	DNA sợi kép thẳng
Phage MS2 (ở <i>E. coli</i> )	3.569	4	RNA
SV40 (gây khối u ở khỉ)	5.226		DNA sợi kép vòng
Virus SARS (ví dụ, H5N1)	29.600		RNA
Epstein-Barr virus (EBV)	172.282	80	DNA sợi kép thẳng
<b>Prokaryote</b>			
<i>Haemophilus influenzae</i>	1.830.138	1.738	Gây nhiễm tai giữa
<i>Chlamydia trachomatis</i>	1.042.519	936	Bệnh lây qua tình dục
<i>Streptococcus pneumoniae</i>	2.160.837	2.236	<i>Pneumococcus</i>
<i>Vibrio cholerae</i>	4.033.460	3.890	Ở 2 NST; gây cholera
<i>Mycobacterium tuberculosis</i>	4.411.532	3.959	Gây bệnh lao
<i>Bacillus subtilis</i>	4.214.814	4.779	
<i>Escherichia coli</i> *	4.639.221	4.377	Có 4290 cistron
<i>Agrobacterium tumefaciens</i> **	4.674.062	5.419	Vector hữu ích...
<i>E. coli</i> O157:H7	$5,44 \times 10^6$	5.416	Gây bệnh ở người
<b>Eukaryote</b>			
<i>Saccharomyces cerevisiae</i>	12.495.682	5.770	Men bia nảy chồi
<i>Schizosaccharomyces pombe</i>	12.462.637	4.929	Nấm men phân cắt
<i>Plasmodium falciparum</i> ***	22.853.764	5.268	Gây sốt rét nguy hại
<i>Neurospora crassa</i>	38.639.769	10.082	+ 498 gene RNA
<i>Caenorhabditis elegans</i> ****	100.258.171	~19.000	Giải tr.tự đầu tiên...
<i>Arabidopsis thaliana</i>	115.409.949	25.498	Thực vật có hoa
<i>Drosophila melanogaster</i>	122.653.977	13.379	Ruồi giấm
Người ( <i>Homo sapiens</i> )*****	$\sim 3,2 \times 10^9$	~25.000	Giải tr.tự 4/2003 ...
Lúa ( <i>Oryza sativa</i> )	$4,3 \times 10^8$	~60.000	
Loa kèn ( <i>Lilium longiflorum</i> )	$3 \times 10^{11}$	?	
Chuột	$2,2 \times 10^9$	?	
Lưỡng thê	$10^9 - 10^{11}$	?	

*Chú thích:* \*Có 4290 gene mã hóa protein, còn lại là các gene RNA; \*\* Vector hữu ích để chuyển gene ở thực vật; \*\*\* Cộng với 53 gene RNA; \*\*\*\* Eukaryote đa bào đầu tiên được xác định trình tự; \*\*\*\*\* Được xác định đầy đủ trình tự vào 4/2003 bởi HGP, với tổng cộng 3.164.700.000 cặp base; và theo ước tính mới nhất công bố ngày 21/10/2004 trên tạp chí *Nature*, bộ gene người chúng ta chứa khoảng 20.000-25.000 gene mã hóa protein, chiếm khoảng 2% toàn bộ gene.

Người ta gọi tổng hàm lượng DNA trong bộ gene đơn bội là *giá trị C* (C-value). Với phân tích ở trên cho thấy không hề tồn tại một mối quan hệ nhất quán giữa giá trị C và tính phức tạp của một sinh vật (như lưỡng thể với thú); cái đó được gọi là *nghịch lý giá trị C* (C-value paradox).

#### 4. Về kích thước DNA các bào quan

Kích thước DNA của một số bào quan (bảng 5.4) cho thấy chúng có vẻ đơn giản và không có dấu hiệu tiến hóa rõ rệt. Nói chung, kích thước mỗi phân tử *DNA ty thể* (mitochondrial DNA = mtDNA) của người và các động vật có vú thường nằm trong khoảng 15.000-17.000 bp; ví dụ mtDNA người là 16.569 bp. Trong khi đó, kích thước một *DNA Lạp thể* (chloroplast DNA = ctDNA hay cpDNA) ở phần lớn tế bào các thực vật thường vào khoảng 130.000 -150.000 bp. Chẳng hạn, ctDNA ở lúa trồng (*O. sativa*) thuộc hai nhóm indica và japonica có kích thước tương ứng là 134.494 bp và 135.525 bp, ở lúa mỳ (*Triticum aestivum*) là 134.545 bp và ở ngô (*Zea mays*) là 140.384 bp... Còn các plasmid hiện diện ở một số tế bào thực vật thường có kích thước rất bé khoảng 1-2 ngàn cặp base.

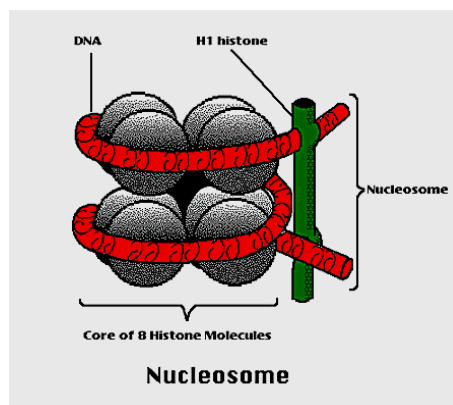
**Bảng 5.4 Kích thước DNA bào quan ở một số sinh vật**

DNA ty thể	Số cặp base	DNA Lạp thể	Số cặp base
Người	16.569	Lúa (ind; jap)	134494 ; 134525
<i>D. melanogaster</i>	19.517	Ngô	140.384
<i>S. cerevisiae</i>	85.779	Mía	141.182
Plasmid		Plasmid	
Lúa (indica; japonica)	1485 ; 2.135	<i>N. crassa</i>	3581; 3675; 7050
Ngô	1.913	<i>Brassica</i>	11.640

## IV. Tổ chức phân tử của các nhiễm sắc thể

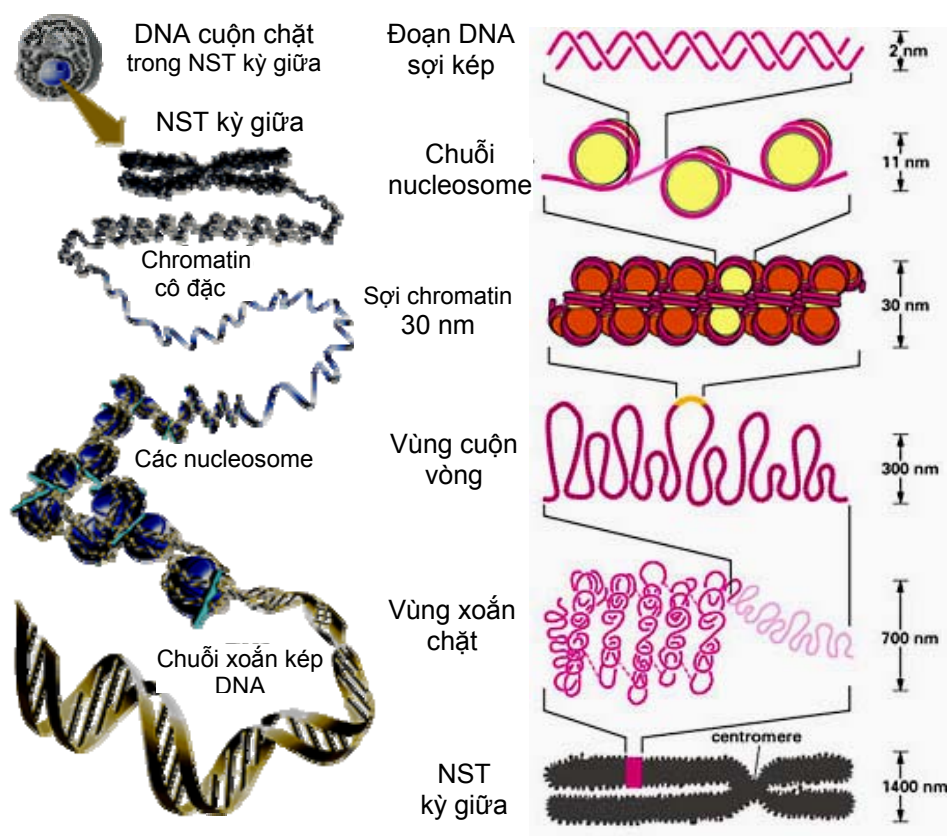
### 1. Cấu trúc chất nhiễm sắc

Ở đây ta chỉ tìm hiểu tổ chức của DNA trong các nhiễm sắc thể eukaryote. Mỗi nhiễm sắc thể eukaryote là một phức hợp *nucleoprotein* bao gồm một phân tử DNA sợi kép mạch thẳng kết hợp với các phân tử protein cơ sở gọi là histone (giàu các amino acid lysine và arginine). Phức hợp như thế gọi là *chất nhiễm sắc* (chromatin).



**Hình 5.14 Sơ đồ cấu trúc một nucleosome**

Đơn vị tổ chức của cơ sở các nhiễm sắc thể eukaryote là các *nucleosome* (hình 5.14). Mỗi nucleosome có đường kính khoảng 11 nm, gồm một khối cầu tám phân tử histone,  $(H2A + H2B + H3 + H4)_2$ , gọi là *lõi octamer* và đoạn DNA có kích thước 146 cặp base quấn  $1\frac{3}{4}$  vòng xung quanh nó (thường được mô tả đơn giản: một đoạn DNA dài 160-200 bp quấn hai vòng quanh octamer). Một phân tử H1 bám vào đoạn DNA "nối" (linker DNA) bên ngoài khối cầu, giữ vững sự tương tác của DNA với các histone lõi. Đoạn DNA nối giữa hai nucleosome dài khoảng 150 bp.



**Hình 5.15 Tổ chức DNA trong nhiễm sắc thể eukaryote.** DNA cuộn chặt trong nhiễm sắc thể kỳ giữa được mở xoắn dần qua các mức độ khác nhau (hình trái). DNA đóng xoắn dần qua các mức cho đến khi hình thành nhiễm sắc thể kỳ giữa nguyên phân, với chiều dài rút ngắn khoảng 50.000 lần (hình phải).

Như vậy, bậc cấu trúc đầu tiên của chromatin được hình dung dưới dạng một chuỗi các nucleosome, hay *sợi nucleosome* (nucleosome fiber) có độ dày 11 nm (1 nanomet =  $10 \text{ \AA}$ ). Bậc thứ hai của sự cuộn gấp chromatin có liên quan tới sự xoắn lại của sợi nucleosome thành ra một



sợi dày 25 nm gọi là một *solenoid*. Các histone H1 được coi là tham gia vào sự xoắn lại này bằng cách tương tác với các phân tử H1 khác. Bước thứ ba của sự hóa xoắn có lẽ là cuộn vòng của sợi 25 nm thành một cấu trúc kiểu như bàn chải (brushlike) với các vòng được neo dính vào một giá trung tâm (dày ~300 nm). Đây chính là vùng tháo giãn xoắn của nhiễm sắc thể, tương ứng với *chất đồng nhiễm sắc* (euchromatin). Sau đó các dây vòng được sắp xếp trong các không gian ba chiều này xoắn chặt tạo thành các vùng gọi là *chất dị nhiễm sắc* (heterochromatin) trên một *chromatid* với độ dày khoảng 700 nm. Tại kỳ giữa của nguyên phân, mỗi nhiễm sắc thể có cấu trúc điển hình gồm hai chromatid chị em (two sister chromatids) dính chung nhau ở *tâm động* (centromere) với độ dày toàn bộ chừng 1400 nm (hình 5.15). Như vậy, chất nhiễm sắc trong các tế bào eukaryote tồn tại ở hai dạng: (1) Chất dị nhiễm sắc là các phần cuộn xoắn chặt và không có hoạt tính phiên mã; và (2) Chất đồng nhiễm sắc là các vùng giãn xoắn và chỉ ít cũng có tiềm năng hoạt tính.

## 2. Sơ lược các thành phần DNA trong bộ gene eukaryote

Nói chung, trong mỗi bộ gene eukaryote bao gồm các kiểu DNA chính sau đây, và để cho tiện ta lấy bộ gene người làm thí dụ :

- DNA *bản sao đơn* (single-copy or unique), nghĩa là các gene có mặt chỉ một lần trong bộ gene; loại này chiếm khoảng 75% bộ gene và bao gồm hầu hết các gene.

- DNA *lặp lại* (repetitive) bao gồm một số gene được lặp lại vài lần cho đến hàng ngàn lần trong bộ gene. Nhóm này được chia làm hai loại:

- + DNA *lặp lại kiểu phân tán* (interspersed): loại này chiếm khoảng 15% và phân bố rải rác khắp bộ gen giữa các gene cũng như bên trong các gene; bao gồm các *trình tự Alu* (là các đoạn DNA dài khoảng 300 bp và được lặp lại khoảng 300.000 lần trong bộ gene; chúng có thể có mặt ở chỗ tiếp giáp hoặc bên trong các gene trong các intron hoặc các vùng không được dịch mã) và cả các đoạn lặp nối tiếp có số lượng biến thiên (VNTRs = variable numbers of tandem repeats), vốn là các đoạn lặp ngắn chỉ vài cặp base nhưng có chiều dài sai khác, hay các *tiểu vệ tinh* (mini- hay microsatellite). Ví dụ: các đoạn lặp phân tán (5'→3') trong các đoạn Alu:

GCTGCGG.....GCTGAGG.....GCTGAGG

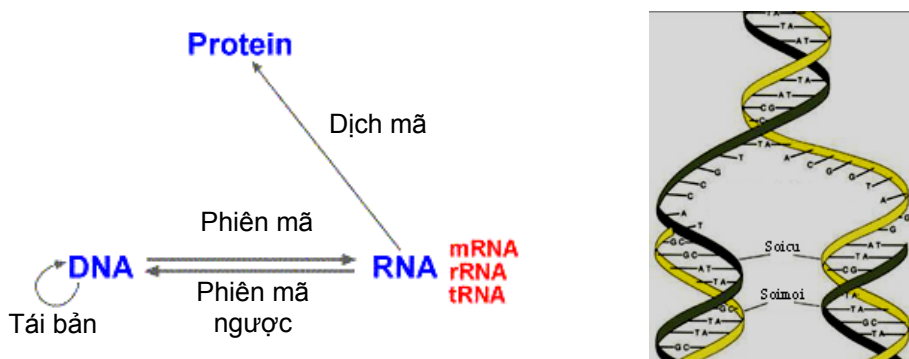
- + DNA *vệ tinh* (satellite) hay DNA lặp lại kiểu *nối tiếp* (tandem): loại này chiếm khoảng 10% và được lặp lại rất nhiều lần, bao gồm các trình tự đơn giản thường khu trú ở các *tâm động* (centromere) và các *đầu mút* (telomere) của các nhiễm sắc thể. Ví dụ: các đoạn lặp nối tiếp (5'→3') thuộc các microsatellite hay telomere:

5'-...TTAGGGTTAGGGTTAGGGTTAGGG...-3'

Nhìn chung, bộ gene người có các đặc điểm sau: (1) Về bộ gene nhân, kích thước bộ gene đơn bội là ~3,2 tỷ bp; hầu như tất cả mức độ phức tạp của nó đều nằm trong lớp DNA bản sao đơn (~75%). Bộ gene người được coi là có chứa khoảng 25.000 *gene mã hóa protein*. Trong đó đa phần là các *gene phân đoạn* với cấu trúc phức tạp (xem chương 6). Các gene đều chứa từ 1 cho đến trên 75 *exon* (đoạn mã hóa protein) và có chiều dài biến thiên từ dưới 100 cho đến khoảng 2.400.000 bp (gene gây rối loạn dưỡng cơ DMD được xem là dài nhất với 2,4 triệu bp này chiếm ~ 1,5% nhiễm sắc thể X mà nó định khu). Trình tự *Alu* có mặt khắp bộ gene. (2) Về bộ gene ty thể, đây là bộ gene mạch vòng với 16.569 bp và chứa ~ 40 gene.

## V. Tái bản DNA

Sau khi khám phá ra cấu trúc DNA, Watson và Crick đã đề xuất ba kiểu truyền thông tin di truyền có thể có trong các tế bào: (1) *Tái bản* (replication): DNA→DNA; (2) *Phiên mã* (transcription): DNA→RNA; và (3) *Dịch mã* (translation): RNA→Protein. Các kênh truyền thông tin di truyền này được gọi là *Giáo lý Trung tâm* (Central Dogma) của sinh học phân tử. Sau này đến năm 1970, Baltimore và Temin qua nghiên cứu cơ chế hoạt động của bộ gene RNA ở virus Sarcoma đã bổ sung thêm kênh RNA→DNA, gọi là *phiên mã ngược* (reverse transcription) vì nó ngược hướng với kênh phiên mã phổ biến. Điều này được minh họa ở hình 5.16.



**Hình 5.16** Các kiểu truyền thông tin di truyền (trái), và mô hình tái bản DNA theo kiểu bán bảo toàn do Watson và Crick đề xuất.

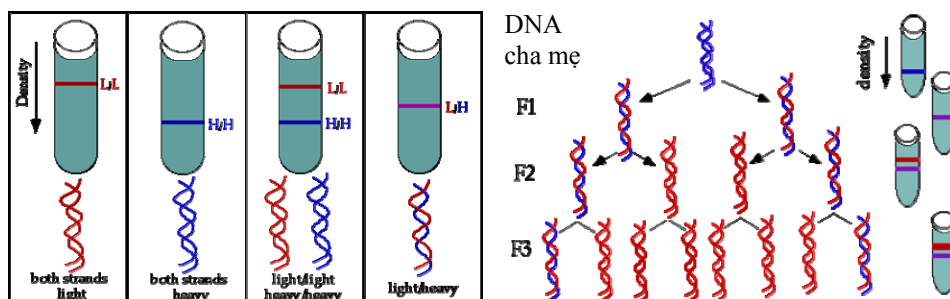
### 1. Các nguyên tắc và đặc điểm chung của tái bản DNA

*Tái bản* (replication) là một đặc tính quan trọng nhất của vật chất di truyền, nhờ đó sự sống được duy trì liên tục, các loài bảo tồn được tính chất đặc trưng của mình, và con cái thường giống bố mẹ. Vậy DNA và các bộ gene nói chung được tái bản như thế nào?



Trong khi khám phá ra mô hình cấu trúc DNA, chính Watson và Crick đã đưa ra dự đoán chính xác (dựa trên nguyên tắc bổ sung của các cặp base) rằng sự tái bản DNA phải diễn ra theo kiểu *bán bảo toàn* (*semi-conservative*) như ở hình 5.16. Đến 1956, S. Luria và Max Delbruck đề nghị ba kiểu tái bản có thể có: *bán bảo toàn*, *bảo toàn* (conservative) và *phân tán* (dispersive). Tuy nhiên, nhiều thí nghiệm sau đó đã nhanh chóng khẳng định sự tái bản DNA diễn ra theo kiểu bán bảo toàn.

Điển hình là thí nghiệm của Meselson và Stahl năm 1958 trên đối tượng là *E. coli* bằng phương pháp đánh dấu đồng vị phóng xạ  $N^{15}$  (nitơ nặng) kết hợp với *ly tâm siêu tốc* (ultra-centrifugation). Đầu tiên cho vi khuẩn này sinh trưởng trên môi trường  $N^{15}$ ; rồi đưa trở lại môi trường  $N^{14}$  (nitơ nhẹ) và sau một, hai hoặc ba thế hệ đều có lấy các mẫu DNA. Để tách DNA nặng và nhẹ, người ta cho trộn lẫn các mẫu này với cesium chloride (CsCl) trước khi đem ly tâm. Khi ly tâm, DNA có các tỷ trọng nặng (heavy), nhẹ (light) và trung bình (intermediate) sẽ tách thành các vạch tương ứng khác nhau trong ống nghiệm (hình 5.17).



**Hình 5.17** Thí nghiệm Meselson-Stahl về sự tái bản bán bảo toàn của DNA.

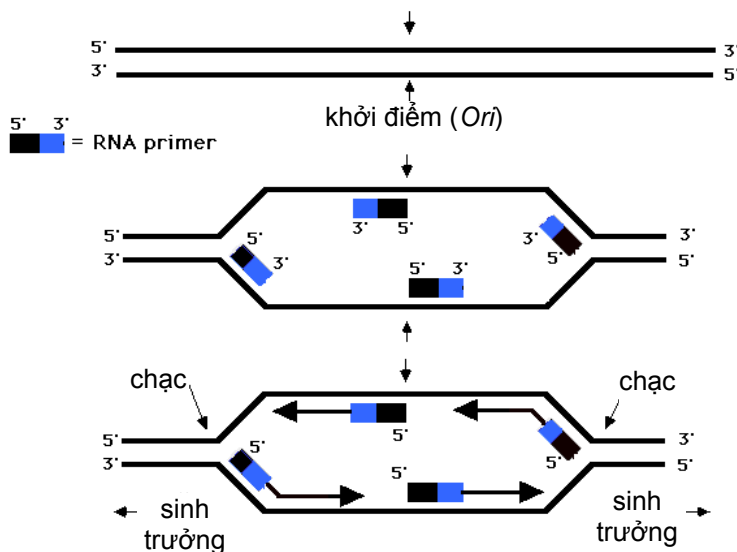
Các kết quả cho thấy rằng sau một thế hệ, 100% DNA sợi kép có tỷ trọng trung bình, nghĩa là một sợi nặng (từ dạng cha mẹ) và một sợi nhẹ (được tổng hợp mới). Kết quả này cho phép dự đoán và kết luận sự tái bản xảy ra theo kiểu bán bảo toàn đúng như dự đoán của Watson và Crick rằng, các sợi DNA làm khuôn cho sự tái bản của riêng chúng.

Dưới đây là các nguyên tắc và đặc điểm chung của tái bản DNA.

(i) Tái bản theo kiểu *bán bảo toàn* và *gián đoạn* (discontinuous);

(ii) Sự tái bản được bắt đầu tại một hoặc nhiều vị trí đặc thù trên phân tử DNA và diễn ra đồng thời theo hai hướng ngược nhau gọi là *khởi điểm tái bản hai hướng* (bidirectional origin of replication). Nghĩa là, từ khởi điểm này DNA sợi kép mở xoắn thành hình vòng mở sinh trưởng theo hai hướng đối lập nhau tạo ra hai *chạc tái bản* (replication fork). Mỗi khởi điểm tái bản cùng với hai chạc như vậy gọi là một *đơn vị tái bản*

(replicating unit hay replicon) được minh họa ở hình 5.18. Nói chung, đối với các DNA mạch vòng (bộ gene một số virus, vi khuẩn, các plasmid và các DNA bào quan của eukaryote), mỗi phân tử chỉ có một khởi điểm tái bản; trong khi đó mỗi DNA trong nhiễm sắc thể eukaryote có nhiều khởi điểm tái bản hoạt động theo một trình tự đặc thù (xem hình 5.20).



**Hình 5.18** Một khởi điểm và hai chạc tái bản sinh trưởng theo hai hướng đối lập nhau. Ở đây cũng cho thấy đoạn mỗi RNA được tổng hợp trước.

(iii) Quá trình tái bản DNA phụ thuộc vào một hệ thống gồm nhiều protein và enzyme khác nhau (xem mục 2 bên dưới);

(iv) Tại mỗi chạc tái bản, trước tiên xảy ra sự tổng hợp các đoạn mỗi RNA (primer) bởi vì các DNA polymerase tự nó không thể bắt đầu tổng hợp mới được. Mặt khác, do hai sợi đơn của mỗi chạc phân cực ngược chiều nhau trong khi các enzyme DNA polymerase và RNA polymerase chỉ xúc tác theo chiều  $3' \rightarrow 5'$ , cho nên phương thức tái bản DNA diễn ra trên hai sợi khuôn là không giống nhau: một sợi liên tục hay còn gọi là *sợi dẫn đầu* (leading strand) và một sợi không liên tục hay còn gọi là *sợi ra chậm* (lagging strand). Kiểu tái bản như thế gọi là *tái bản nửa gián đoạn* (semi-discontinuous).

Sự tổng hợp DNA không liên tục dưới dạng các đoạn ngắn do R. Okazaki phát hiện đầu tiên năm 1969; sau này chúng được gọi là các *đoạn Okazaki* (Okazaki fragments). Sau đó, các đoạn mỗi sẽ được cắt bỏ và chỗ trống được thay bằng DNA, và các đoạn Okazaki được nối lại bởi enzyme DNA ligase.

## 2. Các enzyme tham gia tái bản DNA

Mặc dù mỗi nhóm sinh vật có một hệ thống các enzyme và protein tái bản riêng, song chúng có các enzyme với vai trò chung nhất sau đây:

(1) Protein nhận biết và bám vào khởi điểm để từ đó hình thành nên "phức hợp mở" (open complex). Ở *E. coli*, đó là protein *dnaA*.

(2) *DNA gyrase*: mở cuộn DNA siêu xoắn phía trước mỗi chạc tái bản.

(3) *Helicase*: tháo xoắn DNA sợi kép tại mỗi chạc tạo thành các vùng sợi đơn. Ở *E. coli*, nó cũng gọi là protein *dnaB* hay protein *rep*.

(4) *Protein SSB* (single strand binding protein): bám vào các vùng DNA sợi đơn do helicase tách ra, giữ cho tạm thời không dính trở lại và nhờ đó mỗi sợi đơn mới có thể làm khuôn (template) cho tái bản.

(5) *Primase*: tổng hợp RNA mồi. Ở *E. coli*, nó còn gọi là protein *dnaG*.

(6) Các DNA polymerase xúc tác chính cho việc tổng hợp DNA mới nhờ có hoạt tính xúc tác: polymerase 5'→3', một số còn có hoạt tính đọc sửa: exonuclease 3'→5'. Ở *E. coli*, đó là *DNA polymerase III*.

(7) DNA polymerase vừa cắt bỏ dần đoạn mồi đi trước nhờ hoạt tính cắt bỏ exonuclease 5'→3', vừa kéo dài đoạn Okazaki theo sau lấp chỗ trống. Ở *E. coli*, đó là *DNA polymerase I*.

(8) *DNA ligase*: hàn liền khe hở giữa các đoạn DNA mới bằng cách hình thành liên kết 3',5'-phosphodiester.

**Bảng 5.5 Đặc tính của các DNA polymerase ở *E. coli* và người**

DNA polymerase <i>E. coli</i>	pol I	pol II	pol III		
Polymerase 5'→3'	có	có	có		
Exonuclease 3'→5'	có	có	có		
Exonuclease 5'→3'	có	không	không		
DNA polymerase người	$\alpha$	$\beta$	$\gamma$	$\delta$	$\epsilon$
Định vị	nhân	nhân	ty thể	nhân	nhân
Tái bản	có	không	có	có	có
Sửa chữa	không	có	không	có	có
Chức năng					
Polymerase 5'→3'	có	có	có	có	có
Exonuclease 3'→5'	không	không	có	có	có
Exonuclease 5'→3'	không	không	không	không	không
Primase	có	không	không	không	không
Kết hợp với PCNA*	không	không	không	có	có
Mẫu trượt (Processivity)	thấp			cao	
Tổng hợp sợi	ra chậm	sửa chữa	cả hai	đẫn đầu	ra chậm

*Lưu ý:* (i) Ở *E. coli*, cả ba loại protein *dnaB*, *dnaC* và *dnaG* hợp thành một phức hợp có tên là *primasome*; trong đó protein *dnaC* bám protein *dnaB*. Bảng 5.5 mô tả các đặc tính của các DNA polymerase ở *E. coli* và người đại diện cho các nhóm tương ứng, prokaryote và eukaryote.

(ii) Tất cả các DNA polymerase đều cần có môi với một nhóm 3'-OH tự do; xúc tác tổng hợp chuỗi theo một hướng 5'→3'; và chỉ một số enzyme này có *hoạt tính đọc sửa* (proofreading activity) 3'→5'.

(iii) Ngược với *E. coli*, các tế bào người có ít nhất năm loại DNA polymerase, trong đó ba loại chịu trách nhiệm tái bản DNA nhân là các DNA polymerase *alpha*, *delta* và *epsilon*. Polymerase  $\delta$  chịu trách nhiệm chính cho tổng hợp trên *sợi dẫn đầu* (leading strand) và thậm chí cả *sợi ra chậm* (lagging strand). Polymerase  $\alpha$  kết hợp với một RNA primase và được coi là có các hoạt tính cho tổng hợp một đoạn môi RNA-DNA ngắn. Có điều không chắc chắn lắm khi nói về chức năng của alpha trên sợi ra chậm. Vì dường như nó thiếu mất hoạt tính đọc sửa, nên không thể tổng hợp DNA với độ chính xác cao và như vậy có thể nó không phải là polymerase chính trong tổng hợp sợi ra chậm. Polymerase  $\epsilon$  có thể hoạt động như DNA polymerase I của vi khuẩn. Loại *kháng nguyên trong nhân làm tăng sinh tế bào* (proliferating cell nuclear antigen = PCNA) có vai trò trong cả tái bản lẫn sửa chữa. Một trong các chức năng của nó là dùng làm *nhân tố trượt* (processivity factor) cho DNA polymerase  $\delta$  và  $\epsilon$ . PCNA giữ DNA polymerase với sợi khuôn để tổng hợp DNA nhanh.

(iv) Ngoài ra còn có một số enzyme và protein đặc thù tham gia vào khâu kết thúc tái bản, tổng hợp các *đầu mút* (telomere) hoặc tham gia cắt nối trong quá trình tái tổ hợp.

### 3. Cơ chế tái bản DNA

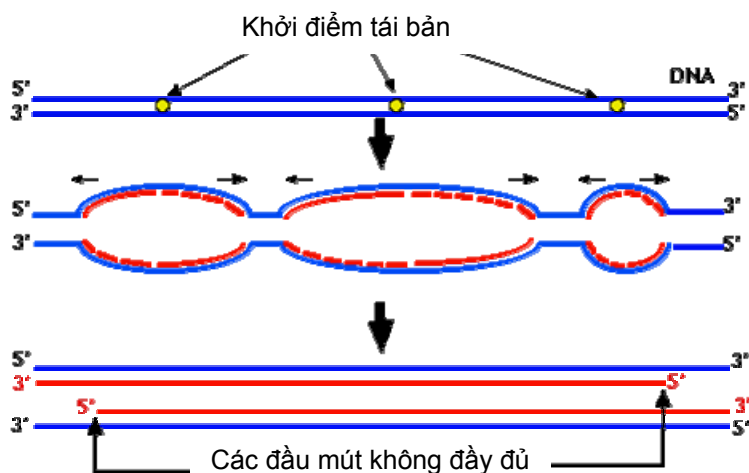
Trong phần này, chúng ta tìm hiểu kỹ cơ chế tái bản ở vi khuẩn *E. coli*, và nêu một ít vấn đề liên quan ở eukaryote mà đại diện là DNA người.

#### 3.1. Giai đoạn khởi đầu của sự tái bản (initiation of replication)

Đối với nhiễm sắc thể *E. coli*, sự tái bản bắt đầu tại một khởi điểm đặc thù gọi là *Ori* (hình 5.18). Trong khi đó, trên mỗi nhiễm sắc thể eukaryote có nhiều khởi điểm tái bản hai hướng như thế (hình 5.19). Quá trình diễn biến tại khởi điểm cho đến lúc hình thành hai chạc có thể tóm tắt (theo Kelman và O'Donnell, 1994) như sau: (1) Các protein bám khởi điểm *dnaA* được tổng hợp để bám *Ori* tạo ra cấu trúc nucleoprotein chuyên hoá của khởi điểm; (2) Sau đó, cấu trúc này mở xoắn vùng DNA giàu AT để hình thành "*phức hợp mở*" (open complex); và (3) Hai phân tử helicase *dnaB* chui vào phức hợp mở làm mở xoắn khởi điểm theo cả hai

hướng, tạo thành hai *chạc tái bản* (replication fork). Khi cả hai sợi đơn của mỗi chạc được tách ra thì các protein SSB bám vào. Nhờ vậy các sợi đơn được dùng làm khuôn hiệu quả cho các enzyme tái bản hoạt động.

Khởi đầu trong tái bản DNA là sự tổng hợp một đoạn mồi ngắn bởi primase, mà ở *E.coli* là phức hợp *primasome* (thể mở đầu) như đã nói ở trên. Kích thước đoạn mồi ở các sinh vật khác nhau là không giống nhau, và thường không vượt quá 12 ribonucleotide. Ở *E. coli*, theo kết quả nghiên cứu của bà Okazaki và cs (1985), đoạn này dài 10-12 nucleotide.



**Hình 5.19** Nhiều khởi điểm tái bản (origins of replication) trên một nhiễm sắc thể eukaryote. Ở đây cũng cho thấy các đầu mút (telomeres) không được tổng hợp đầy đủ, cụ thể là ở các đầu 5'.

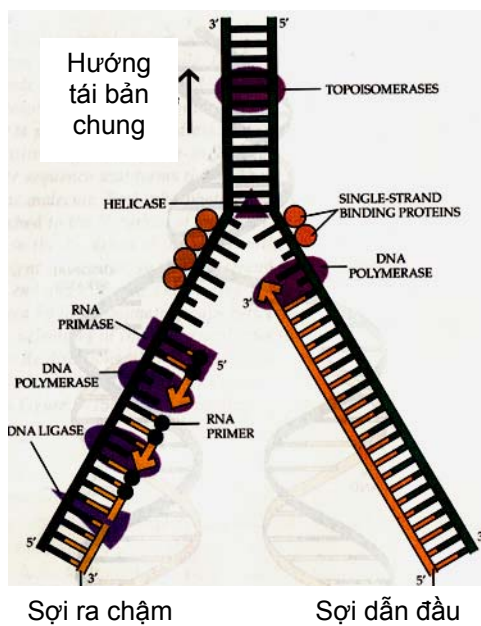
### 3.2. Giai đoạn kéo dài (elongation)

Một khi primosome tổng hợp xong một mồi, đó là lúc sự kéo dài chuỗi DNA mới bắt đầu. Ở *E. coli*, enzyme then chốt cho quá trình này là DNA polymerase III hoàn chỉnh (*holoenzyme*), một phức hợp gồm mười polypeptide khác nhau. Mỗi phân tử cho một sợi khuôn. Sự kéo dài trong suốt quá trình tổng hợp DNA ở *E. coli* rõ ràng là cần tới một *replisome* (thể tái bản) do kết hợp giữa primosome và DNA polymerase III hoàn chỉnh. Tốc độ tổng hợp trung bình là 1.000 nucleotide mỗi giây.

Sự tái bản DNA trên mỗi chạc, như đã nói ở trên, xảy ra theo kiểu *nửa gián đoạn* (semi-discontinuous). Cụ thể quá trình đó như sau (hình 5.20):

Trên sợi khuôn dẫn đầu (3'→5'): Trước tiên, enzyme primase hay primosome chỉ tổng hợp một đoạn mồi RNA với đầu 3'-OH tự do. Sau đó, enzyme hoàn chỉnh DNA polymerase III (*replisome*) bắt đầu kéo dài chuỗi DNA mới sinh trưởng theo chiều 5'→3' một cách liên tục.

Trên sợi khuôn ra chậm (5'→3'): Sự kéo dài diễn ra không liên tục dưới dạng các đoạn Okazaki. Kích thước trung bình mỗi đoạn Okazaki ở *E. coli* là 1.000 - 2.000 nucleotide; ở eukaryote là 100-200; và nói chung là 100-1.000 nucleotide. Quá trình này đòi hỏi sự "mồi hóa" nhiều lần và có tính chu kỳ, với sự tham gia lần lượt của bốn enzyme như sau: (i) primase tổng hợp một đoạn mồi RNA; (ii) DNA polymerase III hoàn chỉnh kéo dài đoạn Okazaki; (iii) DNA polymerase I vừa cắt bỏ dần từng nucleotide của đoạn mồi vừa lấp khoảng trống bằng cách kéo dài dần đoạn Okazaki theo sau;



**Hình 5.20 Sự tái bản nửa gián đoạn trên một chạc.**

(iv) DNA ligase hàn liền khe hở còn lại giữa hai đoạn Okazaki kề nhau bằng một liên kết 3',5'-phosphodiester. Quá trình cứ diễn ra luân phiên như vậy, và cuối cùng sợi DNA mới được tổng hợp trở nên dài ra và liên tục. Còn ở eukaryote, vai trò của các enzyme và protein tham gia vào giai đoạn kéo dài trên cả hai sợi khuôn đã được phân tích ở trên.

### 3.3. Giai đoạn kết thúc (termination)

Do đặc điểm cấu trúc nhiễm sắc thể ở hai nhóm prokaryote và eukaryote là hoàn toàn khác nhau, đặc biệt là cấu trúc các đầu mút nhiễm sắc thể eukaryote được phát hiện gần đây (hình 5.21), nên cơ chế kết thúc tái bản của chúng cũng khác nhau.

#### 3.3.1. Vấn đề kết thúc tái bản ở prokaryote

Để hoàn thành việc tái bản DNA, tế bào phải lấp đầy các khoảng trống do các RNA mồi bị cắt bỏ để lại. Đối với các DNA mạch vòng như của vi khuẩn chẳng hạn, không có vấn đề lấp tất cả các khoảng trống bởi vì bao giờ cũng có đầu 3' DNA khác nằm phía trước dùng làm mồi. Thật vậy, cả hai chạc tái bản được bắt đầu từ một khởi điểm duy nhất (*ori*), và di chuyển hầu như cùng tốc độ, theo hai hướng đối lập nhau xung quanh nhiễm sắc thể mạch vòng cho tới khi chúng gặp nhau tại một điểm kết thúc chung đối diện với *ori*. Thực ra, theo Bastia và cs (1997) cũng như

nhiều tác giả khác, đây là vùng chứa các trình tự đặc thù gọi là các *điểm kết thúc tái bản* (replication termini). Và tại các trình tự đặc thù này có các *protein kết thúc tái bản* (replication terminator protein = RTP) bám vào và các phức hợp protein-DNA này ngăn cản sự di chuyển của các chạc tái bản theo một cách phân cực hoặc định hướng đặc thù. Sự ngừng lại của các chạc tái bản tại vùng kết thúc tạo nên bước đầu tiên trong quá trình hoàn thành một vòng tái bản và sau đó, tách hai nhiễm sắc thể con rời ra một cách có trật tự nhờ xúc tác của *topoisomerase IV*.

*Lưu ý:* (1) Các nghiên cứu gần đây cho thấy RTP của *E. coli* có trọng lượng phân tử ~36kD, được mã hóa bởi gene *tus* (*ter*) và trong mỗi tế bào có ~80 bản sao của protein này được duy trì hầu như ổn định trong suốt chu kỳ tế bào. Còn RTP của *Bacillus subtilis* là một dimer có trọng lượng phân tử mỗi tiểu đơn vị là 14.500 kD, được mã hóa bởi gene *rtp*. (2) Các trình tự nhất trí (*consensus sequences*) của vùng kết thúc tái bản ở *E. coli* (R6K) và *B. subtilis*, theo Bastia và cs (1997), như sau:

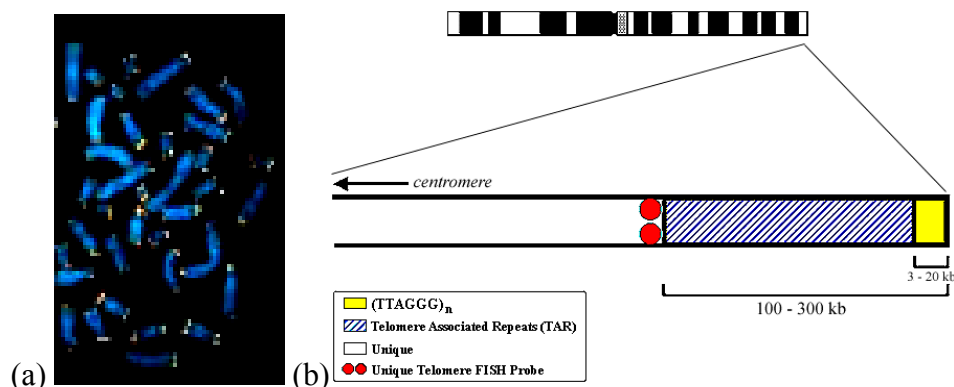
*E. coli* (R6K) 5'NN(A/T)(A/T)(A/T)G(A/T)(A/G)TGTTGTA ACTA(A/C)NN3'

*B. subtilis*

5'ACT(A/G)AN(T/A)(A/G)(A/C)(A/T)(C/T)(T/A)(A/G)TG(T/A)AC/CTTCAN3'

### 3.3.2. Vấn đề kết thúc tái bản ở eukaryote

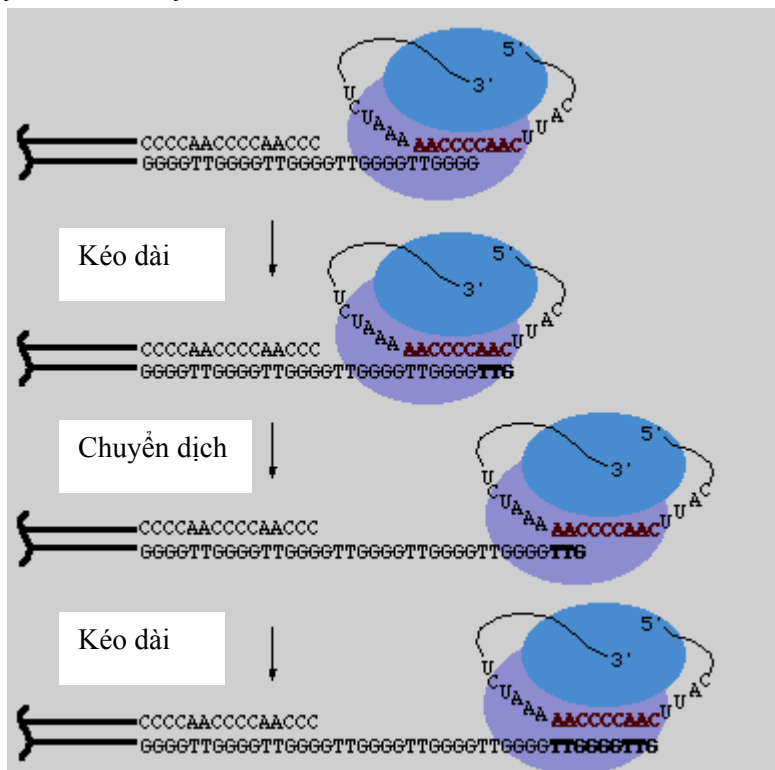
Như chúng ta đều biết, mỗi nhiễm sắc thể eukaryote chứa một phân tử DNA sợi kép mạch thẳng kết hợp với nhiều loại protein, có các đầu mút đặc trưng gọi là *telomere*. Một khi đoạn môi đầu tiên trên mỗi sợi được loại bỏ thì nó không có cách nào để bù đắp lại khoảng trống đó, bởi vì DNA không thể nào nối rộng theo chiều 3'→5', và cũng chẳng có đầu 3'



**Hình 5.21** (a) Ảnh chụp cho thấy các đầu mút nhiễm sắc thể - telomeres có màu vàng đặc trưng, và (b) sơ đồ minh họa tổ chức của telomere ở người.

nằm phía trước như trong các DNA mạch vòng. Nếu quả thực như vậy thì các sợi DNA sẽ ngắn bớt đi sau mỗi lần tái bản. Vậy các tế bào eukaryote

giải quyết vấn đề này như thế nào?



**Hình 5.22** Mô hình của Carol Greider về cơ chế tổng hợp các đoạn lặp telomere trên sợi giàu G nhờ telomerase ở *Tetrahymena*. Từ đây có thể hình dung các bước tiếp theo: lấp khoảng trống trên sợi đối diện giàu C được thực hiện bởi primase và DNA polymerase, và cuối cùng là cắt bỏ đoạn môi.

Các kết quả nghiên cứu đầu tiên của Elizabeth Blackburn và các đồng sự của bà đã giải đáp cho vấn đề này (hình 5.22). Các telomere không chứa các gene; thay vì thế chúng có cấu trúc đơn giản gồm những trình tự ngắn (6-8 cặp base) lặp lại nối tiếp cả ngàn lần và đặc thù cho từng loài. Chẳng hạn, ở *Tetrahymena*, một nhóm động vật nguyên sinh có lông tơ, nó là  $(TTGGGG)_n$ ; ở *Caenorhabditis*:  $(CCCTCCC)_n$ ; ở bọ *Oxytrichia* thuộc Euplotes:  $(CCCCAAA)_n$ ... Còn ở các động vật có vú và người, nó là 5'-TTAGGG-3' được lặp lại khoảng 1.000-2.000 lần (theo Yakoob và cs 1999, khoảng biến thiên phát hiện được trong các tế bào ở các giai đoạn khác nhau là 150-2000 lần). Các đoạn lặp này được gắn thêm vào đầu 3' của các sợi DNA, không phải bằng tái bản bán bảo toàn, mà bằng một enzyme gọi là *telomerase*. Nếu như telomerase không cho phép một cơ chế tái bản bán bảo toàn, thì nó không thể sử dụng một sợi DNA làm



khuôn cho sợi kia. Blackburn cho thấy rằng tính chất đặc thù này nằm ngay trong chính bản thân telomerase, và do một RNA nhỏ trong enzyme. Như vậy, telomerase là một ribonucleoprotein và là một *enzyme phiên mã ngược* (reverse transcriptase), tổng hợp DNA từ một khuôn RNA. Chẳng hạn, telomerase của *Tetrahymena* có một RNA dài 160 nucleotide có chứa trình tự 3'-CAACCCCAA-5' làm khuôn cho tổng hợp các đoạn lặp 5'-TTGGGG-3' trong các telomere của *Tetrahymena* (hình 5.22); ở người, phân tử RNA trong telomerase dài khoảng 450 nucleotide có chứa trình tự 3'-AAUCCC-5' (nằm trong đoạn ứng với vị trí 46-56) làm khuôn cho tổng hợp các đoạn lặp telomere 5'-TTAGGG-3' trên các sợi đơn có đầu mút 3'. Sau đó, ở sợi đối diện, DNA polymerase có thể hoàn thành việc tổng hợp "các đầu mút không đầy đủ" (incomplete ends) trên sợi đối diện sau khi đã được methyl hóa bên trong các telomere đó; và cuối cùng đoạn methyl sẽ bị cắt bỏ và khe hở được nối lại.

Các nghiên cứu gần đây cho thấy rằng: enzyme telomerase chỉ có mặt trong các *tế bào mầm* (germ cells), kể cả các *tế bào gốc của phôi* (embryonic stem cells); các *tế bào ung thư* (cancer cells); và các eukaryote đơn bào như *Tetrahymena thermophila*. Trong khi đó, các tế bào soma bình thường của động vật có vú không thấy có telomerase.

Điều đó cho phép lý giải tại sao các tế bào mầm cũng như các tế bào ung thư có khả năng phân chia gần như là vô hạn; hay nói cách khác, telomerase và sự duy trì chiều dài telomere chính là chìa khóa cho sự bất tử của tế bào. Ngược lại, hậu quả của sự vắng bóng telomerase trong các tế bào soma (do gene mã hóa nó bất hoạt) là ở chỗ: cứ sau mỗi lần *nguyên phân* (mitosis), tất cả 92 telomere của các tế bào soma người đồng loạt mỗi cái bị mất đi chừng 100 cặp base, nghĩa là khoảng 16 đoạn lặp TTAGGG. Theo lý thuyết, với tốc độ này thì sau 125 lần nguyên phân các telomere sẽ biến mất hoàn toàn. Trên thực tế, các thí nghiệm khác nhau kể từ khám phá đầu tiên của Leonard Hayflick vào đầu thập niên 1960 cho đến nay (Greider và Blackburn 1996; Hayflick 1997; Shay và cs 2004; ...) đã xác nhận rằng: chỉ sau khoảng 30-50 lần nguyên phân, các tế bào soma mất hẳn khả năng phân chia và đi vào giai đoạn *lão hóa* (senescence) và chết tự nhiên. Cái đó được gọi là *giới hạn Hayflick* (Hayflick limit).

Chính quá trình rút ngắn telomere theo thời gian (nghĩa là số lần nguyên phân) như thế khiến người ta liên tưởng tới sự tồn tại của cái gọi là "*đồng hồ di truyền cho sự lão hóa*" (genetic clock for aging), và bắt đầu tiến hành các nghiên cứu về "*hiệu ứng vị trí telomere*" (telomere position effect = TPE) (Borek 2002). Đó là lý do tại sao các tế bào soma có số lần phân chia hữu hạn trước khi chúng chết và cũng là lý do tại sao tất cả

chúng ta cũng như mọi sinh vật đều phải chết. Như thế, cái chết tất yếu trên phương diện sinh học, cái chết tự nhiên của thân xác đã được lý giải khá rõ ràng và đầy đủ. Và từ các nghiên cứu này đã mở ra triển vọng to lớn cho *liệu pháp ung thư* (cancer therapy) cũng như vấn đề lão hóa và bệnh tật phát sinh từ đó ở con người (Greider và Blackburn 1996; Yakoob và cs 1999; Borek 2002; Shay và cs 1998, 2004).

## **VI. Tái bản của các bộ gene RNA (RNA genomes)**

### *1. Đặc điểm tái bản của các bộ gene RNA virus và hậu quả của chúng*

Ở các virus có bộ gene RNA, phương thức tái bản của chúng rõ ràng là khác với các hệ thống tái bản DNA đã biết. Trước hết, các enzyme tổng hợp RNA không cần có môi (primer), vì vậy không có cơ hội cho việc *đọc sửa* (proofreading). Ngoài ra, về mặt di truyền RNA là một phân tử kém bền vững so với DNA bởi vì nhóm hydroxyl có mặt ở nguyên tử C<sub>2'</sub> của gốc đường cho phép cắt đứt (thủy phân) liên kết phosphodiester giữa hai ribonucleotide và tạo thành dạng phosphodiester vòng giữa các nhóm hydroxyl của C<sub>2'</sub> và C<sub>3'</sub> trong cùng một gốc đường; sau đó cấu trúc vòng này mở ra và để lại nhóm phosphate ở vị trí C<sub>3'</sub>. Sự kết hợp của hai đặc điểm trên là nguyên nhân làm hạn chế kích thước các bộ gene RNA. Các bộ gene này không thể sinh trưởng quá lớn, hoặc không thể tránh khỏi tỷ lệ sai sót cao trong quá trình tái bản ( $10^{-3}$ - $10^{-4}$ ). Vì vậy, trên thực tế, bộ gene RNA lớn nhất được biết là một RNA sợi đơn với chừng 29.600 base. Đó là trường hợp của virus gây bệnh sốt viêm phổi cấp (SARS) thuộc họ coronavirus được phát hiện hồi tháng 3/2002 tại một số nước châu Á, Canada...Đó cũng là lý do tại sao các virus RNA cho nhiều biến thể khác nhau đến như vậy, mặc dù kích thước bộ gene đều có lớn. Điều này làm cho các virus RNA thực sự trở thành mối hiểm họa bởi vì chúng có thể tiến hóa nhanh để xâm nhập vào các hệ thống miễn dịch của vật chủ. Chẳng hạn, các bộ gene HIV có nhiều biến thể gây khó khăn cho việc bào chế các vaccine và thuốc chống lại tất cả các biến thể của HIV.

Dưới đây nêu khái quát về hai kiểu tái bản của các virus RNA.

### *2. Tái bản của bộ gene RNA*

Kiểu truyền thông tin trực tiếp từ RNA sang RNA xảy ra trong các tế bào lây nhiễm virus RNA như: virus đốm thuốc lá và nhiều virus thực vật khác, kể cả các phage RNA như MS2... Bộ gene RNA của các virus này có mang gen mã hoá enzyme tái bản đặc thù. Sau khi được tổng hợp trong tế bào chủ, enzyme này sử dụng bản thân RNA virus làm khuôn để tổng hợp các phân tử RNA bổ sung. Đến lượt mình các RNA lại làm khuôn để tổng hợp trở lại các phân tử RNA của các virus thế hệ con .

### 3. Phiên mã ngược

*Phiên mã ngược* (reverse transcription) là kiểu truyền thông tin từ RNA sang DNA, chỉ xảy ra trong các tế bào động vật và người bị lây nhiễm bởi một số virus mang một sợi RNA có khả năng gây khối u hoặc hai sợi RNA như trường hợp HIV chẳng hạn. Trên mỗi sợi RNA lõi của các virus này có mang một *enzyme phiên mã ngược* (reverse transcriptase). Khi xâm nhập vào tế bào chủ, enzyme này sử dụng RNA của virus làm khuôn để tổng hợp sợi *DNA bổ sung* (complementary DNA = cDNA). Sau đó, sợi cDNA này có thể làm khuôn để tổng hợp trở lại bộ gene của virus (cDNA→RNA), hoặc tổng hợp ra sợi DNA thứ hai bổ sung với nó (cDNA→DNA) như trong trường hợp virus gây khối u mà kết quả là tạo ra một cDNA sợi kép. Phân tử DNA sợi kép được tổng hợp trước tiên trong quá trình lây nhiễm có thể xen vào DNA của vật chủ và ở trạng thái *tiền virus* (provirus). Vì vậy, provirus được truyền lại cho các tế bào con thông qua sự tái bản của DNA vật chủ, nghĩa là các tế bào con cháu của vật chủ cũng bị chuyển sang tình trạng có mầm bệnh ung thư. Các tế bào ung thư này mất khả năng kiểm soát sự sinh trưởng - phân chia điển hình của tế bào bình thường; chúng tăng sinh rất nhanh và tạo ra *khối u* (tumor). Đó chính là cơ chế gây ung thư bởi virus.

Ngày nay, người ta có thể tinh chiết các enzyme phiên mã ngược để phục vụ cho kỹ thuật tạo dòng cDNA tái tổ hợp (chương 10).

## Câu hỏi và Bài tập

1. Trên cơ sở các đặc điểm của mô hình Watson-Crick, hãy cho biết: (a) Mô hình này cho phép giải thích các quy luật Chargaff như thế nào? (b) Cơ sở của đặc tính đối song song trong chuỗi xoắn kép DNA là gì? (c) Mô hình này gọi lên khả năng tự tái bản của DNA ra sao?

2. Hãy xác định hàm tương đối (%) của các nucleotide trong DNA của người, cá hồi (*Salmo salar*) và gà (*Gallus domesticus*). Biết rằng tỷ lệ  $(A+T)/(G+C)$  tương ứng của các loài này là 1,52; 1,43; và 1,34.

3. Dựa vào cấu trúc hóa học lập thể của các base (A, T, G và C), hãy giải thích: (a) Tại sao trong ADN chỉ tồn tại hai kiểu kết cặp A-T và G-C mà không có các kiểu khác? (b) Các kiểu kết cặp A-C và G-T có thể xảy ra khi nào và gây ra hậu quả gì? Giải thích và cho các sơ đồ minh họa.

4. Giả sử tỷ lệ  $(A+T)/(G+C)$  ở một sợi của chuỗi xoắn kép ADN là 0,25. (a) Hãy cho biết tỷ lệ này trên sợi bổ sung và trên cả phân tử? (b) Nếu cho rằng 0,25 là của tỷ lệ  $(A+G)/(T+C)$ , thì tỷ lệ này trên sợi bổ sung

và trên cả phân tử sẽ như thế nào?

5. Hãy chỉ ra những điểm giống nhau và khác nhau giữa prokaryote và eukaryote về các vấn đề sau: (a) Khởi điểm tái bản; (b) Các enzyme tham gia tái bản; và (c) Các cơ chế mở đầu, kéo dài, và kết thúc tái bản.

6. Dựa vào hiểu biết về thuyết di truyền nhiễm sắc thể, hãy giải thích thí nghiệm Griffith về sự biến nạp ở phé cầu khuẩn *S. pneumoniae*.

7. So sánh cấu trúc của các nucleotide và các chuỗi polynucleotide cũng như cấu trúc phân tử của DNA và RNA.

8. Thế nào là biến tính và hồi tính của DNA? Các hiện tượng này có ý nghĩa như thế nào đối với các bộ gene sinh vật và kỹ thuật di truyền?

9. Thế nào là giá trị C (C-value) và nghịch lý giá trị C (C-value paradox)? Cho thí dụ và nêu ý nghĩa về mặt lý luận của hiểu biết này.

10. Sự tái bản của các bộ gene RNA ở một số virus khác với các hệ thống tái bản DNA ở những điểm nào? Tại sao kích thước các bộ gene RNA thường nhỏ, nhưng chúng lại có khả năng gây nguy hiểm thực sự đối với các vật chủ? Hãy giải thích một cơ chế gây ung thư bởi virus.

## Tài liệu Tham khảo

### Tiếng Việt

Hồ Huỳnh Thùy Dương. 1997. *Sinh học Phân tử*. NXB Giáo Dục.

Phạm Thành Hồ. 2000. *Di truyền học*. Tái bản lần II, NXB Giáo Dục.

Phan Cự Nhân. 2001. *Di truyền học Động vật*. NXB Khoa học và Kỹ thuật, Hà Nội.

Phan Cự Nhân. 1999. Công nghệ ADN tái tổ hợp. Trong: *Di truyền học tập II* (Phan Cự Nhân, chủ biên). Trang: 257-303. NXB Giáo Dục, Hà Nội.

Hoàng Trọng Phán. 1995. *Một số vấn đề về Di truyền học hiện đại* (Tài liệu BDTX cho giáo viên THPT chu kỳ 1993-1996). Trường ĐHSP Huế.

Hoàng Trọng Phán. 1997. *Di truyền học Phân tử*. NXB Giáo Dục.

Watson JD. 1968. *Chuỗi xoắn kép: Hồi ký về việc phát minh ra cấu trúc của DNA*. (Bản dịch của Lê Đình Lương và Thái Doãn Tĩnh). NXB Khoa học và Kỹ thuật, Hà Nội, 1984.

### Tiếng Anh

Alberts B, Johnson A, Lewis J, Raff M, Robert K, Walter P. 2002. *Molecular Biology of the Cell*. Garland Science, NY.

Bastia D, Manna AC, Sahoo T. 1997. Termination of DNA replication in

prokaryotic chromosomes. In: *Genetic Engineering, Vol. 19.* (Setlow JK ed.) pp 101-119. Plenum Press, New York, USA.

Borek C. 2002. *Telomere Control and Cellular Aging*:

[http://www.lef.org/magazine/mag2002/oct2002 report telo 01.html](http://www.lef.org/magazine/mag2002/oct2002%20report%20telo%2001.html).

Donovan S and Diffley JFX. 1995. Replication origins in eukaryotes. In: *Current Opinion in Genetics & Development* (Herr W and Kingston R, eds.), Vol.5(2): 203-207. Current Biology Ltd, UK.

Greider CW and Blackburn EH. 1996. Telomere, Telomerase and Cancer. *Scientific American*, 2/96, p.92: <http://www.genethik.de/telomerase.htm>.

Hamlin JL and Dijkwel PA. 1995. On the nature of replication origins in higher eukaryotes. In: *Current Opinion in Genetics & Development* (Herr W and Kingston R, eds.), Vol.5(2): 153-161. Current Biology Ltd, UK.

Harley CB, Villeponteau B. 1995. Telomeres and telomerase in aging and cancer. In: *Current Opinion in Genetics & Development* (Herr W and Kingston R, eds.), Vol.5(2): 249-255. Current Biology Ltd, UK.

Hayflick L. 1997. *Mortality and Immortality at the Cellular Level. A Review*. Copyright 2005BioProtNetwork:[webmaster@ protein.bio.msu.su](mailto:webmaster@protein.bio.msu.su).

Kelman Z, O'Donnell M. 1994. DNA replication: enzymology and mechanisms. In: *Current Opinion in Genetics & Development* (Stillman B and Green M, eds.), Vol.4(2): 185-195. Current Biology Ltd, UK.

Lingner J and Cesh TR. 1998. Telomerase and chromosome end maintenance. In: *Current Opinion in Genetics & Development* (Allis CD and Gasser SM, eds.), Vol.8(2): 226-232. Current Biology Ltd, UK.

Lewis R. 2003. *Human Genetics: Concepts and Applications*. 5<sup>th</sup> ed, McGraw-Hill, Inc, NY.

Price CM. 1999. Telomeres and telomerase: broad effects on cell growth. In: *Current Opinion in Genetics & Development* (Kadonaga JT and Grunstein M, eds.), Vol.8(2): 218-224. Current Biology Ltd, UK.

Russell PJ. 2003. *Essential Genetics*. Benjamin/Cummings Publishing Company, Inc, Menlo Park, CA.

Shay JW and Wright WE. 1998. *Teleomerase, Cellular Senescence and Cancer*. Page maintained by: [Cell Biology and Neuroscience](#).

Shay JW. 2002. Telomeres, Aging, and Tumorigenesis. *Gac Méd Méx* Vol.138 Suplemento No.1, 2002, pp.S2-S3.

Shay J, Schneider E, Masoro EJ. 2004. *Is there a genetic clock for aging?:* [www.infoaging.org/b-tel-home.html](http://www.infoaging.org/b-tel-home.html).

Tamarin RH. 1999. *Principles of Genetics*. 6<sup>th</sup> ed, McGraw-Hill, Inc, NY.

Twyman RM. 1998. *Advanced Molecular Biology*. BIOS Scientific Publishers Ltd/ Springer-Verlag Singapore Pte Ltd.

Watson JD, Hopkins NH, Roberts JW, Steitz JA, Weiner AM. 1987. *Molecular Biology of the Gene*. 4<sup>th</sup> ed, Benjamin/Cummings Publishing Company, Inc, Menlo Park, CA.

Weaver RF, Hedrick PW. 1997. *Genetics*. 3<sup>rd</sup> ed, McGraw-Hill Companies, Inc. Wm.C.Brown Publishers, Dubuque, IA.

Yakoob J, Hu G-L, Fan X-G, Zhang Z. 1999. Telomere, telomerase and digestive cancer. *World Journal of Gastroenterology*, 5(4):334-337.

### **Một số trang web**

[http://www.lef.org/magazine/mag2002/oct2002 report telo 01.html](http://www.lef.org/magazine/mag2002/oct2002%20report%20telo%2001.html).

<http://www.genethik.de/telomerase.htm>.

<http://www.infoaging.org/b-tel-home.html>.

<http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/>

<http://www.ornl.gov/hgmis>

<http://www.horizonpress.com/pcr/>

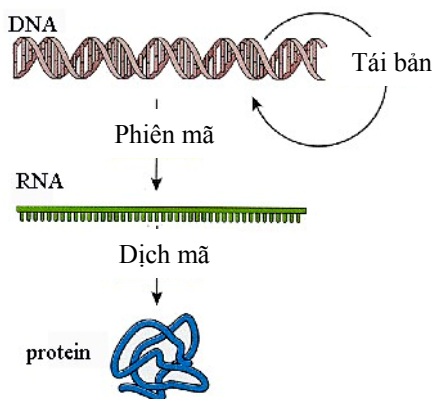
<http://dir.niehs.nih.gov/dirlmg/repl.html>

## Chương 6

# Gene và Quá trình Sinh tổng hợp Protein

Gene, đơn vị thông tin được truyền từ cha mẹ cho con cái, là khái niệm then chốt của di truyền học. Nói đến gene tức là nói đến DNA và các quan hệ của nó với RNA và protein dưới dạng sơ đồ sau đây, được gọi là *Lý thuyết trung tâm* (Central Dogma) của Sinh học Phân tử (hình 6.1).

Trong đó các sợi đơn của DNA được dùng làm khuôn cho *tái bản* (replication; như đã xét ở chương 5). Mặt khác, từng đoạn xác định của nó (tức các gene) có thể làm khuôn cho sự tổng hợp các RNA trong một quá trình gọi là *phiên mã* (transcription). Đến lượt, các phân tử RNA này lại làm khuôn cho sự tổng hợp các chuỗi polypeptide mà từ đó tạo thành các protein; quá trình này được



**Hình 6.1** Lý thuyết trung tâm của Sinh học Phân tử

gọi là *dịch mã* (translation), bởi vì nó chuyển bức thông tin dưới dạng các nucleotide thành ra sản phẩm được xây dựng bằng các amino acid. Hai quá trình sau được coi là hai giai đoạn chính trong sự biểu hiện của *gene mã hóa protein* (protein coding gene). Thực ra, sự biểu hiện của một gene chịu sự kiểm soát ở nhiều cấp độ khác nhau (chương 7).

Trong chương này, chúng ta sẽ lần lượt tìm hiểu sự phát triển của khái niệm gene, cấu trúc và chức năng protein - sản phẩm của gene, bản chất của mã di truyền, và các quá trình phiên mã và dịch mã.

### I. Sự phát triển của khái niệm gene

#### 1. Các quan niệm của Mendel và Morgan về gene

Mendel là người đầu tiên nêu lên định nghĩa về *gene* năm 1865 (thuật ngữ này được Johannsen đưa ra năm 1909). Theo đó, gene là đơn vị di truyền tồn tại ở dạng hạt riêng biệt, xác định một tính trạng cụ thể trong cặp tính trạng tương phản. Đây mới chỉ là sự suy luận thuần túy, không có cơ sở vật chất đặc thù.

Quan niệm chính xác hơn về cơ sở vật chất và chức năng của gene nảy

sinh từ nhiều nguồn nghiên cứu độc lập nhau trong suốt 50 năm đầu của thế kỷ XX. Trường phái Morgan sau khi xác định các gene nằm trên nhiễm sắc thể và đề xuất phương pháp lập bản đồ gene bằng tái tổ hợp, đã khẳng định rằng các gene là những đơn vị cơ sở và không chia nhỏ của vật chất di truyền về cả cấu trúc lẫn chức năng; chúng liên kết với nhau theo kiểu thẳng hàng trên nhiễm sắc thể.

### 2. Giả thuyết một gene - một enzyme của Beadle và Tatum

Một hướng nghiên cứu khác tập trung vào phương diện chức năng sinh hóa của gene. Năm 1902, Archibald Garrod gợi ý rằng rối loạn chuyển hóa *alkapton niệu* (alcaptonuria) bắt nguồn từ một sai hỏng của một enzyme đặc thù và được di truyền theo kiểu lặn nhiễm sắc thể thường, mà ông gọi là sai sót chuyển hóa bẩm sinh. Đến năm 1941, Beadle và Tatum mới làm sáng tỏ ý tưởng trên bằng các thí nghiệm gây đột biến bằng tia X ở *Neurospora*. Để giải thích các tổn thương sinh hóa đặc thù do đột biến, họ đã đề xuất "*giả thuyết một gene - một enzyme*" nổi tiếng; nó được xem như là mô hình về chức năng của gene, mở đường cho sự ra đời của di truyền sinh hóa. Về sau, quan niệm "một gene - một enzyme" được mở rộng thành "*một gene - một protein*", và tiếp tục chính xác hóa bằng mệnh đề "*một gene - một polypeptide*".

Thật vậy, từ khi Avery và các đồng sự chứng minh DNA là vật chất mang thông tin di truyền vào năm 1944, và đặc biệt là sau khi Watson và Crick khám phá ra cấu trúc phân tử DNA năm 1953, quan niệm về gene không ngừng được phát triển và chính xác hóa. Về mặt cấu trúc, gene là một đoạn xác định của bộ gene (DNA ở hầu hết sinh vật và RNA ở một vài virus). Về phương diện chức năng, như chúng ta đã rõ, không phải mọi gene đều mã hóa các enzyme mà một số mã hóa các polypeptide với các chức năng khác nhau, và một số mã hóa các phân tử RNA chức năng như RNA ribosome (rRNA) và RNA vận chuyển (tRNA). Hơn nữa, thông tin trong gene có thể được sử dụng một cách có chọn lọc để sinh ra nhiều hơn một loại sản phẩm (các gene phân đoạn). Trước tiên, ta hãy tìm hiểu công trình nghiên cứu của Benzer về cấu trúc tinh vi của gene.

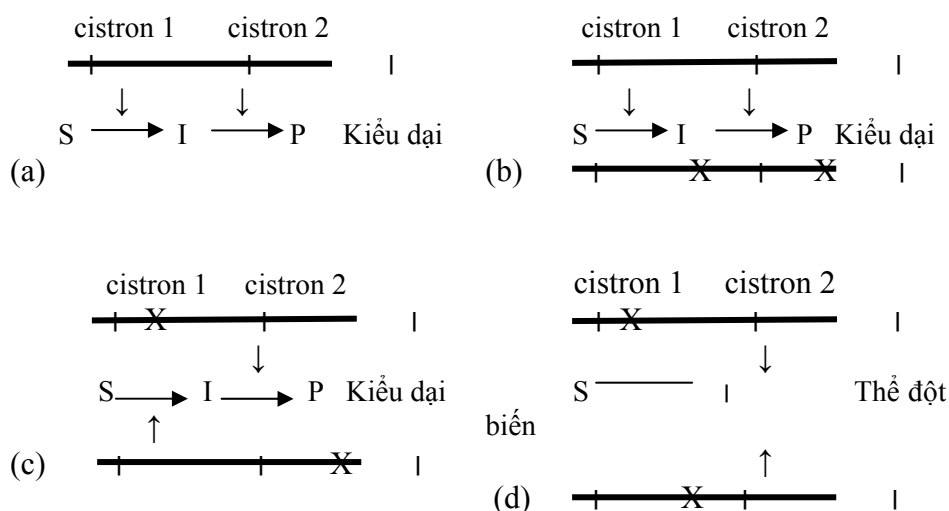
### 3. Quan niệm của Benzer về các đơn vị cấu trúc và chức năng di truyền

Các công trình nghiên cứu của Seymour Benzer (từ 1957 đến 1961) về tái tổ hợp ở phage T4 đã cho thấy rằng, gene theo quan niệm của Morgan có thể chia nhỏ thành các đơn vị nhỏ hơn. Ông đã đưa ra các thuật ngữ *muton*, *recon* và *cistron* để định nghĩa các đơn vị không chia nhỏ tương ứng là đột biến, tái tổ hợp và chức năng. Bằng cách lai các thể đột biến của cùng một gene có nguồn gốc độc lập nhau trong khi cho lây nhiễm phage, đã làm xuất hiện phage kiểu đại. Điều này chỉ có thể xảy ra bởi sự



tái tổ hợp bên trong gene (intragenic recombination), nếu như các phần nhỏ riêng biệt của gene đều bị đột biến. Điều này chứng tỏ rằng gene bị phân chia thành các đơn vị nhỏ hơn thông qua tái tổ hợp và đột biến. Tuy nhiên, vì kích thước của muton và recon được coi là tương đương với một cặp nucleotide, cho nên ngày nay tự thân hai đơn vị này không còn giá trị sử dụng nữa.

Thuật ngữ *cistron* của Benzer có nghĩa là đơn vị chức năng di truyền không chia nhỏ. Điều này có thể xác định bằng sự *phân tích bổ sung* (complementation analysis), trong đó gene mà cụ thể là sản phẩm của nó được trắc nghiệm về khả năng bù đắp cho một đột biến tại một gene tương đồng trong cùng tế bào. Sự bổ sung liên tiếp làm phục hồi kiểu hình đại.



**Hình 6.2** Sơ đồ minh họa trắc nghiệm *cis-trans*: (a) con đường chuyển hóa bình thường; (b) trắc nghiệm *cis*; (c) và (d) trắc nghiệm *trans*. Chú thích: *S*-cơ chất (substrate); *I*- sản phẩm trung gian (intermediate); *P*- sản phẩm cuối cùng (product), ở đây là sắc tố đặc trưng cho kiểu hình đại; các mũi tên ( $\downarrow$ ) chỉ các enzyme sản phẩm sinh ra từ các cistron 1 và cistron 2.

Cơ sở của phân tích bổ sung là trắc nghiệm *cis-trans* (*cis-trans test*), mà từ đây nảy sinh ra thuật ngữ *cistron*, trong đó các cặp đột biến bất nguồn độc lập được xét ở các cấu hình *cis* (đều) và *trans* (lệch). Trắc nghiệm *cis* được dùng làm đối chứng, vì nếu như cả hai đột biến đều có mặt trong một bộ gene thì bộ gene kia phải là kiểu đại ở cả hai locus và sinh ra các sản phẩm gene bình thường, do đó cho ra kiểu hình đại (hình 6.2b). Trắc nghiệm *trans* là phép thử bổ sung và xác định giới hạn của đơn vị chức năng. Nếu như các đột biến nằm trong các gene khác nhau, khi chúng có mặt ở cấu hình *trans*, mỗi một bộ gene có thể bổ sung sản phẩm

mà gene kia không tạo ra được. Khi có đủ tất cả các sản phẩm gene cần thiết thì tế bào là kiểu đại (hình 6.2c), nghĩa là có sự *bổ sung dương tính* (positive complementation). Nếu như cả hai đột biến thuộc cùng một gene, khi chúng có mặt ở cấu hình *trans*, thì mỗi một bộ gene có thể mang một bản sao đột biến của gene đó và không có sản phẩm hoạt động chức năng được tạo ra trong tế bào, nghĩa là không có sự bổ sung (hình 6.2d).

Sự phân tích bổ sung ở vi khuẩn và nấm men bia cũng chỉ ra rằng gene là một cistron, nghĩa là gene được định nghĩa như là một đơn vị chức năng. Phương pháp này tỏ ra hữu ích cho việc khẳng định chức năng của gene, xác định số lượng cũng như trật tự hoạt động của các gene trong một con đường chuyển hóa nào đó.

Vậy *cistron* là gì? Cistron là một đoạn xác định của DNA (hay bộ gene nói chung) mang thông tin cấu trúc của một polypeptide cụ thể mà giới hạn của nó được xác định bằng trắc nghiệm *cis-trans*. Kích thước trung bình của một cistron là 1.200 cặp base. Như vậy, cistron chính là gene cấu trúc theo nghĩa hẹp hay gene mã hóa protein.

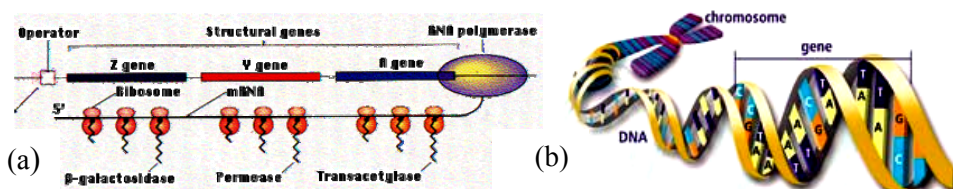
Một cách tương đối, theo nghĩa rộng, có thể định nghĩa *gene là một đoạn xác định của bộ gene mã hóa thông tin của một polypeptid hoặc một phân tử RNA chức năng (như tRNA, rRNA...)*.

Tuy vậy, định nghĩa này không thể bao gồm đầy đủ chức năng và cấu trúc của gene trong toàn bộ sinh giới, bởi vì các chiến lược cho sự biểu hiện gene và tổ chức bộ gene ở các vi khuẩn và eukaryote là rất khác nhau.

#### 4. Mối quan hệ gene - cistron ở các prokaryote và eukaryote

##### 4.1. Sự tương đương gene - cistron ở các bộ gene đơn giản

Ở các prokaryote và eukaryote bậc thấp, thường có một mối quan hệ đơn giản giữa gene và sản phẩm của nó. Trong hầu hết trường hợp, có một sự tương ứng một gene - một sản phẩm, và *sự đồng tuyến tính giữa gene và chuỗi polypeptide* của nó đã được Ch.Yanofsky xác nhận năm 1961. Vì vậy ở các sinh vật này, gene và cistron là tương đương: gene là đơn vị chức năng di truyền, mang thông tin di truyền được biểu hiện trọn vẹn.

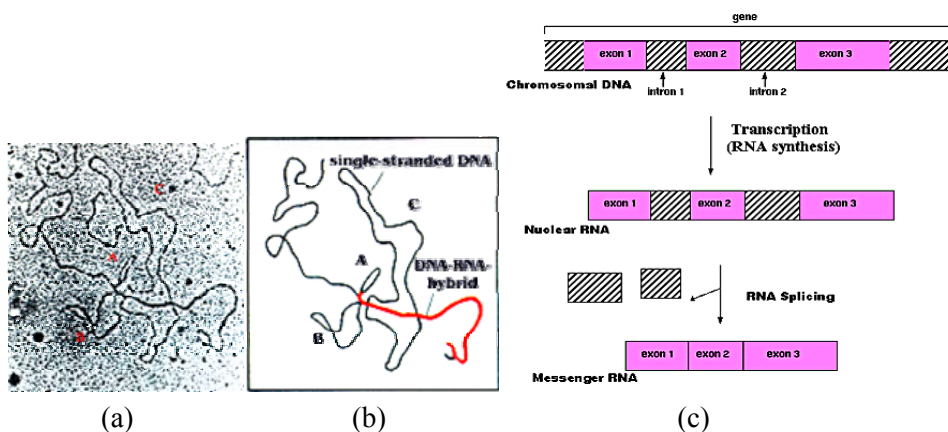


**Hình 6.3** (a) Ở vi khuẩn, các gene thường được sắp xếp trong một operon và được phiên mã thành một phân tử mRNA đa cistron. (b) Ở eukaryote, các gene tồn tại riêng biệt dưới dạng đơn cistron.

Cũng cần lưu ý rằng, ở vi khuẩn, các gene đồng nghĩa với vùng mã hóa hay *khung đọc mở* (open reading frame = ORF), trong khi đó ở các eukaryote nó đồng nghĩa với *đơn vị phiên mã* (transcription unit). Đó là do các gene vi khuẩn thường được sắp xếp trong một *operon* (chương 7), vì thế có nhiều sản phẩm được dịch mã từ một *mRNA đa cistron* (polycistronic mRNA; hình 6.3a). Trái lại, ở các eukaryote, hầu hết các gene được phiên mã dưới dạng *mRNA đơn cistron* (monocistronic mRNA; hình 6.3b).

#### 4.2. Sự không tương đương gene - cistron ở các bộ gene phức tạp

Ở các bộ gene eukaryote bậc cao, thường thường có một mối quan hệ phức tạp giữa gene và sản phẩm của nó (hình 6.4). Hầu hết các gene của eukaryote bậc cao đều có chứa các *intron* (intervening sequences), tức các đoạn không mã hóa protein, nằm xen giữa các *exon* (expressed sequences), các đoạn mã hóa protein. Các gene như vậy được gọi là *gene phân đoạn* (split gene) hay *gene đứt quãng* (interrupted gene); nó được phát hiện lần đầu tiên bởi Phillip Sharp vào năm 1977 (hình 6.4).



**Hình 6.4** Ảnh hiển vi điện tử và hình mô phỏng việc sử dụng vật dò mRNA tế bào chất có đánh dấu để phát hiện gene phân đoạn (a và b); mô tả vắn tắt sự tổng hợp pre-mRNA và cắt bỏ các intron để tạo ra mRNA trưởng thành từ một gene phân đoạn có chứa hai intron (c).

Các khám phá về sau này còn cho thấy những sự kiện rắc rối nảy sinh trong các gene phân đoạn này, ở chỗ: thông tin trong gene được sử dụng một cách chọn lọc để sinh ra nhiều sản phẩm khác nhau, gọi là *cắt-nối có chọn lọc* (alternative splicing) v.v. Các sản phẩm có quan hệ về cấu trúc (ví dụ, calcitonin/CGRP) thường có các chức năng khác nhau. Vì lẽ đó, cistron đôi khi được xem là tương đương với exon của gene eukaryote, và gene phân đoạn được xem như là một chuỗi các cistron gộp nhau (Twyman 1998). Ngoài ra, có vài trường hợp trong đó cần tới hai gene để sinh ra một sản phẩm mRNA đơn thông qua kiểu *cắt-nối chéo* (trans-splicing)

hoặc *biên tập RNA* (RNA editing). Chẳng hạn, khám phá mới nhất cho thấy *glucose 6-phosphate dehydrogenase* là một enzyme có mặt trong các tế bào hồng cầu người; nó bao gồm hai dạng nhỏ/thứ yếu và lớn/chính yếu (minor and major form); dạng đầu có trình tự các amino acid thuộc gene trên nhiễm sắc thể X; và dạng sau gồm hai peptide được mã hóa từ thông tin của hai nhiễm sắc thể, các amino acid trong đoạn 1-53 được mã hóa trên nhiễm sắc thể số 6 và các amino acid ở đoạn tiếp theo 54-479 được mã hóa trên nhiễm sắc thể X (theo McClean 1998). Tất cả những trường hợp này nói lên một điều rằng, để đưa ra một định nghĩa chính xác về gene không hề đơn giản tí nào. Tuy nhiên, nhìn toàn cục thì một khái niệm thống nhất nổi bật là, tất cả các sinh vật từ vi khuẩn *E. coli* cho đến con người đều có chung hệ thống mật mã di truyền và có chung phương thức 'chuyển tải' thông tin trong gene thành ra protein.

#### 4.3. Các thành phần cấu trúc hay là tổ chức của một gene

Tổ chức của một gene có thể bao gồm các vùng riêng biệt với các chức năng đặc thù (Bảng 6.1; theo Twyman 1998, có sửa đổi).

**Bảng 6.1 Các thuật ngữ được dùng để chỉ các phần chức năng của các gene**

<b>Thuật ngữ</b>	<b>Định nghĩa</b>
Allele	Một biến thể về trình tự của một gene (hoặc marker di truyền khác, ví dụ: trình tự RFLP, VNTR).
Cistron	Một đơn vị chức năng di truyền, một vùng của DNA mã hóa một sản phẩm đặc thù.
Vùng mã hóa, khung đọc mở (ORF)	Một vùng của DNA được dịch mã thành protein. Ở vi khuẩn, đó là một gene. Ở eukaryote, vùng mã hóa có thể bị gián đoạn bởi các intron.
Gene phân đoạn	Một gene mã hóa protein gồm các đoạn không mã hóa (intron) nằm xen kẽ giữa các đoạn mã hóa (exon).
Gene	Ở vi khuẩn, một đơn vị chức năng di truyền mã hóa hoặc là 1 polypeptide riêng hoặc phân tử RNA. Ở eukaryote, 1 đơn vị phiên mã có thể mã hóa 1 hay nhiều sản phẩm hoặc đóng góp vào 1 sản phẩm.
Locus gene	Vị trí của một gene trên một nhiễm sắc thể, kể cả các yếu tố điều hòa kề bên. Thuật ngữ locus được dùng theo cách riêng để chỉ vị trí của gene-marker bất kỳ, yếu tố điều hòa, khởi điểm tái bản, v.v..
Operon	Một locus của vi khuẩn có chứa nhiều gene (mà được phiên mã như là một bản sao polycistron đơn) và các yếu tố điều hòa chung của chúng.
Pseudogene	Một trình tự không hoạt động chức năng vốn có cấu trúc tương tự một gene hoạt động chức năng.
Đoạn đệm được	Bất kỳ phần nào của đơn vị phiên mã của 1 gene RNA

phiên mã	hay operon gene RNA sẽ bị loại bỏ trong khi tạo ra các phân tử RNA trưởng thành.
Đơn vị phiên mã, vùng được phiên mã	Một vùng của DNA được phiên mã thành RNA. Ở các eukaryote, đó là một gene. Ở vi khuẩn, nó có thể bao quát nhiều gene.
Vùng không được dịch mã (UTR)	Bất kỳ phần nào của đơn vị phiên mã mà không được dịch thành protein. Các UTR kề bên một vùng mã hóa hay operon được gọi là các UTR 5' và 3'.
Gene bị phân chia (divided gene)	Một gene phân đoạn với các exon ở các locus riêng biệt được phiên mã riêng rẽ và khâu nối lại bởi sự cắt-nối chéo. Thực ra đó là cách gọi sai vì mỗi locus đúng ra phải được coi như là 1 gene riêng.

Ở bất kỳ locus nào, một vùng DNA được phiên mã có thể gọi là một *đơn vị phiên mã* (transcription unit). Như đã đề cập, ở prokaryote, một đơn vị phiên mã có thể gồm nhiều gene; trong khi đó ở eukaryote, các đơn vị phiên mã hầu như bao giờ cũng tương đương với một gene đơn.



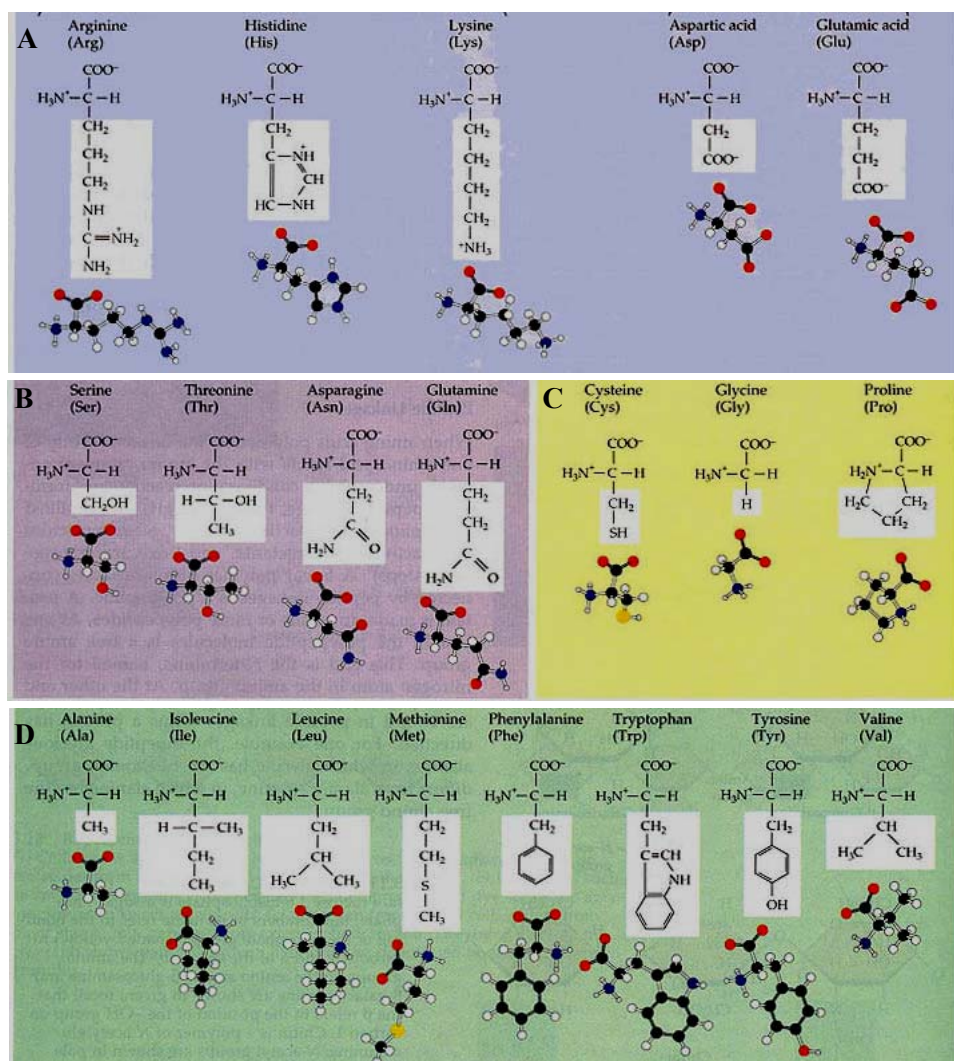
**Hình 6.5** Cấu trúc điển hình của một gene eukaryote.

Đối với các gene mã hóa protein, rõ ràng là có một sự tách biệt giữa vùng được dịch mã thành chuỗi polypeptide và vùng không được dịch mã. Ở vi khuẩn, vùng được dịch mã là *khung đọc mở* (ORF), trong đó các gene phân cách nhau bằng các *đoạn đệm* (spacer) được gọi là các *vùng không mã hóa bên trong* (internal noncoding region). Các gene nằm ở hai đầu của một operon cũng được kèm bởi một vùng không mã hóa có tên là *vùng không dịch mã 5'* (5' untranslated region = 5' UTR) hay *đoạn dẫn đầu* (leader sequence) và *vùng không được dịch mã 3'* (3' UTR) hay *đoạn kéo sau* (trailer sequence). Về bản chất, chúng là các đoạn điều hòa; chẳng hạn, vùng 5'UTR kiểm soát sự bám vào của ribosome, còn vùng 3'UTR thường đóng vai trò quan trọng trong sự ổn định mRNA. Ở các gene eukaryote, vùng mã hóa cũng được kèm bởi các vùng UTR điều hòa, và cả hai vùng UTR 5' và 3' cũng như khung đọc mở có thể bị gián đoạn bởi các đoạn không mã hóa (tức các intron) mà chúng sẽ được cắt bỏ trước khi xuất mRNA trưởng thành ra khỏi nhân. Như thế, bất kỳ đoạn nào mà rốt cuộc bị loại bỏ khỏi pre-RNA thì được gọi là các *đoạn đệm được phiên mã* (transcribed spacer). Cấu trúc điển hình của các gene mã hóa protein ở prokaryote và eukaryote được chỉ ra tương ứng ở hình 6.3a và hình 6.5.

## II. Cấu trúc và chức năng của protein

### 1. Cấu trúc protein

Các protein là những polymer sinh học được cấu tạo bằng các *amino acid* nối kết với nhau bằng các *liên kết peptide*. Có 20 loại L- $\alpha$ -amino acid được phát hiện trong các protein của các tế bào (hình 6.6). Về cấu trúc, nói chung, mỗi amino acid gồm có một nguyên tử *carbon alpha* ( $C\alpha$ ) ở vị trí



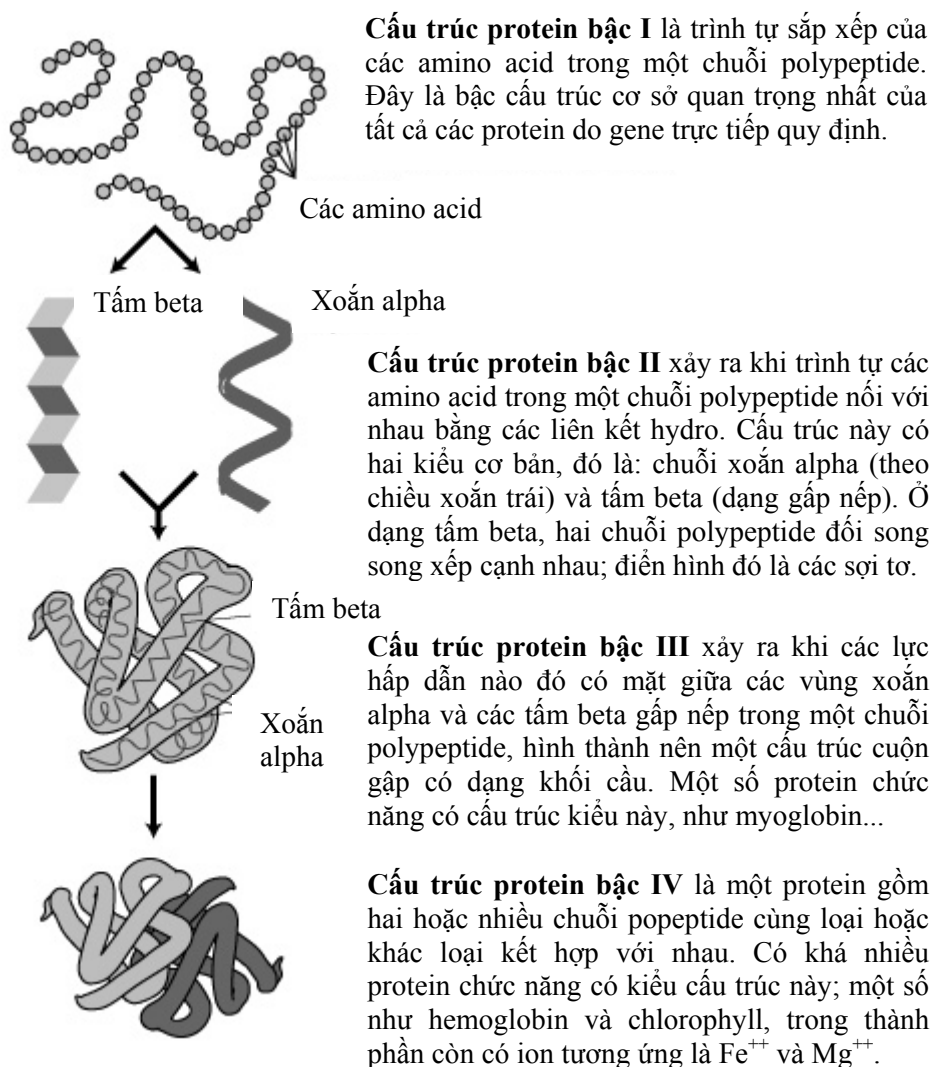
**Hình 6.6** Hai mươi loại amino acid phát hiện được trong các protein, với bốn nhóm: A. Các amino acid có chuỗi bên tích điện dương (3 bên trái) và âm (2 bên phải); B. Các amino acid có chuỗi bên không tích điện; C. Các trường hợp đặc biệt; và D. Các amino acid có chuỗi bên kỵ nước.





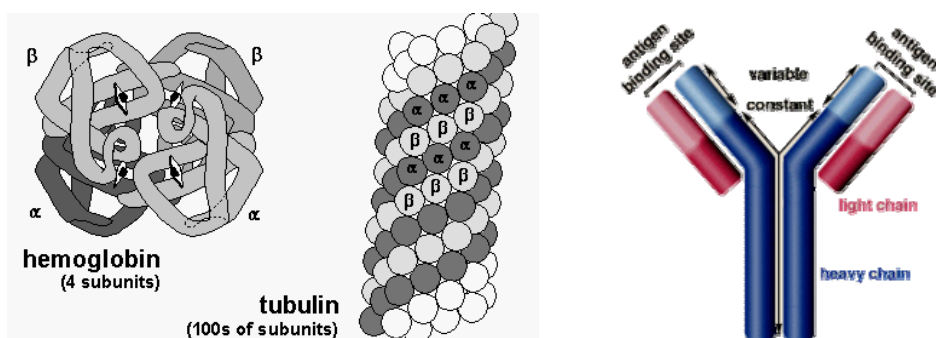
các màng tế bào, các bào quan, bộ máy di truyền của chúng. Đó cũng là các protein dạng sợi làm thành các cơ quan bộ phận trên cơ thể các động vật, như: collagen làm nên xương, sụn, gân và da; keratin cấu tạo nên các lớp ngoài cùng của da và tóc, móng, sừng và lông;

(ii) Các *enzyme* đóng vai trò xúc tác cho tất cả các phản ứng hóa học trong tế bào và cơ thể đều là những protein hình cầu. Quan trọng nhất là các enzyme tham gia vào các con đường chuyển hóa và các enzyme tham gia vào các quá trình truyền thông tin di truyền trong tế bào.



**Hình 6.8** Bốn bậc cấu trúc của protein.





**Hình 6.9 Cấu trúc bậc IV điển hình của hemoglobin, tubulin và immunoglobulin.** Ở đây cho thấy số chuỗi polypeptide và đặc biệt là các vùng chức năng đặc trưng của hemoglobin và kháng thể immunoglobulin kiểu IgG.

(iii) Các *hormone* protein bắt nguồn từ các tuyến nội tiết thì không hoạt động như các enzyme. Thay cho sự kích thích các cơ quan đích, chúng lại khởi đầu và kiểm soát các hoạt động quan trọng, ví dụ như tốc độ chuyển hóa và sản xuất ra các enzyme tiêu hóa và sữa. Insulin được tiết ra từ các đảo Langerhans tuyến tụy, điều hòa sự chuyển hóa carbohydrate bằng cách kiểm soát các mức glucose trong máu. Thyroglobulin (từ tuyến giáp) điều hòa các quá trình chuyển hóa nói chung; calcitonin cũng từ thyroid làm hạ thấp mức calcium trong máu v.v.

(iv) Các *kháng thể* (antibodies) trong hệ thống miễn dịch, còn gọi là các immunoglobulin, làm ra hàng ngàn protein khác nhau vốn được sinh ra trong huyết thanh máu phản ứng lại với các *kháng nguyên* (antigens). Chúng đóng vai trò bảo vệ cơ thể chống lại sự xâm nhập của các vật lạ.

(v) Ngoài ra, các protein còn là *nguồn dinh dưỡng* chính cung cấp năng lượng cho tế bào và cơ thể duy trì các hoạt động trao đổi chất và lớn lên; các protein như *hemoglobin* mang các sinh chất theo máu đi khắp cơ thể; các *fibrinogen* và fibrin được biến đổi từ nó vốn có trong máu cần thiết cho quá trình đông máu. Bên cạnh đó, các *protein cơ* mà chủ yếu là myosin phối hợp với actin tạo thành actomyosin, chịu trách nhiệm cho hoạt động cơ cơ v.v.

### III. Mã di truyền

Gene (DNA) được cấu tạo từ bốn loại nucleotide, trong khi đó protein được cấu tạo bởi 20 loại amino acid. Vậy vấn đề đặt ra là, các gene mã hóa cho các sản phẩm protein của chúng bằng cách nào?

Bằng suy luận, ta có thể suy đoán rằng mỗi amino acid không thể được xác định bởi đơn vị mã gồm một, hai hoặc bốn nucleotide (vì  $4^1 = 4$  hoặc  $4^2 = 16$  thì chưa đủ để mã hóa cho 20 amino acid, trong khi  $4^4 = 256$  thì lại

du thừa quá nhiều) mà có lẽ phải là một nhóm gồm ba nucleotide ( $4^3 = 64$ ). Với 64 kiểu tổ hợp bộ ba hoá ra là đủ thừa để mã hoá cho 20 loại amino acid. Nếu như thế, một amino acid được xác định bởi trung bình ba bộ ba khác nhau. Vậy phải chăng mã di truyền là mã bộ ba?

### 1. Bằng chứng di truyền học về mã bộ ba

Năm 1961, S.Brenner, F.Crick và L.Barnett đã phân tích chi tiết nhiều thể đột biến của phage T4 nhận được bằng cách xử lý acridin, tác nhân gây các đột biến mất hoặc thêm một cặp base (chương 8), đã khẳng định mã di truyền là mã bộ ba (triplet code) đúng như dự đoán. Đơn vị mã (coding unit) gồm ba nucleotide xác định một amino acid như vậy được gọi là *codon*.

### 2. Giải mã di truyền

Việc tiếp theo là xác định xem mỗi amino acid cụ thể được mã hoá bởi một hoặc một số bộ ba nào. Cũng trong năm 1961, M.Nirenberg và H. Matthaei lần đầu tiên sử dụng mRNA nhân tạo có thành phần base biết trước được tổng hợp bằng enzyme *polynucleotide phosphorylase* (do Ochoa tìm ra năm 1959) và hệ thống tổng hợp là dịch chiết tế bào *E. coli* bao gồm đầy đủ các yếu tố (các ribosome, tRNA, amino acid, enzyme, ATP...) cần thiết cho tiến hành giải mã di truyền *in vitro*. Với mRNA chỉ chứa toàn U, poly(U), chuỗi polypeptide sinh ra chỉ chứa toàn phenylalanine (Phe). Điều đó chứng tỏ UUU là bộ ba mã hoá của Phe.

Sau đó, Har Gobind Khorana đã tiến hành các thí nghiệm sử dụng các mRNA tổng hợp có chứa hai, ba hoặc bốn nucleotide được kết nối theo kiểu lặp lại như sau:

(1) Với mRNA nhân tạo chứa hai base là poly(UC) hay UCUCUC..., sẽ chứa hai codon xen kẽ UCU và CUC (chú ý rằng sự dịch mã *in vitro* khởi đầu tại vị trí ngẫu nhiên). Kết quả là thu được một polypeptide gồm hai amino acid xếp xen kẽ nhau là serine và leucine, poly(Ser-Leu).

(2) Với mRNA tổng hợp gồm các bộ ba lặp lại sẽ được dịch thành các homopolypeptide. Ví dụ, poly(UUC) có thể được đọc là (UUC-UUC), hoặc (UCU-UCU), hoặc (CUU-CUU) tùy thuộc vào vị trí bắt đầu dịch mã. Và kết quả là có ba loại polypeptide được tổng hợp, poly(Phe) hoặc poly(Ser) hoặc poly(Leu).

(3) Với mRNA gồm bốn nucleotide lặp lại, ví dụ poly(UAUC), thì nó được dịch thành polypeptide chứa bốn amino acid lặp lại là poly(Tyr-Leu-Ser-Ile). Tuy nhiên, khi sử dụng các poly(GAUA) và poly(GUAA) lại không cho kết quả. Qua so sánh với hàng loạt kết quả nhận được cho phép xác định được các codon "vô nghĩa" hay còn gọi là các tín hiệu kết thúc. Từ các kết quả thu được như vậy Khorana đã xác định được 'nghĩa' của

phần lớn codon có thành phần base không đồng nhất, và việc giải toàn bộ hệ thống mã di truyền (genetic code) được hoàn thành vào tháng 6/1966. Với công lao to lớn đó Khorana và Nirenberg được trao giải thưởng Nobel năm 1968. Từ đây cho phép xây dựng nên bảng mã di truyền (Bảng 6.2) và rút ra được các đặc tính của mã như được trình bày dưới đây.

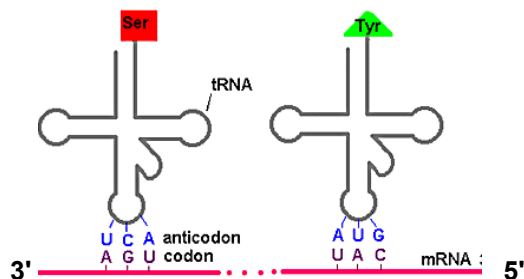
**Bảng 6.2 Mã di truyền** (cho các codon trên mRNA theo chiều 5'→3')

	U	C	A	G		
U	UUU UUC	UCU UCC UCA UCG	UAU UAC	UGU UGC	U C A G	
	UUA UUG		UAA UAG	UGA UGG		
	UUU UUC		UCU UCC UCA UCG	UAU UAC		UGU UGC
C	CUU CUC CUA CUG	CCU CCC CCA CCG	CAU CAC	CGU CGC CGA CGG	U C A G	
	CUU CUC CUA CUG		CCU CCC CCA CCG	CAU CAC		CGU CGC CGA CGG
	CUU CUC CUA CUG		CCU CCC CCA CCG	CAU CAC		CGU CGC CGA CGG
A	AUU AUC AUA	ACU ACC ACA ACG	AAU AAC	AGU AGC	U C A G	
	AUU AUC AUA		ACU ACC ACA ACG	AAU AAC		AGU AGC
	AUU AUC AUA		ACU ACC ACA ACG	AAU AAC		AGU AGC
G	GUU GUC GUA GUG	GCU GCC GCA GCG	GAU GAC	GGU GGC GGA GGG	U C A G	
	GUU GUC GUA GUG		GCU GCC GCA GCG	GAU GAC		GGU GGC GGA GGG
	GUU GUC GUA GUG		GCU GCC GCA GCG	GAU GAC		GGU GGC GGA GGG

### 3. Các đặc tính của mã di truyền

- Mã di truyền là mã bộ ba (triplet code), nghĩa là một bộ ba nucleotide kế tiếp mã hoá cho một amino acid. Các bộ ba trên gene và mRNA gọi là *codon* (mã) và bộ ba đặc trưng của tRNA có thể khớp với codon của mRNA theo nguyên tắc bổ sung gọi là *anticodon* (đôi mã).

- Mã di truyền không gối lên nhau (non-overlapping): Mỗi codon là một đơn vị độc lập, và thông tin của mRNA được đọc theo một chiều 5'→3' (hình 6.10) bắt đầu từ codon khởi đầu.



**Hình 6.10** Các phân tử tRNA mang amino acid Ser (trái) và Tyr (phải) đọc mã trên mRNA bằng cách khớp anticodon của chúng với codon của mRNA.

- Mã di truyền có tính liên tục, *không bị ngắt quãng* (unpunctuated): Nghĩa là không có khoảng hở giữa các codon.

- Mã di truyền có các codon *khởi đầu* (initiation) và *kết thúc* (termination) nằm ở gần hai đầu 5' và 3' của mRNA đóng vai trò là tín hiệu khởi đầu và kết thúc tổng hợp chuỗi polypeptide. Codon khởi đầu AUG quy định amino acid mở đầu chuỗi polypeptide là methionine (ở vi khuẩn là N-formyl-methionine). Các codon kết thúc (UAA, UAG hoặc UGA) không xác định amino acid nào nên còn gọi là các *codon vô nghĩa* (nonsense codon).

- Mã di truyền có tính *đơn trị*, rõ ràng (unambiguous): Mỗi codon xác định một amino acid duy nhất, hoặc xác định sự kết thúc dịch mã.

- Mã di truyền có tính *thoái hóa* (degenerate): Có 61 *codon có nghĩa* (sense codon) trong khi chỉ có 20 loại amino acid; vì vậy mỗi amino acid (hay tín hiệu kết thúc) có thể được xác định bởi nhiều hơn một codon. Các codon cùng xác định một amino acid như thế gọi là các *codon đồng nghĩa*; chúng thường khác nhau ở base cuối, base 3' đó được gọi là base thoái hóa. Ví dụ, các amino acid Arg, Ser và Leu mỗi cái có tới sáu codon đồng nghĩa; Ala, Gly, Pro, Thr và Val mỗi cái có bốn codon đồng nghĩa (xem Bảng 6.2).

- Mã di truyền có tính *phổ biến* (universal), nghĩa là thống nhất cho toàn bộ sinh giới.

#### 4. Những ngoại lệ so với mã di truyền "phổ biến"

Bên cạnh tính *phổ biến* (universal) của hệ thống mã di truyền nói trên, các nghiên cứu gần đây cho thấy một vài chệch hướng mà hầu hết là xảy ra trong các bộ gene ty thể (Bảng 6.3). Tuy nhiên, ở các bào quan thực vật,

**Bảng 6.3 Các ngoại lệ so với mã "phổ biến"**

<b>Nguồn</b>	<b>Codon</b>	<b>Nghĩa phổ biến</b>	<b>Nghĩa mới</b>
Ty thể ruồi giấm	UGA	Kết thúc	Tryptophan
	AGA & AGG	Arginine	Serine
	AUA	Isoleucine	Methionine
Ty thể động vật có vú	AGA & AGG	Arginine	Kết thúc
	AUA	Isoleucine	Methionine
	UGA	Kết thúc	Tryptophan
Ty thể nấm men *(N = U, C, A hoặc G)	CUN*	Leucine	Threonine
	AUA	Isoleucine	Methionine
	UGA	Kết thúc	Tryptophan
Ty thể thực vật bậc cao	UGA	Kết thúc	Tryptophan
	CGG	Arginine	Tryptophan
Các nhân Protozoa <i>Mycoplasma</i>	UAA & UAG	Kết thúc	Glutamine
	UGA	Kết thúc	Tryptophan

sự *biên tập RNA* (RNA editing) là phổ biến và không rõ ràng, ở chỗ: Liệu phải chăng tất cả các trường hợp sai lệch so với mã phổ biến ở thực vật là các biến đổi thật hay là hậu quả của sự biên tập RNA trước khi dịch mã. Một vài thay đổi thỉnh thoảng cũng xảy ra ở các bộ gene vi khuẩn và bộ gene nhân của các eukaryote, nhưng thường thì liên quan với các codon kết thúc. Sự phân bố phát sinh chủng loại của các thay đổi này chỉ ra rằng mã di truyền vẫn còn đang tiến hóa.

#### 5. Sự linh hoạt trong việc kết cặp anticodon-codon

Mặc dù có 61 codon có nghĩa nhưng trên thực tế trong mỗi tế bào prokaryote và eukaryote chỉ có khoảng 45 phân tử tRNA khác nhau. Tuy nhiên, trong tế bào chỉ có 20 loại amino acid được xác định bởi mã di truyền, nên một số loại tRNA phải mang cùng một loại amino acid. Các bản sao tRNA như thế có thể có trình tự anticodon giống nhau, trong trường hợp đó chúng có thể thay thế chức năng cho nhau. Các tRNA khác thì mang các trình tự anticodon khác nhau và do vậy nhận biết các codon khác nhau; chúng được gọi là *isoaccepting tRNA*, và độ giàu tương đối của chúng có thể ảnh hưởng tới cách thức sử dụng codon.

**Bảng 6.4 Các nguyên tắc kết cặp linh hoạt anticodon-codon**

Base 5' của anticodon	Base 3' của codon
G	C hoặc U
C	G
A	U
U	A hoặc G
I	U, C hoặc A

Năm 1966, F.Crick đưa ra một cách để giải thích về sự kết cặp "lỏng lẻo" có thể xảy ra ở vị trí thứ ba của các codon đồng nghĩa. Theo Crick, hai base đầu tiên của một codon phải có sự kết cặp chính xác với anticodon (theo nguyên tắc bổ sung), còn base cuối của codon thì có thể "*linh hoạt*" (wobble), ít đặc thù hơn so với vị trí bình thường của nó để hình thành nên sự kết cặp base bất thường với anticodon. Đề nghị này được gọi là *giả thuyết linh hoạt* (wobble hypothesis). Đặc biệt, Crick đề nghị rằng một base G trong anticodon có thể cặp không chỉ với C ở vị trí thứ ba của một codon (vị trí linh hoạt), mà còn với U. Hơn nữa, Crick còn lưu ý rằng một trong số các nucleoside bất thường có mặt trong tRNA là *inosine* (I), vốn có cấu trúc tương tự với guanosine. Nucleoside này thông thường có thể kết cặp như G, vì vậy ta kỳ vọng nó sẽ cặp với C (kết cặp base bổ sung) hoặc U (kết cặp base linh hoạt) ở vị trí thứ ba của một codon. Nhưng ông còn lưu ý rằng inosine vẫn còn có thể có kiểu kết cặp linh hoạt khác, bây giờ ta biết đó là cặp với A ở vị trí thứ ba của codon. Điều đó có nghĩa là,

một anticodon có I ở vị trí thứ nhất về tiềm năng có thể cặp với ba codon khác nhau có base cuối là C, U hoặc A. Các thí nghiệm sau này khẳng định điều dự đoán của Crick là hoàn toàn đúng và liệt kê các khả năng kết cặp base linh hoạt như ở Bảng 6.4.

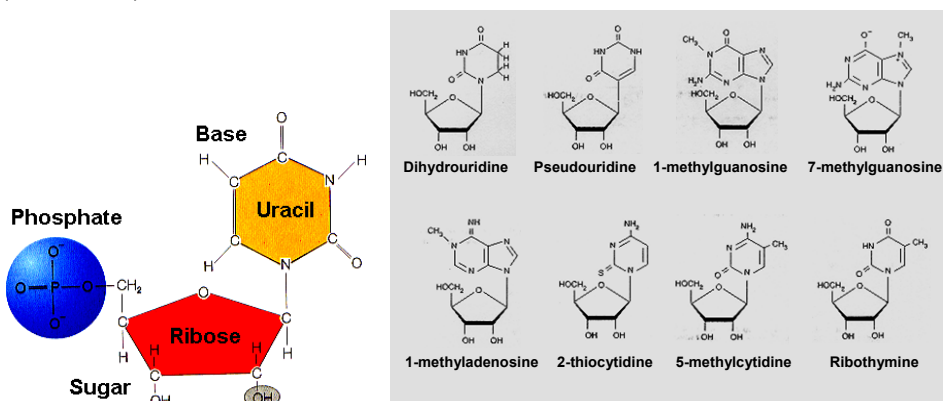
Rõ ràng là, hiện tượng kết cặp linh hoạt này làm giảm đáng kể số lượng các tRNA cần thiết để dịch mã di truyền. Ví dụ, để dịch mã các codon (5'→3') UUU và UUC mã hoá cho phenylalanine chỉ cần một tRNA<sup>Phe</sup> mang anticodon (3'→5') AAG.

#### IV. Cơ chế phiên mã và sửa đổi sau phiên mã

##### 1. Các RNA và đặc điểm chung của phiên mã

###### 1.1. Sơ lược về cấu trúc các RNA

Có bốn loại ribonucleotide nối kết với nhau bằng các liên kết phosphodiester tạo thành các chuỗi polynucleotide của RNA (về nguyên tắc chung, đã được đề cập ở chương 5). Trong thành phần base của các RNA, ngoài bốn loại cơ bản adenine (A), uracil (U), guanine (G) và cytosine (C), còn phát hiện một số kiểu base được sửa đổi phổ biến là trong các tRNA (Hình 6.11).



**Hình 6.11** Cấu trúc một ribonucleotide Uracil đặc trưng của RNA (trái) và một số base sửa đổi có thể có trong thành phần của các RNA.

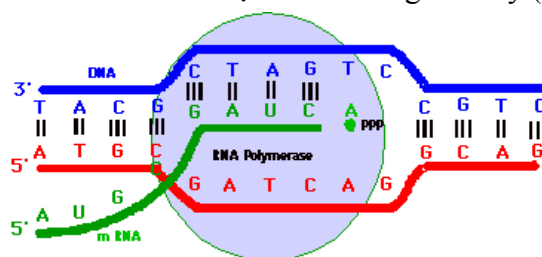
Có ba loại RNA cơ bản tham gia vào quá trình sinh tổng hợp protein của tế bào ở các sinh vật, đó là: *RNA thông tin* (messenger RNA = mRNA), *RNA vận chuyển* (transfer RNA = tRNA) và *RNA ribosome* (ribosomal RNA = rRNA). Riêng các rRNA, ở prokaryote có ba loại với các *hệ số lắng* (sedimentation, được đo bằng đơn vị Svedberg với ký hiệu: S) là 5S, 16S và 23S; trong khi đó ở các tế bào eukaryote có bốn loại: 5S, 5,8S, 18S và 28S. Ngoài ra, trong các tế bào eukaryote còn có các RNA nhân kích thước lớn và sai khác nhau gọi là hnRNA (*heterogenous nuclear RNA*) vốn là tiền

thân của các mRNA, các RNA nhân kích thước bé snRNA (*small nuclear RNA*) có mặt trong thành phần của các enzyme splicing (xem ở phần sau), và các RNA tế bào chất kích thước bé scRNA (*small cytoplasmic RNA*). Cấu trúc và chức năng của các RNA này sẽ được thảo luận ở mục V.

## 1.2. Đặc điểm chung của phiên mã ở prokaryote và eukaryote

*Phiên mã* (transcription) là quá trình tổng hợp các RNA khác nhau từ thông tin di truyền chứa đựng trong DNA. Trừ các gene mã hóa protein trong các operon ở vi khuẩn, nói chung, các RNA mới được tổng hợp chỉ là các *bản sao sơ cấp* (primary transcript) gọi là các pre-RNA. Các pre-RNA này phải trải qua một quá trình sửa đổi để trở thành các RNA trưởng thành (mature) trước khi tham gia vào quá trình sinh tổng hợp protein của tế bào.

Quá trình phiên mã DNA các đặc điểm chung sau đây (Hình 6.12).



**Hình 6.12** Sự tổng hợp RNA trên một sợi khuôn của gene (DNA) dưới tác dụng của RNA polymerase.

(i) Diễn ra dưới tác dụng của các enzyme *RNA polymerase*.

(ii) Vùng DNA chứa gene được mở xoắn cục bộ, và chỉ một sợi đơn gọi là *sợi có nghĩa* (sense strand) được dùng làm *khuôn* (template) cho tổng hợp RNA.

(iii) Phản ứng tổng hợp RNA diễn ra theo *nguyên tắc bổ sung* và được kéo dài theo chiều 5'→3', ngược với chiều của sợi khuôn.

(iv) Nguyên liệu cho tổng hợp là bốn loại *ribonucleoside triphosphate*: ATP, UTP, GTP và CTP.

(v) Sản phẩm của phiên mã là các *RNA sợi đơn* (single RNAs).

(vi) Sự khởi đầu và kết thúc phiên mã phụ thuộc vào các tín hiệu điều hoà là các trình tự DNA đặc thù nằm trước và sau gene được phiên mã.

(vii) Quá trình phiên mã có thể chia làm ba bước, vẫn tắt như sau: *Mở đầu* (initiation) là sự tương tác giữa RNA polymerase với vùng promoter nhằm xác định sợi khuôn của gene và tổng hợp vài nucleotide; *Kéo dài* (elongation) là giai đoạn sinh trưởng tiếp tục của chuỗi RNA dọc theo sợi khuôn cho đến cuối gene; và *Kết thúc* phiên mã (termination) đặc trưng bằng sự giải phóng sợi RNA và RNA polymerase ra khỏi khuôn DNA.





Các *vùng khởi động* (promoter) nói chung nằm kề trước gene và có chứa các đoạn trình tự đặc thù cho phép RNA polymerase nhận biết và bám chặt vào để khởi đầu phiên mã tại vị trí chính xác trên sợi khuôn của gene. Vấn đề này tương đối phức tạp ở các eukaryote, vì vậy ở đây ta chỉ xét các promoter của các gene mã hóa protein mà không đề cập các loại promoter của các gene mã hóa các RNA khác. Các promoter của các gene mã hóa protein eukaryote và của operon vi khuẩn nói chung có cấu trúc khá tương đồng nhau. Đoạn trình tự quan trọng nhất của promoter được gọi là *hộp TATA* (TATA box) hay *hộp Pribnow* (Pribnow box). Đối với vi khuẩn, đó là trình tự TATAAT (hoặc tương tự như thế) nằm ở vị trí "-10" (Hình 6.13); còn đối với các gene mã hóa protein của eukaryote, đó là trình tự TATAAA nằm gần vị trí "-30" và nó đặc trưng riêng cho các RNA polymerase II. (Cần lưu ý là, tất cả các đoạn tín hiệu đều được quy ước trên sợi đối khuôn của gene, vì chúng có trình tự giống như RNA được tổng hợp, chỉ khác là U thay cho T; các ký hiệu '-' và '+' để chỉ các vị trí nằm trước và sau vị trí bắt đầu phiên mã, hay còn gọi là các yếu tố *upstream* và *downstream*).

Ngoài ra, trong các promoter vi khuẩn còn có trình tự TTGACA ở gần vị trí "-35", gọi là *đoạn nhận biết* (recognition sequence). Đối với các gene mã hóa protein eukaryote, nằm phía trước điểm bắt đầu phiên mã chừng 75 nucleotide có trình tự GGCCAAATCT, thường được gọi là *hộp CCAAT* (CCAAT box) - đọc là "hộp cat"; nó đóng vai trò điều hòa tốc độ phiên mã.

Nói chung, các vùng này được bảo tồn cao và đặc thù cho từng loại RNA polymerase nhất định, được gọi là các *trình tự điều hòa* (consensus sequences). Các đột biến thay thế base tại các hộp TATA, nghĩa là làm cho nó bớt giống với trình tự được bảo tồn (ví dụ, TATAAAT → TGTAAAT) do đó sẽ làm yếu khả năng phiên mã của promoter, gọi là *down mutation*. Ngược lại, các đột biến làm cho các trình tự promoter trở nên giống với các trình tự điều hòa (ví dụ, TATCTT → TATAAAT), sẽ làm mạnh khả năng phiên mã của promoter, gọi là *up mutation*.

#### 4. Các giai đoạn của quá trình phiên mã

Ở đây chỉ đề cập một mô hình đơn giản về quá trình phiên mã ở *E. coli*.

- Giai đoạn bám và khởi đầu: Sau khi RNA polymerase holoenzyme nhận biết và bám chặt vào promoter, làm tháo xoắn một đoạn chừng 12 cặp base tại đây. Sau khi tổng hợp được một vài nucleotide, nhân tố sigma tách ra để đi vào một chu kỳ phiên mã khác, gọi là *chu kỳ sigma* (sigma cycle).

- Giai đoạn kéo dài: Enzyme lõi tiến hành kéo dài sợi RNA dọc theo sợi khuôn. RNA polymerase lõi tiến đến đâu thì DNA được mở xoắn và phiên mã đến đấy; và vùng DNA đã được phiên mã đóng xoắn trở lại.

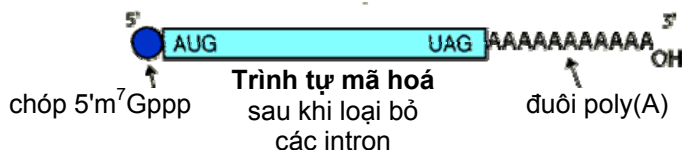
- Giai đoạn kết thúc: Khi quá trình phiên mã tổng hợp xong hai đoạn kết thúc giàu GC và AT nằm đằng sau gene thì tại vùng đuôi sợi RNA hình thành cấu trúc "nút cài tóc" làm dừng sự phiên mã của lõi RNA polymerase. Sau đó, dưới tác dụng của nhân tố *rho* ( $\rho$ ) có bản chất protein, sợi RNA vừa được tổng hợp và enzyme lõi được giải phóng ra khỏi DNA khuôn.

### 5. Sự sửa đổi sau phiên mã đối với các mRNA eukaryote

Như đã đề cập, trừ mRNA prokaryote ra, tất cả các RNA còn lại dù ở pro- hay eukaryote đều phải trải qua quá trình sửa đổi sau phiên mã với rất nhiều cơ chế tinh vi và phức tạp khác nhau để tạo ra các RNA trưởng thành tham gia vào quá trình dịch mã; ở eukaryote, các quá trình này xảy ra trong nhân. Để có cái nhìn hệ thống, ở đây ta hãy tìm hiểu một ít về các cơ chế hoàn thiện các bản sao sơ cấp mRNA (pre-mRNA) ở các tế bào eukaryote.

#### 5.1. Gắn thêm "mũ" m<sup>7</sup>Gppp và "đuôi" poly(A)

Để trở thành phân tử mRNA trưởng thành trước khi đi ra tế bào chất làm khuôn cho dịch mã, tất cả các pre-mRNA của các gene mã hóa protein khác nhau ở tế bào eukaryote, đều được lắp thêm cái "chóp" 7-*methylguanosinetriphosphate* (m<sup>7</sup>Gppp cap) vào đầu 5' và "đuôi" poly(A) vào đầu 3' (Hình 6.14). Đối với các gene mã hóa protein có vùng mã hóa là liên tục (không bị gián đoạn bởi các exon) như các gene histone chẳng hạn, quá trình hoàn thiện mRNA dừng lại ở đây. Loại này chiếm khoảng 10%.



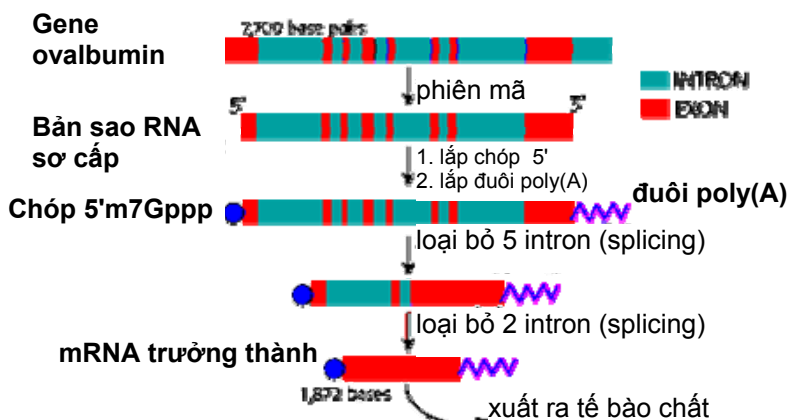
**Hình 6.14** Cấu trúc điển hình của một mRNA trưởng thành ở eukaryote.

Cần lưu ý rằng, ở đầu 3' của hầu hết các gene mã hóa protein eukaryote có chứa trình tự AATAAA đóng vai trò là tín hiệu cho việc gắn "đuôi" poly(A) vào đầu 3' của mRNA. Sự phiên mã thông thường vẫn còn tiếp diễn sau khi đi qua vị trí polyadenyl hoá này. Trình tự tương ứng ở vùng cuối 3' mRNA được phiên mã là AAUAAA báo hiệu cho *endonuclease* nhận biết và cắt chuỗi RNA tại một điểm xác định nằm sau nó khoảng 10-30 base. Sau đó, một enzyme khác là *poly(A)-polymerase* sẽ lắp thêm vào đầu 3' của mRNA một dãy adenine dài khoảng 150-200 base gọi là *đuôi poly -A* (poly [A] tail; Hình 6.14). Đuôi poly(A) có chức năng bảo vệ mRNA khỏi bị suy thoái và trong nhiều trường hợp nó còn kích thích sự dịch mã.

#### 5.2. Sự cắt-nối đối với pre-mRNA của các gene phân đoạn (*split gene*)

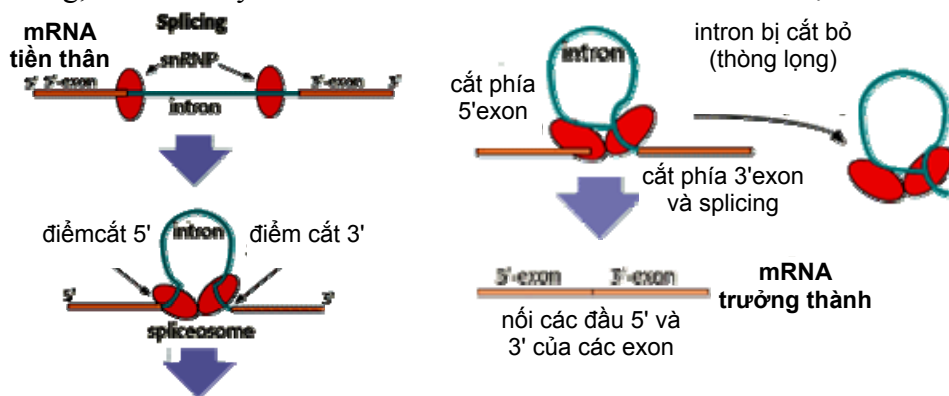
Ví dụ, gene ovalbumin gồm bảy intron xen kẽ giữa tám exon có độ dài 7.700 cặp base đã được E.Chambon phân tích trình tự năm 1981. Sau khi

enzyme splicing cắt bỏ các intron và nối tất cả các exon trong một quá trình gọi là *xử lý RNA* (RNA processing) thì mRNA trưởng thành có vùng mã hóa protein dài 1.872 base (Hình 6.15).



**Hình 6.15** Phiên mã gene ovalbumin và sự tạo thành mRNA trưởng thành.

Việc lý giải cơ chế *cắt-nối* (splicing) trong quá trình xử lý pre-mRNA dựa chủ yếu trên hai sự kiện sau: (i) Ở hai đầu mút của mỗi intron có hai nucleotide rất ổn định, gọi là các "trình tự chuẩn", đó là 5'-GU.....AG-3'. Mỗi intron nói chung có độ dài khoảng  $10^2$ - $10^4$  nucleotide; và (2) Ở một số snRNA chứa trong thành phần của enzyme splicing cũng có các trình tự dinucleotide bổ sung với các trình tự chuẩn trong intron. snRNA trong phức hợp snRNP (*small nuclear ribonucleoprotein*) tương tác với các đầu mút của mỗi intron, kéo hai đầu mút xích lại gần nhau tạo ra cấu trúc hình vòng, nhờ đó enzyme tiến hành cắt bỏ intron và nối các exon lại với nhau.



**Hình 6.16** Một mô hình về cơ chế cắt-nối trong quá trình xử lý pre-mRNA.

Vấn đề đặt ra là sự có mặt của các intron trong các gene phân đoạn như thế có ý nghĩa gì? Bởi vì với lối tổ chức như vậy làm cho trình tự mã

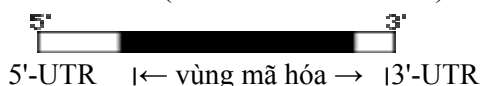
hoá của gene bị gián đoạn, và trong quá trình chế biến pre-mRNA của nó có thể xảy ra dù là một sai sót nhỏ cũng đủ để tạo ra mRNA có khung đọc mã bị thay đổi. Người ta cho rằng có khoảng 90% bản sao mRNA bị suy thoái và chỉ để lại khoảng 10% mRNA trưởng thành đi ra tế bào chất.

Theo quan niệm hiện nay, sự có mặt của các intron trong các gene phân đoạn có thể có các vai trò sau: (i) các intron như là các *đoạn đệm* (spacer) tạo thuận lợi cho sự tái tổ hợp giữa các exon bên trong gene; (ii) các intron như là các vùng đệm tách biệt các vùng chức năng của một số protein; và (iii) các intron như là các vùng phân cách cho phép các exon có thể được *cắt-nối có chọn lọc* (alternative splicing) để tạo ra các mRNA trưởng thành khác nhau và dịch mã thành các polypeptide khác nhau từ một gene; như đã nói trước đây, con đường cắt-nối này mang tính đặc thù cho từng mô ở các eukaryote bậc cao.

## V. Cấu trúc và chức năng của các loại RNA và ribosome

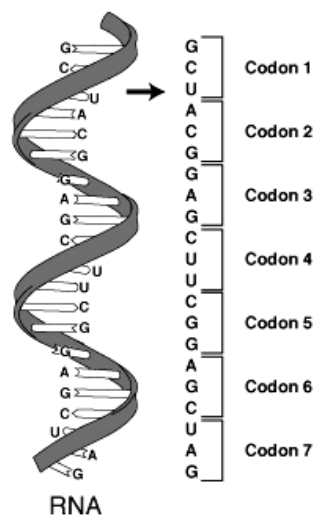
### 1. RNA thông tin (messenger RNA = mRNA)

Các mRNA là loại RNA quan trọng nhất dùng làm khuôn cho tổng hợp polypeptide (Hình 6.17); chúng có cấu trúc 'mở' với kích thước khác nhau tùy từng gene và đều có ba phần chính như sau (xem thêm Hình 6.14):



(i) *Vùng dẫn đầu* (5'UTR) tuy không được dịch mã nhưng cần thiết cho sự bám vào của tiểu đơn vị ribosome bé.

(ii) *Vùng mã hoá* (coding region) nằm kế sau 5'UTR, mang thông tin cấu trúc của một polypeptide, nếu là mRNA của eukaryote hoặc



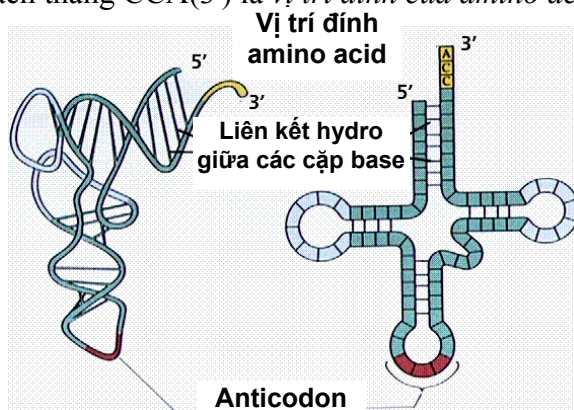
**Hình 6.17** Một đoạn của phân tử mRNA với trình tự các codon của nó. mang thông tin của nhiều polypeptide khác nhau (cách nhau bởi các đoạn đệm không được dịch mã), nếu là mRNA của prokaryote.

(iii) *Vùng theo sau* (3'-UTR) nằm cuối gene, không được dịch mã.

### 2. RNA vận chuyển (transfer RNA = tRNA)

Các tRNA có chức năng chính là mang amino acid và dịch mã cho một codon trên mRNA (Hình 6.10 và 6.18). Trong thành phần của chúng thường có một số base hiếm tập trung ở các vòng thân (stem loop) như: dihydro-uridine (DHU), inosine (I), ribothymidine (T) và pseudouridine

(Ψ) (xem hình 6.11). Các tRNA đều có cấu trúc không gian ba chiều gọn chặc là do sự kết cặp các base đặc thù ở một số đoạn của chúng (Hình 6.18). Nói chung, các phân tử tRNA gồm khoảng 75-85 base, với hệ số lằng chằng 4S; chúng thường rất giống nhau ở nhiều đoạn và sai khác chủ yếu là ở anticodon. Mỗi tRNA thường có 3-4 vòng với chức năng khác nhau: (i) *vòng DHU* nhận biết aminoacyl-tRNA synthetase; (ii) *vòng anticodon* đọc mã trên mRNA bằng sự kết cặp anticodon-codon; (iii) *vòng "phụ"* (extra loop) có thể không có ở một số tRNA; (iv) vòng TΨC nhận biết ribosome để đi vào đúng vị trí tiếp nhận aminoacyl-tRNA (vị trí A); và (v) đoạn mạch thẳng CCA(3') là *vị trí đính của amino acid*.



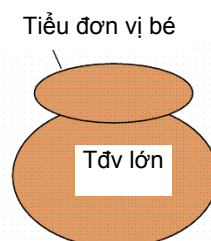
**Hình 6.18** Cấu trúc bậc ba (trái) và bậc hai của một phân tử tRNA.

### 3. RNA ribosome (ribosomal RNA = rRNA)

Các rRNA cùng với các protein đặc thù là những thành phần cấu trúc nên các ribosome -"nhà máy" tổng hợp protein của tế bào (Hình 6.19). Ở vi khuẩn có 3 loại rRNA có các hệ số lằng là 23S, 16S và 5S, với số lượng nucleotide tương ứng là 2904, 1542 và 120. Ở tế bào eukaryote có 4 loại rRNA với các hệ số lằng là 28S, 18S, 5,8S và 5S.

### 4. Ribosome

	<b>R 70S ở vi khuẩn</b>	<b>R 80S ở eukaryote</b>
<b>R 30S/40S</b>	rRNA: 16S Protein: 21 phân tử	rRNA: 18S Protein: >30 phân tử
<b>R 50S/60S</b>	rRNA: 23S + 5S Protein: 34 phân tử	rRNA: 28S + 5S + 5,8S Protein: >50 phân tử
<b>Đường kính</b>	18-20 nm	20-22 nm



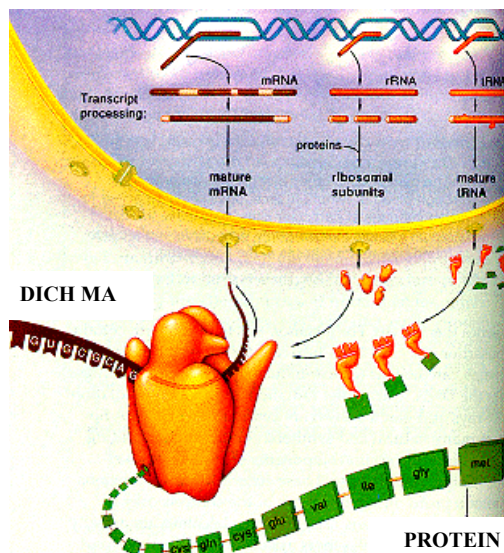
**Hình 6.19** Sơ đồ cấu trúc một ribosome với hai tiểu đơn vị lớn và bé.

Mỗi ribosome có hai *tiểu đơn vị bé* và *lớn* (small and large subunits).

Tiểu đơn vị bé có chức năng bám vào mRNA trước tiên trong dịch mã. Tiểu đơn vị lớn có chứa *peptidyl transferase* xúc tác hình thành các liên kết peptide; nó có chứa hai vị trí: *vị trí A* là nơi bám vào của aminoacyl-tRNA và *vị trí P* là chỗ dừng tạm của peptidyl-tRNA. Hai tiểu đơn vị này chỉ kết hợp với nhau để tạo ra một ribosome hoạt động khi quá trình dịch mã trên mRNA thực sự bắt đầu.

## VI. Cơ chế dịch mã (translation)

*Dịch mã* (translation) hay tổng hợp protein là một quá trình sinh học quan trọng diễn ra trong tế bào chất, và phụ thuộc vào nhiều yếu tố; quan trọng nhất là mRNA, các tRNA và ribosome. mRNA mang thông tin quy định trình tự kết hợp các amino acid vào chuỗi polypeptide, mà việc dịch mã mRNA được thực hiện bởi các aminoacyl-tRNA, còn ribosome đóng vai trò ổn định việc kết hợp giữa mRNA với các tRNA (Hình 6.20). Quá trình này được chia làm hai giai đoạn dưới đây.



**Hình 6.20** Đại cương về các quá trình phiên mã và sửa đổi sau phiên mã diễn ra trong nhân, và dịch mã trong tế bào chất ở tế bào eukaryote.

### 1. Hoạt hoá amino acid

Quá trình này diễn ra trong bào tương và tạo nguồn các tRNA mang các amino acid sẵn sàng tham gia dịch mã. Mỗi amino acid được đính vào tRNA thích hợp nhờ một enzyme *aminoacyl-tRNA synthetase* đặc thù. Trước tiên, enzyme này (E) xúc tác cho phản ứng ATP hoạt hoá amino acid, với sự có mặt của  $Mg^{2+}$ , tạo ra phức hợp  $[E\text{-aminoacyl}\sim\text{AMP}]$ .



Tiếp theo, cũng dưới tác dụng của enzyme đó, phức hợp này kết hợp với tRNA thích hợp bằng liên kết đồng hoá trị để tạo ra *aminoacyl-tRNA*.

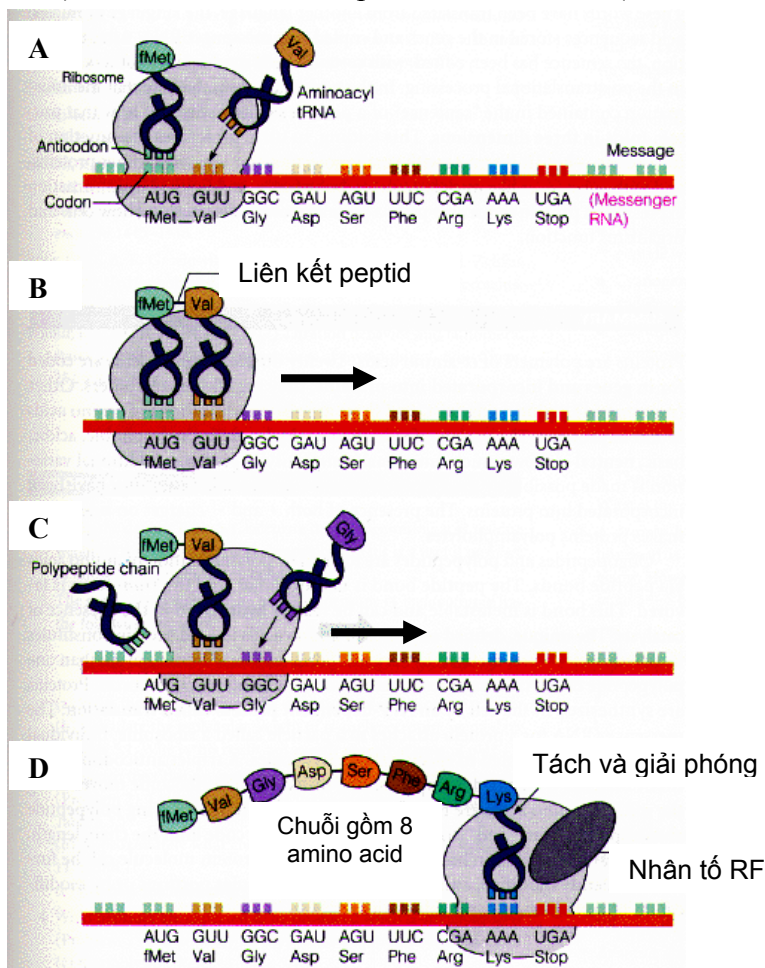


### 2. Cơ chế của quá trình dịch mã (tổng hợp polypeptide)

**Bước 1:** *Mở đầu* (initiation)



Quá trình dịch mã bắt đầu khi một tiểu đơn vị ribosome bé bám vào mRNA tại vị trí của codon khởi đầu AUG. Lúc này một phân tử tRNA khởi đầu đặc thù mang methionine (ở vi khuẩn là formyl-Met) đi vào và khớp anticodon của nó với codon mở đầu của mRNA. Kế đó, tiểu đơn vị ribosome lớn bám vào tiểu đơn vị bé tạo ra một ribosome hoạt động hoàn chỉnh. Lúc này Met-tRNA ở vị trí P và vị trí A để trống; một tRNA thứ hai (ví dụ, tRNA<sup>Val</sup>) đi vào vị trí A và khớp với codon thứ hai (hình 6.21A).



**Hình 6.21** Quá trình sinh tổng hợp chuỗi polypeptide.

Bước 2: Kéo dài (elongation)

Quá trình kéo dài bắt đầu sau khi liên kết peptide đầu tiên được hình thành. Phản ứng này được xúc tác bởi enzyme *peptidyl transferase*, và kết quả là tạo ra một peptidyl-tRNA ở vị trí A (hình 6.21B). Sau đó, ribosome lập tức chuyển dịch sang một codon mới dọc theo mRNA theo chiều

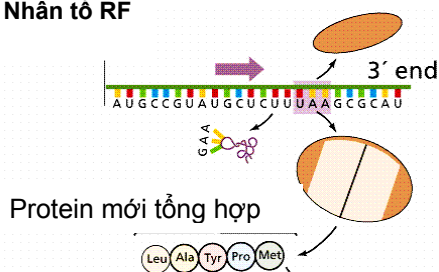
5'→3' (hình 6.21C). Phản ứng này đẩy phân tử tRNA tự do vốn ở vị trí P ra ngoài; lúc này peptidyl-tRNA nằm ở vị trí P và vị trí A lại để trống. Một chu kỳ dịch mã mới lại bắt đầu, một aminoacyl-tRNA thứ ba đi vào và khớp anticodon của nó với codon đang để trống ở vị trí A, một liên kết peptid thứ hai được hình thành, và ribosome lại dịch chuyển sang codon kế tiếp. Quá trình nói trên cứ diễn ra một cách tuần tự dọc theo mRNA làm cho chuỗi polypeptide dài dần ra cho đến dịch mã xong codon 'có nghĩa' cuối cùng.

### Bước 3: Kết thúc (termination)

Quá trình tổng hợp chuỗi polypeptide sẽ dừng lại khi một codon kết thúc được đưa đối diện với vị trí A để trống, vốn được nhận biết bởi một protein kết thúc gọi là *nhân tố giải phóng* RF (release factor). Sự có mặt của nó

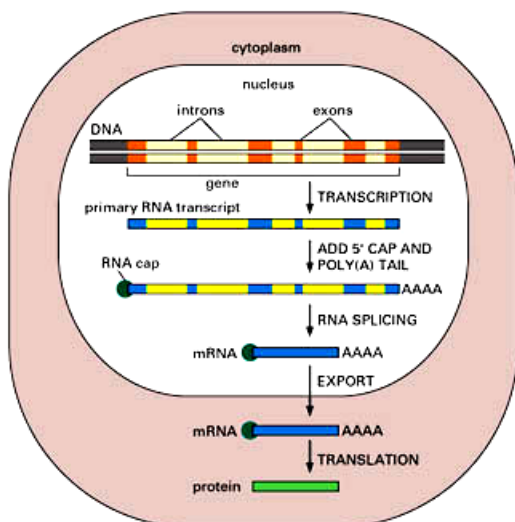
cùng với enzyme transferase cắt rời chuỗi polypeptide ra khỏi tRNA cuối cùng (hình 6.21D) và phóng thích hai tiểu đơn vị ribosome cũng như chuỗi polypeptide và tRNA ra khỏi mRNA (hình 6.22).

### Nhân tố RF

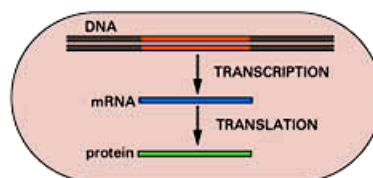


Hình 6.22 Sự kết thúc dịch mã.

### (A) EUKARYOTES



### (B) PROCARYOTES



Hình 6.23 So sánh các cơ chế tổng hợp và dịch mã mRNA ở các tế bào eukaryote (trái) và prokaryote.



Một số điểm cần lưu ý thêm:

(1) Trên đây mới chỉ phân tích hoạt động của một ribosome trên mRNA. Thực ra, trên một mRNA có rất nhiều ribosome cùng hoạt động, gọi là polyribosome hay *polysome*, tạo ra nhiều polypeptide giống nhau.

(2) Trên nguyên tắc, amino acid mở đầu sẽ được cắt bỏ khỏi chuỗi polypeptide (sau khi tổng hợp được vài amino acid như trong trường hợp các vi khuẩn) hoặc trước khi chuỗi được tổng hợp đầy đủ (như ở trường hợp eukaryote). Tuy nhiên, ở các eukaryote không phải lúc nào amino acid mở đầu này cũng bị tách bỏ, mà trong một số protein nó vẫn được giữ lại.

(3) Sau khi được tổng hợp, các chuỗi polypeptide sơ cấp này sẽ được sửa đổi và chuyển sang các bậc cấu trúc cao hơn theo cách đặc thù để trở thành các protein hoạt động chức năng (xem các hình 6.8 và 6.9).

(4) Tham gia vào các bước mở đầu, kéo dài và kết thúc còn có các yếu tố protein, với tên gọi tương ứng là các *nhân tố mở đầu* (IF: initiation factor), *nhân tố kéo dài* (EF: elongation factor), và *nhân tố giải phóng* (release factor) cùng với ATP, GTP và các ion như  $Mg^{2+}$ ,  $K^+$  và  $NH_4^+$ .

(5) Trong các tế bào prokaryote, do không có màng nhân và các mRNA đa cistron vốn dĩ không phải qua sửa đổi sau phiên mã, cho nên các ribosome và các aminoacyl-tRNA sẽ bám vào đầu 5' của mRNA để bắt đầu quá trình dịch mã ngay trong khi ở đầu 3' của nó quá trình phiên mã đang còn tiếp diễn. Ngược lại, ở các tế bào eukaryote vì có màng nhân phân cách và các pre-mRNA còn phải trải qua các công đoạn sửa đổi phức tạp sau phiên mã (tất cả đều diễn ra trong nhân), còn dịch mã diễn ra sau đó ở trong tế bào chất (hình 6.23). Vì thế cho nên phiên mã và dịch mã rõ ràng là hai quá trình tách biệt nhau cả về cả không gian lẫn thời gian.

## Câu hỏi và Bài tập

1. Tại sao nói *gene* là khái niệm căn bản của di truyền học và thật khó để có thể đưa ra một định nghĩa chính xác? Hãy chứng minh điều đó qua các giai đoạn phát triển của di truyền học cho đến nay.

2. Bạn hiểu gì về các gene phân đoạn (*split gene*)? Cho thí dụ và sơ đồ minh họa. Các *intron* là gì và có vai trò như thế nào? Phân tích một trường hợp sửa đổi (sau phiên mã) sản phẩm của gene phân đoạn để làm sáng tỏ mệnh đề “một gene có thể xác định nhiều hơn một loại polypeptide”.

3. (a) Hãy làm sáng tỏ nhận định sau: Khác với các tế bào prokaryote, sự phiên mã và dịch mã ở các tế bào eukaryote là hai quá trình tách biệt cả về không gian lẫn thời gian. (b) Quá trình tổng hợp và hoàn thiện các

RNA ở prokaryote và eukaryote giống và khác nhau ở những điểm nào?

4. Trình bày vai trò chính của các yếu tố tham gia vào quá trình sinh tổng hợp protein của tế bào và phân tích cơ chế sinh tổng hợp protein dựa trên sự tương tác giữa mRNA, các aminoacyl-tRNA và ribosome.

5. Bằng suy luận, hãy cho biết tại sao mật mã di truyền chỉ có thể được mã hóa theo nguyên tắc bộ ba mà không thể là kiểu khác? Bằng thực nghiệm điều đó đã được giải quyết như thế nào? Tính thoái hóa của mã và nguyên tắc kết cặp linh hoạt có quan hệ với nhau ra sao? Cho ví dụ.

6. Phân tích sự phù hợp giữa cấu trúc và chức năng của các loại RNA và mối quan hệ giữa gene và protein.

7. Giả sử tổng hợp được một mRNA có thành phần 75%U và 25%G. Khi sử dụng mRNA này để tổng hợp protein *in vitro* đã thu được các amino acid trong các protein với các tần số như sau :

Phe : Val : Leu : Cys : Gly : Trp = 1,00 : 0,44 : 0,33 : 0,33 : 0,15 : 0,11.

Hãy trình bày phương pháp và chỉ ra kết quả của việc giải đoán các codon cho mỗi amino acid nói trên (*không sử dụng bảng mã di truyền*). Biết rằng các codon cùng xác định một amino acid thường có hai nucleotide đầu giống nhau, và Cys được xác định bởi UGU.

8. (a) Hàm lượng GC của DNA phage T3 là 53%. Bạn sẽ kỳ vọng hàm lượng G+C của mRNA T3 ra sao? (b) Nếu biết được hàm lượng purine của DNA phage T3 và không biết mạch nào làm khuôn, có thể dự đoán hàm lượng purine của mRNA T3 hay không? Tại sao, hoặc tại sao không?

9. Nếu sử dụng các phân tử mRNA nhân tạo có thành phần gồm các cụm gồm ba hoặc bốn nucleotide lặp lại dưới đây để tiến hành tổng hợp protein trong ống nghiệm, thành phần amino acid thu được từ các polypeptide sẽ như thế nào? Có trường hợp nào không tổng hợp được protein hay không? Tại sao?

(a) (UUC)<sub>n</sub> ; (b) (UAC)<sub>n</sub> ; (c) (GAUA)<sub>n</sub> ; (d) (GUAA)<sub>n</sub>.

10. Nếu sử dụng ba phân tử mRNA tổng hợp có thành phần gồm các base được kết hợp ngẫu nhiên theo tỷ lệ sau đây để tiến hành tổng hợp protein *in vitro*, thì thành phần và tỷ lệ các amino acid được kết hợp trong các protein sẽ như thế nào? (a) 1U:5C (b) 1A:1C:4U (c) 1A:1C

## Tài liệu Tham khảo

### Tiếng Việt

Hoàng Trọng Phán. 1995. *Một số vấn đề về Di truyền học hiện đại* (Tài liệu BDTX cho giáo viên THPT chu kỳ 1993-1996). Trường ĐHSP Huế.

Hoàng Trọng Phán. 1997. *Di truyền học Phân tử*. NXB Giáo Dục.

### **Tiếng Anh**

Alberts B, Johnson A, Lewis J, Raff M, Robert K, Walter P. 2002. *Molecular Biology of the Cell*. Garland Science, NY.

Cambridge Healthtech Institut. 2005. *Gene Definition*.

[www.healthtech.com](http://www.healthtech.com)

Cooper DN. 2004. *Gene structure, function and expression*. [http://www.cardiff.ac.uk/medicine/medical\\_genetics/study/medical\\_teaching/](http://www.cardiff.ac.uk/medicine/medical_genetics/study/medical_teaching/)

McClellan P. 1998. *Definition of Gene*.

<http://www.ag.ndsu.nodak.edu/plantsci/undrcour.htm>

Lewin B. 1999. *Genes VI*. Oxford University Press, Oxford.

Lewis R. 2003. *Human Genetics: Concepts and Applications*. 5<sup>th</sup> ed, McGraw-Hill, Inc, NY.

Palladino MA. 2002. *Understanding the Human Genome Project*. Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA.

Cambridge Healthtech Institut. 2005. *RNA Glossary*.

[www.healthtech.com](http://www.healthtech.com)

Russell PJ. 2003. *Essential Genetics*. Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA.

Tamarin RH. 1999. *Principles of Genetics*. 6<sup>th</sup> ed., McGraw-Hill, Inc, NY.

Twyman RM. 1998. *Advanced Molecular Biology*. BIOS Scientific Publishers Ltd/ Springer-Verlag Singapore Pte Ltd.

Watson JD, Hopkins NH, Roberts JW, Steitz JA, Weiner AM. 1987. *Molecular Biology of the Gene*. 4<sup>th</sup> ed., Benjamin/Cummings Publishing Company, Inc, Menlo Park, CA.

Weaver RF, Hedrick PW. 1997. *Genetics*. 3<sup>rd</sup> ed, McGraw-Hill Companies, Inc./ Wm.C.Brown Publishers, Dubuque, IA.

### **Một số trang web bổ sung**

*Gene structure, function and expression*.

[http://www.cardiff.ac.uk/medicine/medical\\_genetics](http://www.cardiff.ac.uk/medicine/medical_genetics)

<http://www.oup-usa.org/isbn/019879276X.html>

## Chương 7

# Sự Điều hòa Biểu hiện của Gene

### I. Các nguyên lý điều hòa và mức độ kiểm soát phiên mã

Không phải tất cả các gene đều có biểu hiện liên tục. Mức độ biểu hiện của gene khác nhau giữa các tế bào hoặc khác nhau theo giai đoạn trong chu trình tế bào. Chẳng hạn gene mã hóa cho hemoglobin được biểu hiện ở mức độ cao chỉ ở trong tế bào tiền thể (precursor) của tế bào máu. Hoạt tính của gene khác nhau theo chức năng tế bào. Ở động vật có xương sống như chuột, chứa khoảng 200 loại tế bào được phân hóa chức năng khác nhau. Tất cả các tế bào đều chứa cùng thông tin di truyền, những tế bào khác nhau chỉ ở những gene hoạt động. Trong nhiều trường hợp, hoạt tính của gene được điều hòa ở mức độ phiên mã, cả qua những tín hiệu bắt đầu bên trong tế bào và cả phản ứng với những điều kiện bên ngoài. Tuy nhiên thông tin di truyền được điều hòa theo những cách khác nhau. Các bước điều khiển hoạt động gene bao gồm:

- Cấu trúc lại DNA, trong đó những thay đổi biểu hiện gene phụ thuộc vào vị trí trình tự DNA trong genome.

- Điều hòa phiên mã trong tổng hợp bản phiên mã RNA bằng sự điều khiển sự mở đầu và sự kết thúc

- Quá trình chế biến RNA hoặc điều hòa qua quá trình cắt-nối trên RNA (RNA splicing)

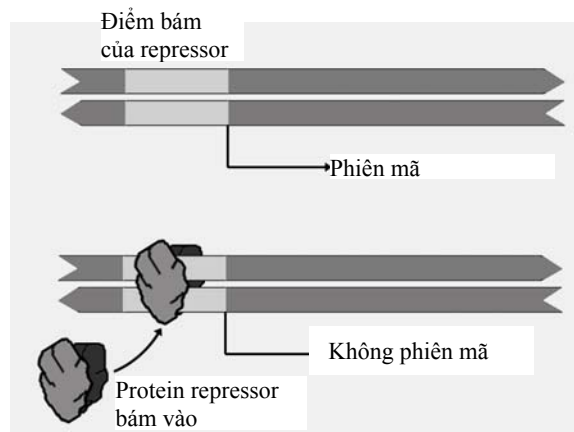
- Điều hòa dịch mã quá trình tổng hợp chuỗi polypeptid

- Sự bền vững của mRNA

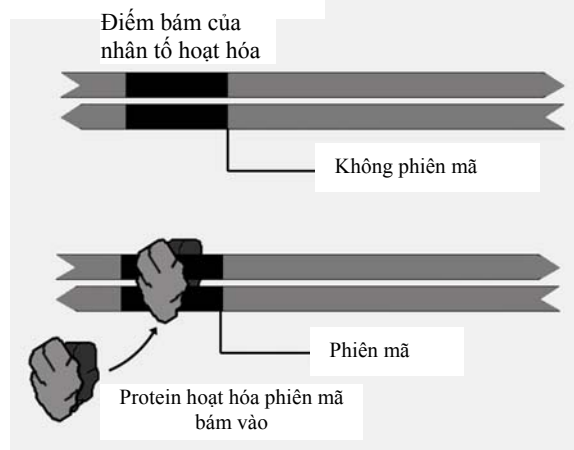
Có sự khác nhau đáng kể trong sự điều hòa biểu hiện của các gene ở eukaryote và prokaryote. Prokaryote là những cơ thể đơn bào sống tự do, sinh trưởng và phân chia trong điều kiện thích hợp và được cung cấp đầy đủ chất dinh dưỡng. Do đó hoạt động của các gene được điều hòa do nhu cầu của tế bào khi cần thiết.

Khác với prokaryote, eukaryote là những cơ thể đa bào. Trong cơ thể đang phát triển, một tế bào không chỉ sinh trưởng và phân chia mà tế bào thể hệ sau trải qua những thay đổi quan trọng về hình thái và hóa sinh và duy trì trạng thái biến đổi của nó. Hơn nữa trong quá trình phát triển phôi, tế bào eukaryote ít chịu ảnh hưởng của môi trường hơn so với vi khuẩn. Cuối cùng, ở cơ thể trưởng thành, sự sinh trưởng và phân chia tế bào của hầu hết các loại tế bào bị ngừng, mỗi tế bào chỉ phải duy trì các tính chất riêng biệt của nó.

### A. Điều hòa âm tính



### B. Điều hòa dương tính



**Hình 7.1** Mô hình điều hòa âm tính (negative regulation) và điều hòa dương tính (positive regulation).

A. Trong điều hòa âm tính, protein ức chế bám vào phân tử DNA, ngăn cản phiên mã

B. Trong điều hòa dương tính, sự bám vào của protein hoạt hóa phiên mã, kích thích phiên mã.

## II. Điều hòa hoạt động gene ở prokaryote

Ở vi khuẩn và phage, hoạt tính đóng mở gene thường được điều khiển qua phiên mã tổng hợp các mRNA xảy ra khi sản phẩm của gene được cần đến.

Cơ chế phân tử cho mỗi mô hình điều hòa hoàn toàn khác nhau, nhưng thường theo một trong hai kiểu chính: điều hòa âm tính và điều hòa

dương tính (Hình 7.1). Trong hệ thống điều hòa âm tính (Hình 7.1A) một protein ức chế có mặt trong tế bào, ngăn cản sự phiên mã. Trong hệ thống cảm ứng được điều hòa âm tính, protein ức chế hoạt động làm ngăn cản phiên mã. Nhân tố cảm ứng kim hãm chất ức chế, cho phép bắt đầu phiên mã. Trong hệ thống ức chế, protein aporepressor gắn với co-repressor tạo ra chất ức chế có hoạt tính, làm ngăn cản phiên mã. Ngược lại trong hệ thống điều hòa dương tính (Hình 7.1B), sự tổng hợp mRNA xảy ra nếu protein điều hòa gắn vào một vùng của gene làm hoạt hóa phiên mã. Những protein này được xem là những nhân tố hoạt hóa phiên mã. Điều hòa âm tính và dương tính không loại trừ lẫn nhau. Một vài hệ thống là cả điều hòa âm tính và dương tính, sử dụng cả 2 hệ thống điều hòa để phản ứng với các điều kiện khác nhau trong tế bào. Điều hòa âm tính là phổ biến cho prokaryote, trong khi điều hòa dương tính lại phổ biến cho eukaryote.

### 1. Cấu trúc của operon

Mô hình operon của điều hòa phiên mã

Cơ chế điều hòa di truyền của hệ thống *lac* (*lac* system) được giải thích bằng mô hình operon của Francois Jacob và Jacque Monod (1960) (Hình 7.2).

Hệ thống sử dụng lactose gồm 2 loại thành phần: gene cấu trúc mã hóa protein cần thiết cho sự vận chuyển và chuyển hóa lactose và các yếu tố điều hòa (gene ức chế *lacI*, promotor *lacP* và operator *lacO*).

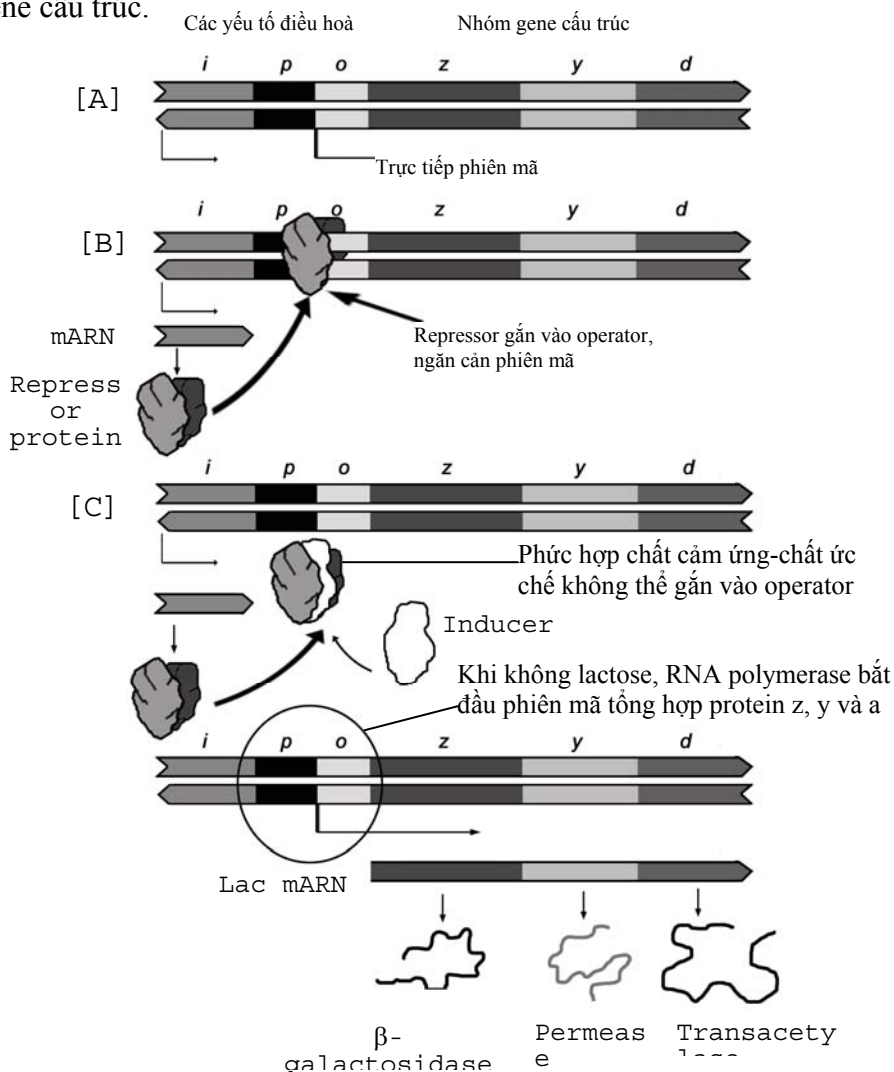
Sản phẩm gene cấu trúc được mã hóa bởi một phân tử mRNA đa cistron (polycistronic). Gene Z mã hóa cho enzyme  $\beta$ -galactosidase (thủy phân đường lactose thành galactose và glucose), gene Y mã hóa cho enzyme permease (cần cho vận chuyển lactose qua màng), gene A mã hóa cho enzyme transacetylase (vai trò chuyển hóa lactose chưa rõ).

Đột biến promotor (*lacP*) làm mất khả năng tổng hợp mRNA.

Sản phẩm của gene *lacI* là chất ức chế, nó bám vào trình tự các base của DNA cấu tạo operator. Chất ức chế bám vào operator, ngăn cản sự khởi đầu phiên mã mRNA nhờ RNA polymerase. Chất cảm ứng (lactose) kích thích sự sinh tổng hợp mRNA bằng cách kết hợp và làm bất hoạt chất ức chế. Sự có mặt của chất cảm ứng làm chất ức chế không gắn vào operator, promotor cho phép khởi đầu tổng hợp mRNA.

Khi môi trường có lactose, lactose được chuyển vào tế bào nhờ permease. Khi vào trong tế bào một số lactose (liên kết  $\beta$ -1,4) được chuyển thành allolactose (liên kết  $\beta$ -1,6) nhờ  $\beta$ -galactosidase. Allolactose là chất cảm ứng, nó gắn vào protein kim hãm, gây biến đổi cấu hình tạo

phức hợp allolactose-repressor. Phức hợp này mất khả năng gắn vào operator. Lúc này operon mở ra, RNA polymerase bắt đầu phiên mã từ gene cấu trúc.



**Hình 7.2 A Bản đồ của operon *lac*;**

**B. Sơ đồ của operon *lac* ở trạng thái bị kìm hãm**

**C. Sơ đồ của operon *lac* ở trạng thái được kích thích**

Khi môi trường không có lactose, protein ức chế có hoạt tính gắn vào operator, làm sự phiên mã của tất cả các gene cấu trúc của operon *lac* bị dừng.

Sự điều hòa của operon yêu cầu promotor nằm chồng lên một phần hoặc kề sát bên promotor của gene cấu trúc, vì nó gắn với chất ức chế

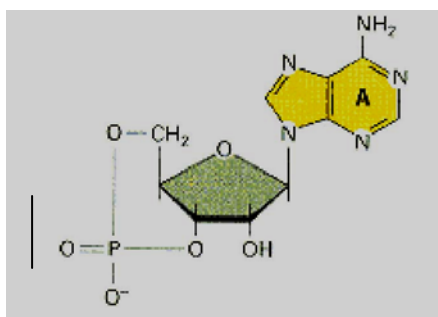
ngăn cản phiên mã.

## 2. Điều hòa dương tính operon lactose

Chức năng của  $\beta$ -galactosidase trong chuyển hóa lactose để tạo ra glucose bằng cách cắt lactose (một sản phẩm cắt khác là galactose cũng có thể được chuyển thành glucose nhờ enzyme của operon galactose). Nếu cả glucose và lactose có mặt trong môi trường sinh trưởng, không cần hoạt động của operon lac. Trong môi trường có glucose, enzyme  $\beta$ -galactosidase không được tạo thành cho đến khi glucose trong môi trường được sử dụng hết. Sự có mặt của glucose làm mRNA không được tổng hợp, do đó không có sự tổng hợp  $\beta$ -galactosidase, vì sự thêm vào nhân tố cảm ứng làm bất hoạt chất kim hãm. Một yếu tố khác cần cho sự bắt đầu tổng hợp mRNA, hoạt tính của yếu tố này được điều hòa bởi nồng độ glucose.

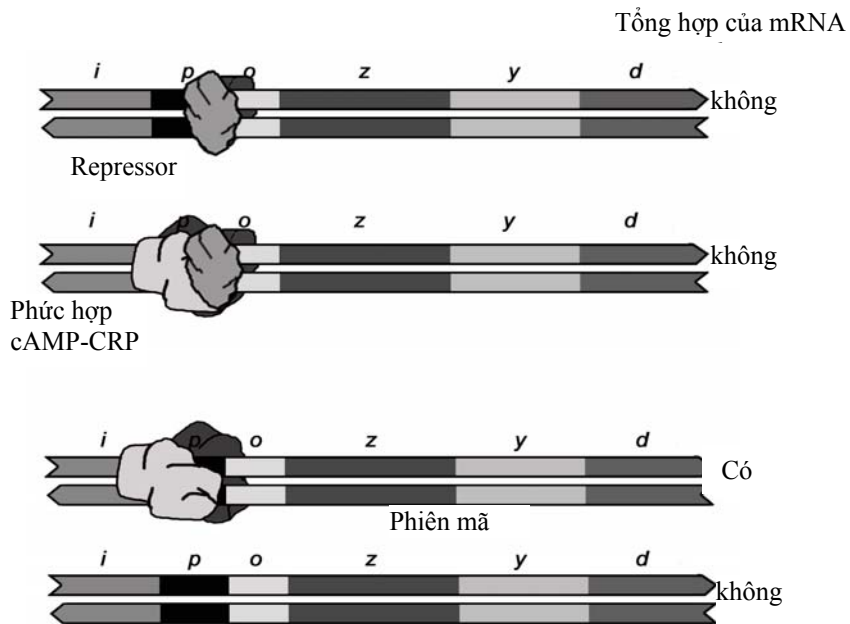
Glucose có ảnh hưởng ức chế gián tiếp lên sự biểu hiện của operon lac. Những phân tử nhỏ adenosine monophosphate vòng (cAMP) phân bố rộng rãi trong mô động vật và trong các cơ thể eukaryote đa bào, có vai trò quan trọng làm chất trung gian hoạt động hormone. Chất này cũng có trong tế bào *E. coli* và nhiều tế bào vi khuẩn khác với chức năng khác nhau. cAMP được tổng hợp bởi enzyme adenyl cyclase, và nồng độ của cAMP được điều hòa gián tiếp qua trao đổi chất glucose. Khi vi khuẩn sinh trưởng ở môi trường chứa glucose, hàm lượng cAMP rất thấp. Trong môi trường chứa glycerol hoặc các nguồn carbon không thể đi vào con đường hóa sinh được sử dụng để trao đổi chất glucose (con đường glycolytic) hoặc khi vi khuẩn bị đói nguồn năng lượng, nồng độ cAMP cao. Hàm lượng glucose giúp điều hòa nồng độ cAMP trong tế bào và cAMP lại điều hòa hoạt tính của operon lac.

*E. coli* chứa protein nhận cAMP hay CRP (cyclic AMP receptor protein) còn được gọi là *protein hoạt hóa dị hóa CAP* (catabolite activator protein), được mã hóa bởi gene *crp*. Đột biến ở gene *crp* và gene *adenyl cyclase* làm ngăn cản sự tổng hợp của mRNA *lac*. Điều này cho thấy chức năng của CRP và cAMP cần thiết cho tổng hợp mRNA *lac*. CRP và cAMP gắn vào một vị trí khác tạo phức hợp cAMP-CRP được biểu hiện. Phức hợp này là một yếu tố điều hòa hoạt hóa ở hệ thống *lac*.



**Hình 7.3** Cấu trúc của cAMP





**Hình 7.4** Bốn trạng thái điều hòa của operon *lac*

Nhu cầu về cAMP-CRP phụ thuộc vào hệ thống ức chế *lac*, vì đột biến *crp* và adenyl cyclase không thể tạo mRNA *lac* ngay cả khi có mặt đột biến ở gene điều hòa (*lacI*) và gene chỉ huy (*lacO*). Sở dĩ như vậy là vì phức hợp cAMP-CRP phải gắn vào trình tự base trên DNA ở vùng promoter để xảy ra phiên mã (Hình 7.4). Khác với chất ức chế trong điều hòa âm tính, phức hợp cAMP-CRP là chất điều hòa dương tính. Các thí nghiệm ở điều kiện *in vitro* cho thấy: mRNA *lac* được tổng hợp chỉ khi có cAMP vòng và không có chất ức chế.

Khi vắng mặt phức hợp cAMP-CRP, RNA polymerase chỉ bám lỏng lẻo vào promoter. Vì vậy hiếm khi dẫn đến phiên mã vì không có sự tương tác đúng giữa RNA polymerase và promoter. Nhưng RNA polymerase được kích thích gắn vào promoter khi cAMP-CRP được gắn vào DNA.

Những kết quả này giải thích chức năng lactose và glucose cùng nhau tham gia điều hòa phiên mã operon *lac* như thế nào.

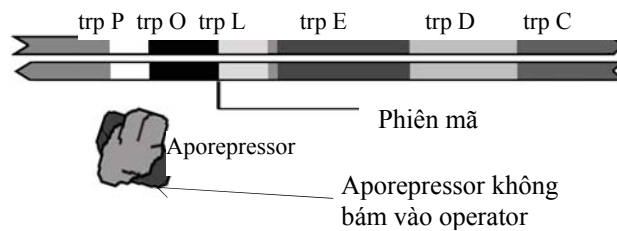
### 3. Điều hòa âm tính operon tryptophan

Operon tryptophan (*trp*) của *E.coli* chứa các gene cấu trúc mã hóa cho các enzyme tổng hợp amino acid tryptophan. Operon này được điều hòa theo cách sau: khi tryptophan có mặt đầy đủ trong môi trường sinh trưởng, sự phiên mã của operon bị ức chế. Khi sự cung cấp tryptophan bị thiếu, sự phiên mã xảy ra. Sự điều hòa của operon lactose tương tự với

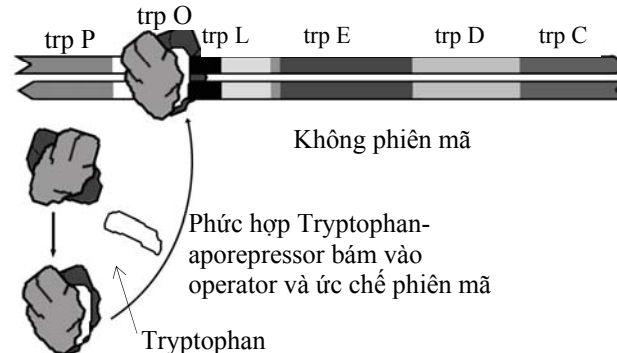
operon lactose, vì sự tổng hợp mRNA được điều hòa âm tính nhờ chất ức chế. Tuy nhiên, khác với điều hòa ở operon *lac*, tryptophan hoạt động như chất đồng kim hãm, kích thích chất ức chế gắn vào operator ngừng sự tổng hợp. Operon tryptophan hoạt động theo kiểu ức chế, điều hòa âm tính.

Tryptophan được tổng hợp qua các giai đoạn khá phức tạp, mỗi giai đoạn có sự xúc tác của một enzyme đặc biệt. Các gene mã hóa cho các enzyme này nằm kề nhau trên nhiễm sắc thể của *E.coli*. Đó là các gene *trpE*, *trpD*, *trpC*, *trpB*, *trpA*. Các enzyme được dịch mã từ một phân tử mRNA đa cistron. Vùng mã hóa gene E được dịch mã trước tiên. Phía trước *trpE* về đầu 5' có promotor, operator và 2 vùng xếp lần lượt là leader (*trpL*) và đoạn kim hãm phiên mã attenuator (*trpA*, không phải là *trpA*). Gene ức chế *trpR* nằm xa operon, tổng hợp protein *aporepressor*, là chất kim hãm mà riêng nó không có hoạt tính. Khi tryptophan dư thừa, nó kết hợp với *aporepressor* tạo chất kim hãm có hoạt tính, gắn vào operator của operon tryptophan làm dừng phiên mã các gene cấu trúc. Khi nồng độ tryptophan thấp, nó tách khỏi phức hợp kim hãm và *aporepressor* mất hoạt tính. Lúc này operator mở ra, RNA polymerase dịch mã 5 gene cấu trúc để tổng hợp 5 enzyme tham gia tổng hợp ra tryptophan. (Hình 7.5).

Phiên mã xảy ra



Phiên mã bị ức chế



Hoạt hóa aporepressor

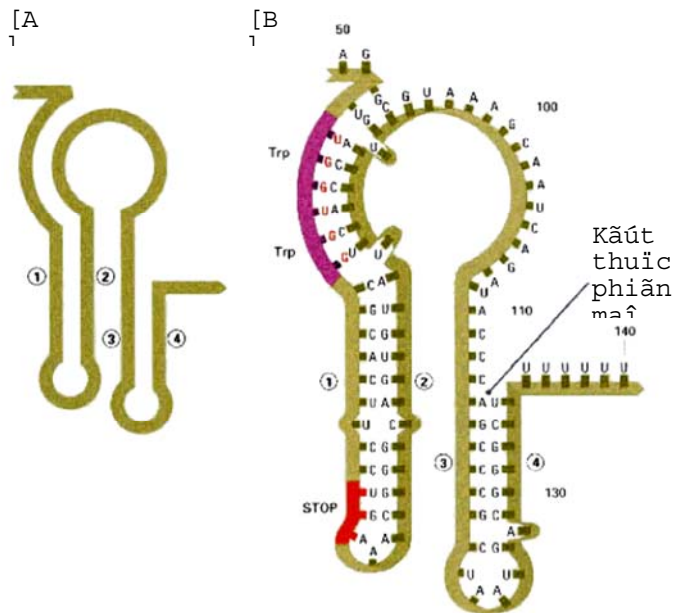
**Hình 7.5 Điều hòa của *trp* operon ở *E. coli***

A. Protein *aporepressor* không bám được vào operator, phiên mã xảy ra.

B. Khi có đủ tryptophan, phức hợp aporepressor và tryptophan làm chất ức chế hoạt động gắn được vào operator, sự phiên mã bị kìm hãm.

#### 4. Phiên mã dờ (Attenuation)

Kiểu điều hòa thứ hai được phát hiện ở operon tryptophan được gọi là attenuation. Nó dùng sử dịch mã để điều khiển sự phiên mã. Khi có mặt tryptophan nội bào, ngay cả với nồng độ thấp, sự dịch mã một phần vùng leader của mRNA ngay khi vừa được tổng hợp, kết quả làm dừng sự phiên mã trước khi gene cấu trúc đầu tiên của operon được sao chép.



**Hình 7.6 (A) Sơ đồ phiên mã của leader *trp*;  
(B) Chi tiết cấu trúc của 2 codon *trp* ở vòng 1-2**

Phiên mã dờ (attenuation) là kết quả sự tương tác giữa các trình tự DNA trong vùng leader của bản phiên mã *trp*. Ở tế bào kiểu dại, sự phiên mã operon *trp* thường được bắt đầu. Tuy nhiên khi có mặt một lượng nhỏ tryptophan, hầu hết phân tử mRNA kết thúc ở vùng 28 base đặc biệt ở trong trình tự leader. Kết quả của sự kết thúc sớm này tạo phân tử mRNA chứa 140 nucleotide chấm dứt một đoạn ngắn của các gene mã hóa cho các enzyme *trp*. Vùng 28 base xảy ra sự kết thúc phiên mã sớm như thế được gọi là *attenuator*. Trình tự base của vùng này thường có các tính chất diêm kết thúc, gồm dạng đoạn và vòng (stem-loop) trên mRNA theo sau là trình tự của 8 uridine.

Trình tự leader có các đặc điểm:

- Một vùng có codon AUG và phía sau là codon kết thúc UGA, mã hóa cho một polypeptide chứa 14 amino acid được gọi là leader polypeptide.

- Hai codon tryptophan ở vị trí 10 và 11 trên mRNA của leader polypeptide. Trình tự lặp lại ngắn này có ý nghĩa trọng điều hòa.

- Bốn đoạn của RNA leader là vùng 1, 2, 3 và 4 tạo thành do khả năng kết cặp của các base với nhau. Các base ở vùng 1 kết cặp với vùng 2, vùng 3 kết cặp với vùng 4 (Hình 7.6).

Khi sự kết cặp xảy ra ở dạng này, sự phiên mã kết thúc ở đoạn đi qua uridine phía trước nucleotide 140. Kiểu kết cặp này xảy ra ở mRNA leader được tinh sạch.

- Một kiểu kết cặp biến đổi có thể xảy ra, trong đó các base vùng 2 kết cặp với vùng 3 nhờ các cặp base ở 2 vùng này gần như bổ sung nhau (Hình 7.6B). Qua mô hình kết cặp base biến đổi này (3-4 hoặc 2-3), sự tổ chức trình tự mRNA có thể điều hòa phiên mã qua dịch mã của leader polypeptide (Hình 7.7). Khi vùng leader được phiên mã, sự dịch mã leader polypeptid cũng bắt đầu. Vì có 2 codon của tryptophan trong trình tự mã hóa, nên sự dịch mã nhạy cảm với số lượng tRNA<sup>trp</sup> đưa vào.

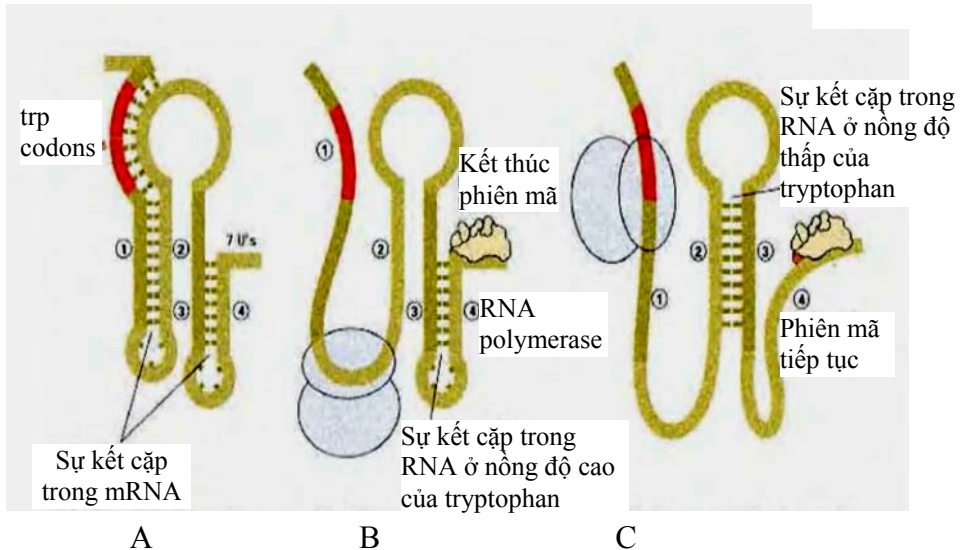
Nếu môi trường cung cấp đầy đủ tryptophan, ribosome trượt qua codon tryptophan và đi vào vùng 2 (Hình 7.7B). Sự có mặt của ribosome loại bỏ khả năng kết cặp của vùng khoảng 10 base ở mỗi phía của codon đang dịch mã. Sự có mặt của ribosome ở vùng 2 ngăn cản nó kết cặp với vùng 3. Trong trường hợp này vùng 3 kết cặp với vùng 4, tạo ra điểm kết thúc phiên mã. Sự phiên mã kết thúc khi qua các uridine nằm phía sau vùng 4.

Khi số lượng tRNA<sup>trp</sup> không đủ, sự dịch mã leader polypeptide bị dừng lại đột ngột ở các codon tryptophan. Sự dừng lại này ngăn cản ribosome tiến vào vùng 2, vì vậy vùng 2 được tự do sẽ kết cặp với vùng 3 làm cản trở sự hình thành cấu trúc kết thúc. Vì vậy phân tử *trp* mRNA hoàn chỉnh được tạo thành, chứa cả trình tự mã hóa cho gene cấu trúc.

Tóm lại, attenuation là cơ chế điều hòa tinh tế trên cơ sở điều hòa âm tính: Khi tRNA<sup>trp</sup> đến đủ cung cấp cho sự dịch mã leader polypeptide, sự phiên mã bị dừng, các *trp* enzyme không được tổng hợp. Khi nồng độ tRNA<sup>trp</sup> quá thấp, sự phiên mã xảy ra cho đến hết, các *trp* enzyme được tạo nên.

Nhiều operon chịu trách nhiệm tổng hợp các amino acid khác (như các operon của leucine, isoleucine, phenylalanine, histidine) cũng được điều hòa nhờ attenuator với chức năng tạo ra vùng kết cặp biến đổi ở bản

phiên mã. Ở operon histidine vùng mã hóa của leader polypeptide chứa 7 codon histidine kế nhau. Ở operon phenylalanine vùng mã hóa cho leader polypeptide chứa 7 codon phenylalanin chia 3 nhóm.



**Hình 7.7 Phiên mã dở (attenuation) của operon *trp* ở *E. coli***

A. Ở mRNA tự do có sự kết cặp base giữa 1-2 và 3-4

B. Ở nồng độ cao của tryptophan, ribosome tiến đến vùng 2 và sự kết cặp 3-4 làm kết thúc phiên mã

C. Ở nồng độ tryptophan thấp, ribosome ở vùng codon *trp* cho phép kết cặp 2-3 và phiên mã không bị kết thúc sau khi qua vùng 4

Điều hòa kiểu attenuation không thể xảy ra ở eukaryote vì ở eukaryote sự phiên mã và dịch mã không xảy ra đồng thời. Sự phiên mã xảy ra trong nhân, còn sự dịch mã xảy ra ở tế bào chất.

Điều hòa ở operon *lac* và operon *trp* là ví dụ về một trong số các cơ chế quan trọng điều hòa hoạt động gene ở mức phiên mã của prokaryote.

### III. Điều hòa biểu hiện gene ở eukaryote

Điều hòa hoạt động gene ở eukaryote phức tạp hơn nhiều so với prokaryote. Liên quan đến điều hòa, giữa prokaryote và eukaryote có một số điểm khác biệt:

- Ở eukaryote thường chỉ có một chuỗi polypeptide đơn được dịch mã từ một phân tử mRNA hoàn chỉnh. mRNA đa cistron (polycistronic) chỉ có ở prokaryote, không có ở eukaryote.

- DNA của eukaryote gắn với protein histone tạo sợi chromatin, và gắn với protein phi histone. Chỉ có một đoạn nhỏ DNA để trần. Ở

prokaryote, một vài protein có mặt trên nhiễm sắc thể bị gấp nếp, còn lại hầu hết DNA trần.

- Một đoạn DNA quan trọng của eukaryote chứa trình tự nucleotide lặp lại trung bình hoặc lặp lại cao. Một vài trình tự lặp lại được xếp nối tiếp thành các bản sao tiếp nối nhau, một vài trình tự khác lại không xếp nối tiếp. Vi khuẩn chứa đoạn DNA lặp lại nhỏ khác các gene của rRNA và tRNA nhân đoạn và một vài yếu tố di động.

Một đoạn lớn DNA của eukaryote không được dịch mã. Hầu hết trình tự nucleotide không được mã hóa thành protein. Các eukaryote đơn bào, chẳng hạn nấm men là trường hợp ngoại lệ đối với tính chất này.

Một vài gene eukaryote được biểu hiện và điều hòa nhờ cơ chế cấu trúc lại những đoạn DNA nhất định theo một phương thức được điều khiển và để tăng số lượng các gene đặc biệt này khi cần thiết.

Các gene eukaryote là gián đoạn được phân thành các exon và intron. Các intron được cắt bỏ trong quá trình chế biến bản phiên mã RNA trước khi bắt đầu dịch mã.

Ở eukaryote, mRNA được tổng hợp trong nhân và được chuyển qua màng nhân, vào tế bào chất mới được sử dụng. Tế bào vi khuẩn không có nhân được tách biệt với tế bào chất.

### *1. Sự biến đổi DNA*

Một số gene của eukaryote được điều hòa bằng sự biến đổi DNA. Chẳng hạn, những trình tự nhất định có thể được khuếch đại hoặc cấu trúc lại trong genome hoặc các base có thể bị biến đổi về mặt hóa học. Một vài biến đổi được phục hồi, những biến đổi khác lại bền vững. Tuy nhiên những thay đổi bền vững này thường xảy ra ở tế bào sinh dưỡng, vì vậy chúng không được truyền lại cho thế hệ sau qua dòng giao tử.

### *2. Các promoter*

Tương tự vi khuẩn, các promoter của eukaryote cũng nằm phía trước điểm xuất phát trên mRNA và có những trình tự được bảo tồn trong tiến hóa. Hộp TATA định hướng cho mRNA bắt đầu phiên mã nằm ở phía trước điểm bắt đầu phiên mã khoảng 30 bp ở động vật có vú và 60 đến 120 bp ở nấm men. Hộp TATA hoạt động có hiệu quả cùng với 2 trình tự tương ứng ở phía trước khoảng 40 bp là CCAAT và 110 bp là trình tự giàu GC. Sự thay đổi hộp TATA làm giảm tốc độ phiên mã. Hiệu quả của tốc độ phiên mã được đo bằng sự thay đổi của từng base trong promoter.

\* Sự cấu trúc lại DNA theo chương trình (programmed DNA rearrangement)

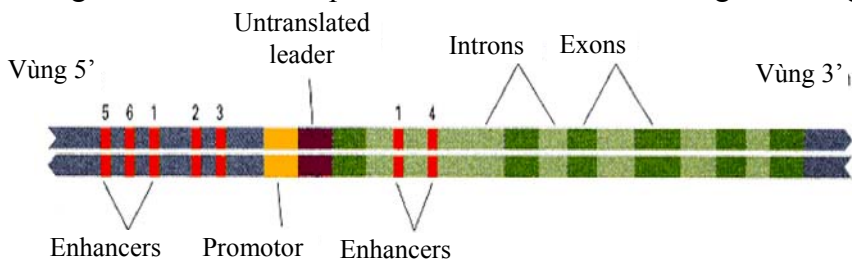
Sự cấu trúc lại trình tự DNA trong genome là cơ chế bất thường

nhưng quan trọng, nhờ đó một số gene được điều hòa.

### 3. Những trình tự tăng cường phiên mã (Enhancer)

Các thụ thể của hormone và những protein hoạt hóa phiên mã khác gắn với trình tự DNA đặc biệt được biết là enhancer. Trình tự enhancer khá ngắn (thường chỉ 20 cặp base) được tìm thấy ở các vị trí khác nhau quanh gene được điều hòa. Hầu hết các enhancer nằm ở phía dưới điểm bắt đầu phiên mã (đôi khi cách xa nhiều kb). Những enhancer khác là các intron nằm trong vùng mã hóa và một vài enhancer thậm chí nằm ở đầu 3' của gene.

Enhancer là thành phần nhạy cảm của tổ chức gene ở eukaryote vì chúng cho phép gene phiên mã chỉ khi nào có nhân tố hoạt hóa phiên mã đúng. Một vài enhancer phản ứng với các phân tử bên ngoài tế bào, chẳng hạn hormone steroid tạo phức hợp receptor-hormone. Những enhancer khác phản ứng với các phân tử được tạo thành ở bên trong tế bào (chẳng hạn trong suốt quá trình phát triển). Những enhancer này cho phép các gene dưới sự điều khiển của nó tham gia vào *biệt hóa tế bào* (differentiation) hoặc được biểu hiện theo cách đặc biệt ở trong mô. Nhiều gene ở dưới sự kiểm soát của các enhancer khác nhau, vì vậy chúng có thể phản ứng với nhiều tín hiệu phân tử khác nhau cả bên trong và bên ngoài.



**Hình 7.8** Sơ đồ tổ chức các gene điển hình ở eukaryote bậc cao

Qua sơ đồ ở (Hình 7.8) cho thấy nhiều yếu tố di truyền được tìm thấy ở trong gene của eukaryote điển hình. Phức hợp phiên mã bám vào promotor bắt đầu tổng hợp mRNA. Vùng mã hóa của gene (exon) bị gián đoạn bởi một hoặc nhiều trình tự gián đoạn (intron), các trình tự này sẽ bị loại bỏ trong quá trình chế biến RNA. Sự phiên mã được điều hòa bởi các yếu tố enhancer, các yếu tố này phản ứng với các phân tử khác nhau có vai trò là tín hiệu cảm ứng. Enhancer có mặt ở cả phía trên và phía dưới promotor. Một vài enhancer có nhiều bản sao.

Nhiều enhancer kích thích phiên mã bằng cách hình thành *vòng DNA* (DNA looping), liên quan đến sự tương tác giữa các vùng cách xa nhau có liên quan dọc sợi DNA. Các nhân tố cần thiết cho phiên mã bao gồm

protein hoạt hóa phiên mã, nó tương tác với ít nhất một yếu tố protein có trong một hoặc nhiều phức hợp protein lớn, nhiều yếu tố. Yếu tố protein trong phức hợp này được biết như là *yếu tố phiên mã chung* (general transcription factors), vì chúng gắn liền với sự phiên mã của nhiều gene khác nhau. Nhân tố phiên mã chung ở eukaryote có tính bảo tồn cao trong tiến hóa. Một trong số các phức hợp này là TFIID, gồm protein gắn với hộp TATA (TATA-box-binding protein) TBP gắn với promotor trong vùng TATA box. Ngoài TBP, phức hợp TFIID cũng gồm một số protein khác được gọi là những *nhân tố gắn liền với TBP* (TBP-associated factors) = TAFs. Chúng hoạt động như các chất trung gian qua đó ảnh hưởng của nhân tố hoạt hóa phiên mã được truyền đi. Sự phiên mã cũng yêu cầu một holoenzyme là RNA polymerase, chứa pol III tổ hợp ít nhất với 9 yếu tố protein khác. Ở nấm men, những yếu tố này chứa nhân tố phiên mã TFIIB, TFIIF và TFIIH cũng như những protein khác.

#### 4. Trình tự bất hoạt gene (*gene silencing*)

Trình tự bất hoạt gene có thể nằm trước hoặc sau gene được điều hòa khoảng vài trăm cặp base. Chúng làm ngừng quá trình phiên mã bằng cách biến đổi histone và DNA.

#### 5. Promoter chọn lọc (*alternative promoter*)

Một vài gene eukaryote có hai hoặc nhiều promoter hoạt động trong những tế bào khác nhau. Những promoter khác nhau này dẫn đến những bản phiên mã khác nhau chứa cùng vùng mã hóa protein. Ví dụ ở *Drosophila*, gene mã hóa cho alcohol dehydrogenase, cấu tạo của nó trong genome có ba vùng mã hóa protein bị gián đoạn bởi hai intron. Sự phiên mã ở ấu trùng sử dụng một promoter khác với sự phiên mã ở ruồi trưởng thành. Bản phiên mã ở ruồi giấm trưởng thành có trình tự dẫn đầu 5' dài hơn. Nhưng hầu hết trình tự này bị loại bỏ trong splicing. Promoter biến đổi làm sự điều hòa phiên mã có thể không phụ thuộc vào ấu trùng hay ruồi giấm trưởng thành.

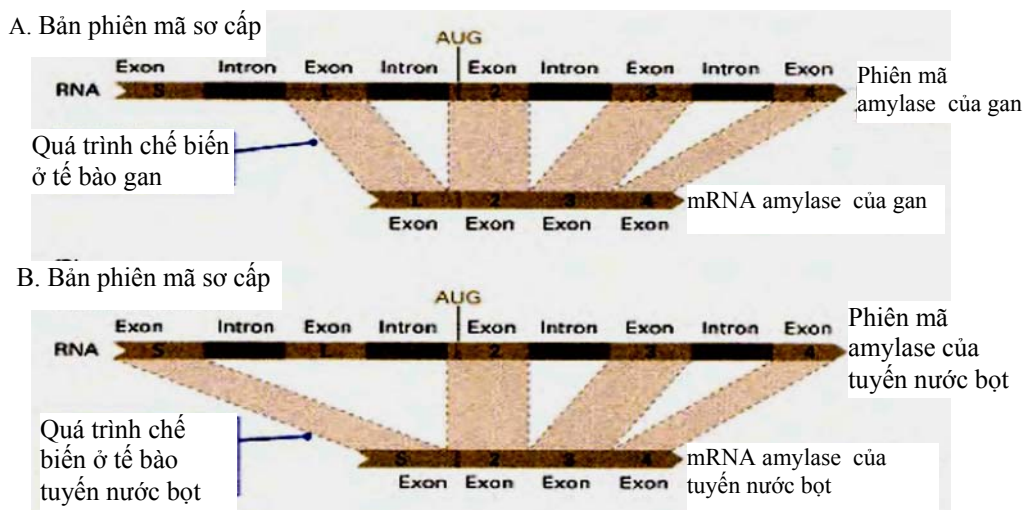
#### 6. Splicing chọn lọc

Cùng promoter được sử dụng để phiên mã một gene, ở các loại tế bào khác nhau có thể tạo sản phẩm có số lượng khác nhau hoặc tạo ra những protein khác nhau. Điều này là do cùng một bản phiên mã có thể có quá trình chế biến ở các loại tế bào là khác nhau.

Sự tổng hợp  $\alpha$ -amylase ở chuột, những phân tử mRNA khác nhau được tạo ra từ cùng một gene vì những phần intron khác nhau bị loại bỏ trong quá trình chế biến RNA. Tuyến nước bọt của chuột tạo nhiều enzyme hơn gan mặc dù cùng một trình tự mã hóa được phiên mã. Ở mỗi



loại tế bào, cùng một *bản phiên mã sơ cấp* (primary transcript) được tổng hợp, nhưng có hai quá trình chế biến khác nhau (Hình 7.9). Trình tự mã hóa bắt đầu 50 bp bên trong exon 2 được tạo thành nhờ nối với exon 3 và những exon tiếp theo. Ở tuyến nước bọt bản phiên mã sơ cấp được chế biến để exon S nối với exon 2 (exon L bị loại đi như là intron 1 và 2). Ở gan exon L nối với exon 2, exon S bị loại đi cùng với intron 1. Exon S và exon L trở thành những trình tự 5' biến đổi của amylase mRNA và mRNA này được dịch mã ở những tỷ lệ khác nhau.



**Hình 7.9** Sản phẩm các phân tử mRNA amylase khác nhau do quá trình splicing khác nhau xảy ra ở tế bào tuyến nước bọt và tế bào gan của chuột.

## Câu hỏi và Bài tập

1. Sự biểu hiện gene ở prokaryote và eukaryote có đặc điểm gì?
2. Hãy nêu các kiểu điều hòa chủ yếu.
3. Trình bày cơ chế hoạt động của operon lactose.
4. Operon là gì? Nêu ý nghĩa của operon đối với sự biểu hiện của gene.
5. Gene mã hóa cho chất ức chế (repressor) của operon vi khuẩn có cần thiết phải ở gần gene cấu trúc không? Tại sao?
6. So sánh điều hòa âm tính và điều hòa dương tính của phiên mã. Cho ví dụ.
7. Khi có mặt glucose trong tế bào *E. coli*, nồng độ cAMP-CRP như

thế nào? Sự gắn của cAMP-CRP vào DNA có ảnh hưởng đến sự gắn của repressor không?

8. Nêu ý nghĩa của hiện tượng phiên mã dờ (attenuation).
9. Phân biệt giữa repressor và aporepressor và cho ví dụ.
10. Thế nào là splicing chọn lọc? Cho ví dụ minh họa.

## **Tài liệu Tham khảo**

### **Tiếng Việt**

Phạm Thành Hồ (2000). *Di truyền học*. NXB Giáo Dục.

Lê Đình Lương, Phan Cự Nhân (1998). *Cơ sở Di truyền học*. NXB Giáo Dục.

Hoàng Trọng Phán (1995). *Di truyền học Phân tử*. Trung tâm Đào tạo Từ xa, Đại học Huế.

### **Tiếng Anh**

Anthony J. F. Griffiths, Susan R. Wessler, Richard C. Lewontin, William M. Gelbart, David T. Suzuki, Jeffrey H. Miller. 2004. *An introduction to genetics analysis*. W.H. Freeman Publishers.

Harlt D.L., Jones E.W. (1998). *Genetics - Principle and analysis*. Jone and Bartlett Publshers. Toronto, Canada.

Stansfield W.D. 1991. *Schaum's outline of theory and problems of genetics*. McGraw-Hill, Inc., New York.

## Chương 8

# Đột biến gene, Tái tổ hợp và các Yếu tố Di truyền Vận động

### I. Đột biến gene

Đột biến gene là những biến đổi xảy ra bên trong cấu trúc gene. Mỗi đột biến gene dẫn đến sự thay đổi trình tự nucleotide tạo ra các allele khác nhau. Đột biến gene có thể xảy ra do biến đổi của trình tự nucleotide trong gene. Đột biến gene không phát hiện được khi quan sát tế bào học.

Trong tự nhiên, tất cả các gene đều có đột biến được gọi là *đột biến tự nhiên hay ngẫu phát* (spontaneous mutation). Các đột biến tự nhiên thường xuất hiện rất ít.

#### 1. Các kiểu đột biến gene

Đột biến gene hay đột biến điểm: là các biến đổi rất nhỏ trên một đoạn DNA, thường liên quan đến một cặp base đơn của DNA hoặc một số ít cặp base kề nhau. Đột biến điểm làm thay đổi gene kiểu dại (wild-type gene). Thực tế đột biến điểm hầu như làm giảm hoặc làm mất chức năng của gene hơn là làm tăng cường chức năng của gene.

Về nguồn gốc, đột biến điểm được phân ra làm đột biến *ngẫu phát* (spontaneous) và đột biến *cảm ứng* (induced).

Đột biến cảm ứng: là dạng đột biến xuất hiện với tần số đột biến tăng lên khi xử lý có mục đích bằng tác nhân đột biến hoặc tác nhân môi trường đã được biết. Đột biến ngẫu nhiên là đột biến xuất hiện khi không có sự xử lý của tác nhân đột biến. Đột biến ngẫu nhiên được tính là tỉ lệ cơ sở của đột biến và được dùng để ước chừng nguồn biến dị di truyền tự nhiên trong quần thể. Tần số đột biến ngẫu nhiên thấp nằm trong khoảng  $10^{-5}$  -  $10^{-8}$ , vì vậy đột biến cảm ứng là nguồn đột biến quan trọng cho phân tích di truyền.

Tác nhân đột biến được sử dụng phổ biến là nguồn chiếu xạ năng lượng cao (high-energy radiation) hoặc các hóa chất đặc biệt.

Các dạng đột biến điểm: có hai dạng đột biến điểm chính trong phân tử DNA:

- + Đột biến *thay thế cặp base* (base substitution)
- + Đột biến *thêm-bớt cặp base* (base insertion - base deletion)

Các đột biến này có thể phát sinh do ảnh hưởng của môi trường như ảnh hưởng của các tác nhân gây đột biến.

### 1.1. Đột biến thay thế cặp base

Kiểu đột biến đơn giản nhất là thay thế một base, trong đó một cặp nucleotide trong gene được thay thế bằng một cặp nucleotide khác.

Ví dụ: A được thay thế bằng G trong sợi DNA. Sự thay thế này tạo ra sự cặp base G-T. Ở lần sao chép tiếp theo tạo ra cặp G-C trong một phân tử DNA con và cặp A-T ở phân tử DNA con kia.

Tương tự, đột biến thay thế A bằng T trên một sợi, tạo ra sự kết cặp tạm thời T-T. Kết quả sao chép tạo ra T-A trên một phân tử DNA con và A-T trên phân tử DNA con kia. Trong trường hợp này, cặp base T-A là đột biến và cặp A-T không đột biến. Nếu sợi gốc DNA không đột biến có trình tự 5'-GAC-3', trên sợi đột biến có trình tự 5'-GTC-3' và sợi kia không đột biến có trình tự 5'-GAC-3'.

Đột biến thay thế cặp base được chia làm hai loại:

+ *Đột biến đồng hoán* (transition mutations): Nếu một đột biến mà base pyrimidine được thay thế bằng một pyrimidine và một purine thay bằng một purine.

Đột biến đồng hoán có thể là:

$T \rightarrow C$  hoặc  $C \rightarrow T$

(Pyrimidine  $\rightarrow$  pyrimidine)

$A \rightarrow G$  hoặc  $G \rightarrow A$

(purine  $\rightarrow$  purine)

Đột biến *đảo hoán* (Transversion): Đột biến làm thay một pyrimidine thành một purine hay một purine được thay thế bằng một pyrimidine. Các đột biến đảo hoán:

$T \rightarrow A$ ,  $T \rightarrow G$ ,  $C \rightarrow A$  hoặc  $C \rightarrow G$

(Pyrimidine  $\rightarrow$  purine)

$A \rightarrow T$ ,  $A \rightarrow C$ ,  $G \rightarrow T$  hoặc  $G \rightarrow C$

(Purine  $\rightarrow$  pyrimidine)

Như vậy có thể có 4 thay thế kiểu đột biến đồng hoán và có đến 8 thay thế kiểu đột biến đảo hoán. Nếu các thay thế này xảy ra với ngẫu nhiên xác suất như nhau, sẽ có tỷ lệ đột biến: 1 đồng hoán : 2 đảo hoán. Tuy nhiên trong thực tế, đột biến thay thế base có xu hướng nghiêng về đột biến đồng hoán, cho nên trong số các đột biến thay thế base tự phát thì tỷ lệ xảy ra đột biến là: 2 đồng hoán : 1 đảo hoán

1.2. Đột biến *thêm hoặc bớt base* (base-pair addition/deletion), còn gọi là indel mutation (insertion-deletion). Trường hợp đơn giản nhất của đột biến

này là thêm hoặc mất một cặp base đơn. Đôi khi đột biến làm thêm hoặc mất đồng thời nhiều cặp base.

Hậu quả của đột biến điểm đến cấu trúc và sự biểu hiện của gene

Đột biến điểm xuất hiện trong vùng mã hóa chuỗi polypeptide của gene (a polypeptide-coding part of a gene), chẳng hạn đột biến thay thế base đơn có thể gây nhiều hậu quả, nhưng tất cả đều có tác động lên mã di truyền theo 2 hướng: làm thoái hóa mã di truyền hoặc xuất hiện mã kết thúc quá trình dịch mã. Có các dạng:

*Đột biến đồng nghĩa* (synonymous mutations): đột biến thay đổi một codon mã hóa acid amine thành codon mới mã hóa cho cùng acid amin đó. Đột biến đồng nghĩa cũng có thể xem là đột biến im lặng (silent mutations)

*Đột biến nhầm nghĩa* (missense mutations), đôi khi còn gọi là đột biến không đồng nghĩa (nonsynonymous mutations): codon mã hóa cho một acid amin này bị thay đổi thành codon mã hóa cho một acid amin khác.

*Đột biến vô nghĩa* (nonsense mutations): codon mã hóa cho một acid amin bị thay đổi thành codon kết thúc dịch mã (translation termination/stop codon).

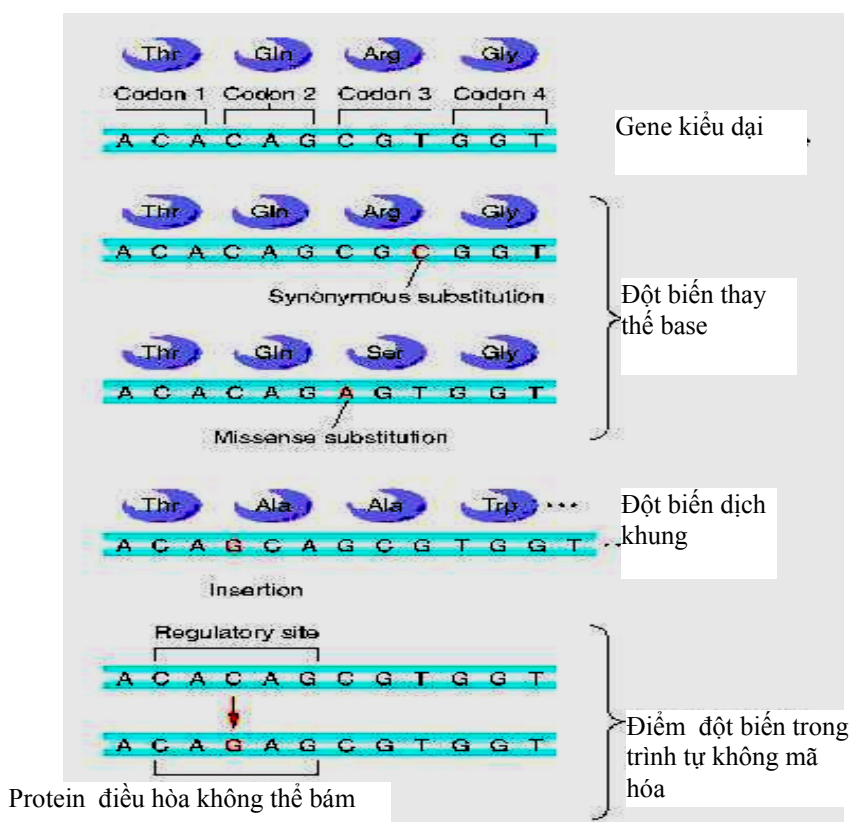
Mức độ ảnh hưởng của đột biến nhầm nghĩa và vô nghĩa lên chuỗi polypeptide khác nhau tùy trường hợp.

Nếu đột biến nhầm nghĩa thay thế một acid amin này bằng một acid amin khác tương tự về mặt hóa học, được xem là *đột biến thay thế bảo thủ* (conservative substitution). Sự thay đổi này hầu như ít ảnh hưởng đến cấu trúc và chức năng protein. Ngược lại, nếu thay thế bằng một acid amin khác về phương diện hóa học gọi là nonconservative substitution, hầu hết đều gây ra sự thay đổi lớn ở cấu trúc và chức năng protein.

Đột biến vô nghĩa sẽ dẫn đến sự kết thúc dịch mã sớm. Vì vậy chúng gây ra hậu quả tương ứng trên chức năng protein. Nếu đột biến vô nghĩa xảy ra càng ở gần đầu 3' của khung đọc mã, kết quả ít ảnh hưởng đến protein. Tuy nhiên nhiều đột biến vô nghĩa ở vùng này vẫn tạo ra các sản phẩm hoàn toàn bị mất hoạt tính.

**Bảng 8.1 Đột biến điểm ở mức độ phân tử**

Kiểu đột biến	Kết quả và ví dụ
<ul style="list-style-type: none"> <li>Ở mức độ DNA</li> </ul> Đồng hoán	Purine được thay thế bằng một purine khác, pyrimidine được thay thế bằng một pyrimidine khác: AT → GC, GC → AT, CG → TA, TA → CG
Đảo hoán	Purine được thay thế bằng một pyrimidine hoặc một pyrimidine được thay thế bằng một purine: AT → CG, AT → TA, GC → TA, GC → CG TA → GC, TA → AT, CG → AT, CG → GC
Thêm/Mất (Insertion-deletion)	Thêm vào hoặc mất đi một hoặc một số cặp base của DNA (thêm/mất base được gạch dưới) AAGACTCCT → AAGAG <u>C</u> TCCT AAGACTCCT → AA <u>A</u> CTCCT
<ul style="list-style-type: none"> <li>Ở mức độ protein</li> </ul> Đột biến đồng nghĩa	Codon đặc biệt mã hoá cho cùng một acid amin: AGG → CGG Arg Arg
Đột biến nhầm nghĩa Loại bảo thủ	Codon tạo thành mã hoá cho amino acid khác Mã hoá cho acid amin có cùng bản chất hoá học: AAA → AGA Lys Arg (kiềm) (kiềm)
Loại không bảo thủ	Mã hoá cho amino acid khác về bản chất hoá học: UUU → UCU Phenylalanin Serine Kỵ nước Phân cực
Đột biến vô nghĩa	Codon kết thúc chuỗi: CAG → UAG Gln Stop
Đột biến dịch khung	Thêm vào một cặp base: AAG ACT CCT → AAG <u>AGC</u> TCC T... Mất một cặp base: AAG <u>ACT</u> CCT → AAA CTC CT...



**Hình 8.1** Hậu quả của đột biến điểm trong gene. Codon 1-4 nằm trong vùng mã hóa của gene.

Giống với đột biến vô nghĩa, đột biến thêm bớt base gây hậu quả trên trình tự polypeptide kể từ điểm bị đột biến (Hình 8.1). Trình tự trên mRNA được đọc theo từng khung gồm ba base (codon) một lúc. Mất hoặc thêm base sẽ làm thay đổi khung đọc trong quá trình dịch mã từ điểm bị đột biến cho đến kết thúc theo khung mới. Vì vậy loại đột biến này được gọi là *đột biến dịch khung* (frameshift mutations). Đột biến này tạo ra trình tự acid amin kể từ điểm bị đột biến cho đến kết thúc khác với trình tự acid amin gốc. Đột biến dịch khung gây ra sự mất hoàn toàn cấu trúc và chức năng của protein bình thường.

Trường hợp đột biến xảy ra ở trình tự điều hòa và các trình tự không mã hóa khác (Hình 8.1). Những phần đó của gene không trực tiếp mã hóa cho protein mà chứa nhiều điểm bám DNA chủ yếu cho protein xen vào, đó là những trình tự không nhạy cảm cho sự biểu hiện của gene hoặc cho hoạt tính của gene.

Ở mức độ DNA, những điểm mất đi (docking) gồm những điểm mà RNA polymerase và những nhân tố gắn kết của nó bám vào, cũng như những điểm mà protein điều hòa phiên mã đặc trưng gắn vào. Ở mức độ RNA, những docking quan trọng thêm vào gồm *điểm bám của ribosom* (ribosome-binding site) trên mRNA vi khuẩn, những điểm nối đầu 5' và 3' để gắn các exon ở eukaryote và các điểm có vai trò cho điều hòa dịch mã và định vị mRNA đến vùng đặc biệt trong tế bào. Nhìn chung hậu quả chức năng của bất kì đột biến điểm nào ở vùng như thế đều phụ thuộc vào việc làm gián đoạn (hoặc tạo ra) một điểm bám. Đột biến làm gián đoạn ở những điểm đó có khả năng làm thay đổi phần biểu hiện của gene dựa vào sự thay đổi số lượng sản phẩm được biểu hiện ở một thời điểm nhất định hoặc ở một mô nhất định. hay bằng sự thay đổi phản ứng với những tín hiệu (cue) của môi trường nhất định. Ngược lại, đột biến ở một vài điểm bám có thể hoàn toàn phá hủy một giai đoạn cần cho sự biểu hiện bình thường của gene, như điểm bám của mRNA polymerase hoặc là nhân tố splicing. Vì vậy nó làm bất hoạt sản phẩm của gene hoặc ngăn cản sự hình thành sản phẩm.

Cần phân biệt giữa những thay đổi xảy ra của một đột biến gene đó là sự thay đổi trình tự DNA của gene với sự thay đổi ở mức độ kiểu hình. Nhiều đột biến điểm trong trình tự không mã hóa làm ít thay đổi hoặc không thay đổi trên kiểu hình như đột biến giữa điểm bám DNA cho protein điều hòa hoặc thay đổi những điểm khác trong gene làm thay đổi chức năng của chúng.

## 2. Các tác nhân gây đột biến gene

### 2.1. Các tác nhân gây đột biến vật lý

- Phóng xạ ion hóa: các tia phóng xạ  $\alpha$ ,  $\beta$ ,  $\gamma$  hoặc tia X, các chùm tia neutron hoặc proton làm bắn các điện tử gây ion hóa và biến đổi DNA.

- Phóng xạ không ion hóa: tia tử ngoại (UV) hoặc nhiệt làm tăng mức năng lượng của nguyên tử. Tia tử ngoại thường tạo dimer thymine gắn hai thymine kề nhau của một mạch.

### 2.2. Các tác nhân gây đột biến hóa học

- Tạo sai hỏng khi sao chép: các chất đồng đẳng của base (bromouracil, 2-amino purine...) gắn vào sai khi sao chép có chúng. Các chất màu acridine có thể gắn vào giữa các base của mạch xoắn kép làm mất hay thêm vào một base.

- Tác động trực tiếp lên gene khi không có các sao chép có thể gây đột biến điểm hay các biến đổi A-T  $\rightarrow$  G-C hoặc G-C  $\rightarrow$  A-T.

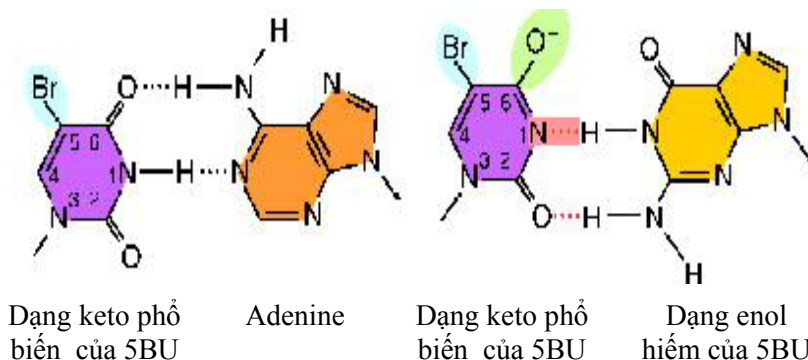
## 3. Cơ chế phân tử của các đột biến gene



Khi kiểm tra dãy đột biến được gây tạo bởi các tác nhân đột biến khác nhau cho thấy mỗi tác nhân đột biến được đặc trưng bởi một đặc tính đột biến khác nhau hay "preference" về cả một dạng đột biến nhất định và một điểm đột biến nhất định, được gọi là *điểm dễ xảy ra đột biến* (mutational hot spots). Đặc tính đột biến như thế được chú ý lần đầu tiên ở locus rII của bacteriophage T4.

Tác nhân đột biến hoạt động ít nhất qua ba cơ chế khác nhau: chúng có thể làm thay thế một base trong DNA; làm biến đổi một base gây kết cặp nhầm với một base khác; làm sai hỏng một base, dẫn đến không thể kết cặp với bất kỳ base nào trong điều kiện bình thường.

Đột biến thay thế base: một vài hợp chất hóa học tương tự nitrogen base bình thường của DNA, đôi khi chúng có thể gắn vào DNA thay cho base bình thường. Những chất như thế được gọi là các *chất tương đương với base* (base analogs). Các chất tương đương này kết cặp không như sự kết cặp của các base bình thường. Vì vậy chúng có thể gây ra đột biến do gắn vào một nucleotide không đúng trong quá trình sao chép.



**Hình 8.2** Chứng minh một vài kết cặp nhầm có thể xảy ra do kết quả của sự thay đổi một tautomer thành một tautomer khác.

Để hiểu hoạt động của các chất tương đương base, trước hết cần phải xem xét khuynh hướng tự nhiên của các base đối với sự hình thành các dạng khác nhau. Mỗi base trong phân tử DNA có thể xuất hiện ở một trong số nhiều dạng được gọi là tautomers, chúng là các đồng phân khác nhau ở vị trí vị trí nguyên tử và những liên kết giữa các nguyên tử. Dạng keto của mỗi base thường có trong DNA (Hình 8.2), trong khi dạng imino và enol của base là hiếm. Tautomer imino hoặc enol có thể kết cặp sai với base tạo một *kết cặp nhầm* (mispair). Khả năng kết cặp nhầm như thế gây ra đột biến trong quá trình sao chép được chú ý đầu tiên bởi Watson và Crick khi các tác giả này nghiên cứu công thức về mô hình cấu trúc DNA.

Sự kết cặp nhầm có thể sinh ra ngẫu nhiên, nhưng cũng có thể sinh ra khi base bị ion hóa. Tác nhân gây đột biến 5-bromouracil (5-BrU) là chất tương đương với thymine, có brome ở vị trí carbon số 5 thay cho nhóm -CH<sub>3</sub> của thymine. Hoạt tính của nó dựa trên quá trình inolization và ionization. Ở dạng keto, 5-BrU kết cặp với adenin như trường hợp thymine. Tuy nhiên, sự có mặt của nguyên tử bromine làm thay đổi một cách có ý nghĩa sự phân bố electron ở vòng base. Vì vậy 5-BrU có thể chuyển sang dạng enol và dạng ion, và nó có thể kết cặp với guanine như trường hợp cytosine tạo ra cặp 5-BrU-G. Trong lần nhân đôi tiếp theo G kết cặp với C, tạo cặp G-C thay cho cặp A-T. Kết quả gây ra đột biến đồng hoán. Tương tự 5-BrU cũng có thể gây ra đột biến đồng hoán A-T thay cho cặp G-C.

Một hóa chất gây đột biến khác là 2-amino-purine (2-AP), là hóa chất tương đương adenine, có thể kết cặp với thymine. Khi bị proton hóa, 2-AP có thể kết cặp nhầm với cytosine, có thể gây ra thể hệ sau đột biến đồng hoán G-C thay cho A-T do kết cặp nhầm với cytosine trong lần sao chép tiếp theo.

- Thay thế base (base alteration)

Một vài tác nhân đột biến không gắn vào DNA, mà lại làm biến đổi base gây ra sự kết cặp sai. Tác nhân alkyl được sử dụng phổ biến như là tác nhân đột biến, chẳng hạn như ethylmethanesulfonate (EMS) và nitrosoguanidine (NG) gây đột biến theo cách này.

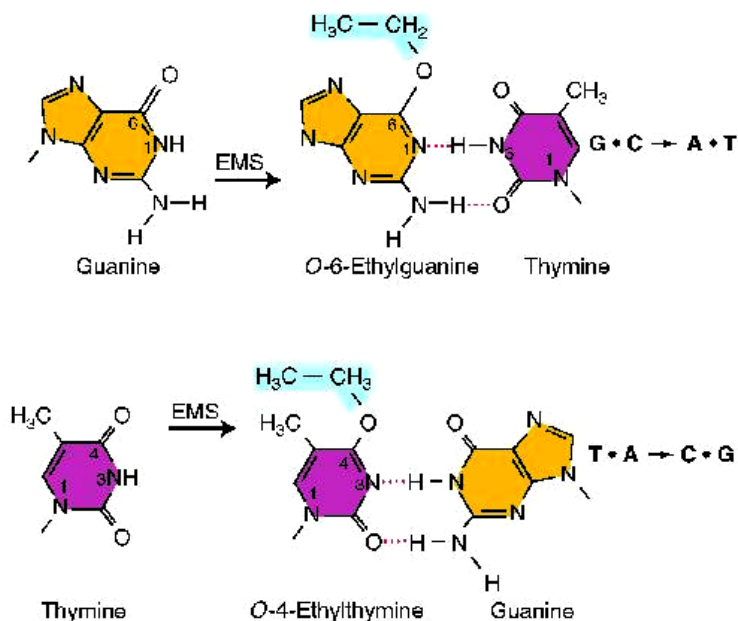
Những tác nhân như thế sẽ thêm nhóm alkyl (nhóm ethyl trong trường hợp EMS và nhóm methyl trong trường hợp NG) ở nhiều vị trí trên cả 4 base. Tuy nhiên, đột biến hầu như chỉ xảy ra khi nhóm alkyl được thêm vào ở oxy số 6 của guanine tạo ra O-6-alkylguanine. Sự alkyl hóa này dẫn đến sự kết cặp nhầm với thymine (Hình 8.3). Kết quả sinh ra đột biến đồng hoán G-C→A-T trong lần sao chép tiếp theo.

*Tác nhân xen vào giữa* (intercalating agents) là nhóm tác nhân quan trọng khác gây biến đổi DNA. Nhóm của các hợp chất này bao gồm proflavin, acridin cam và một nhóm các hợp chất hóa học khác. Các tác nhân này là nhóm các phân tử bắt chước các cặp base và có thể xen vào giữa các base nitơ ở lõi chuỗi xoắn kép DNA. Ở vị trí xen vào này chúng gây sự thêm vào hoặc mất đi một cặp nucleotide.

- Sai hỏng base

Một số lớn tác nhân đột biến gây sai hỏng một hoặc nhiều base. Vì vậy không thể kết cặp với base đặc trưng. Kết quả làm cản trở sự sao chép vì DNA polymerase không thể tiếp tục quá trình tổng hợp DNA qua

những base sai hỏng. Ở *E.coli* quá trình này xảy ra đòi hỏi hoạt tính của hệ thống SOS. Hệ thống này được kích thích như là một phản ứng khẩn cấp ngăn cản sự chết tế bào khi DNA bị sai hỏng nặng.



**Hình 8.3 Sự kết cặp nhầm chuyên biệt do đột biến cảm ứng alkyl hoá.**

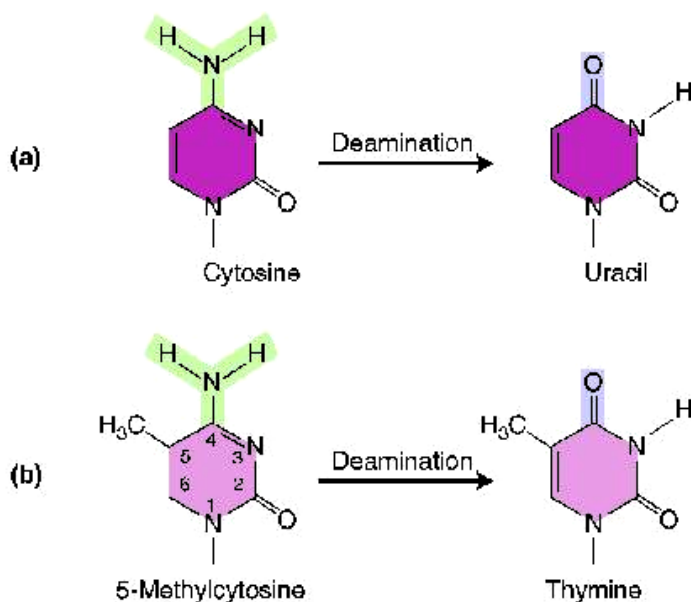
\* Cơ chế của đột biến ngẫu nhiên

Đột biến ngẫu nhiên xảy ra do nhiều nguyên nhân: gồm sai hỏng trong quá trình sao chép DNA, các tổn thương ngẫu nhiên, sự chen vào của yếu tố di động. Đột biến ngẫu nhiên hiếm nên khó xác định cơ chế cơ bản. Tuy nhiên, một vài hệ thống chọn lọc cho phép thu được đột biến ngẫu nhiên và phân tích ở mức độ phân tử. Từ bản chất của những thay đổi trình tự có thể suy ra quá trình dẫn đến đột biến ngẫu nhiên.

Những sai hỏng ngẫu nhiên (spontaneous lesions) đến DNA có thể sinh ra đột biến. Hai tổn thương ngẫu nhiên thường xuất hiện nhất: depurine hoá và deamin hoá, trong đó depurine hoá phổ biến hơn.

Depurination do tác dụng của aflatoxin, làm mất một base purine. Ngoài ra, quá trình mất purine cũng xảy ra một cách tự nhiên. Một tế bào động vật mất ngẫu nhiên khoảng 10.000 purine của DNA trong một thể hệ tế bào khoảng 20 giờ ở 37°C. Nếu tổn thương này được giữ lại, dẫn đến sai hỏng di truyền đáng kể vì trong quá trình sao chép, vị trí mất purine không thể định rõ được loại base nào. Trong những điều kiện nhất định một base có thể chen vào tạo ra đột biến.

Deamination của cytosine tạo ra uracil. Uracil sẽ kết cặp với adenin trong quá trình sao chép, kết quả tạo ra đột biến đồng hoán G-C → A-T. Deamination 5-methylcytosine tạo ra thymine (Hình 8.4). Quá trình sao chép tạo ra đột biến đồng hoán chuyển C thành T.



**Hình 8.4** Deamination của Cytosine (a) và 5-methylcytosine.

Ngoài 2 quá trình gây sai hỏng như trên, sự oxy hóa tạo ra các base bị sai hỏng là dạng tổn thương thứ ba.. Dạng oxygen hoạt động như gốc superoxid ( $O_2^-$ ), hydrogen peroxide ( $H_2O_2$ ) và gốc hydroxyl ( $-OH$ ) được tạo ra do sản phẩm của quá trình chuyển hóa (aerobic metabolism). Các dạng này có thể gây tổn thương oxy hóa đến DNA, kết quả tạo ra đột biến.

Các sai hỏng trong sao chép DNA cũng là nguồn đột biến khác.

- Thay thế base: sai hỏng trong sao chép DNA có thể xảy ra khi có một cặp nucleotide ghép không chính xác (như A-C) tạo ra trong quá trình tổng hợp DNA dẫn đến sự thay thế một base.

- Đột biến thêm vào và mất base: Một loại sai hỏng sao chép khác dẫn đến thêm vào hoặc mất đi một hoặc một số cặp base. Trong trường hợp số base thêm vào hoặc mất đi không chia hết cho 3, sẽ tạo ra đột biến dịch khung trong vùng mã hóa protein.

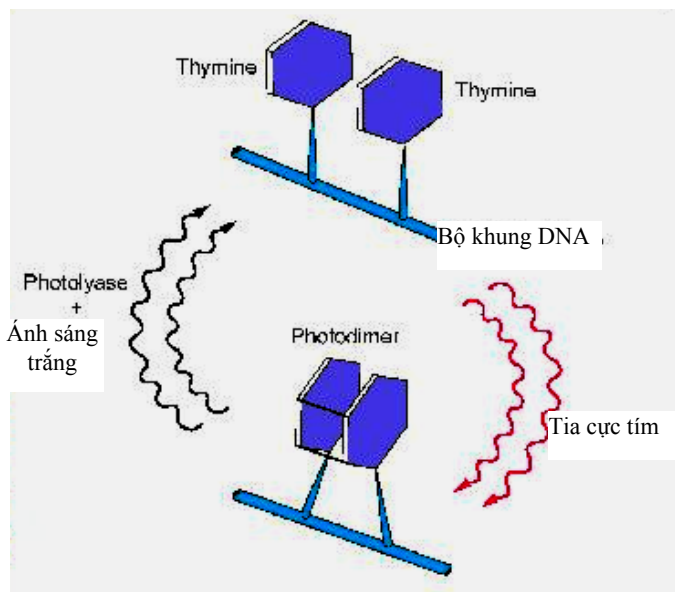
## II. Sửa chữa và bảo vệ DNA

\* Cơ chế sửa sai sinh học

Tế bào sống có hàng loạt hệ thống sai hỏng DNA theo nhiều cách

khác nhau. Tỷ lệ đột biến tự nhiên thấp do nhờ tính hiệu quả của hệ thống sửa sai này. Sai hỏng của hệ thống sửa sai này dẫn đến tỷ lệ đột biến cao.

- *Quang phục hoạt* (photoreactivation) hay *sửa sai nhờ ánh sáng* (light repair)



**Hình 8.5 Sự tạo thành và sự loại bỏ dimer thymine.**

Sau khi xử lý tia tử ngoại gây đột biến, nếu đưa ra ánh sáng thì phần lớn sai hỏng được phục hồi nhờ enzyme photolyase. Enzyme này gắn vào photodimer cắt nó thành các monomer dưới tác dụng của ánh sáng mặt trời có bước sóng 320-370 nm. Sau đó phục hồi các base ban đầu. (Hình 8.5).

- Sửa sai bằng làm mất nhóm alkyl (*dealkylation*)

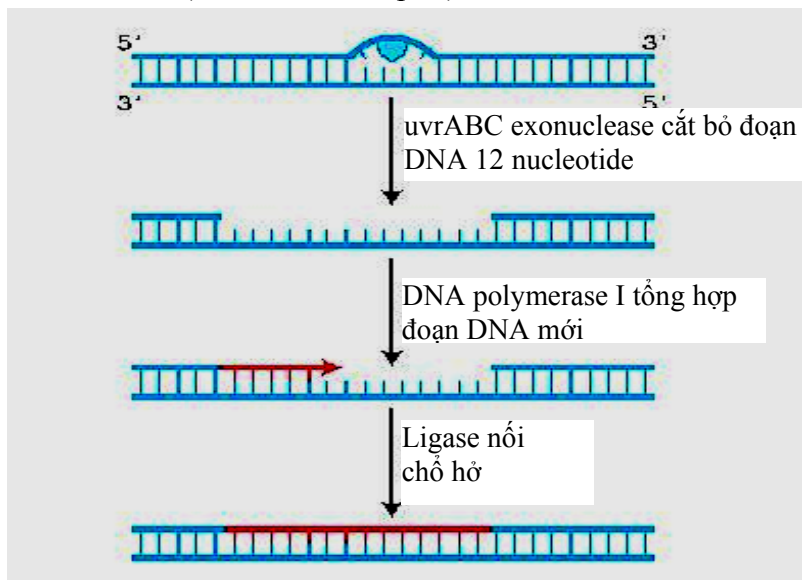
Enzyme alkyltransferases có thể sửa trực tiếp các sai hỏng. Chúng cắt nhóm alkyl từ chất nitrosoguanine và ethylmethanesulfonate và gắn vào vị trí O-6 guanine.

Enzyme methyltransferase của *E. coli* có khả năng chuyển nhóm methyl từ chất O-6 methylguanin sang gốc cistein trên phân tử protein. Tuy nhiên hệ thống sửa sai này có thể bảo hòa nếu mức độ alkyl hóa đủ cao.

- *Sửa sai bằng cắt bỏ* (excision repair pathway)

Phần lớn các cơ chế sửa sai khác thực hiện theo lối cắt bỏ (excision repair) không cần ánh sáng nhờ các nuclease, sau đó thay vào các base đúng. Có thể xảy ra theo nhiều cách:

- *Cắt các base* (base excision repair)



**Hình 8.6** Sửa sai bằng cắt bỏ nucleotide.

Sự cắt bỏ các base sai hỏng nhờ các enzyme DNA glycosylase. Các enzyme này nhận biết các base bị biến đổi và các điểm mất purine hay mất pyrimidine và thủy giải liên kết N-glycosilic nối base với đường. Rồi enzyme AP endonuclease cắt liên kết đường và phosphate gần base bị biến đổi. Sau đó enzyme thứ ba, deoxyribophosphodiesterase loại bỏ từng nucleotide kế tiếp nhau ở đoạn bị hỏng. Sau đó, DNA polymerase lấp đầy khoảng trống với các nucleotide bổ sung với sợi khuôn còn lại. Enzyme DNA ligase sẽ gắn các khe hở giữa 2 đầu 3'-5' (Hình 8.6). Trong tế bào tồn tại một số DNA glycosylase. Chẳng hạn, enzyme uracil-DNA glycosylase cắt uracil khỏi DNA. Uracil tạo thành do đột biến mất nhóm amin ngẫu nhiên ở cytosine, dẫn đến đột biến đồng hoán thay C bằng T. Enzyme này phát hiện ra uracil trên DNA như là một bất thường, chúng sẽ cắt bỏ và sửa sai.

- *Cắt các nucleotide:* Sự cắt bỏ vùng có nhiều pyrimidine dimer được thực hiện nhờ enzyme exonuclease (enzyme rạch mạch hay enzyme tạo khác trên DNA) như phức hợp 3 enzyme được mã hóa bởi gene uvr ABC của *E. coli*. Phức hợp này cắt đoạn 12 nucleotide trên một mạch: 8 nucleotide từ một đầu bị sai hỏng và 4 nucleotide của đầu còn lại. Khoảng trống của 12 nucleotide này sẽ được lấp đầy nhờ enzyme DNA polymerase I dựa vào mạch đơn bổ sung kia của trình tự DNA gốc. DNA ligase sẽ gắn vào các khe hở.

- Đọc sửa đổi với các base bắt cặp sai

Cơ chế *đọc sửa đổi với các base bắt cặp sai* (proofreading for base-pair matching) được thực hiện trong sao chép DNA. Trong quá trình sao chép, trước khi thực hiện phản ứng polymer hóa nối các nucleotide, các nucleotide triphosphate mới phải bắt cặp bổ sung với mạch khuôn. Nếu sự bắt cặp sai xảy ra, DNA polymerase sẽ loại bỏ nucleotide bắt cặp sai. Ngay cả trước khi nucleotide mới ráp vào, enzyme dò lại cặp base cuối, nếu chúng không bắt cặp thì sự polymer hóa tiếp theo bị dừng. Cặp nucleotide ở đầu cuối 3' bắt cặp sai sẽ bị loại bỏ nhờ hoạt tính exonuclease $3' \rightarrow 5'$  của DNA polymerase. Khi đã bắt cặp đúng, quá trình polymer hóa mới được tiếp tục.

Hoạt tính đọc sửa đổi với các base bắt cặp sai là đặc tính của nhiều DNA polymerase đảm bảo cho sự kéo dài chính xác của mạch đang được tổng hợp.

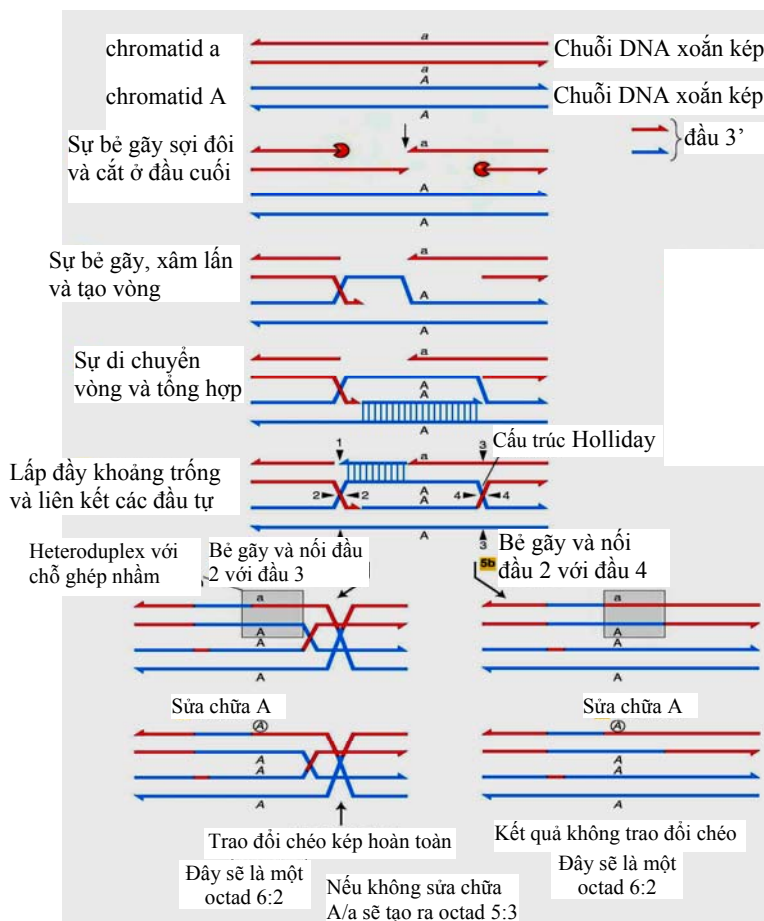
+ *Sửa sai dựa vào tính tương đồng* (Homology-dependent repair system)

Một hệ thống sửa sai quan trọng đã phát hiện tính chất bổ sung đối song song của 2 mạch đơn DNA để phục hồi đoạn sai hỏng trở lại trạng thái bình thường ban đầu. Trong hệ thống này, đoạn DNA sai hỏng bị cắt bỏ và thay bằng một đoạn nucleotide mới được tổng hợp bổ sung với sợi khuôn đối diện. Sự sửa sai xảy ra qua sợi khuôn và nguyên tắc của sao chép DNA bảo đảm sự sửa sai hoàn thành với độ chính xác cao - đó là sự *giải phóng sai hỏng* (error-free). Có 2 hệ thống chủ yếu để loại bỏ sai hỏng: Hệ thống sửa chữa sai hỏng phát hiện ra trước khi sao chép và hệ thống sửa chữa sai hỏng phát hiện trong quá trình diễn biến sao chép (sửa sai sau sao chép).

+ *Sửa sai đứt mạch kép* (repair of double-strand break)

Khi cả 2 sợi của chỗi xoắn kép bị đứt ở cùng một vị trí, được gọi là đột biến đứt mạch đôi, có thể gây ra sai hình nhiễm sắc thể, làm chết tế bào hoặc tạo ra trạng thái tiến ung thư. Tế bào sử dụng nhiều protein và con đường sửa sai đứt gãy mạch đôi là thực hiện tái tổ hợp trong giảm phân. Quá trình sửa chữa do trao đổi chéo trong giảm phân xảy ra như sau (Hình 8.7):

Trên một nhiễm sắc thể xảy ra sự đứt mạch đôi và kết quả ăn mòn các đầu mút ở đoạn ngắn của DNA sợi đơn. Đầu 3' của một trong những sợi này "xâm lấn" vào một chromatid.



**Hình 8.7** Mô hình bẻ gãy sợi đôi nhờ trao đổi chéo.

Đoạn xâm lấn làm môi cho tổng hợp các base bị mất của nó nhờ sử dụng sợi đối song song của chromatid như là sợi khuôn. Sự tổng hợp mới này sẽ tạo ra một vòng sợi đơn lai với một sợi đơn không xâm lấn. Vì vậy tạo ra một vùng dị hợp tử nhỏ "Aa" và sử dụng như mạch khuôn để khôi phục các base bị mất trên sợi đó. DNA polymerase sẽ lấp đầy chỗ trống và enzyme ligase sẽ nối các đầu mút xảy ra trong cấu trúc đặc biệt giống với trao đổi chéo 2 sợi đơn. Cấu trúc này cũng chứa các đoạn bắt cặp không tương đồng đơn giản.

Trao đổi chéo sợi đơn được gọi là *cấu trúc Holliday* (Holliday structure) do Holliday phát hiện vào những năm 1960.

- Hệ thống SOS



Ở tế bào vi khuẩn hoặc tế bào eukaryote bị sai hỏng nặng do chiếu tia uv, tia X hoặc do tác dụng của các hóa chất gây đột biến, hệ thống sửa sai khẩn cấp được khởi động. Ở *E. coli*, hệ thống này có liên quan với hai protein được mã hóa bởi gene *lexA* và *recA*. Protein *lexA* là một chất ức chế, nó gắn vào hộp SOS, chồng lấp các promotor của các gene SOS, ngăn cản sự phiên mã nhóm các gene của hệ thống SOS. Một vài sản phẩm của DNA bị tổn thương sẽ làm hoạt hóa enzyme protease *recA*. Protein *recA* bị hoạt hóa sẽ cắt bỏ protein *lexA*, cho phép các gene của hệ thống SOS phiên mã. Phản ứng của hệ thống SOS xảy ra trong thời gian ngắn nhưng phức tạp. Nó bao gồm các quá trình làm tăng hoạt tính tái tổ hợp, thay đổi trong khởi sự sao chép, ức chế nuclease và kích thích phục hồi sao chép và chuyển sai hỏng thành sửa sai úp sấp (error-prone replication). Tế bào bây giờ sẽ xảy ra sự sao chép DNA nhanh hơn bình thường. Nếu sửa sai không kịp, tế bào phải chấp nhận hoặc bị đột biến hoặc bị chết.

### III. Các yếu tố di truyền vận động (Transposable genetic elements)

#### 1. Các yếu tố di truyền vận động ở prokaryote

*Trình tự đoạn xen* (insertion sequence) của vi khuẩn

Trình tự xen đoạn là một đoạn DNA của vi khuẩn di chuyển từ một vị trí trên nhiễm sắc thể đến vị trí mới trên cùng nhiễm sắc thể hoặc trên nhiễm sắc thể khác. Khi xen vào giữa gene, yếu tố IS làm gián đoạn trình tự mã hóa và làm bất hoạt sự biểu hiện của gene. Một số trường hợp, có tín hiệu kết thúc phiên mã và dịch mã, yếu tố IS làm cản trở sự biểu hiện ở sau promotor trong cùng operon..

Yếu tố IS được tìm thấy đầu tiên ở operon *gal* của *E. coli*, chia làm bốn nhóm: IS1, IS2, IS3 và IS4. Chúng có thể phân bố rải rác trên nhiễm sắc thể chính của vi khuẩn và trên các plasmid. Ví dụ yếu tố IS1 có khoảng 5-8 bản sao trên nhiễm sắc thể với chiều dài 768 bp.

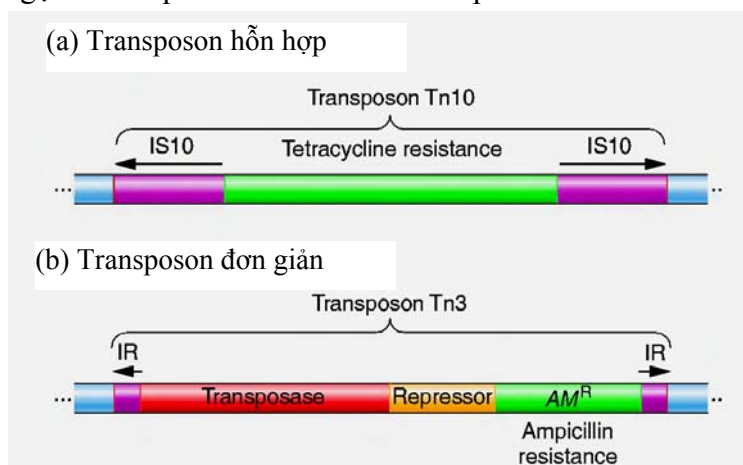
Tất cả các yếu tố IS đều chứa đoạn DNA mã hóa cho protein, được gọi là transposase, là enzyme cần thiết cho sự di chuyển của yếu tố IS từ một vị trí trên nhiễm sắc thể đến vị trí khác. Đoạn gene này nằm giữa 2 *đoạn lặp lại đảo ngược* (inverted repeat - IR) ngắn. Ví dụ yếu tố IS1 có khoảng 5-8 bản sao trên nhiễm sắc thể với chiều dài 768 bp, đoạn lặp lại đảo ngược có kích thước 18-23 bp, yếu tố IS2 có 5 bản sao và các yếu tố IS khác. Yếu tố IS là những vùng của các trình tự xác định, chúng là những vị trí xảy ra trao đổi chéo. Ví dụ: Sự tái tổ hợp của plasmid và nhiễm sắc thể của *E. coli* tạo ra những chủng Hfr xảy ra qua trao đổi chéo đon giữa yếu tố IS trên plasmid và yếu tố IS trên nhiễm sắc thể.

**Bảng 8.2. Một vài trình tự xen vào và kích thước của chúng**

Trình tự xen vào	Số bản sao trong <i>E. coli</i>	Chiều dài (bp)	Đoạn lặp lại đảo ngược (bp)
IS1	5-8 bản sao trên nhiễm sắc thể	768	18-23
IS2	5 bản sao trên nhiễm sắc thể, 1 bản sao trên plasmid	1327	32-41
IS3	5 bản sao trên nhiễm sắc thể, 2 bản sao trên plasmid	1400	32-38
IS4	1-2 bản sao trên nhiễm sắc thể,	1400	16-18
IS5	Chưa biết	1250	Ngắn

### 1.1. Gene nhảy của prokaryote

Các yếu tố IS riêng lẻ không chỉ có khả năng tự di chuyển mà khi hai yếu tố này nằm đủ gần nhau thì chúng có thể vận động như một đơn vị hoàn chỉnh và mang theo các gene nằm giữa chúng. Cấu trúc phức tạp này được gọi là transposon. Có hai kiểu transposon ở vi khuẩn:



**Hình 8.8** Đặc trưng về cấu trúc của *transposon hỗn hợp* (composite transposon) và *transposon đơn giản* (simple transposon).

*Transposon hỗn hợp* (composite transposon) chứa nhiều gene nằm giữa 2 trình tự IS gần nhau, có hướng ngược nhau tạo ra trình tự lặp lại đảo ngược (inverted repeat = IR). Một trong 2 yếu tố IS mã hóa cho transposase xúc tác cho sự chuyển vị của cả transposon. Chẳng hạn Tn10 là transposon hỗn hợp mang gene mã hóa cho tính kháng kháng sinh tetracycline. Gene này nằm giữa hai yếu tố IS10 có hướng ngược nhau.

*Transposon đơn giản* (simple transposon) ở giữa các trình tự IR, nhưng những trình tự này ngắn (<50 bp) và không mã hóa cho transposase. Sự chuyển vị của chúng không phải là kết quả của sự liên kết với yếu tố IS. Các transposon đơn giản mã hóa transposase riêng thêm vào để mang các gene của vi khuẩn. Tn3 là một transposon đơn giản (Hình 8.8).

Transposon hỗn hợp và transposon đơn giản đều chứa các gene thêm vào liên quan đến chức năng mới ở tế bào vi khuẩn. Cả hai loại này thường được gọi chung là transposon. Transposon dài hơn yếu tố IS, thường chứa vài kb), chúng chứa các gene mã hóa cho protein thêm vào.

## 1.2. Cơ chế của sự chuyển vị

Đầu tiên, transposase cắt vết hình chữ chi qua 5 cặp base (khác với sự cắt của enzyme restriction endonuclease) ở vị trí *DNA mục tiêu* (target site DNA) (Hình 8.9). Tiếp theo là sự hội nhập của transposon qua trung gian của transposase, transposon xen vào giữa các đầu mút của chữ chi. Đầu lồi ra của sợi đơn được sử dụng như là khuôn để tổng hợp sợi bổ sung thứ hai. Sự gắn vào tạo sự sao chép 5 cặp base, được gọi là sự *sao chép điểm mục tiêu* (target site duplication).

Hầu hết các yếu tố di động của prokaryote đều sử dụng một trong 2 cơ chế chuyển vị là *sao chép* (replicative) và *bảo thủ* (conservative) hay không sao chép. Trong con đường sao chép (như ở Tn3), một bản sao mới của yếu tố di động tạo ra khi chuyển vị, kết quả là một bản sao ở vị trí mới và bản sao còn lại ở vị trí cũ. Trong con đường bảo thủ (như ở trường hợp Tn10) không có sự sao chép. Thay vào đó, yếu tố được cắt ra từ nhiễm sắc thể hoặc plasmid và được gắn vào vị trí mới. Con đường này còn được gọi là con đường "cắt và dán" (cut and paste)

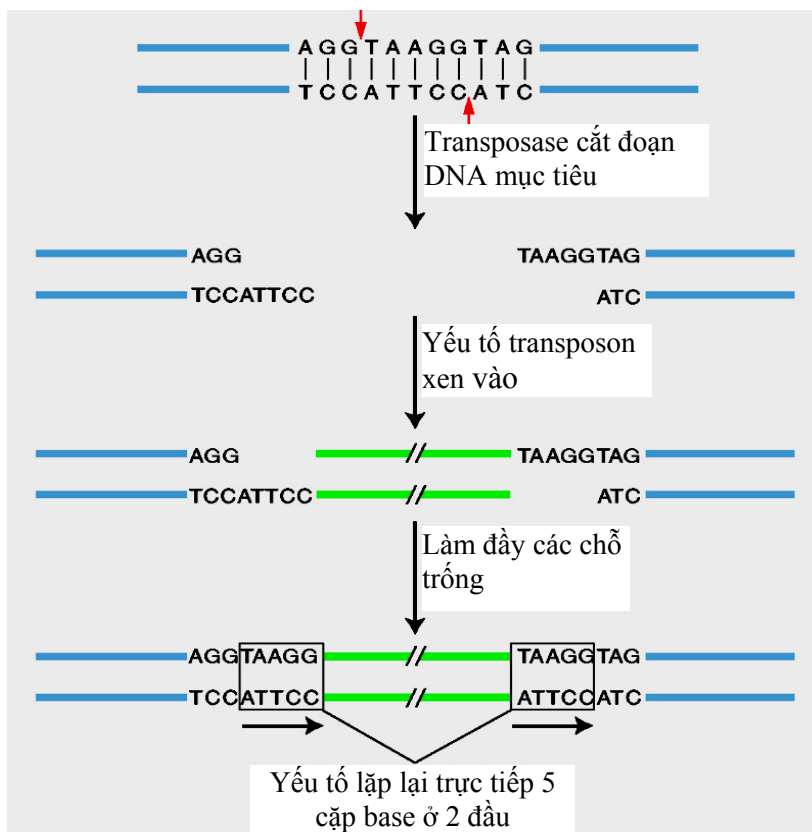
## 2. Các yếu tố di truyền vận động ở eukaryote

Các yếu tố di truyền vận động ở eukaryote chia làm 2 nhóm: nhóm 1 là các retrotransposon và nhóm 2 là các DNA transposon.

### 2.1. Các retrotransposon

Gery Fink và cs. là những người đầu tiên sử dụng nấm men để nghiên cứu điều hoà hoạt tính gene ở eukaryote. Các tác giả này đã phân lập được hàng ngàn đột biến gene HIS4 mã hóa enzyme tham gia tổng hợp histidine. Trong số hơn 1.500 đột biến ngẫu nhiên HIS4 được tìm thấy có 2 đột biến có kiểu hình không bền vững. Các đột biến không bền vững này có tần số phục hồi lại dạng kiểu dại cao 1.000 lần hơn các đột biến HIS4 khác. Những đột biến này cho đoạn DNA lớn xen vào gene HIS4, sự xen vào này được thực hiện do một trong các yếu tố là Ty của nấm men. Có 35

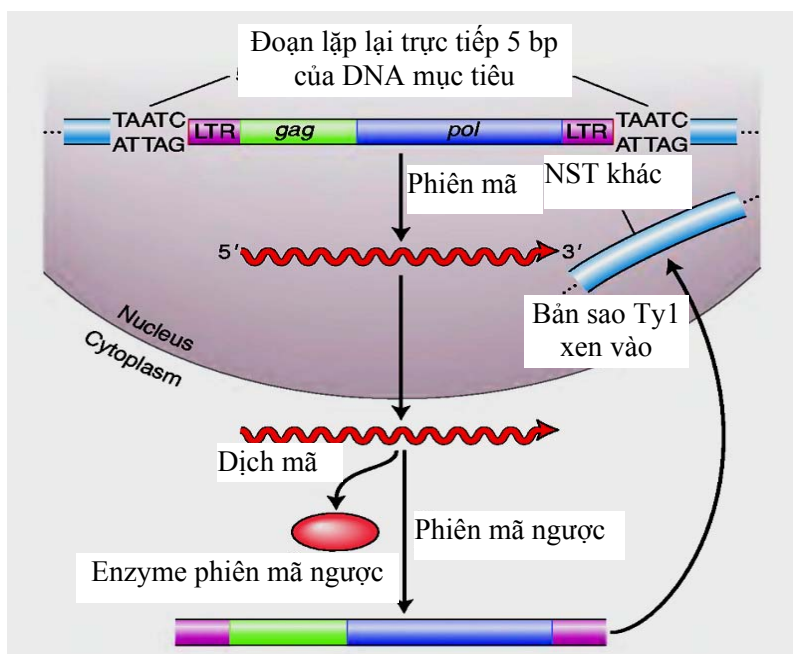
bản sao của yếu tố xen đoạn gọi là Ty1 ở genome của nấm men.



**Hình 8.9** Sự nhân đôi đoạn trình tự DNA ngắn ở điểm xen vào.

Việc tạo dòng những yếu tố này từ các allele đột biến cho thấy xen đoạn này không giống với yếu tố IS hoặc transposon của vi khuẩn. Thay vào đó chúng có đặc tính của retrovirus (virus của động vật). Có sự giống nhau trong cấu trúc và thành phần gene của retrovirus và yếu tố Ty1 được phân lập từ đột biến HIS4. Giống với retrovirus, transposon của nấm men có lặp đoạn cuối dài (long terminal repeat sequence) LTRs, chứa hàng trăm cặp base, được gọi là trình tự  $\delta$  nằm ở 2 phía đoạn mã hóa, cả hai đều chứa gene *gag* và gene *pol*. Retrovirus có ít nhất 3 gene mã hóa cho 3 protein trong quá trình sao chép: gene *gag* mã hóa cho một protein có vai trò làm biến tính RNA genome. Gene *pol* mã hóa enzyme reverse transcriptase. Gene *env* mã hóa cho protein vỏ. Yếu tố Ty chỉ chứa gene *gag* và gene *pol*, không chứa gene *env*. (Hình 8.10).

Mô hình về sự chuyển vị nhờ retrotransposon. Một bản phiên mã RNA từ retrotransposon dưới tác dụng của enzyme phiên mã ngược tạo thành DNA nhờ enzyme reverse transcriptase được mã hóa bởi retrotransposon. Bản sao DNA được chèn vào vị trí mới trên bộ gene.



**Hình 8.10** Sự chuyển vị nhờ retrotransposition.

Vào năm 1985, J. Bocke và G. Fink đã chứng minh, yếu tố Ty1, giống với retrovirus, thực hiện việc di chuyển qua trung gian RNA. Chúng bắt đầu bằng biến đổi yếu tố Ty1 của nấm men được tạo dòng trên plasmid. Trước tiên ở một đầu mút của yếu tố, có sự xen vào một promoter được hoạt hóa nhờ thêm galactose vào môi trường. Thứ hai, một intron từ một gene khác của nấm men được đưa vào vùng mã hóa của transposon Ty. Sự thêm vào galactose làm tăng tần số chuyển vị của yếu tố Ty bị biến đổi. Điều này làm tăng số lượng RNA, vì galactose kích thích phiên mã RNA Ty bắt đầu từ promoter nhạy cảm galactose.

Nhiều đột biến ngẫu nhiên được phân lập ở ruồi dấm cho thấy cũng chứa retrotransposon. Yếu tố copia của ruồi dấm cấu trúc tương tự với yếu tố Ty của nấm men. Chúng xuất hiện từ 10-100 vị trí trên bộ gene của ruồi dấm. Những đột biến nhất định của ruồi dấm là kết quả của sự xen vào của yếu tố copia và những yếu tố khác. Chẳng hạn, đột biến white-apricot ( $w^a$ ) về màu mắt của ruồi dấm tạo ra do sự xen vào yếu tố của họ copia ở locus white.

Sự xen retrotransposon LTR vào các gene ở thực vật như ở ngô, cũng góp phần tạo ra các đột biến ngẫu nhiên.

## 2.2. DNA transposon

Nhiều yếu tố di động được tìm thấy ở eukaryote chuyển vị bằng cơ chế giống với vi khuẩn. Bản thân yếu tố IS và transposon hoặc là bản sao của chúng có thể xen vào vị trí mới trên genome. Các yếu tố di chuyển theo cách này được gọi là yếu tố nhóm 2 hoặc DNA transposon. Yếu tố di động được McClintock phát hiện đầu tiên ở ngô là các DNA transposon. Tuy nhiên tính đặc trưng phân tử của DNA transposon đầu tiên lại là yếu tố p ở ruồi dấm.

Yếu tố p được phát hiện khi nghiên cứu thể lai mất khả năng sinh sản (dysgenesis), hiện tượng xảy ra khi lai ruồi cái thuộc chủng phòng thí nghiệm với ruồi đực ở quần thể tự nhiên. Trong phép lai như thế, chủng phòng thí nghiệm được xem là có kiểu tế bào M (M cytotype) và giống gốc tự nhiên được xem là có kiểu tế bào P (P cytotype). Phép lai của M (cái) x P (đực) cho thế hệ sau kiểu hình ở dòng mầm với gồm dạng bất thụ với tỷ lệ đột biến, tần số sai hình nhiễm sắc thể và không chia ly nhiễm sắc thể cao. Phép lai nghịch không xảy ra hiện tượng *thoái hóa giống* (dysgenics). Đột biến dysgenics không bền, chúng có thể phục hồi dạng kiểu dại hoặc biến đổi thành allele đột biến khác với tần số cao.

Sự giống nhau về đột biến không bền vững ở ruồi dấm và ở ngô đưa đến giả thuyết đột biến dysgenic được gây ra do sự xen yếu tố di động vào những gene đặc biệt, làm chúng bị bất hoạt. Chẳng hạn hầu hết các đột biến trên ở ruồi dấm do xen yếu tố di động vào gene *white*. Những yếu tố như vậy gọi là yếu tố p, có từ 30-50 bản sao trên genome ở chủng P và hoàn toàn vắng mặt ở chủng M. Yếu tố P có chiều dài từ 0,5-2,9 kb. Kích thước khác nhau này do do nhiều yếu tố P không đầy đủ mà bị mất đoạn ở giữa. Yếu tố P với kích thước đầy đủ tương đương với transposon đơn giản ở vi khuẩn, có đầu mút là đoạn lặp đảo ngược ngắn (31 bp) và nó mã hóa cho transposase. Gene mã hóa cho transposase ở eukaryote chứa 3 intron và 4 exon.

Ở ngô, yếu tố Ac và Ds có thể di chuyển trên nhiễm sắc thể, tác động lên các gene kiểm tra màu sắc của các hạt aleuron. Các yếu tố này được phân lập cho thấy có liên quan với DNA transposon của vi khuẩn và của các eukaryote khác. Giống với yếu tố P của ruồi dấm, yếu tố Ac có đoạn cuối lặp lại đảo ngược mã hóa cho transposase, yếu tố Ds lại không mã hóa cho transposase. Khi Ac trên genome, transposase của nó có thể bám vào đầu mút của cả yếu tố Ac và Ds và khởi động sự chuyển vị của chúng..

Yếu tố Ac và Ds là thành viên của họ transposon đơn giản. Ngoài ra, còn có các họ yếu tố di động khác ở ngô. Mỗi họ chứa yếu tố tự động mã hóa cho transposase có thể chuyển được yếu tố trong cùng một họ, nhưng không thể chuyển đến các họ khác vì transposase chỉ có thể bám vào đầu mút của các thành viên trong họ.

Ở người, yếu tố di động chiếm một nửa genome người. Đa số yếu tố di động thuộc 2 dạng retrotransposon là *yếu tố nhân rải rác kích thước dài* (long interspersed nuclear element) LINEs và *yếu tố nhân rải rác kích thước ngắn* (short interspersed nuclear element) SINEs. LINE di chuyển nhờ retrotransposition đã sử dụng yếu tố mã hóa reverse transcriptase, nhưng lại thiếu một vài tính chất về cấu trúc của yếu tố giống retrovirus bao gồm LTR.

SINE được mô tả như là LINE không tự động, vì chúng có tính chất cấu trúc đặc trưng của LINE nhưng không mã hóa cho reverse transcriptase riêng. Người ta cho rằng chúng di chuyển được nhờ enzyme reverse transcriptase được mã hóa bởi LINE trong genome.

SINE ở người được gọi là trình tự Alu vì nó chứa trình tự điểm cắt của enzyme cắt hạn chế Alu. Genome của người chứa hơn 1 triệu trình tự Alu chỉ một phần hoặc toàn bộ, chiếm hơn 10% genome của người. Trình tự Alu đầy đủ có kích thước 200 nucleotide. DNA genome mang yếu tố di động lớn hơn 20 lần DNA mã hóa cho tất cả protein người.

Tần số đột biến ngẫu nhiên do xen vào yếu tố nhóm 2 là thấp chưa đến 0,2% trong tổng các đột biến ngẫu nhiên. Trong khi đó ở những tế bào động vật khác như chuột do xen retrotransposon lên đến 10% đột biến ngẫu nhiên, cao hơn 50 lần ở người có lẽ liên quan với hoạt tính của yếu tố retrotransposon cao hơn ở chuột.

## Câu hỏi và Bài tập

1. Vì sao phần lớn các đột biến ảnh hưởng đến các gene cấu trúc thường là lặn so với các allele hoang dại?
2. Sự hình thành đột biến dịch khung diễn ra như thế nào?
3. Các hóa chất gây đột biến có đặc điểm gì?
4. Giải thích cơ sở đột biến của các tác nhân gây đột biến sau: 5-bromuracil, acid nitơ và acridin. Xác định xem chúng tạo ra dạng đột biến nào (đồng hoán, đảo hoán hay dịch khung)?
5. Hãy mô tả một loại sai hỏng ngẫu nhiên dẫn đến đột biến.
6. Phân tích sự giống nhau và khác nhau giữa các kiểu transposition.

7. Các trình tự đảo ngược có vai trò gì trong transposition?

8. Cho một chuỗi trình tự nucleotide trên mRNA như sau:

Dạng hoang dại: ... 5' AAUCCUUACGGA 3' ...

Dạng đột biến: .... 5' AAUCCUACGGA 3' ...

Hãy cho biết: Sai hỏng trên xảy ra do loại đột biến nào?

9. Đột biến xảy ra trong trình tự nucleotide do kết cặp nhầm như sau:

5' AGCTGCCTT 3'

3' ACGATGGAA 5' (mạch khuôn)

↓

mRNA

Hãy cho biết: Amino acid nào sẽ được tìm thấy ở codon có nucleotide bị thay đổi như trên?

10. Ở bắp, tần số đột biến của locus R (màu cây) rất cao: 492 đột biến trên  $10^6$  giao tử. Gene tạo màu đỏ của nội nhũ Pr có tần số đột biến là 11 đột biến trên  $10^6$  giao tử. Hãy cho biết: Cần phải phân tích bao nhiêu cây mới tìm được 1 đột biến kép của 2 gen trên?

## Tài liệu Tham khảo

### Tiếng Việt

Phạm Thành Hồ. 2000. *Di truyền học*. NXB Giáo Dục.

Lê Đình Lương, Phan Cự Nhân (1998). *Cơ sở di truyền học*. NXB Giáo dục.

Hoàng Trọng Phán. 1995. *Di truyền học phân tử*. Trung tâm Đào tạo Từ xa, Đại học Huế

### Tiếng Anh

Anthony J. F. Griffiths, Susan R. Wessler, Richard C. Lewontin, William M. Gelbart, David T. Suzuki, Jeffrey H. Miller. 2004. *An introduction to genetics analysis*. W.H. Freeman Publishers.

Harlt D.L., Jones E.W. 1998. *Genetics - Principle and analysis*. Jone and Bartlett Publishers, Toronto, Canada.

Stansfield W.D. 1991. *Schaum's outline of theory and problems of genetics*. McGraw-Hill, Companies, Inc., United States of America.

Watson D.J, Baker T.A., Bell S.P., Gann A., Levine M., Losick R. 2004. *Molecular biology of the gene*. Benjamin/Cummings, San Francisco, USA.



## Chương 9

# Di truyền Tế bào chất

### I. Sự di truyền tế bào chất

#### 1. Sự di truyền của các gene lạp thể

Trong sự thụ phấn của thực vật bậc cao, một tế bào trứng có kích thước lớn có nhiều tế bào chất phối hợp với nhân của hạt phấn không có tế bào chất ở chung quanh. Do đó hợp tử nhận được phần lớn tế bào chất của tế bào trứng. Nếu hai bố mẹ có thành phần nguyên liệu di truyền trong tế bào chất khác nhau thì thế hệ con sẽ nhận được nhiều nguyên liệu di truyền trong tế bào chất của mẹ. Do đó sẽ xảy ra sự di truyền theo thế hệ mẹ.

Hiện tượng di truyền lá đốm được phát hiện rất sớm ở *Mirabilis jalapa* (Correns, 1908), ở *Pelargonium zonale* (E. Bauner, 1909). Các cây có lá đốm có thể có nguyên nhân với lá trắng không có chlorophyll.

Thí nghiệm: Tạp giao giữa cây *Mirabilis jalapa* có những cành khảm trắng xanh theo các phép lai như sau:

- Thụ phấn cho hoa trên cành lá trắng bằng hạt phấn của hoa trên cành lá xanh lục, cho thế hệ con những cây giống cá thể mẹ có lá trắng không có chlorophylle. Các cây này chết vì không có khả năng quang hợp

- Tạp giao hoa trên cành lá xanh lục bằng hạt phấn của hoa trên cành lá trắng, tất cả thế hệ con có lá màu xanh lục bình thường.

- Nếu thụ phấn các hoa của cành lá đốm bởi phấn hoa của cây xanh lục thì ở đời con có các cá thể lá trắng, lá đốm và lá xanh lục.

- Nếu thụ phấn cho hoa của cành lá xanh lục với phấn hoa cây lá đốm thì ở đời con gồm toàn cá thể lá xanh lục.

\* Giải thích: Trong trường hợp này người ta thấy những chất cơ sở hình thành lạp thể có ở trong tế bào trứng sẽ hình thành tiền lạp thể, sau đó hình thành lục lạp. Hạt phấn không có chất cơ sở để hình thành lạp thể nên hạt phấn không thể truyền lục lạp được. Sự khác nhau giữa con cái và bố mẹ về một hoặc nhiều tính trạng khi tạp giao thuận nghịch chứng tỏ có sự tham gia của nguyên liệu di truyền ở trong tế bào chất. Sự di truyền theo hệ mẹ quy định sự thể hiện tính trạng phụ thuộc vào cá thể mẹ.

Ở thực vật *Pelargonium zonale* có trường hợp di truyền theo dòng cha. Nếu hoa của cây lá đốm được thụ phấn của cây lá xanh lục thì 30% cây lai có lá đốm, 70% lá xanh lục. Khi lai hoán đổi cha mẹ thì 70% cây

lai lá đốm và 30% lá xanh lục.

## 2. Sự di truyền của các gene ty thể

### 2.1. Đặc điểm di truyền của các gene ty thể

Theo Mendel, khi tạp giao những sinh vật lưỡng bội thì có sự phân ly tính trạng theo đúng định luật của Mendel vì những gene ở trong nhân đều nằm trên nhiễm sắc thể và trong giảm phân được phân chia cho các giao tử cùng với nhiễm sắc thể. Đối với những tính trạng ở trong tế bào chất không có một hệ thống phân chia nào đảm nhận nên không có sự phân ly theo một quy luật nhất định.

Ở nấm men có một thể đột biến có thể hình thành những khuẩn lạc petite kích thước nhỏ hơn bình thường, đường kính chỉ bằng  $1/2 - 1/2$  khuẩn lạc bình thường. Các tế bào tạo nên khuẩn lạc petite có kích thước giống kích thước tế bào bình thường. Nguyên nhân tạo nên khuẩn lạc kích thước nhỏ là do các tế bào đột biến petite bị hỏng hệ thống hô hấp, tức là những enzyme oxy hóa trong ty thể là các cytochrom b, c, a, a<sub>3</sub> và cytochrom oxydase bị phá hủy. Đây là những enzyme của màng trong ty thể. Khác với kiểu dại, các đột biến petite không thực hiện được phản ứng phosphoryl hóa để sản ra năng lượng, vì vậy tốc độ sinh trưởng và phân bào của chúng thấp hơn.

Ở ty thể của nấm men (*Saccharomyces cerevisiae*) có 3 kiểu đột biến chủ yếu: petite, antR và mit<sup>-</sup>.

Một ví dụ về ty thể là đột biến thiếu năng hô hấp ở nấm men. Vào những năm 1940, Boris Ephrussi và cs. đã mô tả các đột biến đặc biệt ở nấm men. Các đột biến này được gọi là petite, có khuẩn lạc nhỏ hơn nhiều so với khuẩn lạc hoang dại. Theo phương thức di truyền, các đột biến petite chia làm 3 loại khác nhau:

- *Petite phân ly* (Segregation petites): khi lai với dạng hoang dại khuẩn lạc bình thường thì tỷ lệ phân ly trong các nang bào tử (ascospore) là 1 khuẩn lạc to: 1 petite.

- *Petite trung tính* (Neutral petites): khi lai với khuẩn lạc to thì sự phân ly trong nang bào tử chỉ có dạng khuẩn lạc to bình thường, thể hiện sự di truyền theo một cha mẹ (Uniparental)

- *Petite ức chế* (Suppressive petites): khi lai tạo các nang bào tử, một số mọc thành khuẩn lạc to bình thường, một số khác tạo khuẩn lạc petite. Tỷ lệ giữa khuẩn lạc to và nhỏ dao động nhưng có tính đặc hiệu của chủng, một số petite ức chế chỉ tạo thể hệ con khuẩn lạc petite. Qua các petite ức chế cho thấy có sự di truyền ngoài nhân tế bào và một số có sự di truyền theo một cha mẹ.

Khi lai nấm men 2 tế bào cha mẹ, hai tế bào cha mẹ kết hợp với nhau và góp tế bào chất như nhau vào tế bào con lưỡng bội. Sự di truyền của các petite trung tính và ức chế độc lập với kiểu bất cặp thể hiện rõ sự di truyền ngoài nhân nên được gọi là petite tế bào chất. Qua nghiên cứu chúng có các đặc điểm kiểu hình như sau:

- Chuỗi chuyền điện tử của ty thể bị sai hỏng ở các petite tế bào chất. Do sai hỏng này, chúng lên men để tạo ATP kém nên mọc chậm.

- Không có sinh tổng hợp protein ở các petite tế bào chất. Các ty thể có hệ thống sinh tổng hợp riêng gồm tRNA, các ribosome khác với tế bào chất.

- mtDNA ở các đột biến petite có biến đổi lớn. Ty thể của tất cả các Eukaryote có mtDNA riêng tuy số lượng nhỏ, nhưng khác với DNA của nhân tế bào. Ở các petite trung tính, mtDNA bị mất hoàn toàn, còn ở các petite ức chế có sự thay đổi đáng kể tỷ lệ base so với mtDNA của dạng khuẩn lạc to bình thường.

Nhóm các đột biến thứ hai của nấm men là *ant<sup>R</sup>* (*ant<sup>R</sup>* mutants), có kiểu hình đề kháng với các kháng sinh khác nhau. Ví dụ: *cap<sup>R</sup>* (chloramphenicol resistance) kháng chloramphenicol, *ery<sup>R</sup>* kháng erythromycine, *spi<sup>R</sup>* kháng spiromycine, *par<sup>R</sup>* kháng paranomycine và *oli<sup>R</sup>* kháng oligomycine. Các đột biến này khi lai (ví dụ *ery<sup>R</sup>* × *ery<sup>S</sup>*) cho tỷ lệ phân ly không theo quy luật Mendel, giống như các petite ức chế nhưng sự di truyền có khác. Khi các tế bào cha mẹ kết hợp, sản phẩm lưỡng bội là hợp tử hai cha mẹ cytohet (cytoplasmically heterozygote). Các diploid này có thể sinh sản vô tính bằng mọc chồi. Trong nguyên phân, quá trình phân ly tế bào chất và tái tổ hợp xảy ra và các tế bào con trở thành *ery<sup>S</sup>* hay *ery<sup>R</sup>*.

Nhóm đột biến quan trọng thứ ba là *mit<sup>-</sup>* (*mit<sup>-</sup>* mutants) được phát hiện sau cùng nhờ kỹ thuật chọn lọc đặc biệt. Các đột biến này, tương tự các đột biến petite ở chỗ có khuẩn lạc nhỏ và các chức năng bất thường của chuỗi chuyền điện tử, nhưng điểm khác căn bản là sinh tổng hợp protein bình thường và có khả năng hồi biến. Như vậy, các kiểu đột biến *mit<sup>-</sup>* là đột biến điểm. Sự di truyền của kiểu đột biến *mit<sup>-</sup>* giống với kiểu *ant<sup>R</sup>*, có sự phân ly tế bào chất và sự di truyền theo một cha mẹ trong giảm phân.

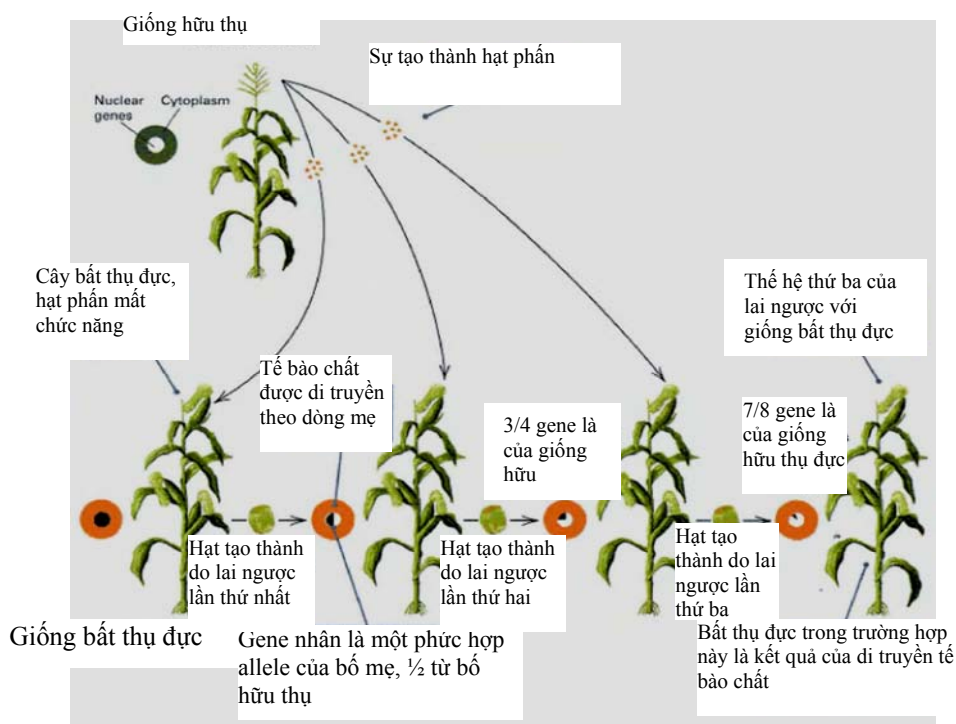
Trong thế hệ con của những tế bào thuộc khuẩn lạc bình thường, có khoảng vài phần trăm tế bào hình thành những khuẩn lạc petite. Những tế bào khuẩn lạc petite luôn luôn phát triển thành những khuẩn lạc petite. Điều đó chứng tỏ có sự thay đổi về cấu trúc di truyền. Ngoài đột biến xảy ra ở kiểu bào gene nói trên dẫn đến sinh ra những khuẩn lạc petite, còn có những khuẩn lạc petite do những gene ở trong nhân quy định.

Thí nghiệm: Tạp giao của một nòi nấm men kích thước khuẩn lạc bình thường với một nòi có kích thước khuẩn lạc petite, thế hệ con hình thành khuẩn lạc bình thường. Còn đối với những gene trong nhân (gene ade), thì sự phân ly ở thế hệ con về những gene này cho tỷ lệ 1:1, do chúng nằm trên NST và được chia đều cho các tế bào con.

Ở đây, nguyên liệu di truyền trong tế bào sẽ được trộn lẫn nhau trong hợp tử và khi tạo thành bào tử thì mỗi bào tử đều nhận được các gene ở trong ty thể như nhau, nên chúng đều có chức năng hô hấp bình thường.

Thí nghiệm cho thấy sự di truyền khuẩn lạc không theo quy luật Mendel.

## 2.2. Hiện tượng bất dục bào chất đực



**Hình 9.1** Sơ đồ thí nghiệm chứng minh sự di truyền theo hệ mẹ của hiện tượng bất thụ đực.

Tính bất dục do nhiều nguyên nhân, bất dục đực (không tạo phần hoa hay tạo phần hoa không có khả năng thụ tinh) ở thực vật có các trường hợp sau:

- Do gene nhân quy định, như gene *ms* ở cây ngô
- Do ảnh hưởng của điều kiện môi trường như độ ẩm, quang chiếu,

khả năng cung cấp chất dinh dưỡng không đáp ứng đúng nhu cầu sinh lý của cây

*Ví dụ:* gene *ms* ở cây ngô

- Do lai xa cũng đưa đến các cơ thể lai không có hạt phấn vì NST có nguồn gốc khác nhau không thể tiếp hợp nhau trong giảm phân. Những hiện tượng bất dục này đều có ý nghĩa hạn chế chỉ có bất dục bào chất đực là có vai trò quan trọng. Đó là trường hợp bất dục của hạt phấn bất nguồn từ tế bào chất, còn nhân thì có thể có điều chỉnh được nhờ đó có thể dùng các cây bất dục bào chất đực để phát huy ưu thế lai ở các đối tượng ngô, cao lương, củ cải đường...

*Ví dụ:* Xét mối quan hệ giữa kiểu gene, kiểu bào gene và kiểu hình của bắp được sử dụng trong lai một tính mà không cần khử đực ở cây mẹ

STT	Kiểu gene	kiểu bào gene	kiểu hình (hạt phấn)
1	Rfrf	S (bất dục)	hông
2	Rfrf	N (hữu dục)	tốt
3	RfRf hoặc Rfrf	N	tốt
4	RfRf hoặc Rfrf	S	tốt

Vậy hạt phấn của ngô chỉ bị mất hoạt tính khi có yếu tố bất dục trong tế bào chất mà lại thiếu thiếu gene phục hồi hữu dục (Rf) ở trong nhân, alen của gene này là rf là gene cũng cố tính bất dục.

Cây được phục hồi hữu dục RfrfS cho tự thụ phấn thì ở đời sau sẽ có 1/4 rfrf có hạt phấn hông. Nếu lấy bắp của dạng rfrfS thụ phấn của rfrfN, thì phần hoa của toàn bộ đời sau sẽ bị hông, cây này chỉ còn bắp mang nhị cái. Đó là cách dùng phương pháp di truyền để khử cò ngô. Khi sản xuất hạt giống, những cây này muốn có hạt thì phải thụ phấn hữu dục của những cây bình thường. Nếu muốn dùng những hạt đó để sau này trồng lại thì cây bố phải có kiểu gene RfRf và kiểu bào gene N hoặc S.

Trong sản xuất giống ngô có thể dùng tổ hợp dòng thuần dạng 1 làm cây mẹ và dạng 3 hoặc dạng 4 đồng hợp tử làm cây bố. Như thế sẽ đỡ mất công khử đực ở cây mẹ và hạt lai thu được từ cây mẹ sẽ có kiểu gene Rfrf, kiểu bào gene S. Kiểu gene này đảm bảo được sự thụ phấn bình thường lúc trồng trong sản xuất.

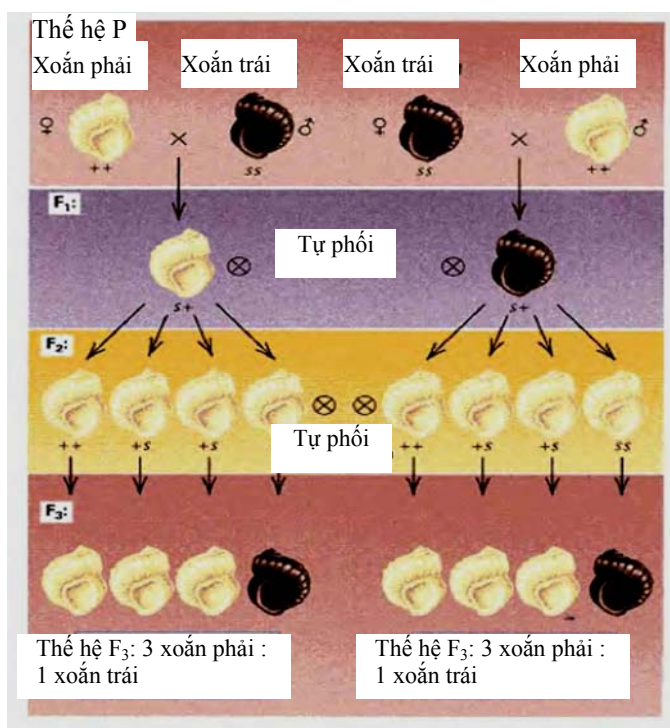
Bất thụ đực tế bào chất ở ngô liên quan đến 2 plasmid dạng thẳng S1 và S2. Chúng ở trong ty thể cùng với mtADN. Một trong những tính chất khó hiểu của plasmid này là chúng có thể thực hiện tái tổ hợp với mtADN.

### 3. Hiệu quả dòng mẹ lên chiều xoắn vỏ ốc

Trứng và phôi chịu ảnh hưởng của môi trường cơ thể mẹ nhiều hơn

cơ thể bố. Ngay cả khi bị tách khỏi cơ thể mẹ từ một giai đoạn rất sớm, chúng cũng đã nhận được tế bào chất, các chất dinh dưỡng trong trứng từ mẹ và những ảnh hưởng đặc biệt lên hoạt động của gene. Những tiềm năng nhất định của trứng được xác định trước khi thụ tinh và trong một số trường hợp chúng đã chịu ảnh hưởng của môi trường mẹ bao quanh. Sự quyết định trước như vậy do các gene của mẹ hơn là các gene của con, được gọi là hiệu ứng mẹ. Sự tồn tại của hiệu ứng mẹ nói chung được chứng minh bằng phương pháp lai thuận nghịch, khi đó kết quả phép lai thuận nghịch sẽ khác nhau.

Ví dụ về hiệu quả dòng mẹ: chiều xoắn của vỏ ốc *Limnaea peregra*. Chiều xoắn này được xác định bởi một cặp alen. D là xoắn phải, d là xoắn trái. Các phép lai cho thấy D là trội so với d và chiều xoắn luôn luôn được xác định bởi kiểu gene của mẹ.



**Hình 9.2** Hiệu quả dòng mẹ lên chiều xoắn vỏ ốc

Tiến hành phép lai thuận nghịch giữa đồng hợp tử DD xoắn phải và đồng hợp tử dd xoắn trái:

Khi trứng bắt nguồn từ ốc xoắn phải thì F1 là Dd về kiểu gene và ốc có kiểu hình xoắn phải. Cho các con ốc này tự phối thì sinh ra thế hệ sau có tỉ lệ phân li về kiểu gene là 1DD : 2 Dd : 1 dd, tất cả các con ốc thậm

chí cả con có kiểu gene dd đều có chiều xoắn phải. Nhưng khi mỗi con ốc của thế hệ này tự phối riêng biệt thì chỉ có những con có kiểu gene DD và Dd cho thế hệ sau xoắn phải, còn những con dd mặc dù có kiểu hình xoắn phải nhưng lại cho thế hệ sau xoắn trái.

Ngược lại nếu dùng ốc dd xoắn trái làm mẹ thì F<sub>1</sub> Dd đều xoắn trái giống mẹ. Cho các ốc con này tự phối thì tất cả ốc thế hệ sau có xoắn phải vì kiểu gene của mẹ chúng là Dd. Nếu cho mỗi con của thế hệ này tự phối thì kết quả nhận được như phép lai thuận trên, nghĩa là thế hệ sau có tỷ lệ phân ly 3 xoắn phải : 1 xoắn trái.

Như vậy hiệu quả dòng mẹ chỉ kéo dài một thế hệ và trường hợp này không thể xem là di truyền qua tế bào chất bởi vì ở đây các đặc tính của tế bào chất đã được xác định trước bởi tác dụng của các gene trong nhân chứ không phải bởi các gene trong tế bào chất. Nói cách khác ở đây cơ chế di truyền nhiễm sắc thể làm biến đổi tế bào chất của trứng trước khi nó thụ tinh.

## II. Lập bản đồ ở ty thể và lục thể

### 1. Lập bản đồ gene của DNA lục thể

Các gene của lục thể ở *Chlamydomonas reinhardtii*

Sự di truyền lục thể được nghiên cứu chi tiết hơn cả ở vi tảo *Chlamydomonas reinhardtii*. Tế bào của vi tảo này có một lục thể lớn với đường kính trung bình 5 μm chứa 50 đến 80 bản sao của phân tử DNA vòng tròn dài 196 kb.

Theo Sager (1975), ở *Chlamydomonas reinhardtii* có các đột biến trong nhóm liên kết gene của lục thể. Các đột biến có biểu hiện kiểu hình như sau:

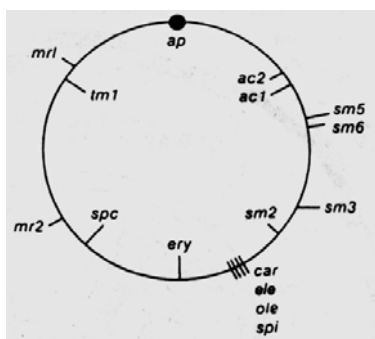
- Mất khả năng quang hợp để mọc được ngoài ánh sáng và trong tối cần bổ sung đường khử là acetat
- Nhạy cảm với nhiệt độ cao hoặc thấp
- Tính đề kháng với thuốc kháng sinh hoặc có nhu cầu được cung cấp thuốc kháng sinh.

Tất cả các đột biến trên có sự di truyền theo một cha mẹ, có kiểu bất cặp mt<sup>+</sup> (có thể coi là dòng mẹ). Điều này liên quan đến sự hình thành lục thể trong hợp tử, bằng cách nào đó, chỉ nhận DNA từ lục thể mt<sup>+</sup>.

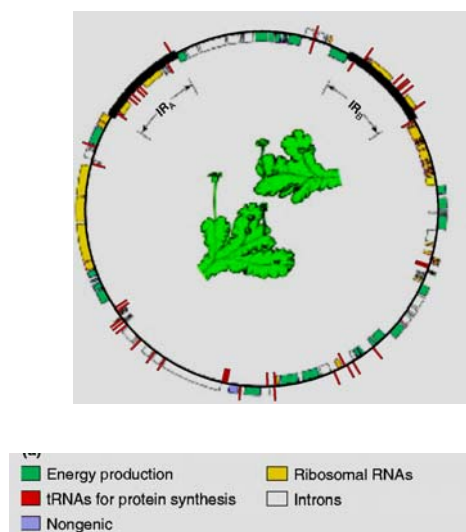
Năm 1954, R. Sager đã nghiên cứu các đột biến kháng streptomycin ở *C. reinhardtii* từ dạng hoang dại nhạy cảm sm-s. Một số đột biến sm-r có sự di truyền nhiễm sắc thể với sự phân ly 1 : 1. Tuy nhiên một số đột biến có sự di truyền khác thường như sau:

$sm-r\ mt^+ \times sm-s\ mt^- \rightarrow$  tất cả đời con đều  $sm-r$  với tỷ lệ  $1\ mt^+ : 1\ mt^-$ .

$sm-s\ mt^+ \times sm-r\ mt^- \rightarrow$  tất cả đời con đều  $sm-s$  với tỷ lệ  $1\ mt^+ : 1\ mt^-$ .



**Hình 9.3** Bản đồ vòng tròn của cpDNA ở *Chlamydomonas*



**Hình 9.4** Bản đồ DNA chloroplast của *Marchantia polymorpha*

$IR_A$  và  $IR_B$  là những trình tự đảo ngược (theo K.Umesono và H.Ozeki, 1987)

Như vậy ở đây khi có sự hoán đổi cha mẹ trong lai, thế hệ con đều có kiểu hình streptomycin của  $mt^+$ . Sự truyền thụ tính trạng này được gọi là sự di truyền theo một cha mẹ. Sager coi  $mt^+$  như dòng mẹ và trường hợp trên giống như di truyền theo dòng mẹ. Các kiểu bất cặp  $mt$  có tỷ lệ phân ly của gene trong nhân là 1 : 1.

Trong tổ hợp lai  $mt^+ sm-r \times mt^- sm-s$  có khoảng 0,1% thế hệ hợp tử con mang cả  $sm-r$  và  $sm-s$ . Các hợp tử như vậy gọi là hợp tử hai cha mẹ cytohet (cytoplasmically heterozygote).



Tần số các cytohet có thể tăng lên 40-100% ở thế hệ hợp tử con nếu  $mt^+$  được chiếu tia tử ngoại trước khi lai.

Ở *Chlamydomonas*, các hợp tử 2 cha mẹ được dùng làm điểm xuất phát cho tất cả các nghiên cứu về sự phân ly và tái tổ hợp của các gene lục lạp.

Trên cơ sở nhiều tổ hợp lai, R. Sager đã nêu ra bản đồ vòng tròn của cpDNA với các gene tương ứng.

## 2. Lập bản đồ gene của DNA ty thể

\* Lập bản đồ bộ gene ty thể của nấm men

Có các phương pháp khác nhau xây dựng bản đồ bộ gene ty thể:

- *Lập bản đồ tái tổ hợp* (recombination mapping): Ở nấm men sự phân li tế bào chất và tái tổ hợp xảy ra trong quá trình mọc chồi ở cytohet lưỡng bội. Sự phân li và tái tổ hợp có thể phát hiện trực tiếp ở các dạng lưỡng bội mọc chồi hay quan sát sản phẩm giảm phân khi chồi được kích thích tạo bào tử.

Ví dụ khi lai  $ery^R spi^R \times ery^S sp^S$  có thể nhận được các kiểu bộ bốn với số lượng như sau:

$ery^R spi^R$	63 bộ bốn
$ery^S spi^S$	48 bộ bốn
$ery^S spi^R$	7 bộ bốn
$ery^R spi^S$	1 bộ bốn

Các kiểu gene  $ery^S spi^R$  và  $ery^R spi^S$  là các dạng tái tổ hợp.

- *Lập bản đồ bằng phân tích petite*: Một số kỹ thuật lập bản đồ có hiệu quả dựa trên sự phối hợp cả 3 loại đột biến petite,  $ant^R$  và  $mit^-$ . Phần lớn các tiếp cận này dựa vào sự mất đoạn của mtDNA của các đột biến petite. Sự kết hợp các kiểu phân tích di truyền đặc biệt với kỹ thuật tái tổ hợp DNA đã dẫn đến xây dựng bản đồ di truyền hoàn chỉnh của mtDNA.

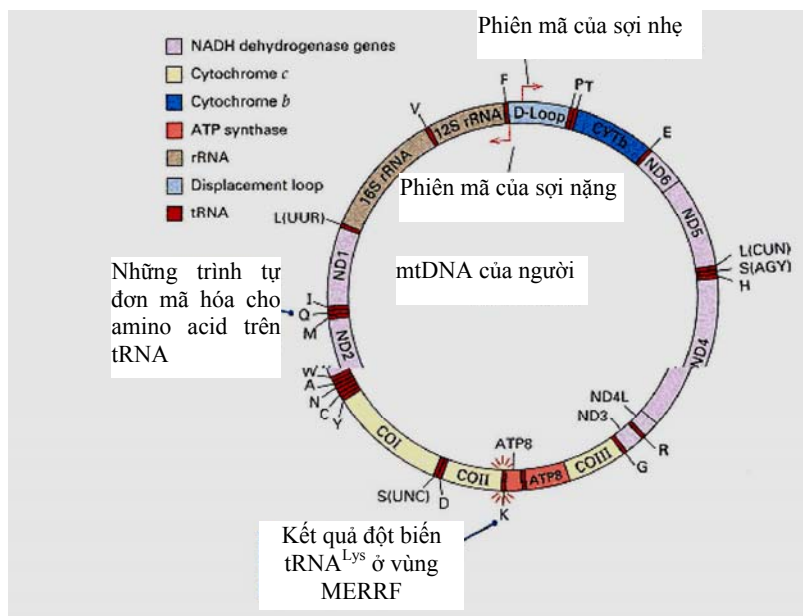
- *Lập bản đồ bằng phân tích enzyme cắt giới hạn*: các enzyme cắt giới hạn (xem chương 10) là công cụ hữu hiệu mới để phân tích di truyền. Nó được sử dụng không những để phân tích mtDNA của nấm men, mà cả mtDNA của bất kỳ sinh vật nào miễn chiết tách và tinh sạch được DNA.

\* Bản đồ mtDNA của nấm men và người

Việc xây dựng bản đồ mtDNA hoàn chỉnh của nấm men và người là những thành tựu đáng kể của nghiên cứu di truyền tế bào chất.

- Một số trình tự có codon khởi sự và codon kết thúc ở cuối nhưng chưa biết chức năng

- mtDNA của người rất ít dư thừa



**Hình 9.5 Bản đồ mtDNA của người** (Theo Larson và Clayton, 1995)

### III. Di truyền học phân tử các bào quan

#### 1. Các bộ gene lục thể (cpDNA)

Là bào quan có khả năng tự tái sinh ở tế bào thực vật. Sự phân chia của các bào quan này về về các tế bào con trong phân bào là không đều như sự phân chia của nhiễm sắc thể trong nguyên phân và giảm phân. Chúng có số lượng lớn và phân chia ngẫu nhiên về các tế bào con nên mỗi tế bào có thể chứa nhiều hoặc ít lục lạp.

DNA của lục lạp được ký hiệu là cpDNA (Chloroplast DNA). Bộ gene này ở dạng DNA vòng tròn, thường dài hơn DNA của ty thể 8-9 lần. Trong lục lạp còn tìm thấy bộ máy sinh tổng hợp protein khác rất nhiều với hệ thống trong tế bào chất của eukaryote nhưng giống với bộ máy sinh tổng hợp protein của prokaryote.

Mặc dù sự di truyền của lục lạp được phát hiện rất sớm, nhưng trong một thời gian dài sự hiểu biết chi tiết về các gene của lục lạp không có bước tiến đáng kể. Các nghiên cứu phân tử đã góp phần chủ yếu cho sự phân tích chi tiết các gene ở các bào quan. Ngoài các nghiên cứu ở *Mirabilis jalapa* và *Chlamydomonas*, bản đồ chi tiết cpDNA của thực vật *Marchantia polymorpha* đã được xây dựng.

cpADN điển hình dài khoảng 120-200 kb tùy loài thực vật. Ở *Marchantia*, kích thước phân tử là 121 kb.

Trên cpDNA của *Marchantia* có tất cả 136 gene gồm 4 loại mã hóa tổng hợp rRNA, 31 loại mã hóa tổng hợp tRNA và khoảng 90 gene tổng hợp protein. Trong số 90 gene mã hóa tổng hợp protein, có 20 gene mã hóa tổng hợp enzyme cho quang hợp và chuỗi chuyền điện tử. Các gene mã hóa cho các chức năng dịch mã chiếm khoảng một nửa bộ gene của lục lạp và bao gồm các protein và các RNA cần thiết cho dịch mã bên trong lục lạp.

Thực tế DNA của lục lạp, ty thể và nhân tế bào có sự phối hợp chặt chẽ trong việc tạo ra các tiểu phần của những protein được sử dụng bên trong lục lạp. Ribulose-1,5-biphosphate carboxylase/ oxygenase là enzyme dồi dào nhất của lục lạp. Nó xúc tác 2 phản ứng cạnh tranh nhau, cố định CO<sub>2</sub> và bước đầu tiên của *quang hô hấp* (photorespiration) với sự tạo ra glycolate. Enzyme gồm 8 *tiểu phần lớn* LS (large unit) giống nhau và 8 *tiểu phần nhỏ* giống nhau được mã hóa tương ứng bởi các gene của lục lạp và nhân tế bào. Tiểu phần lớn LS mang trung tâm xúc tác, còn vai trò của các tiểu phần nhỏ chưa rõ. Gene LS nằm trên cpDNA của một số thực vật như bèo, *Chlamydomonas reinhardtii*, thuốc lá, *Euglena*... Trong tất cả các trường hợp, gene LS hiện diện 1 bản sao cho 1 DNA của lục lạp. Ngược lại, các gene của tiểu phần nhỏ được tìm thấy ở các trình tự DNA của nhân tế bào với số bản sao ít.

## 2. Các bộ gene ty thể (mtDNA)

Bào quan ty thể có ở tất cả các tế bào của eukaryote. Bộ gene của ty thể được ký hiệu là mtDNA (Mitochondrial DNA). mtDNA mã hóa cho sự tổng hợp nhiều thành phần của ty thể như hệ thống 2 loại rRNA, 22-25 loại tRNA và nhiều loại protein có trong thành phần màng bên trong ty thể. Trong khi đó, phần lớn protein của ribosom của ty thể thì do các gene ở trong nhân xác định.

Bộ gene của ty thể có hai chức năng chủ yếu:

- Mã hóa cho một số protein tham gia chuỗi chuyền điện tử
- Mã hóa cho hệ thống sinh tổng hợp protein gồm một số protein, tất cả các tRNA và cả 2 loại rRNA.

Tuy nhiên trong cả hai trường hợp, những cấu phần còn lại của hệ thống được mã hóa do các gene nhân và được dịch mã ở bào tương (cytosol) rồi chuyển vào ty thể.

Ở Việt Nam đã có công trình phân tích mtDNA ở 50 người Việt dân tộc Kinh, phát hiện được các hình thái khác nhau khi cắt mtDNA bằng 6

enzyme hạn chế.

Như vậy, việc nghiên cứu các gene của ty thể cho thấy tế bào eukaryote không lục lạp có ít nhất 2 hệ thống sinh tổng hợp protein độc lập tương đối nhưng luôn hợp tác chặt chẽ với nhau. Ở các eukaryote có lục lạp thì 3 hệ thống sinh tổng hợp protein độc lập tương đối nhưng hợp tác với nhau. Cả 2 bào quan ty thể và lục lạp tham gia trực tiếp vào chuyển hóa năng lượng của tế bào.

Di truyền tế bào chất là hiện tượng di truyền do các gene nằm trên nhiễm sắc thể ở ngoài nhân quy định.

### Câu hỏi và Bài tập

1. Hãy nêu các kiểu đột biến của ty thể.
2. Trình bày chứng minh thực nghiệm về di truyền lục lạp thể.
3. So sánh sự giống nhau và khác nhau của mtDNA và cpDNA.
4. Nếu có hạt bắp từ dòng bất thụ đực, làm sao xác định sự bất thụ do gene trong nhân hay ngoài nhân?
5. Nêu các đặc điểm riêng của di truyền ngoài nhiễm sắc thể và nêu các sai khác cơ bản so với di truyền của gene nhân.
6. Hãy phân biệt hiệu quả dòng mẹ với sự di truyền ngoài nhân.
7. Hãy phân biệt các loại đột biến petite ở nấm men.
8. Cho 2 dòng bắp bất thụ đực: một dòng bất thụ bào chất, một dòng bất thụ do gene nhân di truyền theo Mendel. Hãy trình bày phương pháp xác định hai dòng bất thụ trên.
9. Một con ốc loài *Limnaea peregra* có vỏ xoắn trái qua tự phối cho tất cả thế hệ sau xoắn phải. Hãy xác định genotype của con ốc này. Nếu thế hệ sau cho tự phối riêng rẽ, hỏi chiều xoắn vỏ của các con ốc thế hệ này như thế nào?
10. Tỷ lệ tiến hóa DNA của ty thể được tính bằng sự biến đổi 1 nucleotide của 1 ty thể trong thời gian 1.500 đến 3.000 năm. DNA ty thể người có khoảng 16.500 nucleotide. Hỏi tỷ lệ biến đổi của 1 nucleotide trong  $10^6$  năm là bao nhiêu?

### Tài liệu Tham khảo

#### Tiếng Việt

Phạm Thành Hồ. 2000. *Di truyền học*. NXB Giáo Dục.

Lê Đình Lương, Phan Cự Nhân. 1998. *Cơ sở di truyền học*. NXB Giáo dục.

Phan Cự Nhân, Nguyễn Minh Công, Đặng Hữu Lanh. 1999. *Di truyền học*. NXB Giáo dục

Hoàng Trọng Phán. 1995. *Di truyền học phân tử*. Trung tâm Đào tạo Từ xa, Đại học Huế

### **Tiếng Anh**

Anthony J. F. Griffiths, Susan R. Wessler, Richard C. Lewontin, William M. Gelbart, David T. Suzuki, Jeffrey H. Miller. 2004. *An introduction to genetics analysis*. W.H. Freeman Publishers.

Harlt D.L., Jones E.W. 1998. *Genetics - Principle and analysis*. Jone and Bartlett Publishers, Toronto, Canada.

Stansfield W.D. 1991. *Schaum's outline of theory and problems of genetics*. McGraw-Hill, Companies, Inc., United States of America.

Vu Trieu An. 1999. *Mitochondrial DNA Polymorphism in the Vietnamese population*. *Eur. J. of Immunogenetics*, 26: 471- 422.

Watson D.J, Baker T.A., Bell S.P., Gann A., Levine M., Losick R. 2004. *Molecular biology of the gene*. Benjamin Cummings, San Francisco, United States of America.

## Chương 10

# Đại cương về Công nghệ DNA Tái tổ hợp

Việc tinh chế enzyme cắt giới hạn đầu tiên (1970) và sử dụng nó để tạo ra các phân tử DNA tái tổ hợp đầu tiên trong ống nghiệm (1972-1973) là nền tảng cho sự ra đời của *kỹ thuật di truyền* (genetic engineering) và *công nghệ DNA tái tổ hợp* (recombinant DNA technology). Chính sự phát triển nhanh chóng của lĩnh vực này không những đã đưa lại khối lượng tri thức khổng lồ về cấu trúc và cơ chế hoạt động của các gene, các bộ gene prokaryote và eukaryote mà còn trở thành lực lượng sản xuất trực tiếp của xã hội: tạo ra hàng loạt các chế phẩm y-sinh học hữu ích từ các tế bào vi khuẩn, nấm men; tạo các giống sinh vật mới... góp phần giải quyết những vấn đề thực tiễn đặt ra trong y học và trong công tác chọn tạo giống.

Trong chương này chúng ta sẽ tìm hiểu các đặc tính của enzyme cắt giới hạn, các nguyên lý cơ bản của kỹ thuật DNA tái tổ hợp và một số ứng dụng của lĩnh vực công nghệ này.

## I. Các công cụ chính của kỹ thuật tạo dòng DNA tái tổ hợp

### 1. Các enzyme cắt giới hạn

#### 1.1. Enzyme cắt giới hạn là gì?

*Enzyme cắt giới hạn* (restriction endonuclease) hay gọi tắt là *enzyme giới hạn* (restrictase) là loại enzyme có khả năng nhận biết đoạn trình tự nucleotide đặc hiệu trên các phân tử DNA và cắt cả hai sợi DNA bổ sung tại các vị trí đặc thù.

#### 1.2. Vai trò của các enzyme cắt giới hạn

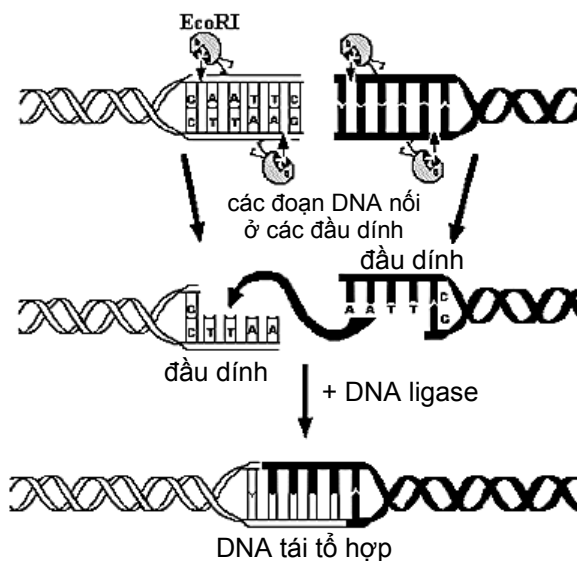
Từ 1953 người ta đã phát hiện thấy rằng, khi đưa DNA của một nòi vi khuẩn *E. coli* này vào tế bào thuộc một nòi khác thường thì DNA được đưa vào, gọi là DNA ngoại lai hay DNA lạ, mất hẳn hoạt tính di truyền và hầu như bao giờ cũng bị phân cắt thành các đoạn ngắn. Chỉ trong một số ít trường hợp DNA lạ đó mới không bị phân cắt và do đó nó có thể tái bản trong tế bào chủ. Điều đó chứng tỏ DNA lạ được sửa đổi bằng cách nào đó dưới sự kiểm soát của tế bào chủ. Các hiện tượng nói trên xảy ra chủ yếu khi các thể thực khuẩn (*phage*) xâm nhiễm các tế bào vi khuẩn.

Cho đến đầu thập niên 1970 người ta mới biết rõ rằng các tế bào vi khuẩn là những hệ thống chứa cả hai loại enzyme: các enzyme sửa đổi và các enzyme cắt giới hạn. Chúng đều có đối tượng nhận biết là các đoạn trình tự của DNA vật chủ và DNA ngoại lai, nhưng có vai trò khác nhau. Cụ thể, các *enzyme sửa đổi* (methylase) đóng vai trò bảo vệ DNA vật chủ

bằng cách gắn thêm nhóm methyl (-CH<sub>3</sub>) ở một số base nhất định trong *đoạn nhận biết* (recognition sequence) hay *đoạn đích* (target sequence). Hiện tượng *methyl hoá* (methylation) này thường xảy ra đối với adenine và biến đổi nó thành N-6 methyladenine. Trong khi đó, các enzyme giới hạn lại đóng vai trò vô hiệu hoá hoạt tính di truyền của các DNA lạ bằng cách phân cắt ở các vị trí đặc thù chừng nào nó chưa được sửa đổi cho giống với DNA vật chủ. Như vậy, *các enzyme giới hạn đóng vai trò là hàng rào bảo vệ tự nhiên của các vi khuẩn nhằm chống lại sự xâm nhập của các phage lạ*.

### 1.3. Tính chất chung của các enzyme giới hạn

Trước tiên, cần lưu ý rằng các enzyme giới hạn chỉ phát hiện thấy ở các vi khuẩn mà không có ở các eukaryote. Vì vậy, tên gọi của các enzyme giới hạn thông dụng là tên hệ thống, được biểu thị bằng ba hoặc bốn chữ cái viết tắt của vi khuẩn mà từ đó enzyme được chiết xuất. Chữ cái đầu tiên được viết hoa để chỉ *chi* (genus) và hai chữ cái tiếp theo viết thường để chỉ *loài* (species), và khi cần thiết thêm chữ cái thứ tư để chỉ *nòi* hoặc *chủng* (strain, type). Ngoài ra, để phân biệt các enzyme cùng một nòi người ta dùng số La Mã kèm theo sau tên hệ thống (xem bảng 10.1).



**Hình 10.1** Enzyme giới hạn (EcoRI) cắt các DNA khác nhau, nhờ đó có thể kiến tạo phân tử DNA tái tổ hợp *in vitro* (bằng enzyme DNA ligase).

Tính chất quan trọng nhất của các enzyme giới hạn là *tính đặc hiệu vị trí*, nghĩa là chúng có thể nhận biết đoạn trình tự DNA đặc thù để cắt ở vị trí xác định. Tùy theo vị trí cắt so với đoạn nhận biết mà chia ra hai loại:

loại I bao gồm các enzyme giới hạn cắt bên ngoài phạm vi đoạn nhận biết và loại II bao gồm các enzyme cắt đặc hiệu bên trong đoạn nhận biết. Ở đây chúng ta chỉ xét các enzyme giới hạn loại II vốn được xem là công cụ hiệu năng (giống như con dao mổ tinh vi) cho phép thao tác trên các gene trong kỹ thuật DNA tái tổ hợp (hình 10.1).

Đặc trưng nổi bật của các đoạn đích là có kích thước ngắn, 4-8 cặp base, và có tính *đối xứng xuôi ngược* (palindrome).

Nhìn chung, các enzyme giới hạn khác nhau có hai kiểu cắt sau đây: cắt lệch và cắt thẳng. Với kiểu cắt lệch tức là các vị trí cắt trên hai sợi của DNA sợi kép là so le, tạo ra các đoạn DNA có các đầu sợi đơn gồm một số base bổ sung gọi là các *đầu dính* (cohesive/sticky ends). Các enzyme giới hạn như thế có vai trò to lớn trong việc kiến tạo DNA tái tổ hợp *in vitro* (hình 10.1). Điển hình ở đây là EcoRI và BamHI (bảng 10.1). Với kiểu cắt thẳng, tức cắt cùng vị trí trên cả hai sợi của DNA sợi kép, do đó tạo ra các đoạn DNA có các *đầu bằng* (blunt ends); ví dụ, SmaI... (bảng 10.1).

Các enzyme giới hạn khác nhau có đoạn đích giống nhau, mặc dù vị trí và kiểu cắt có thể giống hoặc khác nhau, gọi là *các enzyme giới hạn tương ứng* (isoschizomers); ví dụ, SmaI và XmaI (bảng 10.1).

**Bảng 10.1 Các trình tự nhận biết và vị trí cắt của các enzyme giới hạn được chọn lọc** (mũi tên chỉ vị trí cắt; các trình tự ở đây được chỉ ra trên sợi 5'→3')

<b>Nguồn vi sinh vật</b>	<b>Tên enzyme</b>	<b>Trình tự nhận biết</b>
<i>Arthrobacter luteus</i>	<i>AluI</i>	AG↓CT
<i>Bacillus amyloliquefaciens H</i>	<i>BamHI</i>	G↓GATCC
<i>Escherichia coli RY13</i>	<i>EcoRI</i>	G↓AATTC
<i>Haemophilus influenzae Rd</i>	<i>HindII</i>	GTPy↓PuAC
<i>Haemophilus influenzae Rd</i>	<i>HindIII</i>	A↓AGCTT
<i>Nocardia otitidis-caviarum</i>	<i>NotI</i>	GC↓GGCCGC
<i>Providencia stuartii</i>	<i>PstI</i>	CTGCA↓G
<i>Serratia marcescens Sb</i>	<i>SmaI</i>	CCC↓GGG
<i>Xanthomonas malvaccarum</i>	<i>XmaI</i>	C↓CCGGG

## 2. Các vector thông dụng trong kỹ thuật di truyền

### 2.1. DNA tái tổ hợp là gì?

*DNA tái tổ hợp* (recombinant DNA) là phân tử DNA được tạo ra trong ống nghiệm bằng cách kết hợp các DNA từ các nguồn (loài) khác nhau, theo một quy trình kỹ thuật nhất định, gọi là kỹ thuật tái tổ hợp DNA.

Thông thường một phân tử DNA tái tổ hợp bao gồm một phân tử DNA có bản chất là *plasmid* hoặc *phage* nguyên vẹn gọi là *vector* (thể tải) và một đoạn DNA từ nguồn khác mang một gene hoặc yếu tố điều hòa mong

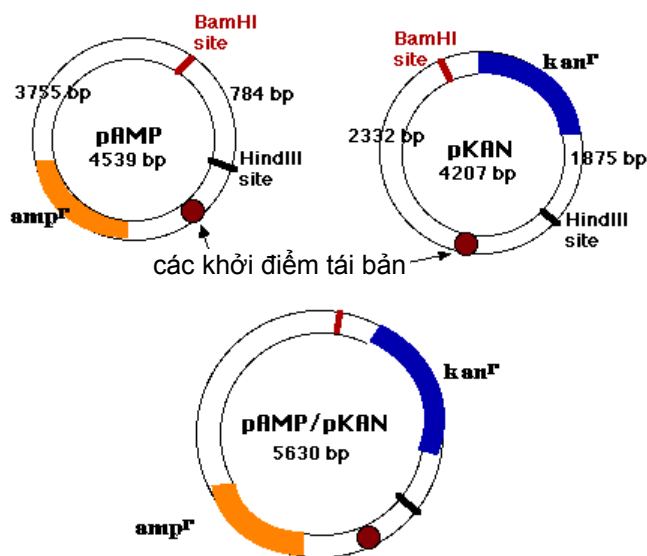


muốn được cho xen vào; nó được gọi là *DNA ngoại lai* (foreign DNA).

## 2.2. Hai loại vector thông dụng

Vector là phân tử DNA có kích thước bé hoặc vừa phải, đóng vai trò là vật trung gian mang truyền đoạn DNA ngoại lai nghiên cứu vào trong tế bào thể nhận (tế bào khả biến) bằng con đường *biến nạp* (transformation) hoặc *tải nạp* (transduction).

Có hai loại vector thông dụng là các plasmid hoặc các phage. Các plasmid vi khuẩn (hình 10.2) được sử dụng rộng rãi hơn cả, bởi vì chúng có các đặc điểm sau: (i) có khả năng xâm nhập vào tế bào vật chủ mà vẫn hoạt động (tái bản, biểu hiện gene) bình thường; (ii) có trọng lượng phân tử thấp nên dễ dàng tinh chiết; (iii) số bản sao trong mỗi tế bào vi khuẩn thường khá cao; và (iv) đặc biệt là, một số plasmid có chứa các gene kháng thuốc tiện lợi cho việc theo dõi và phát hiện sự có mặt của plasmid tái tổ hợp trong vi khuẩn chủ.



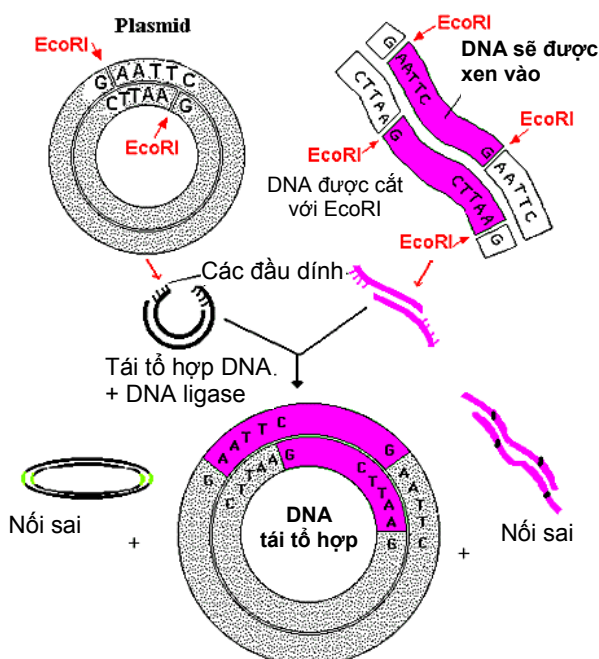
**Hình 10.2** Các plasmid có chứa khởi điểm tái bản (*ori*), các điểm cắt của một số relictase và các gene kháng ampicillin ( $Amp^R$ ), kanamycin ( $Kan^R$ ).

Trong số các phage dùng làm vector thì phage lambda ( $\lambda$ ) có nhiều ưu thế nhất, bởi lẽ ở phần giữa của bộ gene có chứa một số gene không quan trọng và không liên quan với sự tái bản của nó, nên thuận lợi cho việc xen đoạn DNA mong muốn vào đây. Các phage không chứa các gene kháng thuốc cho nên việc theo dõi phage tái tổ hợp được xác định dựa vào các vết tan dương tính (positive plaques) trên nền vi khuẩn.

### 3. Thiết lập phân tử DNA tái tổ hợp in vitro

### 3.1. Phương pháp sử dụng các đầu dính

Bất kỳ đoạn DNA nào nếu được cắt bởi cùng một loại enzyme giới hạn (ví dụ, *EcoRI*) cho các đầu dính thì có thể dính lại với nhau và được nối bởi DNA ligase (hình 10.3). Phương pháp thành lập phân tử DNA tái tổ hợp kiểu này lần đầu tiên được đưa ra bởi J.Mert và R.Davis năm 1972 bằng thực nghiệm trên các virus. Và sau đó, lần đầu tiên năm 1973, H.Boyer và nhóm nghiên cứu của S.Cohen đã tạo ra được phân tử DNA tái tổ hợp gồm vector là plasmid nhỏ pSC101 của *E. coli* và DNA "ngoại lai" là một plasmid khác. Chính sự kiện này đã đặt nền móng và mở ra triển vọng to lớn cho kỹ thuật DNA tái tổ hợp sau này.



**Hình 10.3** Hai phân tử DNA khác nhau được cắt bởi cùng một enzyme giới hạn thì có thể nối với nhau nhờ xúc tác của DNA ligase.

### 3.2. Phương pháp nối trực tiếp hoặc tổng hợp các đầu bổ sung

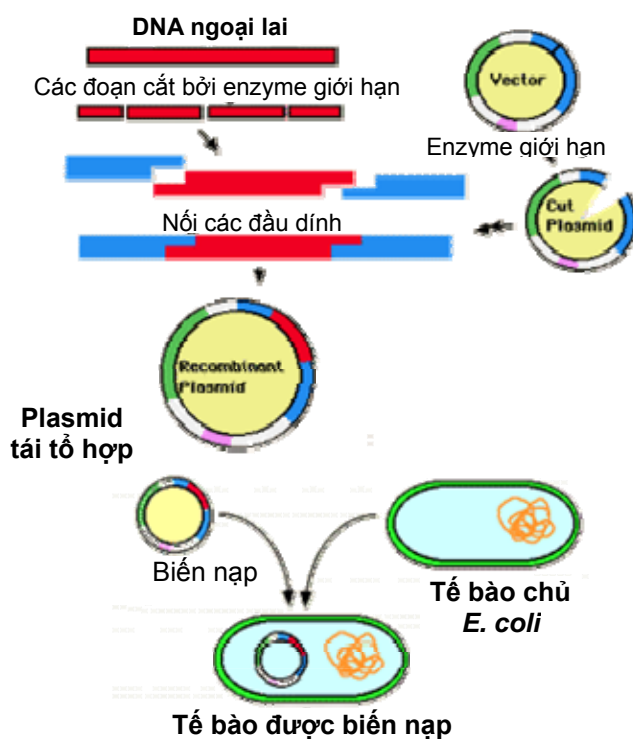
Đối với các đoạn DNA được tạo ra bằng cách xử lý enzyme giới hạn cắt thẳng như HindII chẳng hạn, thì việc nối các đoạn DNA có đầu bằng được tạo ra có thể thực hiện theo hai cách sau: Nối trực tiếp bằng DNA ligase của phage T4 hoặc tổng hợp thêm các đầu dính vào các đầu 3' một số nucleotide bổ sung bằng cách sử dụng các enzyme *end transferase*, rồi sau đó các đoạn DNA như thế sẽ được nối với nhau bởi DNA ligase của vi khuẩn. Cơ sở của phương pháp kết hợp DNA này được thực hiện lần đầu tiên giữa DNA của virus SV40 với DNA của phage  $\lambda$  bởi L.Lobban và

D.Kaiser (1972), và D.Jackson và P.Berg (1972)

## II. Tạo dòng gene hay DNA tái tổ hợp

### I. Nguyên tắc chung

Về nguyên tắc, kỹ thuật DNA tái tổ hợp hay *tạo dòng* (cloning) gồm các bước chung nhất như sau: (1) Tách chiết và tinh sạch DNA thuộc các nguồn khác nhau (gồm vector và DNA mang gene mong muốn); (2) tạo ra phân tử DNA tái tổ hợp *in vitro*; (3) đưa phân tử DNA tái tổ hợp vào trong tế bào nhận, thường là *E. coli* hoặc nấm men. Hình 10.4 mô tả một quy trình kỹ thuật đơn giản như thế. Tuy nhiên, trên thực tế, sự phức tạp là ở bước (4), phát hiện và phân lập các dòng DNA tái tổ hợp đặc hiệu.



**Hình 10.4** Một quy trình biến nạp DNA tái tổ hợp với vector là plasmid, enzyme giới hạn *EcoRI*, enzyme nối - DNA ligase, và tế bào nhận là *E. coli*.

Trong tế bào chủ, phân tử DNA tái tổ hợp có thể biểu hiện gene mong muốn (cho sản phẩm protein) hoặc tái bản độc lập nhiều lần để tạo ra hàng loạt bản sao của nó, và khi tế bào chủ phân chia sẽ kéo theo sự *tạo dòng phân tử* (molecular cloning). Mặt khác, do tốc độ phân chia rất nhanh của các vi khuẩn nên có thể tạo hàng triệu bản sao mong muốn trong một thời gian ngắn. Vì thế nhà khoa học có thể tách dòng bất kỳ một gene nào để

dùng cho nghiên cứu hoặc cho sản xuất trên quy mô công nghiệp một số lượng lớn các protein vốn là những chế phẩm y-sinh học nào đó.

## 2. Quy trình tạo dòng gene tái tổ hợp

Bây giờ ta xét một quy trình kỹ thuật tạo dòng mà việc phát hiện DNA tái tổ hợp dựa trên khả năng kháng thuốc do vector plasmid mang lại.

### • Bước 1: Tinh chiết DNA

Giả sử đã tinh chiết được plasmid có chứa hai gene kháng ampicillin và tetracyclin, ký hiệu là  $Amp^R$  và  $Tet^R$ ; và cũng giả thiết rằng gene  $Tet^R$  có chứa điểm cắt của EcoRI, và phân tử DNA người có mang gene insulin.

### • Bước 2: Kiến tạo phân tử DNA tái tổ hợp in vitro

Trước tiên, dùng enzyme giới hạn đầu dính EcoRI (xem bảng 10.1) để cắt vòng plasmid tại giữa gene  $Tet^R$  và cho cắt DNA người, trong số các đoạn bị cắt có một đoạn mang gene insulin. Sau đó đem trộn lẫn hai loại DNA trên trong ống nghiệm với DNA ligase. Kết quả là có thể xảy ra ba trường hợp: (1) Plasmid tự nối lại thành mạch vòng như lúc đầu; (2) Đoạn DNA tự nối lại thành mạch vòng; và (3) Plasmid tái tổ hợp có mang gene insulin, và có thể mang một đoạn DNA không phải gene đó.

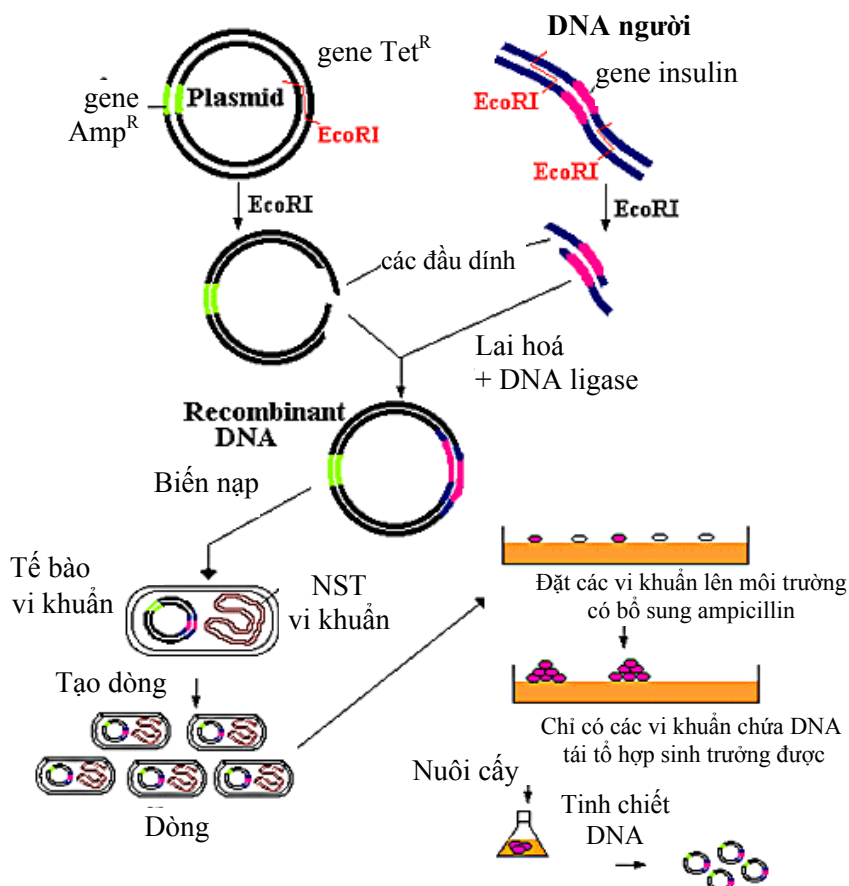
### • Bước 3: Biến nạp và phát hiện dòng DNA tái tổ hợp chung

Đưa các DNA được xử lý vào các tế bào *E. coli*. Nếu phân tử có kích thước lớn người ta phải xử lý vi khuẩn 'thể nhận' bằng chlorid calcium ( $CaCl_2$ ) để làm cho màng trở nên thấm được dễ dàng. Sau đó đem cấy riêng rẽ các vi khuẩn trên môi trường có ampicillin và theo dõi.

+ Nếu có xuất hiện khuẩn lạc (các vi khuẩn trong cùng một khuẩn lạc thì thuộc một dòng vì chúng bắt nguồn từ một vi khuẩn ban đầu), chứng tỏ vi khuẩn có mang gene  $Amp^R$ , tức là chúng có mang plasmid ban đầu (trường hợp 1) hoặc plasmid tái tổ hợp (trường hợp 3). Ngược lại, nếu chỗ cấy không xuất hiện khuẩn lạc, chứng tỏ vi khuẩn mang DNA tự nối (trường hợp 2).

+ Kế đó, đem cấy riêng rẽ các vi khuẩn thu được sang môi trường có tetracyclin. Nếu có xuất hiện khuẩn lạc, chứng tỏ vi khuẩn có mang gene  $Tet^R$  nguyên vẹn (trường hợp 1). Nếu không có khuẩn lạc, chứng tỏ vi khuẩn đem cấy có mang DNA tái tổ hợp (trường hợp 3); vì gene  $Tet^R$  bị bất hoạt do đoạn DNA xen vào. Bằng cách theo dõi như vậy cho phép xác định được dòng vi khuẩn mang DNA tái tổ hợp, nhưng vẫn chưa biết được đâu là các dòng đặc hiệu, nghĩa là có mang gene insulin.

Hình 10.5 dưới đây mô tả một quy trình đơn giản về tạo dòng vi khuẩn mang gene insulin người.

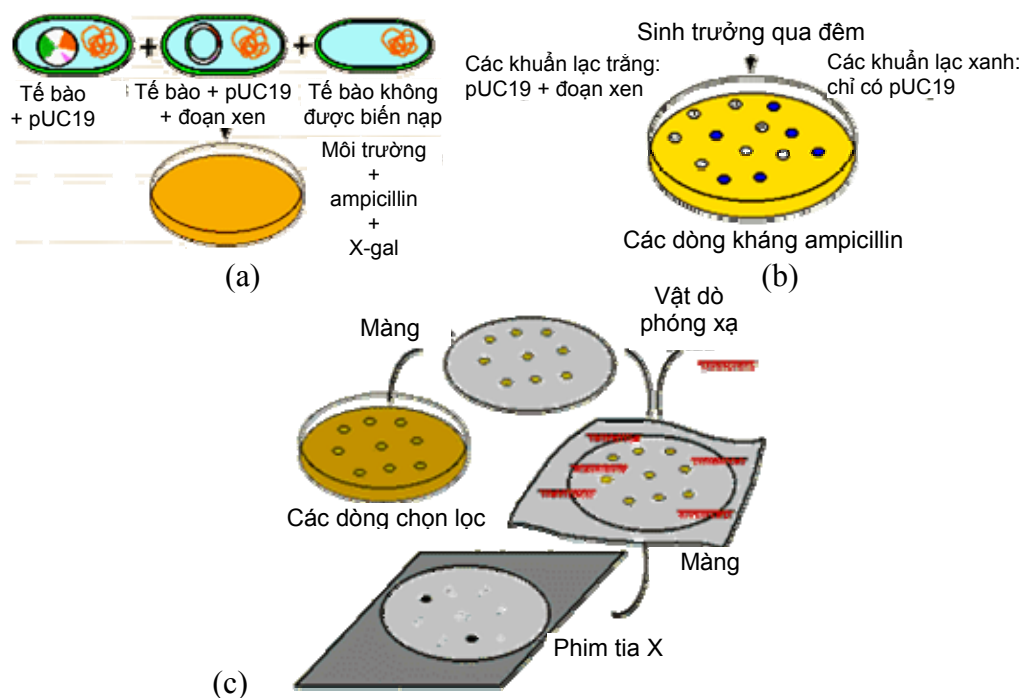


**Hình 10.5** Sơ đồ thí nghiệm tạo dòng vi khuẩn mang DNA tái tổ hợp, ở đây là gene insulin người.

• Bước 4: Chọn dòng DNA tái tổ hợp đặc hiệu

Về nguyên tắc, trong cả triệu phép thử mới có một tế bào mang gene mong muốn. Với trường hợp trên đây chẳng hạn, người ta có thể sử dụng phương pháp miễn dịch học bằng cách dùng kháng thể chống lại protein được sinh ra bởi dòng vi khuẩn tương ứng (tức huyết thanh tìm gene kháng insulin). Nói chung, để tìm dòng lai đặc hiệu người ta sử dụng các mẫu dò là mRNA hoặc rRNA đặc hiệu. Chẳng hạn, trong trường hợp nếu cần chọn dòng lai mang đoạn mRNA cụ thể, người ta đem cấy đều các dòng vi khuẩn có chứa DNA tái tổ hợp lên trên mặt thạch của hộp petri chứa môi trường nuôi cấy. Sau đó đóng dấu lên màng lọc *nitrocellulose*, và thu được bản sao. Việc xử lý bản sao bằng NaOH sẽ làm cho các tế bào vi khuẩn tan vỡ tại chỗ (*in situ*), và các DNA thoát ra từ chúng sẽ bị biến tính (các sợi đơn tách rời nhau) và dính vào màng lọc. Sau đó đem nhúng

màng lọc này vào mẫu mRNA tương ứng đã được tinh khiết và đánh dấu phóng xạ ( $P^{32}$ ); mẫu RNA này được gọi là *vật dò phóng xạ* (radioactive probe). Nếu dòng nào có chứa DNA mã hoá cho mRNA thì sẽ xảy ra hiện tượng lai giữa mRNA và vùng sợi đơn tương ứng trên DNA đó. Sau khi loại bỏ các mRNA không lai được, người ta đặt một miếng phim ảnh lên trên màng lọc; những vết ảnh xuất hiện trên ảnh phóng xạ tự ghi cho thấy vị trí của dòng mang DNA bổ trợ với mẫu RNA. Từ đây có thể tách riêng các dòng lai đặc hiệu để sử dụng cho các nghiên cứu tiếp theo.



**Hình 10.6** Xác định các dòng vi khuẩn mang plasmid có xen một đoạn DNA (gene) đặc hiệu bằng vật dò phóng xạ là mRNA - sản phẩm của nó.

Hình 10.6 minh họa một công đoạn của quy trình thí nghiệm DNA tái tổ hợp ở vi khuẩn *E. coli* khi sử dụng môi trường nuôi cấy có bổ sung ampicillin đối với ba kiểu tế bào: tế bào có mang plasmid tái tổ hợp, tế bào chỉ mang plasmid pUC19 không tái tổ hợp, và tế bào không biến nạp được (hình 10.6a). Qua đêm sinh trưởng, các tế bào nào có mang plasmid tái tổ hợp và plasmid không tái tổ hợp sẽ mọc thành các khuẩn lạc (màu sắc tương ứng ở đây là trắng và xanh; hình 10.6b). Sau khi chọn ra các dòng có xen plasmid tái tổ hợp, đưa lên màng lọc và cho tiến hành lai hóa giữa RNA và DNA (gene) của nó bằng các vật dò phóng xạ như đã nói ở trên. Sau đó đưa sản phẩm lai phân tử này lên tấm phim X quang để định vị

gene quan tâm từ dòng tái tổ hợp (hình 10.6c).

### 3. Tổng hợp và tạo dòng cDNA

Đối với trường hợp cần cho biểu hiện một gene lạ (sản xuất một protein) mong muốn trong vi khuẩn, người ta có thể tổng hợp gene của nó dựa trên khuôn mẫu mRNA và enzyme phiên mã ngược tinh chế từ các virus RNA (xem chương 5). Sau đó cho xen gene này vào plasmid, rồi đem cấy vào vi khuẩn và xác định các dòng cDNA đặc hiệu. Ở đây ta chỉ xét hai bước đầu:

#### Bước 1: Tổng hợp cDNA

Như đã biết, các mRNA eukaryote đều có cái đuôi poly(A) ở đầu 3'. Chính trình tự này tạo điều kiện thuận lợi cho việc tổng hợp sợi DNA bổ sung, cDNA. Khi đem trộn lẫn các đoạn ngắn gồm các nucleotide thymine (oligo(dT)) với mRNA này sẽ xảy ra sự lai hoá giữa nó với vùng đuôi mRNA. Đoạn oligo(dT) làm mồi cho *enzyme phiên mã ngược* (reverse transcriptase) tổng hợp sợi cDNA mà sản phẩm là sợi kép lai RNA-cDNA. Ở đầu 3' của sợi cDNA được tổng hợp có cái 'chóp' (trung tự đầu 5' của mRNA). Tiếp theo, bằng cách xử lý với NaOH, sợi mRNA bị loại ra; kể đó cái 'chóp' ở đầu 3' của cDNA lại làm mồi cho DNA polymerase I tổng hợp sợi thứ hai dọc theo sợi khuôn vốn có của nó. Sản phẩm cDNA bây giờ còn mang cái 'vòng' sợi đơn. Sau đó, cái vòng này được cắt bỏ bằng cách xử lý với nuclease S1 để tạo ra cDNA sợi kép.

#### Bước 2: Xen đoạn cDNA vào plasmid

Để xen đoạn cDNA vào plasmid người ta có thể dùng enzyme *end transferase* để gắn thêm "đuôi" homopolymer (ví dụ, CCCC....) vào các đầu 3' của cDNA. Và plasmid sau khi được mở vòng, cũng phải lắp thêm ở đầu 3' những trình tự tương ứng là GGGG... cũng với enzyme trên. Tuy nhiên, cách phổ biến hơn cả là gắn thêm vào cả hai 'đầu bằng' của sợi kép cDNA này bằng các oligonucleotide gồm 8-10 cặp base, nhờ xúc tác của DNA ligase T4. Sau đó, dùng enzyme giới hạn thích hợp để cắt "đoạn nối" này tạo ra các đầu dính. Đồng thời cắt plasmid bởi cùng một enzyme đó tại gene *Tet<sup>R</sup>*. Hai DNA nối trên được nối với nhau bằng DNA ligase để tạo ra plasmid lai.

### III. Các phương pháp biểu hiện các gene được tạo dòng

Một số các vector đã được sử dụng trong các thí nghiệm tạo dòng tái tổ hợp (như đã đề cập) vẫn có thể được sử dụng như là những *vector biểu hiện* (expression vectors). Các vector này có thể sản sinh ra các sản phẩm protein của các gene được tạo dòng. Chẳng hạn, các vector pUC và pBS được xen vào DNA dưới sự kiểm soát của *lac* promoter, vốn nằm phía

trước so với vị trí tạo dòng phức (multiple cloning site). Nếu như một đoạn DNA được cho xen có mặt trong cùng khung đọc mã thì nó làm gián đoạn gene *lacZ'*, sẽ sinh ra một *protein dung hợp* (fusion protein). Nó sẽ có một phần trong trình tự của protein beta-galactosidase tại đầu amin và trình tự protein khác nữa vốn được mã hóa trong DNA được xen vào, ở đầu carboxyl của nó. Tuy nhiên, nếu ta quan tâm tới sự biểu hiện cao của vector được tạo dòng, thì các vector chuyên biểu hiện thường hoạt động tốt hơn. Có hai yếu tố điển hình cần thiết cho sự biểu hiện gene có hoạt tính: một promoter mạnh và một vị trí bám của ribosome mà bao gồm luôn cả trình tự Shine-Dalgarno nằm gần codon khởi đầu AUG.

Trên thực tế, người ta sử dụng các *vector biểu hiện có các promoter mạnh* (expression vector with strong promoters), chẳng hạn như promoter của operon tryptophan. Nó tạo thành cơ sở cho nhiều vector biểu hiện kể cả *ptrpL1*.

Ngoài ra, người ta còn sử dụng các *vector biểu hiện dạng cảm ứng* (inducible expression vectors). Trường hợp này thường tiện lợi ở chỗ, nó giữ cho một gene được tạo dòng ở trạng thái đóng cho tới khi ta sẵn sàng cho nó biểu hiện. Một lý do là ở chỗ, các protein của eukaryote được sản sinh một số lượng lớn ở vi khuẩn có thể gây độc. Ngay cả các protein vốn không độc thực sự, chúng cũng có thể được tạo ra nhiều đến mức gây rối loạn sự sinh trưởng của vi khuẩn... Promoter của operon lactose (*lac* promoter) là vector biểu hiện kiểu cảm ứng đến một mức độ nào đó, có lẽ là vẫn giữ bất hoạt cho tới khi được kích hoạt bởi chất cảm ứng allolactose hoặc bằng chất tổng hợp tương tự của nó là IPTG. Tuy nhiên, sự biểu hiện vẫn kém bởi chất ức chế *lac* là không đầy đủ hoàn toàn, và sự biểu hiện nào đó của gene được tạo dòng vẫn có thể phát hiện được ngay cả khi không có mặt chất cảm ứng. Một cách xoay quanh vấn đề này là cho biểu hiện gene mong muốn trong một plasmid hay phagemid mà nó mang được gene *lacI* của riêng nó, như là plasmid pBS chẳng hạn.

Bây giờ chúng ta thử tìm hiểu một phương pháp sản xuất insulin người bằng con đường tổng hợp DNA và cho biểu hiện gen ở *E. coli*. Trước tiên cần lưu ý rằng, để thực hiện được điều này người ta phải dựa trên thành tựu mới nhất từ việc nghiên cứu cấu trúc chi tiết và quá trình tổng hợp các chế tiết insulin từ tuyến tụy vào máu. Nói vắn tắt thì sản phẩm sơ cấp của quá trình dịch mã từ phân tử mRNA hoàn chỉnh là *preproinsulin* gồm đoạn peptide "tín hiệu dẫn đầu" và chất tiền thân của insulin là *proinsulin*; đoạn pre- bị tách bỏ trong quá trình tổng hợp. Proinsulin được chế tiết là phân tử gồm ba đoạn A, B và C liền nhau trong một cấu trúc "hình quai" có ba cầu disulfur; khi đoạn peptid C (33 amino acid) bị cắt bỏ bởi enzyme



đặc thù trong các túi của tế bào tuyến tụy sẽ tạo ra các sản phẩm insulin có hoạt tính. Phân tử *insulin* gồm hai chuỗi polypeptid A (21aa) và B (30aa) riêng biệt được duy trì cùng nhau bởi hai cầu disulfur.

Từ đây ta có thể hình dung quá trình tổng hợp gene insulin nhân tạo (cDNA) và cho sản xuất hormone này ở *E. coli* như sau. Trước tiên, dùng mRNA của proinsulin làm khuôn để tổng hợp đầy đủ một DNA sợi kép bằng con đường phiên mã ngược như đã trình bày ở trước. Sau đó lắp thêm bộ ba khởi đầu ATG nhân tạo (mã hoá amino acid đầu chuỗi polipeptide) vào đầu 5' bằng phương pháp hoá học. Tiếp đến, cho nó kết hợp với một phần của operon lactose (gồm một đoạn của gene  $\beta$ -galactosidase và toàn bộ promoter) của *E. coli* để nó có thể hoạt động được trong tế bào thể nhận. Sau đó gene "lai" này được xen vào plasmid pBR322 (Ở đây không đi sâu vào các chi tiết kỹ thuật, chẳng hạn như sử dụng các đoạn nối, linker, và các enzyme giới hạn).

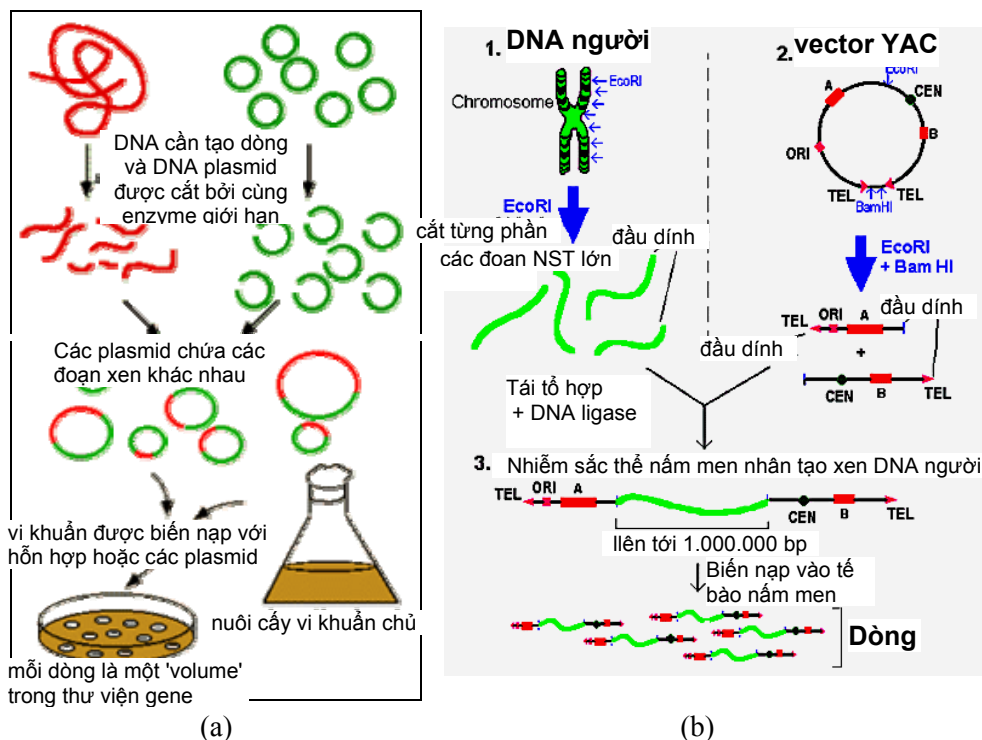
Khi đưa plasmid lai này vào vi khuẩn *E. coli*, nó sẽ sản sinh ra các protein lai gồm một đoạn peptide của  $\beta$ -galactosidase nối liền với phân tử proinsulin qua gốc methionine (MET). Sau khi phân lập protein này và xử lý *in vitro* bằng cyanogen bromide thì gốc MET bị cắt bỏ (kéo theo phần enzyme của vi khuẩn là  $\beta$ -galactosidase tách ra) và thu được proinsulin nguyên vẹn. Cuối cùng, nhờ xử lý với enzyme thích hợp, đoạn peptid C ở giữa proinsulin được tách ra và có thể thu được insulin ở dạng tinh khiết.

Cũng cần lưu ý rằng, hiện nay người ta đã tạo ra được các dòng vi khuẩn và các chủng nấm men mới đặc hiệu có khả năng sản xuất insulin trên quy mô công nghiệp với chỉ số insulin cao, và đặc biệt là, các chủng này có thể trực tiếp bài xuất sản phẩm đặc hiệu vào môi trường nuôi cấy. Trong trường hợp đó, các tế bào chuyên sản xuất insulin vẫn được duy trì và tái sử dụng với hiệu quả kinh tế cao.

## IV. Ứng dụng của công nghệ DNA tái tổ hợp

### 1. Công nghệ DNA tái tổ hợp với việc nghiên cứu bộ gene

Việc tách dòng tái tổ hợp cho phép nhận được một số lượng lớn bất kỳ gene hoặc vùng điều hoà nào để tiến hành phân tích trình tự nucleotide, xác định các vùng chức năng và chỉ ra các cơ chế hoạt động của chúng. Nhờ đó đã đưa lại những hiểu biết mới về tổ chức và hoạt động của các bộ gene prokaryote (như các khởi điểm tái bản, các vùng điều hoà phiên mã, các gene nhảy v.v.) và ở các bộ gene eukaryote (như các centromere, telomere, các gene phân đoạn, các gene già, DNA lặp lại v.v.) như đã được thảo luận ở các chương trước.



**Hình 10.7** Xây dựng thư viện gene hay thư viện bộ gene (gene or genomic library) bằng các thí nghiệm tạo dòng plasmid tái tổ hợp ở vi khuẩn (a) và sử dụng nhiễm sắc thể nấm men nhân tạo, YAC (b).

Bằng các thí nghiệm tạo dòng plasmid tái tổ hợp kinh điển ở vi khuẩn *E. coli*, và bằng cách cải tiến sử dụng *nhiễm sắc thể nấm men nhân tạo* này (yeast artificial chromosome = YAC; hình 10.7), người ta đã thành lập các *thư viện gene* (gene library) hay *thư viện bộ gene* (genomic library) để trên cơ sở đó tiến hành lập *bản đồ vật lý* (physical map) và xác định trình tự DNA bộ gene của nhiều sinh vật khác nhau, kể cả bộ gene người (NHGRI 2005). Nếu như vào năm 1977, F.Sanger xác định đầy đủ các trình tự phage  $\phi$ X174 (5386 nucleotide với 9 gene) và phage G4 (5577 cặp nucleotide với 10 gene), thì đến nay người ta xác định và lập bản đồ cho các DNA có kích thước lớn hơn nhiều; ví dụ: bộ gene phage lambda gồm 48.502 cặp base với khoảng 61 gene (1983), nhiễm sắc thể số 3 của nấm men *Saccharomyces cerevisiae* gồm 370.000 cặp base (1992) v.v...

Đáng kể là, *Dự án Bộ gene Người* (Human Genome Project = HGP; hình 10.8) đã chính thức đi vào hoạt động từ 1990 do James Watson chủ trì cùng với sự cộng tác của nhiều nhà khoa học trên thế giới. Việc phân tích trình tự bộ gene người đã được hoàn thành vào 4/2003, cho thấy bộ

gene (*genome*) của chúng ta có trình tự đầy đủ gồm 3.164.700.000 cặp base, chứa đựng khoảng 25.000 gene mã hóa protein chịu trách nhiệm xây dựng nên một bộ protein (*proteome*) gồm khoảng 50.000 protein khác nhau trong các tế bào. HGP thực sự là một trong những kỳ công thám hiểm vĩ đại nhất trong lịch sử; lần đầu tiên HGP cho phép chúng ta đọc được toàn bộ thông tin di truyền của tự nhiên đã làm nên con người (NHGRI 2005).



**Hình 10.8** Biểu tượng của HGP cho thấy tác động của dự án này lên nhiều lĩnh vực quan trọng của kinh tế - xã hội, các vấn đề về môi sinh, sức khỏe và đời sống con người trên phạm vi toàn cầu.

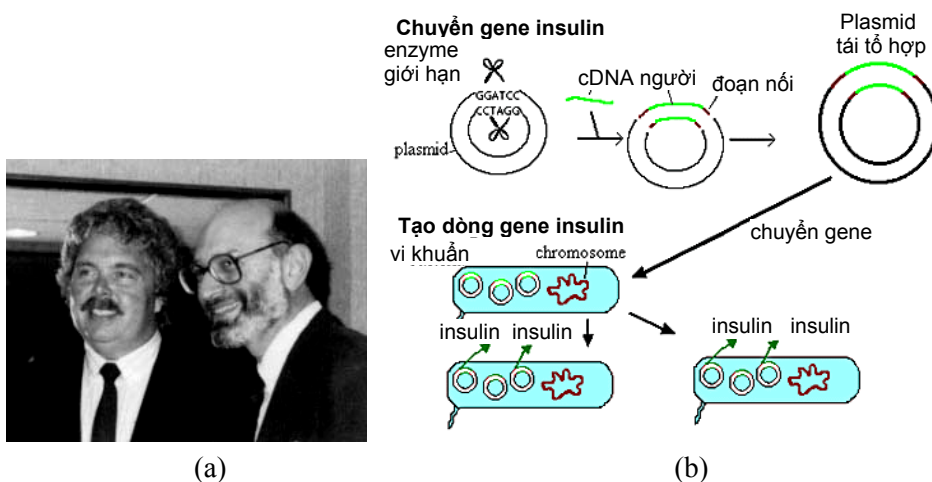
## 2. Công nghệ DNA tái tổ hợp với y-dược học

### 2.1. Sản xuất các chế phẩm y-sinh học bằng công nghệ DNA tái tổ hợp

Nếu như vào năm 1972, Giáo sư Roger Guillemin (France) đã phải dùng tới 500.000 não cừu mới tinh chiết được vài miligram *somatostatin*, một hormone sinh trưởng của vùng dưới đồi thị (với công trình này ông đã được trao giải Nobel năm 1980), thì vào 10/1977 Hebert Boyer (USA) đã chế tạo thành công hormone này từ *E. coli* bằng phương pháp DNA tái tổ hợp. Đến 8/1978 chính Boyer lại là người đầu tiên cho sản xuất thành công insulin người từ *E. coli* (hình 10.9). Các thành tựu đầu tiên này đã mở ra một kỷ nguyên mới phát triển rực rỡ nhất trong lịch sử công nghệ sinh học, công nghệ DNA tái tổ hợp.

Sau đó là hàng loạt các chế phẩm khác như các hormone tăng trưởng của người (human growth hormone = HGH), bò (BGH), lợn (PGH), các vaccine và các interferon phòng chống các căn bệnh nan ở người y như ung thư, viêm gan v.v. và một số bệnh quan trọng ở gia súc như hiện tượng 'lở mồm-long móng' do virus gây ra... tất cả đều được sản xuất từ *E. coli*. Đặc biệt là, sự ra đời của các hãng, các công ty lớn đã đầu tư mạnh mẽ cho sự phát triển của kỹ nghệ này. Nhờ vậy đã góp phần giải quyết một cách có hiệu quả nhiều vấn đề thực tiễn đặt ra trong y học, trong chăn nuôi-thú y, và nông nghiệp nói chung ...

Ở nước ta cũng đã có một số công trình nghiên cứu về các vấn đề này, chẳng hạn như nghiên cứu sự sai khác di truyền ở gene hormon sinh trưởng của một số giống gà Việt Nam. Qua đó cho thấy intron I của gene hormon sinh trưởng ở các giống gà Ri, gà Mía, gà Ấc, gà Hồ dài hơn kích thước dự đoán của gene hormone sinh trưởng gà đã được Tanaka công bố năm 1992 (Trần Xuân Hoàn 2004).



**Hình 10.9** (a) H. Boyer (trái) và S. Cohen; và (b) sơ đồ thí nghiệm tạo dòng và biểu hiện gene insulin người ở vi khuẩn *E. coli*.

Bên cạnh việc sản xuất các hormone, vaccine... nói trên, người ta còn sản xuất được các *kháng thể đơn dòng* (monocloning antibodies) dùng để: (i) xác định vi khuẩn gây bệnh (như thương hàn chẳng hạn); (ii) xác định mức hormone từ đó đánh giá chức năng tuyến nội tiết hoặc sự thay đổi trong quá trình tổng hợp hormone do khối u gây ra; (iii) phát hiện một số protein có ý nghĩa trong chẩn đoán khối u hoặc một số tình trạng trước khi sinh; (iv) phát hiện các thuốc bị cấm có trong máu, hoặc kiểm tra nồng độ thuốc trong máu và tổ chức nhằm đảm bảo liều thuốc sử dụng sao cho không vượt quá ngưỡng gây độc v.v. Nhờ có tính đặc hiệu và chính xác cao và sử dụng dễ dàng, việc sản xuất các kháng thể đơn dòng đã tạo nên một nhánh phát triển mau lẹ nhất của công nghệ sinh học, và cùng với việc sản xuất các vaccine chúng đã trở thành những phương tiện quan trọng nhất trong chính sách y tế cộng đồng của các nước đang phát triển và xét về lâu dài, việc ứng dụng các kháng thể đơn dòng có triển vọng chữa được nhiều bệnh trong đó có các khối u ác tính.

## 2.2. Ứng dụng kỹ thuật di truyền trong chẩn đoán và điều trị gene

Trong chẩn đoán bệnh ở mức phân tử, thành tựu nổi bật nhất là vào năm 1981, lần đầu tiên bệnh thiếu máu hồng cầu hình liềm được chẩn

đoán trước sinh ở mức độ gene nhờ phân tích DNA bằng enzyme giới hạn. Từ đây đã mở ra hai hướng chính trong chẩn đoán gene là: chẩn đoán các bất thường bẩm sinh và chẩn đoán các bất thường DNA soma. Một số thành tựu khác đạt được theo hướng đầu gồm có: bệnh hemophilia A, rối loạn dưỡng cơ Duchenne, hội chứng X-fragile, bệnh retinoblastoma - một dạng ung thư võng mạc... Theo hướng sau, người ta đã áp dụng đối với một số dạng ung thư máu như các lympho Burkitt, lympho nang, bệnh bạch cầu... Đáng kể là gần đây, người ta đã xác định được nhiều gene gây bệnh quan trọng ở người, như: gene gây *bệnh hóa xơ nang* (cystic fibrosis) năm 1990, gene gây bệnh Huntington năm 1993...

*Liệu pháp gene* (gene therapy) cũng là một phương thức sản xuất và điều trị mới bằng các phân tử trị liệu. Đây là hướng có nhiều triển vọng nhất nhưng khó thực hiện nhất vì nó có liên quan đến cả vấn đề phương pháp luận lẫn khía cạnh đạo lý. Sự kiện nổi bật nhất là vào năm 1990-91, W.F. Anderson, R.M. Blaese và Ken Coper thực hiện thành công ca liệu pháp gene đầu tiên trên một bé gái bốn tuổi mắc bệnh suy giảm miễn dịch phối hợp (SCID) với sự thiếu hụt *adenosine deaminase* (ADA), một enzyme cần thiết cho sản xuất các kháng thể trong các tế bào hệ miễn dịch, gọi là hội chứng "bubble-boy" (bệnh do gene ở gần đầu mút vai ngăn nhiễm sắc thể số 20 gây ra). Mặt khác, liệu pháp gene cũng đã mở ra những triển vọng to lớn trong chữa trị các căn bệnh phổ biến nhất, tác động tới hàng triệu triệu người trên hành tinh như: ung thư, SIDA/AIDS, viêm gan do virus, các bệnh tim mạch, các bệnh thoái hóa thần kinh (bệnh Parkinson, bệnh Huntington, bệnh Alzheimer...) hay cả những bệnh mạn tính như đa thấp khớp.

### 2.3. Kỹ thuật di truyền với hình pháp học và một số vấn đề xã hội khác

Kỹ thuật di truyền còn được ứng dụng rất hiệu quả trong các ngành hình pháp học, chẳng hạn bằng cách sử dụng kỹ thuật 'dấu vân DNA' (DNA fingerprinting) thay cho 'dấu vân tay' trước đây cho phép các ngành cảnh sát hình sự xác định chính xác các tội phạm hoặc nhận dạng xác nạn nhân trong chiến tranh hoặc do tai nạn gây ra (hình 10.10).

**Hình 10.10** Dấu vân DNA (*DNA fingerprinting*) có thể giúp các nhà nghiên cứu xác định khả nghi trong trường hợp tội phạm. Kiểu vạch ngang biểu thị bản chất di truyền của một người. Ở mẫu này cho thấy các băng của máu người mang mã số S2 khả nghi trùng khớp với băng chứng, mẫu máu E(vs).

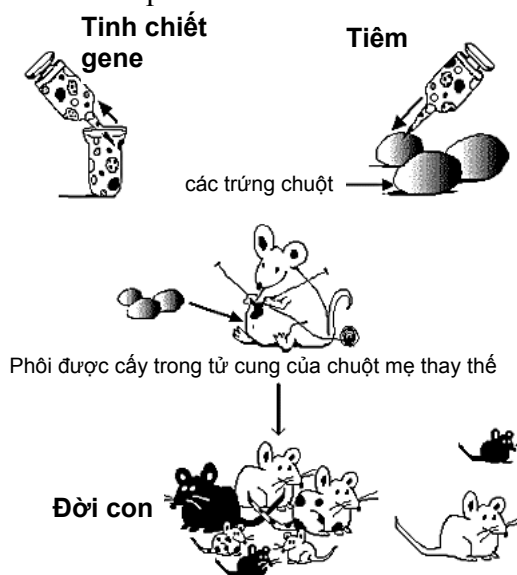


Ở nước ta, trong thời gian gần đây, kỹ thuật này đã và đang được áp dụng một cách hiệu quả trong ngành hình sự. Bên cạnh đó đã có công trình phân tích DNA nhận diện tế bào trên 103 người Việt Kinh bằng kỹ thuật PCR - SSO (Kit Innolipa). Qua phân tích so sánh khoảng cách gene học giữa người Việt Kinh và 11 sắc tộc châu Á và Đại dương khác, đã kết luận rằng người Việt Kinh gần với người Thái rồi đến người Manchu (Vũ Triệu An, 1999).

### 3. Kỹ thuật di truyền với các sinh vật biến đổi gene (genetically modified organisms = GMOs)

- Đối với ngành chăn nuôi, công nghệ sinh học nói chung và công nghệ sinh học nói riêng đã đạt nhiều thành tựu đáng kể, chẳng hạn các kỹ thuật chuyển ghép gene áp dụng cho hợp tử và phôi ở các gia súc nhằm tăng cường khả năng chống bệnh và cải thiện giống nói chung; cũng như các kỹ thuật mới trong xác định giới tính của phôi...

Hình 10.11 cho thấy khả năng ứng dụng kỹ thuật cấy ghép gene trên trứng chuột đã được thụ tinh. Sau đó đem phôi đã ghép gene cấy vào trong tử cung của con chuột làm mẹ khác. Kết quả là tạo ra được chuột con có bộ lông dạng khảm như mong muốn. Điển hình cho các thí nghiệm truyền gene ở động vật là vào năm 1982, R.D.Palmiter, R.L.Brinster và các đồng sự ở Đại học Seattle (Philadelphia, USA) bằng cách truyền gene xác định hormone sinh trưởng của chuột cống vào trứng đã thụ tinh của chuột bình thường, rồi cấy trở lại các tế bào



**Hình 10.11** Mô hình tổng quát về thí nghiệm truyền gene ở động vật (xem giải thích trong bài).

đã được biến đổi gene vào vòi trứng của các chuột cái có thể mang thai cho đến cùng. Và kết quả là, các tác giả này đã thu được dạng chuột nhất có kích thước lớn gấp 2-3 lần chuột bình thường, gọi là *chuột khổng lồ*.

- Đối với trồng trọt, việc sử dụng các phương pháp chuyển ghép gene đã đạt được nhiều thành tựu to lớn. Chẳng hạn, hãng Biogen (USA) năm

1984 đã chuyển thành công plasmid *Ti* vào tế bào thực vật; hãng Calgene và Phytogene (USA, 1984) đã ghép thành công gene kháng glyphosate để bảo vệ cây bông; năm 1985 hãng Molecular Genetics (USA) đã tạo được giống ngô mới cho nhiều tryptophan. Năm 1993, bằng kỹ thuật súng bắn gene vào tế bào thực vật người ta đã đưa được gene sản xuất protein diệt sâu vào cây ngô, và kết quả là đã tạo ra được giống ngô chống chịu cao đối với sâu đục thân. Điều thú vị là việc tách các gene cố định đạm, gene *Nif* (*Nif* = nitrogene fixation) từ các vi khuẩn nốt sần cây họ đậu và đưa vào bộ gene của các cây trồng khác để tạo ra các giống cây trồng mới có khả năng cố định nitơ và cho năng suất cao.

- Trong chọn giống vi sinh vật, người ta đã thực hiện thành công việc chuyển gene *cellulase* vào vi khuẩn (J.P.Aubert, France 1/1983), cải biến *E. coli* để sản xuất L-aspartat (hãng Tanabe, Japan 1985), ghép gene vào xạ khuẩn *S. violaceoniger* để cải tiến việc sản sinh enzyme glucoisomerase (hãng Roquette và Cayla, France 1985), ngoài ra còn tạo được các giống vi sinh vật biến đổi gene có khả năng ăn chặn dầu dùm trong xử lý các phế thải có độc tố nhằm bảo vệ môi sinh. Bên cạnh việc tạo ra giống nấm men mới có thể giết chết các vi khuẩn xuất hiện trong bia (hãng Suntory, 1985), còn tạo được chủng nấm men sản xuất insulin và interferon (A. Kimura, Japan 1986) v.v.

Nhận thức rõ ý nghĩa và tầm quan trọng của công nghệ DNA tái tổ hợp và công nghệ sinh học nói chung đối với sự phát triển kinh tế-xã hội đất nước ta trong thế kỷ XXI, Chính phủ đã ra Nghị quyết 18/CP ngày 11/4/1994 về "*Phương hướng phát triển của công nghệ sinh học ở Việt Nam đến năm 2010*". Đây được xem là một bước ngoặt quan trọng cho sự phát triển của công nghệ sinh học nước nhà trong thời gian qua và sắp tới.

## Câu hỏi và Bài tập

1. Thế nào là enzyme giới hạn? Chúng có những tính chất nào mà được coi là công cụ thiết yếu đối với công nghệ DNA tái tổ hợp? Bằng hai ví dụ về enzyme giới hạn, hãy chứng tỏ rằng ít nhất có hai phương pháp thành lập các phân tử DNA tái tổ hợp *in vitro*.

2. Từ các enzyme giới hạn BamHI, EcoRI và HindIII ở Bảng 10.1 và BglII có đoạn đích được biết là A↓GATCT, hãy cho biết cặp enzyme nào sẽ tạo ra các đầu dính hay đầu bổ sung (sticky/complementary ends) tương thích? Trình tự của đầu dính tương thích này là gì?

3. Giả sử bạn cắt hai DNA khác nhau, một với BamHI và một với BglII, sau đó nối chúng lại với nhau thông qua các đầu dính tương thích.

Một khi đã khâu nối rồi, bạn có thể tách hai DNA này lần nữa bằng một trong hai enzyme giới hạn đó hay không? Tại sao, hoặc tại sao không?

4. Giả sử bạn biến nạp một DNA tái tổ hợp cho một vi khuẩn và do nhầm lẫn, bạn đặt các tế bào được biến nạp đó trên môi trường không có chất kháng sinh nào cả. Kết quả mà bạn quan sát được là gì? Tại sao?

5. Giả sử rằng bạn chiết xuất được một enzyme cắt giới hạn từ loài vi khuẩn có tên khoa học là *Xenobacterium giganticus*. Theo danh pháp đã học, bạn sẽ viết tên enzyme đó như thế nào?

6. Cho biết trình tự của một đoạn DNA có chứa một vị trí nhận biết đối xứng xuôi ngược (palindromic recognition site) cho một enzyme giới hạn là GACGATATCAACT. Hãy tìm trình tự của vị trí nhận biết đó.

7. Thế nào là phân tử DNA tái tổ hợp, vector tách dòng? Hãy nêu các bước của quy trình tạo dòng gene tái tổ hợp và cho sơ đồ minh họa.

8. Giả sử tinh chế được plasmid pBR322 và phân tử DNA người có chứa một gene cần nghiên cứu. Hãy phân tích kỹ thuật tạo dòng gene nói trên ở *E. coli*.

9. Enzyme phiên mã ngược là gì? Nó được tinh chế từ loại sinh vật nào và được ứng dụng vào khâu nào trong công nghệ DNA tái tổ hợp? Giải thích và cho sơ đồ minh họa.

10. Có thể sử dụng các chất kháng sinh để xác định xem liệu plasmid pBR322 đã nhận được một đoạn xen ở trong vị trí EcoRI của nó hay không? Tại sao, hoặc tại sao không?

## Tài liệu Tham khảo

### Tiếng Việt

Trần Xuân Hoàn. 2004. *Nghiên cứu sự sai khác di truyền ở gene hormone sinh trưởng của một số giống gà Việt Nam*. Luận án Tiến sỹ Di truyền học, Thư viện Quốc gia, Hà Nội.

Lê Đình Lương. 2001. *Nguyên lý Kỹ thuật Di truyền*. NXB Khoa học và Kỹ thuật, Hà Nội.

Phan Cự Nhân. 1999. Công nghệ DNA tái tổ hợp. Trong: *Di truyền học tập II* (Phan Cự Nhân, chủ biên). Trang: 257-303. NXB Giáo Dục, Hà Nội.

Hoàng Trọng Phán. 1995. *Một số vấn đề về Di truyền học hiện đại* (Tài liệu BDTX cho giáo viên THPT chu kỳ 1993-1996). Trường ĐHSPT Huế.

Hoàng Trọng Phán. 1997. *Di truyền học Phân tử*. NXB Giáo Dục.

Hoàng Trọng Phán. 1999. "Giới thiệu công nghệ sinh học"; và "Cơ sở



khoa học của công nghệ sinh học". Trong: *Chuyên đề Công nghệ Sinh học* (Tài liệu BDTX giáo viên THPT chu kỳ 1997-2000; biên soạn chung với Nguyễn Bá Lộc và Biền Văn Minh). Trang: 56-96. Trường ĐHSP Huế.

### **Tiếng Anh**

Vu Trieu An 1999. Mitochondrial DNA polymorphism in the Vietnamese population. *Eur. J. of Immunogenetics*, 26: 471-422.

Campbell NA, Reece JB. 2001. *Essential Biology*. Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc, San Francisco, CA.

Collins FS, Green ED, Guttmacher AE DNA Guyer MS. 2003. A Vision for the Future of Genomics Research. *Nature*, Vol.422, No.6934, p.835-847.

Hartwell, Hood, Goldberg, Reynolds, Silver, Veres. 2004. *Genetics - From Genes to Genomes*. 2nd Edition. McGraw Hill, Inc., New York.

Lewis R. 2003. *Human Genetics: Concepts DNA Applications*. 5<sup>th</sup> ed, McGraw-Hill, Inc, NY.

Palladino MA. 2002. *Understanding the Human Genome Project*. Benjamin/Cummings Publishing Company, Inc, Menlo Park, CA.

Watson JD, Hopkins NH, Roberts JW, Steitz JA, Weiner AM. 1987. *Molecular Biology of the Gene*. 4<sup>th</sup> ed, Benjamin/Cummings Publishing Company, Inc, Menlo Park, CA.

Watson JD, Gilman M, Witkowski J, Zoller M. 1992. *Recombinant DNA*. 2<sup>nd</sup> ed, Scientific American Books, New York.

Weaver RF, Hedrick PW. 1997. *Genetics*. 3<sup>rd</sup> ed, McGraw-Hill Companies, Inc. Wm.C.Browm Publishers, Dubuque, IA.

### **Một số trang web**

Agricultural Biotechnology

<http://www.aphis.usda.gov/biotechnology/>

<http://www.agbioworld.org>

Online Mendelian Inheritance in Man (OMIM<sup>TM</sup>):

<http://www.ncbi.nlm.nih.gov/Omim/mimstats.html>

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>

National Human Genome Research Institute (NHGRI):

<http://www.ncbi.nlm.nih.gov/genome/guide/human>

The Institute for Genomic Research: <http://www.tigr.org>

Celera Genomics Corp.: <http://www.celera.com>

## Chương 11

# Di truyền học Người

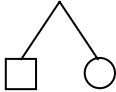
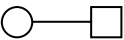
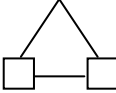





## I. Các phương pháp nghiên cứu di truyền học người

### 1. Phương pháp phân tích phả hệ (Genealogy analysis)

Phương pháp phân tích phả hệ được sử dụng để nghiên cứu sự di truyền các tính trạng người thuộc cùng dòng họ, xác định được tính trạng hoặc bệnh nào đó là trội hay lặn..., do một hay nhiều gene qui định, có tính chất di truyền hay không, di truyền độc lập hay liên kết với giới tính..., khả năng mắc bệnh của các thế hệ tiếp theo. Trong một số trường hợp còn xác định được người dị hợp tử mang gene bệnh. Phương pháp này kết hợp với các xét nghiệm khác cho phép có thể rút ra những lời khuyên về di truyền chính xác và hữu ích cho các gia đình về việc sinh con hoặc kết hôn.

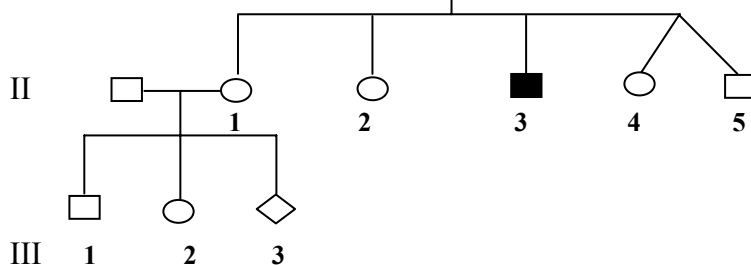
Phương pháp này được áp dụng khi biết được các tổ tiên trực tiếp và con cháu của người bệnh qua nhiều thế hệ.

### Một số ký hiệu dùng để lập phả hệ

- |   |  |   |   |
|---|--|---|---|
| <input type="checkbox"/>  | Nam bình thường  | <input checked="" type="checkbox"/>   | Nam bệnh  |
| <input type="circle"/>  | Nữ bình thường   | <input checked="" type="circle"/>   | Nữ bệnh   |
| <input type="diamond"/>   | Giới tính chưa được xác định                           |  | Cặp sinh đôi khác trứng   |
|  | Cặp vợ chồng   |  | Cặp sinh đôi cùng trứng   |
|  | Chết   |  | Thế dị hợp mang gene bệnh trên nhiễm sắc thể thường                                 |
|  |  |  |   |
| <input checked="" type="circle"/>   | Thế dị hợp mang gene bệnh liên kết với nhiễm sắc thể X |   |  |

Thế hệ I

Hôn nhân



Các con liệt kê từ trái sang phải theo trình tự đẻ

## 2. Phương pháp nghiên cứu trẻ sinh đôi

Có hai loại sinh đôi là sinh đôi cùng trứng (monozygotic twin-MZ) và sinh đôi khác trứng (dizygotic twin-DZ). Người ta dựa vào hàng loạt đặc điểm về số lượng và chất lượng để phân biệt trẻ sinh đôi cùng hay khác trứng: trẻ sinh đôi cùng trứng bắt buộc cùng giới, trẻ sinh đôi khác trứng có thể cùng hay khác giới; trẻ sinh cùng trứng có một màng ối chung, trẻ sinh khác trứng có màng ối khác nhau; trẻ sinh cùng trứng khi ghép mô thì luôn luôn thành công; có sự giống nhau ở trẻ sinh đôi cùng trứng và sự khác nhau ở trẻ sinh đôi khác trứng về nhiều tính trạng. Thông thường, chọn những tính trạng di truyền rõ rệt, ít bị biến đổi dưới ảnh hưởng của các nhân tố môi trường, thuộc những tính trạng này có nhóm máu, sắc tố mắt, da và tóc, nếp vân tay, chân. Trong đó phản ứng cấy ghép mô là phương pháp có thể kết luận một cách chính xác nhất.

So sánh các cặp sinh đôi về một tính trạng hoặc một bệnh nào đó cho phép đánh giá ảnh hưởng của yếu tố di truyền và yếu tố môi trường lên sự hình thành tính trạng, phát hiện các biến dị xảy ra do yếu tố môi trường...

Một bệnh hoặc một tính trạng di truyền nào đó có thể biểu hiện ở cả hai thành viên của cặp sinh đôi (có tương hợp) hoặc cũng có khi chỉ biểu hiện ở một trong hai thành viên của cặp sinh đôi (không tương hợp).

Ở các cặp sinh đôi cùng trứng nếu tương hợp càng lớn thì vai trò của yếu tố di truyền càng lớn. Nếu tương hợp càng nhỏ thì vai trò của yếu tố di truyền càng kém.

## 3. Phương pháp di truyền tế bào học người

Đây là phương pháp được dùng phổ biến hiện nay để phát hiện và quan sát nhiễm sắc thể, qua đó xác định các dị dạng nhiễm sắc thể, các hiện tượng lệch bội, hiện tượng cấu trúc lại nhiễm sắc thể dẫn đến nhiều bệnh di truyền hiếm nghèo ở người.

Dùng mô gồm nhiều tế bào đang phân chia mạnh như mô tủy xương, mô bào thai, mô tinh hoàn, khối u ác tính... làm tiêu bản để phân tích, đánh giá nhiễm sắc thể.

Kỹ thuật lai tế bào soma và kỹ thuật hiện băng nhiễm sắc thể ra đời đã cho phép nghiên cứu cơ chế ung thư, hoạt động của gene trong quá trình phát triển cá thể và góp phần tích cực vào việc lập bản đồ di truyền người.

## 4. Phương pháp nghiên cứu quần thể

Phương pháp này dựa vào phương trình Hardy - Weinberg, đánh giá tần số các kiểu hình để tính mật độ các gene trong quần thể liên quan đến các bệnh di truyền. Nó còn cho phép đánh giá các hậu quả của giao phối cận huyết và theo dõi sự di truyền của các quần thể người về mặt nguồn

gốc.

### 5. Các kỹ thuật sinh học phân tử

Các thành tựu to lớn trong những năm gần đây về lĩnh vực Di truyền học người, đặc biệt là thành tựu giải mã bộ gene người đạt được là nhờ sử dụng các kỹ thuật sinh học phân tử như tách chiết, phân tích định tính và định lượng nucleic acid; các phương pháp lai phân tử: Southern blot, Northern blot, lai tại chỗ (in situ hybridization),... ; các phương pháp xác định trình tự nucleic acid; tạo dòng (cloning); xây dựng thư viện bộ gene, thư viện cDNA; phương pháp PCR (polymerase chain reaction); Sinh tin (Bioinformatics),...

Ngoài các phương pháp trên người ta còn sử dụng các phương pháp nghiên cứu mô phỏng học, phương pháp in và phân tích nếp vân da, phương pháp điều tra dịch tễ học và phương pháp di truyền lâm sàng... trong nghiên cứu Di truyền học người.

## II. Các phương pháp lập bản đồ di truyền người

### 1. Phân tích liên kết (Linkage analysis)

Sự trao đổi chéo giữa các locus trên cùng nhiễm sắc thể có thể tạo ra tái tổ hợp. Việc đánh giá tần số tái tổ hợp có thể thực hiện dựa vào quan sát sự di truyền các allele trong các gia đình. Sự xác định các phase liên kết (có nghĩa là nhiễm sắc thể mà trên đó mỗi một allele được định vị) là một phần quan trọng của phương pháp này.

Mặc dù có sự tương quan giữa centiMorgan và khoảng cách vật lý thật sự giữa các locus, mối quan hệ này bị phức tạp do các sai khác về giới tính trong tái tổ hợp, các tần số tái tổ hợp cao hơn ở gần các telomere và sự tồn tại của các điểm nóng tái tổ hợp (recombination hot spots).

Sự sai khác về thống kê của 2 locus có thể được tính toán bằng cách tính tỉ số của hai hợp lẽ (likelihood): hợp lẽ của liên kết ở tần số tái tổ hợp đã cho chia cho hợp lẽ của không liên kết. Logarithm của tỉ số sai khác này được ký hiệu là LOD (logarithm of odds).  $LOD > 0,3$ : có liên kết;  $LOD < -0,2$ : không liên kết.

Khoảng cách giữa các locus được tính bằng đơn vị centiMorgan (cM). 1 cM tương ứng với tần số tái tổ hợp xấp xỉ 1%.

Phân tích liên kết cho phép chúng ta xác định được khoảng cách tương đối giữa các locus nhưng không xác định được các vị trí đặc trưng với marker hoặc các gene bệnh.

### 2. Các phương pháp lập bản đồ vật lý (physical mapping)

#### 2.1. Phương pháp lập bản đồ mất đoạn (deletion mapping)

Karyotype của các bệnh nhân mắc bệnh di truyền tinh thoảng được phát hiện thấy hiện tượng mất đoạn một vùng đặc trưng trên nhiễm sắc thể. Điều này cho thấy rằng locus gây bệnh có thể nằm trong vùng bị mất. Chiều dài đoạn bị mất có thể biến đổi ở một số bệnh nhân mắc cùng một bệnh. Các đoạn mất của nhiều bệnh nhân được so sánh để xác định vùng bị mất ở tất cả các bệnh nhân, do đó thu hẹp lại vị trí của gene bệnh. Phương pháp lập bản đồ mất đoạn đã được sử dụng để xác định vị trí các gene chịu trách nhiệm đối với ung thư võng mạc (retinoplastoma), hội chứng Prader Willi và Angelman và khối u Wilms. Khối u Wilms là một khối u thận ở giai đoạn đầu do đột biến nhiễm sắc thể số 11 gây ra. Chú ý rằng các mất đoạn này chỉ xảy ra ở một nhiễm sắc thể trong cặp tương đồng, tạo ra bệnh nhân dị hợp tử đối với đoạn bị mất. Nếu đoạn bị mất đủ lớn để có thể quan sát dễ dàng dưới kính hiển vi và xảy ra trên cả hai nhiễm sắc thể của cặp tương đồng thì thường gây chết.

## 2. 2. Phương pháp lai tại chỗ (*In situ* hybridization)

Đoạn DNA thu được từ marker đa hình hoặc gene bệnh mà chúng ta muốn biết vị trí được tạo dòng với mẫu dò (probe) bằng kỹ thuật DNA tái tổ hợp. Mẫu dò được đánh dấu với chất phóng xạ như tritium chẳng hạn. Sau đó mẫu dò được đặt vào một màng lai chứa nhiễm sắc thể kỳ giữa mà DNA của nó đã biến tính. Mẫu dò phóng xạ sẽ bắt cặp bổ sung với DNA đã biến tính của đoạn nhiễm sắc thể đặc trưng. Bởi vì mẫu dò phát ra bức xạ nên vị trí của nó có thể được xác định một cách chính xác bằng cách đặt một phim nhạy cảm tia X lên trên màng lai (phóng xạ tự ghi (autoradiography)).

Phương pháp lai tại chỗ sử dụng các mẫu dò phóng xạ thường diễn ra chậm và mức phân tích khá thấp (2-5 Mb). Kỹ thuật FISH (fluorescence in situ hybridization) sử dụng mẫu dò huỳnh quang thay cho mẫu dò phóng xạ là nhanh hơn và cho phép phân tích vùng của mẫu dò tốt hơn, thường là ở trong khoảng 1 Mb đối với nhiễm sắc thể kỳ giữa. Bằng cách sử dụng nhiễm sắc thể kỳ trung gian, các nhiễm sắc thể ở kỳ này kết xoắn ít hơn các nhiễm sắc thể kỳ giữa, phân tích FISH có thể được tăng lên từ 25-250 kb.

## 2.3. Lai tế bào soma

Các tế bào soma của các loài khác nhau khi sinh trưởng trong cùng môi trường nuôi cấy với sự có mặt của một số tác nhân như polyethylene glycol hoặc virus Sendai, tinh thoảng sẽ dung hợp với nhau tạo thành các tế bào lai. Khi lai các tế bào chuột và tế bào người với nhau sẽ thu được các tế bào lai chứa 86 nhiễm sắc thể : 46 nhiễm sắc thể người và 40 nhiễm sắc thể chuột. Sau đó các tế bào lai được phép sao chép. Các tế bào này bắt đầu mất đi một số nhiễm sắc thể người khi chúng trải qua quá trình nguyên

phân. Cuối cùng các tế bào còn lại bộ nhiễm sắc thể đầy đủ của chuột và chỉ một hoặc một số nhiễm sắc thể người. Các tế bào này được lập karyotype để xác định nhiễm sắc thể nào của người còn lại (nhiễm sắc thể của người và chuột có thể được phân biệt với nhau dựa vào kích thước và các phương pháp hiện băng). Sau đó các tế bào này được nghiên cứu để xác định tế bào nào mang gene thích hợp đang nói đến. Ví dụ nếu gene luôn luôn được tìm thấy trong các tế bào chứa nhiễm sắc thể số 1 của người nhưng chưa bao giờ được phát hiện trong các tế bào không mang nhiễm sắc thể số 1, chúng ta có thể kết luận rằng gene này phải được định vị trên nhiễm sắc thể số 1.

Sự có mặt của gene trong tế bào lai có thể được phát hiện bằng một số phương pháp. Nếu gene mã hoá một enzyme được sản xuất ra bởi tế bào thì một xét nghiệm enzyme có thể được sử dụng. Điện di protein được sử dụng để nhận ra một sản phẩm protein người và để phân biệt nó với sản phẩm protein tương đương của chuột. Tuy nhiên các phương pháp này yêu cầu một sản phẩm protein đã biết phải được phát hiện. Thông thường hơn, hiện nay các nhà nghiên cứu kiểm tra sự lai của các dòng tế bào dung hợp với một mẫu dò đánh dấu phóng xạ mang DNA quan tâm bằng cách sử dụng phương pháp Southern blotting hoặc kỹ thuật PCR.

#### 2.4. Lập bản đồ lai phóng xạ (Radiation hybrid mapping)

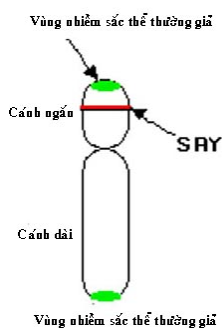
Phương pháp này bắt đầu với một thể lai tế bào soma chứa một nhiễm sắc thể đơn của người. Sau đó các tế bào này được chiếu với tia X hoặc tia  $\gamma$  để làm đứt gãy sợi kép nhiễm sắc thể. Bức xạ này giết chết tế bào vì vậy chúng phải được dung hợp với tế bào động vật gặm nhấm để tồn tại. Một số tế bào lai được tạo ra gọi là các thể lai phóng xạ (radiation hybrid) mang các đoạn nhỏ nhiễm sắc thể người đã lai với nhiễm sắc thể gặm nhấm. Sự có mặt của nhiễm sắc thể người có thể được phát hiện bằng sự sàng lọc đối với các trình tự *Alu*. Các trình tự *Alu* này tìm thấy ở mỗi một đoạn có kích thước vài kb trong nhiễm sắc thể người nhưng không tìm thấy ở trong nhiễm sắc thể của gặm nhấm. Bởi vì bức xạ bẻ gãy nhiễm sắc thể người ở những khoảng cách ngẫu nhiên, các locus được định vị gần với một locus khác hơn sẽ được tìm thấy thường hơn trên cùng đoạn nhiễm sắc thể (có xu hướng cùng đi với nhau). Kỹ thuật PCR có thể được sử dụng để xác định thể lai phóng xạ mang các tổ hợp đặc trưng của các locus người. Sau đó sử dụng phương pháp thống kê để đánh giá mỗi một cặp locus thường được tìm thấy trên cùng đoạn nhiễm sắc thể như thế nào. Kết quả này cung cấp sự đánh giá khoảng cách tương đối giữa các locus.

Ngoài các phương pháp trên người ta còn dùng các phương pháp tạo dòng định vị (positional cloning), phân tích sự bảo tồn của các loài lai

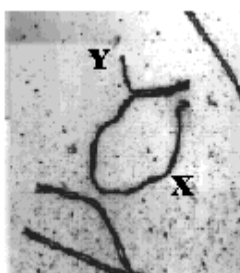
(analysis of cross-species conservation), phân tích trình tự DNA bằng máy tính, sàng lọc các đột biến ở trong trình tự (screen for mutations in the sequence), kiểm tra sự biểu hiện của gene....

### III. Nhiễm sắc thể Y và chất nhiễm sắc giới tính của người

#### 1. Nhiễm sắc thể Y của người



**Hình 11.1**  
Nhiễm sắc thể Y của người



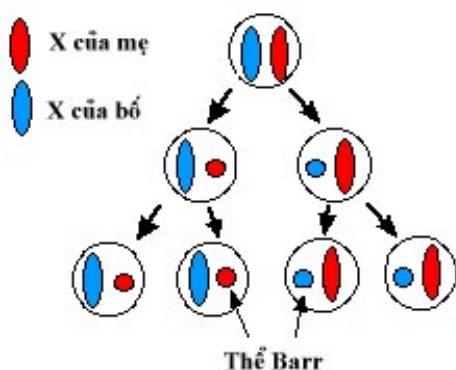
**Hình 11.2**  
Sự tiếp hợp của nhiễm sắc thể X và Y

Trái ngược với nhiễm sắc thể X, nhiễm sắc thể Y hoàn toàn nhỏ và chứa rất ít gene. Phần lớn chiều dài của nhiễm sắc thể Y không xảy ra tái tổ hợp là chất dị nhiễm sắc. Các đoạn lặp DNA của nó đã được xem xét tỉ mỉ.

Các vùng tương đồng của cặp nhiễm sắc thể XY được định vị ở mỗi đầu của nhiễm sắc thể Y. Các vùng này kết cặp và tái tổ hợp với nhiễm sắc thể X trong kỳ đầu của giảm phân I. Chúng được gọi là các vùng nhiễm sắc thể thường giả (pseudoautosomal region-PAR) bởi vì bất kỳ gene nào định vị trong các vùng này (cho đến nay chỉ mới phát hiện được 9 gene) đều được di truyền hoàn toàn giống như gene trên nhiễm sắc thể thường. Nam có hai bản sao của các gene này: một ở trong PAR của Y và một ở trong vùng tương đồng của X.

Khoảng 95% chiều dài của Y nằm giữa các PAR tạo ra vùng không tiếp hợp (non-recombining region-NRY). Gần 80 gene đã được tìm thấy ở đây. Chúng được xếp thành 2 nhóm chính. Các gene thuộc nhóm thứ nhất (mã hoá các protein sử dụng cho tất cả các tế bào (cả hai giới)) được biểu hiện ở nhiều mô. Các gene giữ nhà này (housekeeping gene) có thể tương đồng trên X mà tránh khỏi sự bất hoạt X ở nữ. Các gene thuộc nhóm thứ hai mã hoá các protein hoạt động chức năng chỉ ở trong tinh hoàn và không có thể tương đồng trên X. Một gene có vai trò quan trọng trong nhóm này là SRY (sex-determining region Y). SRY định vị trên cánh ngắn nằm ở bên ngoài PAR và gây ra các khả năng biến đổi phôi thành nam giới.

## 2. Chất nhiễm sắc giới tính của người



**Hình 11.3**

### Sự bất hoạt của nhiễm sắc thể X

Đã từ lâu người ta đã biết rằng nữ có hai nhiễm sắc thể X trong khi đó nam chỉ có một nhiễm sắc thể X. Như vậy mỗi gene liên kết với nhiễm sắc thể X có hai bản sao đối với nữ và một bản sao đối với nam. Nếu tất cả các gene trên nhiễm sắc thể X đều hoạt động thì ở nữ hàm lượng các sản phẩm (như enzym chẳng hạn) do các gene này mã hóa phải nhiều gấp đôi nam. Trong thực tế hàm lượng các sản phẩm trên đều bằng nhau ở cả hai giới. Có thể giải thích điều này như thế nào ?

Vào đầu thập niên 1960, Mary Lyon đã giả thuyết rằng một nhiễm sắc thể X trong mỗi tế bào soma của nữ bị bất hoạt. Hiện tượng này dẫn đến kết quả được gọi là “ sự bù đắp về liều lượng gene” (Gene dosage compensation) làm cho sản phẩm của gene liên kết với nhiễm sắc thể X của nam và nữ ngang nhau. Giả thuyết Lyon cho rằng sự bất hoạt nhiễm sắc thể X xảy ra sớm trong giai đoạn phát triển phôi ở giống cái. Ở một số tế bào, nhiễm sắc thể X bị bất hoạt có nguồn gốc từ bố, trong khi ở những tế bào khác nhiễm sắc thể X bị bất hoạt có nguồn gốc từ mẹ. Quá trình bất hoạt này là ngẫu nhiên. Khi một nhiễm sắc thể X bị bất hoạt trong một tế bào, nó sẽ duy trì sự bất hoạt này ở tất cả các thế hệ tế bào con cháu. Như vậy sự bất hoạt nhiễm sắc thể X là ngẫu nhiên nhưng lại được cố định.

Sự bất hoạt nhiễm sắc thể X sẽ dẫn đến kết quả là ở tất cả cơ thể cái bình thường có hai quần thể tế bào phân biệt: một quần thể có nhiễm sắc thể X hoạt động nhận được từ bố và một quần thể có nhiễm sắc thể X hoạt động nhận được từ mẹ. Với hai quần thể tế bào này, giống cái là dạng khảm đối với nhiễm sắc thể X, giống đực chỉ có một bản sao của nhiễm sắc thể X là dạng bán hợp tử (hemizygous) đối với nhiễm sắc thể X.

Giả thuyết Lyon đã đưa ra một số bằng chứng, phần lớn nhận được từ các nghiên cứu về động vật và người.

Các nghiên cứu của Di truyền học trong thập kỷ 1940 đã cho thấy rằng các tế bào gian kỳ của mèo cái thường chứa một khối chất nhiễm sắc dày đặc bắt màu thuốc nhuộm ở trong nhân tế bào. Các khối này chưa bao giờ tìm thấy ở giống đực. Chúng được gọi là thể Barr, do Murray Barr, một nhà Di truyền học đã mô tả chúng. Barr đã giải thích rằng thể Barr tương



ứng với một nhiễm sắc thể X kết đặc cao. Trạng thái kết đặc này phản ánh thực tế là ADN của nó được tái bản chậm hơn ADN các nhiễm sắc thể khác trong pha S.

Các nghiên cứu tiếp theo đã xác định mARN được sao chép chỉ từ một nhiễm sắc thể X trong tế bào của con cái bình thường. Quá trình bất hoạt xảy ra khoảng hai tuần sau khi thụ tinh. Quá trình được khởi đầu ở một vị trí đơn (single location) trên cánh dài của nhiễm sắc thể X (trung tâm bất hoạt X) và sau đó trải dài dọc theo nhiễm sắc thể. Sự bất hoạt của nhiễm sắc thể X là thường xuyên đối với tất cả các tế bào soma ở con cái, nhưng nó phải trở nên hoạt hóa trong dòng tế bào mầm của cơ thể cái, do vậy mỗi tế bào trứng sẽ nhận được một bản sao của nhiễm sắc thể X hoạt động.

Theo giả thuyết Lyon, số lượng thể Barr trong các tế bào soma là luôn luôn bằng số nhiễm sắc thể X trừ đi 1. Cơ thể cái bình thường có một thể Barr trong tế bào soma trong khi các con đực lại không có. Nam bị hội chứng Klinefelter (XXY) có một thể Bar. Dạng hội chứng này đã dẫn đến một câu hỏi: Nếu nhiễm sắc thể X thừa bị bất hoạt thì tại sao người mang nhiễm sắc thể X thừa có kiểu hình không bình thường? Câu trả lời là sự bất hoạt của nhiễm sắc thể X là không hoàn toàn. Một vài vùng của nhiễm sắc thể X bất hoạt vẫn hoạt động ở tất cả các bản sao. Một nghiên cứu mới đây cho thấy 20% gene của nhiễm sắc thể X bất hoạt có thể vẫn hoạt động. Các gene này là tương đồng với các gene trên nhiễm sắc thể Y và các bản sao của chúng đã góp phần tạo nên sự không bình thường của kiểu hình.

Trung tâm bất hoạt X chứa gene cần thiết cho sự bất hoạt có tên là XIST (X-inactive specific transcript), gene này được sao mã thành mARN 15-17 kb nhưng không được dịch mã.

Một điểm đặc biệt khác của nhiễm sắc thể X bất hoạt là sự methyl hóa, có khả năng là sự methyl hóa chịu trách nhiệm duy trì sự bất hoạt của một nhiễm sắc thể X đặc biệt trong tế bào.

#### **IV. Sự di truyền các gene trội lặn trên nhiễm sắc thể thường và nhiễm sắc thể giới tính**

##### *1. Sự di truyền các gene trội lặn trên nhiễm sắc thể thường*

###### **1.1. Sự di truyền các gene trội trên nhiễm sắc thể thường**

Bệnh di truyền do các gene trội trên nhiễm sắc thể thường được bắt gặp với tần số khoảng 1/200. Nếu tính riêng lẻ thì mỗi bệnh di truyền do gene trội trên nhiễm sắc thể thường là khá hiếm trong quần thể. Tuy nhiên, các bệnh phổ biến nhất có tần số gene khoảng 0,001. Tính trạng hay bệnh di truyền do các gene trội trên nhiễm sắc thể thường có một số các đặc

điểm quan trọng. Thứ nhất, cả hai giới biểu hiện tính trạng với tỉ lệ gần như nhau. Bố và mẹ có vai trò ngang nhau trong việc truyền tính trạng cho con cái. Thứ hai, không có sự ngắt quãng thế hệ: kiểu hình bệnh được truyền từ ông bà sang bố mẹ, rồi từ bố mẹ cho con cái...Nếu bố mẹ không bị bệnh thì sẽ không truyền bệnh cho con cái. Vì vậy kiểu biểu hiện này được gọi là kiểu di truyền phân bố dọc (vertical transmission pattern). Thứ ba, mặc dù sự truyền tính trạng từ bố cho con trai không được yêu cầu để đánh giá sự di truyền gene trội trên nhiễm sắc thể thường nhưng sự hiện diện của nó trong một phả hệ loại trừ chắc chắn các kiểu di truyền khác (đặc biệt là sự di truyền liên kết với nhiễm sắc thể X). Cuối cùng, thể dị hợp tử mang gene bệnh truyền tính trạng này cho gần một nửa con cái của họ.

## 1.2. Sự di truyền các gene lặn trên nhiễm sắc thể thường

Giống như bệnh di truyền do các gene trội trên nhiễm sắc thể thường, các bệnh di truyền do gene lặn trên nhiễm sắc thể thường là khá hiếm trong quần thể. Chỉ khi ở trạng thái đồng hợp thì mới biểu hiện thành kiểu hình. Các thể dị hợp là phổ biến nhiều hơn thể đồng hợp. Sự di truyền các gene lặn trên nhiễm sắc thể thường có một số đặc điểm quan trọng. Thứ nhất, kiểu hình bệnh biểu hiện ở một hoặc nhiều anh chị em nhưng không phát hiện thấy ở các thế hệ trước đó. Do vậy kiểu di truyền này được gọi là kiểu di truyền phân bố ngang (horizontal transmission pattern). Thứ hai, giống như trường hợp di truyền do gene trội trên nhiễm sắc thể thường, nam và nữ đều có khả năng bị mắc bệnh với tỉ lệ ngang nhau. Thứ ba, trung bình khoảng 1/4 con cái của cặp bố mẹ dị hợp sẽ bị mắc bệnh. Cuối cùng, tính đồng huyết (consanguinity) thường hiện diện trong các phả hệ liên quan đến bệnh do gene lặn trên nhiễm sắc thể thường nhiều hơn so với các phả hệ liên quan đến các kiểu di truyền khác.

## 2. Sự di truyền các gene trội lặn trên nhiễm sắc thể giới tính

### 2.1. Sự di truyền gene lặn liên kết với nhiễm sắc thể X

Một số các tính trạng và bệnh nổi tiếng là do các gene lặn liên kết với nhiễm sắc thể X gây ra. Các kiểu di truyền do gene lặn liên kết với nhiễm sắc thể X chắc chắn là khác với gene trên NST thường. Bởi vì nữ được di truyền hai bản sao nhiễm sắc thể X. Ở nữ, tính trạng do gene lặn liên kết với nhiễm sắc thể X rất giống với trường hợp tính trạng do gene lặn trên nhiễm sắc thể thường quy định. Tuy nhiên, chỉ có một nhiễm sắc thể X là bất hoạt trong tế bào, như vậy một nửa số tế bào của nữ dị hợp tử sẽ biểu hiện allele bệnh và nửa còn lại sẽ biểu hiện allele bình thường. Trong lúc đó ở nam được di truyền một allele bệnh lặn liên kết với nhiễm sắc thể X thì sẽ biểu hiện bệnh vì nhiễm sắc thể Y không mang allele tương ứng.

Một bệnh do gene lặn liên kết với nhiễm sắc thể X với tần số  $q$  thì tất cả nam có đột biến sẽ biểu hiện bệnh. Nữ phải cần hai bản sao của allele đột biến để biểu hiện bệnh do đó tần số biểu hiện sẽ là  $q^2$  như trường hợp bệnh do gene lặn trên nhiễm sắc thể thường quy định. Kết quả này cho thấy rằng nam bị tác động một cách thường xuyên và nhiều hơn so với nữ đối với bệnh do gene lặn liên kết với nhiễm sắc thể X.

Một số đặc điểm để phân biệt tính trạng hoặc bệnh do lặn liên kết với nhiễm sắc thể X với tính trạng hoặc bệnh do gene lặn trên nhiễm sắc thể thường:

- Các gene liên kết với nhiễm sắc thể X không được truyền trực tiếp từ bố cho con trai.

- Gene liên kết với nhiễm sắc thể X có thể được truyền từ bố cho tất cả con gái, những con gái này không có biểu hiện bệnh nhưng có mang gene bệnh. Các con gái mang gene bệnh này sẽ truyền lại cho một nửa số con trai của họ, các con trai này bị bệnh. Kiểu di truyền này gây ra sự xuất hiện các thế hệ “ngắt quãng” (skipped generations).

Kiểu ghép đôi phổ biến đối với các gene lặn liên kết với nhiễm sắc thể giới tính X là sự kết hợp của mẹ mang gene bệnh với bố bình thường. Mẹ sẽ truyền gene bệnh cho một nửa số con trai và một nửa số con gái.

Một kiểu ghép đôi phổ biến khác là giữa bố bị bệnh với mẹ bình thường. Tất cả con trai của họ bình thường. Bởi vì tất cả con gái phải nhận nhiễm sắc thể X của bố do đó chúng sẽ trở thành thể mang gene bệnh dị hợp tử. Như vậy không có một đứa con nào biểu hiện bệnh.

Kiểu ghép đôi ít phổ biến hơn là giữa bố bị bệnh với mẹ là thể mang gene bệnh. Một nửa số con gái của họ sẽ trở thành thể mang gene bệnh dị hợp tử, một nửa sẽ đồng hợp tử gene bệnh và sẽ biểu hiện bệnh. Một nửa số con trai bình thường và một nửa số con trai bị bệnh.

Thỉnh thoảng có trường hợp nữ chỉ nhận một bản sao đơn của gene lặn gây bệnh liên kết với nhiễm sắc thể giới tính mà có thể bị bệnh. Các cơ thể nữ này được gọi là *thể dị hợp biểu hiện* (manifesting heterozygotes). Bởi vì các cơ thể nữ này thường duy trì ít nhất một phần nhỏ của nhiễm sắc thể X hoạt động bình thường và có khuynh hướng biểu hiện bệnh tương đối nhẹ. Ví dụ, gần 5% nữ dị hợp tử đối với bệnh ưa chảy máu A (hemophilia A) có mức yếu tố đông máu VIII đủ thấp để được xếp vào loại bệnh ưa chảy máu nhẹ.

Ít phổ biến hơn là nữ chỉ có một nhiễm sắc thể X (hội chứng Turner) đã biểu hiện các bệnh do gene lặn liên kết với nhiễm sắc thể X, như bệnh ưa chảy máu A. Nữ còn có thể bị các bệnh do gene lặn liên kết nhiễm sắc

thể X là kết quả của sự chuyển đoạn hoặc mất đoạn nhiễm sắc thể X. Các trường hợp này là hiếm.

## 2. 2. Sự di truyền các gene trội liên kết với nhiễm sắc thể X

Sự di truyền do gene trội liên kết với nhiễm sắc thể giới tính X giống như sự di truyền do gene trội trên nhiễm sắc thể thường. Một cá thể chỉ cần thừa hưởng một bản sao đơn của gene trội liên kết với nhiễm sắc thể X là có thể biểu hiện tính trạng hoặc bệnh. Bởi vì nữ có hai nhiễm sắc thể X, mà mỗi nhiễm sắc thể X đều có thể có khả năng mang gene bệnh nên nữ bị bệnh gần gấp đôi so với nam (trừ trường hợp rối loạn gây chết ở nam). Bố bị bệnh không thể truyền tính trạng này cho con trai. Tất cả con gái đều di truyền gene bệnh này vì thế đều bị bệnh. Bố bị bệnh thường là dị hợp tử vì vậy sẽ có 50% cơ hội truyền allele bệnh cho con gái và con trai.

## 2.3. Sự di truyền liên kết với nhiễm sắc thể Y

Số tính trạng hoặc bệnh đã biết do gene trên nhiễm sắc thể Y qui định là rất ít, chỉ gặp ở nam giới và được truyền trực tiếp từ bố cho con trai mang tính chất “dòng họ nội”. Tất cả các con trai đều bị bệnh (di truyền thẳng) và con gái không mang bệnh. Các bệnh liên kết với nhiễm sắc thể Y như tật nhiều lông mọc ở vành tai, dày sừng lòng bàn tay, tật dính ngón số 2 và 3...

## V. Di truyền y học

### 1. Các bệnh di truyền do rối loạn chuyển hoá và các bệnh nhiễm sắc thể

#### 1.1. Các bệnh di truyền do rối loạn chuyển hoá

##### 1.1.1. Phenylketonuria (PKU)

PKU là một dị tật bẩm sinh, được di truyền do gene lặn Mendel. PKU được Phelling phát hiện lần đầu tiên vào năm 1934.

Trẻ sơ sinh bị bệnh này là do đột biến gene, gan không có khả năng tổng hợp phenylalanine hydroxylase nên dẫn đến tình trạng không chỉ kìm hãm sự chuyển hoá phenylalanine thành tyrosine mà còn gây ra ứ đọng phenylalanine trong máu, làm tăng sự phân giải phenylalanine thành axit phenylpyruvic, cũng bị tích tụ trong máu. Cả phenylalanine và phenylpyruvic khi lên não nhiều sẽ đầu độc tế bào thần kinh. Bản thân nước tiểu thải ra cũng có phenylalanine.

Nếu đột biến gene lặn xảy ra ở một khâu khác làm kìm hãm hoạt động của enzyme tyrosinase, là enzyme xúc tác phản ứng chuyển hoá tyrosine thành melanine dẫn tới bạch tạng.

Nhiều nước ngày nay đã đặt ra thủ tục để chẩn đoán cho mọi trẻ sơ sinh, phát hiện sớm dạng đồng hợp PKU để cho ăn khẩu phần kiêng đặc

biệt nghèo phenylalanine (cháo sữa ngựa) hoặc khẩu phần ăn không có phenylalanine (rau, mật ong, bơ,...)

### 1.1.2. Alcaptonuria

Bateson, 1902 và Garrod, 1907 đã chứng minh rằng bệnh này do một gene lặn, thường xảy ra ở các gia đình có người cùng dòng máu lấy nhau. Biểu hiện của bệnh này rất dễ nhận biết: nước tiểu của bệnh nhân có màu đen do trong nước tiểu chứa một lượng lớn axit homogenetic. Ở người bình thường, axit homogenetic sẽ được enzyme oxydase phân giải thành axit axetoaxetic. Khi thiếu enzyme oxydase thì axit homogenetic không được phân giải và thải ra theo nước tiểu, bị oxy hoá trong không khí tạo thành một hợp chất có màu đen. Khi còn bé tật bẩm sinh này chưa gây hậu quả nghiêm trọng nhưng khi đứng tuổi bị viêm khớp nặng, sụn tích lũy nhiều sắc tố.

## 1.2. Các bệnh nhiễm sắc thể

Nhờ sự phát triển của các phương pháp nghiên cứu di truyền học người, đặc biệt là phương pháp di truyền tế bào đã làm rõ cơ chế phát sinh nhiều loại bệnh nhiễm sắc thể, do sự phân ly không bình thường của nhiễm sắc thể trong quá trình phân bào hoặc do dị dạng, sai hình nhiễm sắc thể từ các rối loạn trong cấu trúc, hình thái.

### 1.2.1. Hội chứng Down

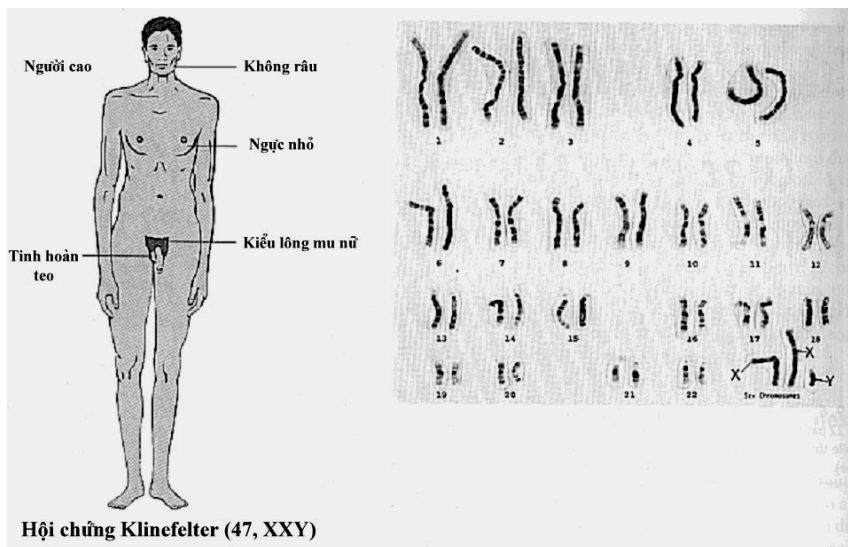
Hội chứng này được Langdom Down phát hiện lần đầu tiên vào năm 1866. Tần số bắt gặp khoảng 1/700 trẻ sơ sinh và là bệnh nhiễm sắc thể bắt gặp cao nhất. Thường thì bộ nhiễm sắc thể của bệnh nhân có 47 chiếc, thừa một nhiễm sắc thể 21. Trong một số trường hợp hội chứng này còn là kết quả của chuyển đoạn giữa nhiễm sắc thể 21 với nhiễm sắc thể của nhóm D hoặc nhóm G. Biểu hiện bệnh lý là ngu đần bẩm sinh, giảm trí lực, nhiều dị tật ở các cơ quan nội tạng, không có khả năng sinh dục, vóc dáng bé, lùn, cổ rụt, đầu bé, cằm dẹt, mắt tròn, khe mắt xếch, môi dày, lưỡi dày có xu thế thè ra thường xuyên...Số liệu về hội chứng Down ở trẻ em Việt Nam đã được trình bày trong chương 3.

### 1.2.2. Hội chứng Klinefelter

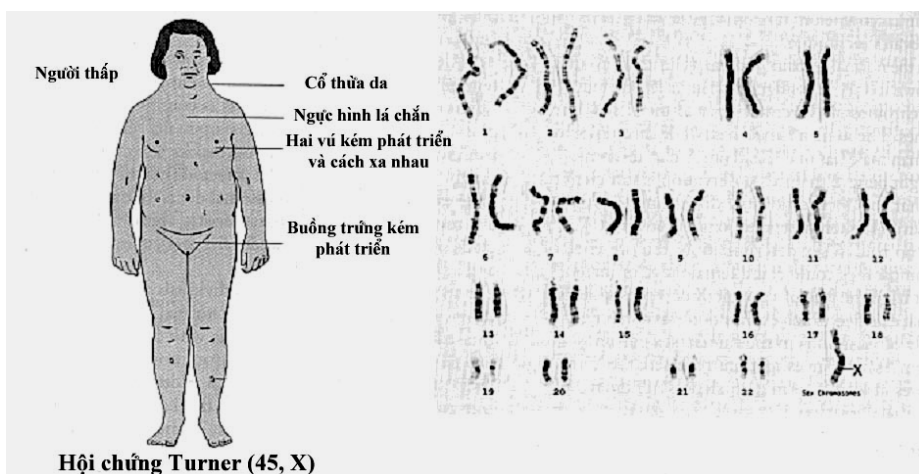
H. F.Klinefelter đã mô tả hội chứng này lần đầu tiên vào năm 1942. Tần số bắt gặp là khoảng 1/1000 bé trai sinh ra. Kiểu nhân của bệnh nhân là 47, XXY. Nguyên nhân có thể là do thụ thai từ một tế bào trứng XX với tinh trùng Y hoặc từ một trứng X với tinh trùng XY. Bệnh nhân là nam không bình thường về tuyến sinh dục, có một số nét giống nữ, có vú như nữ, đặc biệt là các tính trạng giới tính thứ cấp, trí tuệ kém phát triển, không có con, khổ người cao, chân tay dài (Hình 11.4).

### 1.2.3. Hội chứng Turner

Hội chứng này được H.H.Turner phát hiện lần đầu tiên vào năm 1938. Tần số bắt gặp khoảng 1/3000 trẻ em gái sơ sinh. Kiểu nhân của bệnh nhân là 45, X. Nguyên nhân là do một tế bào trứng hoặc một tinh trùng không có nhiễm sắc thể giới tính, hoặc do mất đi một nhiễm sắc thể giới tính trong nguyên phân ở những lần phân cắt đầu tiên sau khi hình thành hợp tử XX hoặc XY. Bệnh nhân có tầm vóc bé, thường bị các dị hình như không có buồng trứng, thiếu các tính trạng giới tính thứ cấp do đó không có con và nhiều dị dạng bề ngoài khác (Hình 11.5).



**Hình 11.4 Hội chứng Klinefelter**



**Hình 11.5 Hội chứng Turner**

Cần nói thêm là, ở Việt Nam, đã có các số liệu nghiên cứu được công bố về các bất thường nhiễm sắc thể trong các thể bệnh Leukemia cấp, một dạng bệnh tăng bạch cầu, hoại huyết (Phạm Quang Vinh, 2003).

## 2. Cơ sở di truyền ung thư

Có một số bằng chứng rõ ràng cho thấy rằng ung thư là một bệnh do sự biến đổi của gene gây ra. Các tế bào khối u có số lượng nhiễm sắc thể đã biến đổi và các nhiễm sắc thể thường có sự sắp xếp lại. Một vài chuyển đoạn là đặc trưng với các dạng ung thư nhất định. Phần lớn các tác nhân đột biến là các tác nhân gây ung thư. Tổ bẩm (predisposition) với các dạng ung thư đã được tìm thấy là di truyền ở một số gia đình. Có ba loại gene liên quan đến ung thư đã được phát hiện là gene ung thư (oncogene), gene ức chế khối u (tumor suppressor gene) và gene sửa chữa DNA (DNA repair gene).

Phần lớn gene ung thư có nguồn gốc từ protooncogene. Protooncogene là các gene mã hoá các sản phẩm kiểm soát sự sinh trưởng và biệt hoá của tế bào. Protooncogene được hoạt hoá để tạo thành gene ung thư do đột biến hoặc biểu hiện quá mức. Lần đầu tiên chúng được phát hiện trong các retrovirus gây ung thư, nhưng cũng có thể được phát hiện trong các khối u không có virus bằng sự chuyển nhiễm vào trong các tế bào 3T3 của chuột. Gene ung thư ở các retrovirus là *v-onc* và khi ở trong các khối u là *c-onc*. Thỉnh thoảng gene ung thư được khuếch đại trong các tế bào khối u. Phần lớn các gene ung thư hoạt động như allele trội làm tăng chức năng đột biến dẫn đến loại bỏ khả năng điều hoà của sự kiểm soát chu trình tế bào. Ngược lại với gene ức chế khối u, gene ung thư không biểu hiện các đột biến dòng mầm (germline mutations) mà các đột biến này gây nên các hội chứng ung thư di truyền (inherited cancer syndromes). Trong khi đó đột biến soma gây ra ung thư đơn phát (sporadic cancer).

Một số khối u do chuyển đoạn đặc hiệu làm tăng sự phiên mã của gene ung thư đã được mô tả đặc điểm. Hiện tượng này đã được tìm thấy ở u lympho của Burkitt (Burkitt's lymphoma).Hoặc sự dung hợp của hai protooncogene có thể tạo thành một gene mới lạ với các đặc tính ung thư đã được quan sát ở bệnh bạch cầu dạng tuỷ mạn tính (chronic myeloid leukemia).

Gene ức chế khối u cần đến cả hai bản sao trở nên bất hoạt trước khi các khối u có thể phát triển. Các đột biến này là đột biến lặn. Thường thường sự bất hoạt xảy ra là do đột biến điếm một allele và mất một allele khác. Sự bất hoạt của gene ức chế khối u xảy ra ở các tế bào soma trong quá trình phát triển của khối u. Gene ức chế khối u được nghiên cứu kỹ nhất là gene p53. Trong hơn 40% các khối u của người, gene này bị đột

biến. p53 đóng vai trò điều hoà chu trình tế bào và quá trình chết theo chương trình của tế bào (apoptosis)

Một số ung thư hiếm gặp và một phần nhỏ ung thư phổ biến đã tập hợp lại trong một số gia đình. Các gia đình này được gọi là có tổ bẩm di truyền (hereditary predisposition) với ung thư. Tổ bẩm di truyền này có thể trở thành một dạng đặc trưng của ung thư ví dụ như ung thư vú hoặc một loại ung thư khác.

Tổ bẩm di truyền có thể tăng lên thông qua các đột biến dòng mầm của gene ức chế khối u. Tất cả các tế bào soma của một cá thể mang một allele đột biến và có nguy cơ hình thành khối u tăng cao do sự bất hoạt của một allele đơn bình thường. Ví dụ ở người gồm có gene võng mạc (retinoblastoma -Rb) nằm trên nhiễm sắc thể 13 và gene tổ bẩm di truyền ung thư vú và ung thư buồng trứng (BRCA1) nằm trên nhiễm sắc thể 17. Allele bình thường bị bất hoạt trong khối u. Các gene ung thư gia đình thường bị đột biến soma trong các trường hợp không thường xuyên (không di truyền) của cùng loại ung thư.

Ung thư di truyền cho thấy kiểu di truyền do gene trội trên nhiễm sắc thể thường bởi vì người bệnh truyền tổ bẩm di truyền cho một nửa số con cái của họ. Tuy nhiên các đột biến ở gene ức chế khối u là đột biến lặn. Nó được giải thích bằng thực tế là nếu một cá thể chỉ có một bản sao hoạt động chức năng của gene ức chế khối u trong mọi tế bào thì nó rất thích hợp cho việc một đột biến thứ hai sẽ gây bất hoạt bản sao hoạt động chức năng bình thường này trong mọi tế bào. Sau đó các tế bào đột biến kép này có thể phát triển thành một khối u.

Các gene sửa chữa DNA cũng kết hợp với ung thư di truyền. Ví dụ bệnh da sùng hoá sắc tố (xeroderma pigmentosum) do một đột biến lặn trên nhiễm sắc thể thường. Người bệnh rất nhạy cảm với ánh sáng mặt trời bởi vì họ thiếu endonuclease xúc tác cho việc sửa chữa DNA. Các allele đột biến của một số gene khác có thể dẫn đến tổ bẩm ung thư ruột kết (colon cancer). Hai trong số các đột biến này cũng liên quan đến sự sửa chữa các DNA không hợp đôi (DNA mismatches). Các tế bào khối u sửa chữa DNA tổn thương ít hiệu quả hơn các tế bào bình thường và tạo ra các đột biến với tần số cao và một số các đột biến sẽ trở thành các gene ức chế khối u hoặc protooncogene. Đây là kiểu hình của gene tăng cường đột biến (mutator phenotype).

## **VI. Tư vấn di truyền y học (Medical genetic counseling)**

Tư vấn di truyền y học hay còn gọi là lời khuyên di truyền, là sự trao đổi ý kiến, giải thích về nguyên nhân, cơ chế về một bệnh tật di truyền nào đó và cho lời khuyên về khả năng mắc bệnh đó ở đời con của các cặp vợ



chồng mà bản thân họ hoặc một số người trong dòng họ mắc phải bệnh ấy. Lời khuyên di truyền nhằm mục đích giúp các cặp vợ chồng đến xin lời khuyên di truyền tự quyết định có nên sinh con hay không và nếu sinh thì phải như thế nào để tránh cho ra đời các đứa trẻ tật nguyền để giảm gánh nặng cho gia đình và xã hội, cũng như cung cấp cho họ các phương pháp, biện pháp đề phòng, điều trị và hạn chế các hậu quả cho bản thân và cho con cái.

Để có được lời khuyên di truyền chính xác, nhà tư vấn di truyền y học phải tìm hiểu tiền sử gia đình, theo dõi phả hệ, xác định đúng là bệnh di truyền và đây là đột biến trội hay lặn và di truyền theo qui luật nào, thăm khám cho người bị bệnh và trong một số trường hợp phải thăm khám cho cả những người có liên quan trong gia đình, làm các xét nghiệm chẩn đoán.

Các đối tượng cần xin lời khuyên di truyền là các cặp vợ chồng đã có con cái bị bệnh di truyền; các nam nữ thanh niên trước khi xây dựng gia đình khi mà một trong hai gia đình hoặc cả hai gia đình đã có người mắc bệnh di truyền; những người bị vô sinh, sảy thai liên tục,...; các cặp vợ chồng muốn sinh con theo ý muốn; cả hai vợ chồng hoặc chỉ vợ hay chồng làm việc trong điều kiện độc hại; đã biết vợ hoặc chồng mang gene bệnh; các cặp vợ chồng đã lớn tuổi, đặc biệt là người vợ và những người kết duyên cùng dòng họ...

Trước đây lời khuyên di truyền hoàn toàn dựa vào sự hiểu biết về qui luật di truyền, tính xác suất xảy ra cho từng trường hợp cụ thể, phân tích phả hệ để xác định bệnh di truyền theo qui luật nào...Ngày nay kỹ thuật chẩn đoán trước sinh đã được thực hiện với những phương pháp phân tích đáng tin cậy ngay khi cơ thể còn ở giai đoạn phôi thai trong cơ thể mẹ như phương pháp chọc dò dịch ối để phân tích di truyền học tế bào và hoá sinh, quan sát gián tiếp bào thai nhờ siêu âm, quan sát trực tiếp bào thai bằng soi phôi thai, lấy các mẫu sinh phẩm từ phôi thai (máu, tua nhau thai...) để xét nghiệm, phương pháp định lượng AFP (alpha fetoprotein) trong máu mẹ. Chẩn đoán trước sinh cho phép đánh giá tình trạng của bộ nhiễm sắc thể, của bộ gene, tình trạng hoạt động của các enzyme trong bào thai giúp tư vấn di truyền y học có cơ sở khoa học chắc chắn để cho lời khuyên chính xác.

## **Câu hỏi và Bài tập**

1. Một cặp vợ chồng có một đứa con bị bệnh u xơ nang. Đây là một bệnh hiếm gặp, được quy định bởi một đột biến lặn di truyền theo định luật Mendel. Cặp vợ chồng này dự định sinh thêm con. Hãy tính xác suất để đứa con tiếp theo: (a) Mang gene bệnh; (b) Không bị bệnh.

2. Ở người, tính trạng tóc quăn trội so với tính trạng tóc thẳng, do gene nằm trên nhiễm sắc thể thường quy định; còn bệnh mù màu do gene lặn nằm trên nhiễm sắc thể giới tính X gây nên. Bố và mẹ có tóc quăn, mắt bình thường sinh ra một con trai tóc quăn, mù màu. Hãy xác định kiểu gene của bố mẹ.

3. Bố mẹ đều có nhóm máu A, sinh con trai có nhóm máu A, con gái có nhóm máu O. Hãy tìm kiểu gene chắc có của những người trong gia đình trên.

4. Kết quả thống kê trên một số lượng lớn các đứa trẻ sinh ra từ nhiều cặp vợ chồng, trong đó các người chồng đều bị bệnh xin men răng còn các người vợ có men răng bình thường cho thấy: 50% bị bệnh xin men răng đều là gái : 50% có men răng bình thường toàn là trai. Hãy xác định tính chất di truyền của bệnh xin men răng và viết sơ đồ lai từ P đến F1. Cho biết tính trạng men răng do một gene quy định.

5. Chồng bình thường, lấy vợ bị bệnh máu khó đông. Con của họ có bị bệnh máu khó đông không?

6. "Bệnh mù màu và bệnh máu khó đông là bệnh của nam giới". Quan niệm như vậy có đúng không? Hãy giải thích tại sao.

7. Trình bày các phương pháp nghiên cứu di truyền người.

8. Trình bày các phương pháp lập bản đồ di truyền người.

9. Nêu đặc điểm di truyền của các gene trội-lặn trên nhiễm sắc thể thường và nhiễm sắc thể giới tính.

10. Trình bày một số bệnh di truyền do rối loạn chuyển hoá và các bệnh nhiễm sắc thể.

## Tài liệu Tham khảo

### Tiếng Việt

Trần Quốc Dung. 2003. *Bài giảng Di truyền học người*. Trường Đại học Sư phạm Huế, Huế.

Phan Cự Nhân, Nguyễn Minh Công, Đặng Hữu Lanh. 1999. Di truyền học người. Trong: *Di truyền học* (Phan Cự Nhân, chủ biên). Tập II. Trang: 161-248. NXB Giáo Dục.

### Tiếng Anh

Doris Bachtrog, Brian Chalesworth. 2001. Towards a complete sequence of the human Y chromosome. *Genome Biol* 2 (5): reviews 1016.1- reviews 1016.5. Published online 2001 April 26.

Jorde, Carey, Bamshad, White. 2000. *Medical Genetics*. Printed in the United States of America.

Winter P. C., Hickey G.I., Fletcher H. L. 1998. *Genetics*. Printed by Biddles Ltd, Guilford, UK.

<http://users.rcn.com/jkiball.ma.ultranet/BiologyPages/S/SexChromosomes.html>

[http://cas.bellarmine.edu/tietjen/HumanBiology/bill\\_developmental\\_abnormalities.htm](http://cas.bellarmine.edu/tietjen/HumanBiology/bill_developmental_abnormalities.htm)

## Chương 12

# Di truyền học Quần thể

Lần đầu tiên vào năm 1908, G.Hardy và W.Weinberg chứng minh rằng tính di truyền tự nó không làm thay đổi tần số allele trong quần thể, sau này được gọi là *nguyên lý Hardy-Weinberg* (Hardy-Weinberg principle). Nó đặt nền móng cho *di truyền học quần thể* (population genetics) - một nhánh của di truyền học - nghiên cứu thành phần di truyền của các quần thể sinh vật và các quá trình ảnh hưởng lên các tần số gene của chúng. Tuy nhiên, đến đầu thập niên 1930 lĩnh vực nghiên cứu này mới thực sự phát triển nhờ các công trình vĩ đại của R.A.Fisher, J.B.S. Haldane và S.Wright mà thực chất đó là các mô hình toán học. Từ đó di truyền học quần thể trở thành nền tảng của các thuyết tiến hoá hiện đại.

Trong chương này, chúng ta sẽ thảo luận một số khái niệm cơ bản của di truyền học quần thể, nguyên lý Hardy-Weinberg và mối quan hệ giữa các tần số allele và kiểu gene trong các trường hợp khác nhau, nội phối và sự gia tăng tần số của các thể đồng hợp, và tìm hiểu sơ lược vai trò của các nhân tố tác động lên thành phần di truyền quần thể.

### I. Các khái niệm cơ bản của Di truyền học quần thể

#### 1. *Quần thể* (population)

Trong tiến hoá, cá thể không được xem là đơn vị thích hợp bởi vì: kiểu gene của một cá thể được giữ nguyên trong quãng đời của nó; hơn nữa, cá thể có tính tạm bợ (dù nó có thể sống tới cả nghìn năm như cây tùng...). Ngược lại, một quần thể thì có tính liên tục qua thời gian và mặt khác, thành phần di truyền của nó có thể thay đổi tiến hoá qua các thế hệ. Sự hình thành các quần thể địa phương tại những vùng lãnh thổ khác nhau chính là phương thức thích ứng của loài trước tự nhiên. Quần thể vì vậy được xem là *đơn vị tiến hóa cơ sở*.

Theo A.V.Yablokov (1986), *quần thể là một nhóm các cá thể cùng loài có khả năng giao phối tự do với nhau, chiếm cứ một khu phân bố xác định và trải qua một khoảng thời gian tiến hoá lâu dài để hình thành nên một hệ thống di truyền độc lập và một ổ sinh thái riêng*.

Nói ngắn gọn, quần thể là một nhóm sinh vật có khả năng giao phối qua lại và cùng chia sẻ một vốn gene chung (Ridley 1993). Nó còn được gọi là *quần thể Mendel*, mà tập hợp lớn nhất là *loài* (species).

#### 2. *Các hệ thống giao phối* (mating systems)

Trên nguyên tắc, cấu trúc di truyền của quần thể ở thế hệ sau được xác định bởi xác suất kết hợp của các giao tử thế hệ trước trong quá trình thụ

ting. Do đó, nó phụ thuộc vào kiểu giao phối của các bố mẹ. Trong di truyền học quần thể, người ta phân biệt ba kiểu giao phối: *Giao phối ngẫu nhiên* hay *ngẫu phối* (random mating hay panmixia), *giao phối chọn lựa* (assortative mating), và *nội phối* (inbreeding).

- *Ngẫu phối* là kiểu giao phối trong đó xảy ra sự bắt cặp ngẫu nhiên giữa các cá thể đực và cái trong quần thể.

Lưu ý rằng định nghĩa quần thể trên đây được áp dụng cho các quần thể thuộc hệ thống ngẫu phối; chúng chiếm vị trí rất quan trọng trong hệ thống các loài và được đề cập chủ yếu trong suốt chủ đề này.

- *Giao phối chọn lựa* là kiểu giao phối trong đó các cá thể đực và cái không bắt cặp ngẫu nhiên mà có sự lựa chọn theo kiểu hình. Có hai trường hợp: (1) Nếu như các cá thể có xu hướng giao phối với các cá thể khác có kiểu hình tương tự, thì gọi là *giao phối chọn lựa dương tính* (positive assortative mating); và (2) Nếu như sự lựa chọn ít được quan tâm nhưng tần số của các cặp giao phối vẫn khác xa với tần số của các cặp ngẫu phối, thì gọi là *giao phối không lựa chọn* (disassortative mating) hay *chọn lựa âm tính* (negative assortative mating). Chẳng hạn, ở người, sự giao phối có lựa chọn xảy ra đối với các tính trạng như chiều cao, màu mắt, màu tóc... Vì vậy nó chỉ ảnh hưởng đến các tần số kiểu gene của locus nào có liên quan đến việc xác định kiểu hình được sử dụng trong giao phối. Còn kiểu giao phối không lựa chọn phổ biến trong các hệ thống *tự bất dục* (self-sterility) ở thực vật.

- *Nội phối* là sự giao phối không ngẫu nhiên xảy ra giữa các cá thể có quan hệ họ hàng gần hoặc điển hình là sự tự thụ tinh (xem mục IV).

### 3. Vốn gene (gene pool)

Vốn gene là tập hợp toàn bộ các allele ở tất cả các gene của mọi cá thể trong quần thể tại một thời điểm xác định.

Vốn gene này được sử dụng chung cho các cá thể trong quần thể. Mỗi quần thể đặc trưng bằng một vốn gene nhất định và nó được mô tả bằng tần số các allele ở từng locus.

### 4. Tần số kiểu gene và tần số allele

Để mô tả thành phần di truyền của một quần thể ta cần phải xác định kiểu gene của các cá thể và số cá thể của mỗi kiểu gene. Giả sử trong một quần thể sinh vật lưỡng bội gồm  $N$  cá thể, xét một locus  $A$  thuộc *nhĩem sắc thể thường* (autosome) với hai allele  $A_1$  và  $A_2$  có mặt trong các cá thể. Lúc đó sẽ có ba kiểu gene:  $A_1A_1$ ,  $A_1A_2$  và  $A_2A_2$  với số lượng tương ứng là  $N_{11}$ ,  $N_{12}$  và  $N_{22}$ ; ( $N = N_{11} + N_{12} + N_{22}$ ). Nếu ký hiệu  $P$ ,  $H$  và  $Q$  là tần số tương ứng với các kiểu gene trên, ta có:

$$P = N_{11} / N; \quad H = N_{12} / N \quad \text{và} \quad Q = N_{22} / N; \quad (P + H + Q = 1)$$

Từ đây ta có thể tính được các tần số gene hay allele (gene or allelic frequencies)  $A_1$  và  $A_2$ , với ký hiệu tương ứng là  $p$  và  $q$  ( $p + q = 1$ ), như sau:

$$p = \frac{2N_{11} + N_{12}}{2N} = P + \frac{1}{2}H$$

$$q = \frac{2N_{22} + N_{12}}{2N} = Q + \frac{1}{2}H \quad (\text{hay } q = 1 - p)$$

*Tóm tắt:*

(1) *Vốn gene* của một quần thể có  $N$  cá thể bao gồm  $2N$  hệ gene đơn bội. Mỗi hệ gene gồm tất cả các thông tin di truyền nhận được từ một cha mẹ. Đối với mỗi locus autosome, trong vốn gene quần thể sẽ có  $2N$  allele.

(2) *Tần số kiểu hình* (phenotypic frequency) bằng số lượng cá thể của kiểu hình cụ thể chia cho tổng số cá thể của quần thể.

(3) *Tần số kiểu gene* (geneotypic frequency) bằng số lượng cá thể của kiểu gene cụ thể chia cho tổng số cá thể của quần thể.

(4) *Tần số allele* (allelic frequency) bằng hai lần số lượng cá thể đồng hợp cộng với số cá thể dị hợp về allele đó chia cho hai lần tổng số cá thể của quần thể; hay tần số của một allele bằng tần số kiểu gene đồng hợp cộng với một nửa tần số kiểu gene dị hợp về allele đó.

*Lưu ý:* (i) Tổng các tần số kiểu hình, kiểu gene hay allele thuộc một locus nào đó luôn luôn bằng đơn vị; (ii) Tần số allele hay tần số gene như một số nhà khoa học thường gọi (Crow 1986; Falconer và Mackay 1996) là khái niệm căn bản nhất của di truyền học quần thể; nó là dấu hiệu đặc trưng của một quần thể cho phép phân biệt với các quần thể khác trong cùng một loài; (iii) Để cho tần số các allele quan sát được là đặc trưng của một quần thể, thì mẫu thu được phải là ngẫu nhiên và có kích thước đủ lớn; (iv) Để tiện cho một số mục đích mô tả các biến dị di truyền ở một locus, người ta sử dụng chủ yếu tần số các allele chứ không phải tần số các kiểu gene, bởi vì ở một locus thường có số allele ít hơn số kiểu gene; (v) Thuật ngữ "thành phần" hay "cấu trúc di truyền của quần thể" dùng để chỉ tần số tương đối của các allele và các kiểu gene trong quần thể tại một thời điểm xác định.

*Ví dụ:* Số liệu phân bố của hệ nhóm máu M-N ở một số quần thể người ở bảng 12.1 cho thấy: (1) Mỗi quần thể ở một vùng địa lý nhất định đều có các tần số allele đặc trưng; (2) Trong khi ở quần thể người Mỹ gốc Âu có các tần số allele M và N hầu như tương đương, mặc dù tần số allele M cao hơn khoảng 8%, thì quần thể thổ dân Úc có tần số allele N rất cao

(nhiều gấp 4,6 lần tần số allele M); còn ở bộ tộc da đỏ Navaho nói riêng và ở vùng Trung-Nam Mỹ nói chung có tần số allele M rất cao.

**Bảng 12.1 Tần số các nhóm máu hệ M-N ở một số quần thể người**

Quần thể	Số lượng			Tần số kiểu gene			Tần số allele	
	M	MN	N	$L^M L^M$	$L^M L^N$	$L^N L^N$	$L^M$	$L^N$
Bộ tộc Navaho	305	52	4	0,845	0,144	0,011	0,917	0,083
Thổ dân Úc	22	216	492	0,030	0,296	0,674	0,178	0,822
Mỹ gốc Âu	1787	3039	1303	0,292	0,496	0,213	0,539	0,461

## II. Nguyên lý Hardy-Weinberg và trạng thái cân bằng của quần thể

### 1. Nguyên lý Hardy-Weinberg

Năm 1908, nhà toán học người Anh Godfrey H. Hardy và bác sĩ người Đức Wilhelm Weinberg đã độc lập chứng minh rằng có tồn tại một mối quan hệ đơn giản giữa các tần số allele và các tần số kiểu gene mà ngày nay ta gọi là định luật hay nguyên lý Hardy-Weinberg (viết tắt: H - W).

#### 1.1. Nội dung nguyên lý H-W

Trong một quần thể ngẫu phối kích thước lớn, nếu như không có áp lực của các quá trình đột biến, di nhập cư, biến động di truyền và chọn lọc, thì tần số các allele được duy trì ổn định từ thế hệ này sang thế hệ khác và tần số các kiểu gene (của một gene gồm hai allele khác nhau) là một hàm nhị thức của các tần số allele, được biểu diễn bằng công thức sau:

$$(p + q)^2 = p^2 + 2pq + q^2 = 1$$

#### 1.2. Chứng minh

Ở một quần thể Mendel, xét một locus autosome gồm hai allele  $A_1$  và  $A_2$  có tần số như nhau ở cả hai giới đực và cái. Ký hiệu  $p$  và  $q$  cho các tần số allele nói trên ( $p + q = 1$ ). Cũng giả thiết rằng các cá thể đực và cái bắt cặp ngẫu nhiên, nghĩa là các giao tử đực và cái gặp gỡ nhau một cách ngẫu nhiên trong sự hình thành các hợp tử. Khi đó tần số của một kiểu gene nào đó chính là bằng tích của các tần số hai allele tương ứng. Xác suất để một cá thể có kiểu gene  $A_1A_1$  là bằng xác suất ( $p$ ) của allele  $A_1$  nhận từ mẹ nhân với xác suất ( $p$ ) của allele  $A_1$  nhận từ bố, hay  $p.p = p^2$ . Tương tự, xác suất mà một cá thể có kiểu gene  $A_2A_2$  là  $q^2$ . Kiểu gene  $A_1A_2$  có thể xuất hiện theo hai cách:  $A_1$  từ mẹ và  $A_2$  từ bố với tần số là  $pq$ , hoặc  $A_2$  từ mẹ và  $A_1$  từ bố cũng với tần số  $pq$ ; vì vậy tần số của  $A_1A_2$  là  $pq + pq = 2pq$  (Bảng 12.2). Điều chứng minh trên được tóm tắt như sau:

- \* Quần thể ban đầu có 3 kiểu gene :  $A_1A_1$     $A_1A_2$     $A_2A_2$    Tổng
- Tần số các kiểu gene :                    P            H            Q
- Tần số các allele :                     $p = P + \frac{1}{2}H$  ;    $q = Q + \frac{1}{2}H$             1
- \* Quần thể thế hệ thứ nhất sau ngẫu phối có :
- Tần số các kiểu gene     $= (p + q)^2 = p^2 + 2pq + q^2$             1
- Tần số các allele:  $f(A_1) = p^2 + \frac{1}{2}(2pq) = p(p+q) = p$
- $f(A_2) = q^2 + \frac{1}{2}(2pq) = q(p+q) = q$

*Nhận xét:*

Từ chứng minh trên cho thấy các tần số allele ở thế hệ con giống hệt ở thế hệ ban đầu, nghĩa là  $f(A_1) = p$  và  $f(A_2) = q$ . Do đó, các tần số kiểu gene ở thế hệ tiếp theo vẫn là  $p^2$ ,  $2pq$  và  $q^2$  (giống như ở thế hệ thứ nhất sau ngẫu phối). Điều đó chứng tỏ rằng các tần số kiểu gene đạt được cân bằng chỉ sau một thế hệ ngẫu phối. Trạng thái ổn định về thành phần di truyền được phản ánh bằng công thức H-W như vậy được gọi là *cân bằng H-W* (Hardy-Weinberg equilibrium).

**Bảng 12.2 Các tần số H-W sinh ra từ sự kết hợp ngẫu nhiên các giao tử**

		Tần số giao tử cái	
		$p(A_1)$	$q(A_2)$
Tần số giao tử đực	$p(A_1)$	$p^2(A_1A_1)$	$pq(A_1A_2)$
	$q(A_2)$	$pq(A_1A_2)$	$q^2(A_2A_2)$

### 1.3. Các mệnh đề và hệ quả

(1) Nếu như không có áp lực của các quá trình tiến hoá (đột biến, di nhập cư, biến động di truyền và chọn lọc), thì các tần số allele được giữ nguyên không đổi từ thế hệ này sang thế hệ khác. Đây là mệnh đề chính của nguyên lý hay định luật H-W.

(2) Nếu sự giao phối là ngẫu nhiên, thì các tần số kiểu gene có quan hệ với các tần số allele bằng công thức đơn giản:  $(p+q)^2 = p^2 + 2pq + q^2 = 1$ .

(3) *Hệ quả 1*: Bất luận các tần số kiểu gene ban đầu (P, H, Q) như thế nào, miễn sao các tần số allele ở hai giới là như nhau, chỉ sau một thế hệ ngẫu phối các tần số kiểu gene đạt tới trạng thái cân bằng ( $p^2$ ,  $2pq$  và  $q^2$ ).

(4) *Hệ quả 2*: Khi quần thể ở trạng thái cân bằng thì tích của các tần số đồng hợp tử bằng bình phương của một nửa tần số dị hợp tử, nghĩa là:

$$p^2 \cdot q^2 = \left[ \frac{2pq}{2} \right]^2$$



Thật vậy, khi quần thể ở trạng thái cân bằng lý tưởng, ta có:  $H = 2pq$

Biến đổi đẳng thức trên ta được:  $pq = \frac{1}{2}H$

Bình phương cả hai vế, ta có:  $p^2 \cdot q^2 = (\frac{1}{2}H)^2$ , trong đó  $H = 2pq$ . Như vậy đẳng thức này cho thấy mối tương quan giữa các thành phần đồng hợp và dị hợp khi quần thể ở trạng thái cân bằng lý tưởng.

(5) *Hệ quả 3*: (i) Tần số của các thể dị hợp không vượt quá 50%, và giá trị cực đại này chỉ xảy ra khi  $p = q = 0,5 \Rightarrow H = 2pq = 0,5$ ; lúc này các thể dị hợp chiếm một nửa số cá thể trong quần thể; (ii) Đối với allele hiếm (tức có tần số thấp), nó chiếm ưu thế trong các thể dị hợp nghĩa là, tần số thể dị hợp cao hơn nhiều so với tần số thể đồng hợp về allele đó. Điều này gây hậu quả quan trọng đối với hiệu quả chọn lọc (xem thêm ở mục 1.5.2 dưới đây).

#### 1.4. Tần số giao phối và sự kiểm chứng nguyên lý H-W

Nguyên lý H-W có thể được chứng minh theo một cách khác dựa trên tần số của các kiểu giao phối. Mặc dù nó công kênh hơn phương pháp đã xét nhưng lại cho thấy rõ hơn bằng cách nào các tần số H-W phát xuất từ quy luật phân ly của Mendel.

Xét cấu trúc giao phối của quần thể ngẫu phối như trên ta thấy có cả thấy là chín kiểu giao phối với tần số giao phối như ở Bảng 12.3. Vì tần số mỗi kiểu gene ở hai giới được xem là như nhau, nên một số kiểu giao phối thuận nghịch là tương đương vì vậy chỉ còn lại sáu kiểu giao phối khác nhau với tần số tương ứng được nêu ở hai cột đầu tiên của bảng 12.4. Bây giờ ta xét các kiểu gene đời con sinh ra từ mỗi kiểu giao phối và sau đó tìm tần số của mỗi kiểu gene trong toàn bộ đời con, với giả thiết rằng tất cả các kiểu giao phối đều hữu thụ ngang nhau và tất cả các kiểu gene đều có sức sống như nhau. Kết quả này được trình bày ở phía bên phải Bảng 12.4. Sau khi rút gọn ta được các tần số kiểu gene đời con tương ứng là  $p^2$ ,  $2pq$  và  $q^2$  (ở dòng cuối cùng của bảng). Các trị số này chính là các *tần số cân bằng H-W* (equilibrium frequencies) đạt được sau một thế hệ ngẫu phối, bất luận các tần số kiểu gene ở đời bố mẹ như thế nào.

**Bảng 12.3 Tần số của các kiểu giao phối ngẫu nhiên**

	Giới đực	$A_1A_1$	$A_1A_2$	$A_2A_2$
Giới cái		(P)	(H)	(Q)
$A_1A_1$ (P)		$P^2$	PH	PQ
$A_1A_2$ (H)		PH	$H^2$	QH
$A_2A_2$ (Q)		PQ	QH	$Q^2$

**Bảng 12.4 Nguyên lý Hardy-Weinberg đối với hai allele**

Bố mẹ		Đời con		
Kiểu giao phối	Tần số	$A_1A_1$	$A_1A_2$	$A_2A_2$
$A_1A_1 \times A_1A_1$	$P^2$	$P^2$	–	–
$A_1A_1 \times A_1A_2$	$2PH$	$PH$	$PH$	–
$A_1A_1 \times A_2A_2$	$2PQ$	–	$2PQ$	–
$A_1A_2 \times A_1A_2$	$H^2$	$\frac{1}{4}H^2$	$\frac{1}{2}H^2$	$\frac{1}{4}H^2$
$A_1A_2 \times A_2A_2$	$2HQ$	–	$HQ$	$HQ$
$A_2A_2 \times A_2A_2$	$Q^2$	–	–	$Q^2$
Tổng	1	$(P+\frac{1}{2}H)^2 = p^2 : 2(P+\frac{1}{2}H)(Q+\frac{1}{2}H) = 2pq : (Q+\frac{1}{2}H)^2 = q^2$		

## 2. Những ứng dụng của nguyên lý Hardy-Weinberg

### 2.1. Xác định tần số của allele lặn

Trong trường hợp trội hoàn toàn, ta không thể phân biệt các thể dị hợp với thể đồng hợp trội. Vì vậy, trên nguyên tắc, ta không thể tính được các tần số allele. Tuy nhiên, có thể giả định các tần số kiểu gene ở dạng cân bằng, qua đó tính được tần số allele lặn và dự đoán tần số của các kiểu gene trong quần thể. Chẳng hạn, *bạch tạng* (albinism) ở người là tính trạng lặn tương đối hiếm gặp. Nếu như ký hiệu  $A$  cho allele xác định sắc tố bình thường và  $a$  cho allele bạch tạng, kiểu gene của người bị bạch tạng là  $aa$ , trong khi những người bình thường thì hoặc là  $AA$  hoặc là  $Aa$ . Giả sử trong một quần thể người tần số của những người bị bạch tạng là  $1/10.000$ . Theo nguyên lý H-W, tần số của thể đồng hợp lặn là  $q^2 = 0,0001$  nên  $q = \sqrt{f(aa)} = \sqrt{0,0001} = 0,01$ . Do đó tần số của allele  $A$  là:  $p = 1 - 0,01 = 0,99$  (vì  $p + q = 1$ ). Từ đây xác định được tần số của hai kiểu gene còn lại:

$$f(AA) = p^2 = (0,99)^2 = 0,9801 \text{ (hay } \sim 98\%)$$

$$f(Aa) = 2pq = 2(0,99)(0,01) = 0,0198 \text{ (hay } \sim 2\%)$$

Lưu ý trong trường hợp tần số allele lặn là rất thấp, nghĩa là kích thước mẫu lớn, ta cần phải lấy số thập phân đầy đủ để đảm bảo chính xác cho các kết quả tính toán sau cùng.

### 2.2. Xác định tần số của các "thể mang" (carrier)

Một điều lý thú của nguyên lý H-W là ở chỗ, các allele hiếm nói chung là các allele lặn gây bệnh trong quần thể thường ẩn tàng trong các thể dị hợp (gọi là "thể mang") và ta có thể tính được tần số của chúng nếu như biết được tần số allele. Nếu cho rằng có sự cân bằng H-W thì tần số của các thể mang allele bệnh lặn trong quần thể được ước tính là  $H = 2q(1-q)$ .

Và tần số của các thể dị hợp trong số những cá thể bình thường, ký hiệu  $H'$ , là tỷ số  $f(Aa)/f(AA+Aa)$ , trong đó  $a$  là allele lặn với tần số  $q$ . Khi đó:

$$H' = \frac{2pq}{p^2 + 2pq} = \frac{2q(1-q)}{(1-q)^2 + 2q(1-q)} = \frac{2q}{1+q}$$

*Ví dụ:* Với trường hợp bạch tạng nói trên, tần số của  $aa$  là 0,0001 thì tần số của những người dị hợp ( $Aa$ ) là 0,02, nghĩa là trong 50 người có một người mang allele bạch tạng. Đây là một tỷ lệ rất cao! Mặt khác, tần số allele  $a$  ở những người dị hợp là 0,02: 2 = 0,01 trong khi ở những người bạch tạng là 0,0001, như vậy allele  $a$  ở những người dị hợp có nhiều hơn ở những người bạch tạng khoảng 100 lần (0,01 : 0,0001 = 100).

Tổng quát, nếu tần số của một allele lặn trong quần thể là  $q$ , thì sẽ có  $pq$  allele lặn trong các thể dị hợp và  $q^2$  allele lặn trong các thể đồng hợp. Tỷ số ấy là  $pq/q^2 = p/q$ , và nếu như  $q$  rất bé thì tỷ số đó sẽ xấp xỉ  $1/q$ . Như vậy, khi tần số của một allele lặn càng thấp bao nhiêu, thì tỷ lệ của allele đó trong các thể dị hợp càng cao bấy nhiêu.

Tương tự, có thể lấy nhiều ví dụ về các allele lặn gây bệnh ở người. Điển hình là bệnh rối loạn chuyển hoá có tên là phenylxetôn-niêu (phenylketonuria = PKU) do một allele lặn đơn, có thể phát hiện sớm vài ngày sau sinh. Một kết quả điều tra ở Birmingham trong hơn ba năm cho thấy có 5 trường hợp bị bệnh trong số 55.715 bé (Raine và cs 1972). Tần số các thể đồng hợp lặn xấp xỉ  $1/11.000$  hay  $90 \times 10^{-6}$ . Tần số allele lặn là

$q = \sqrt{90 \times 10^{-6}} = 0,0095$ . Tần số các thể dị hợp trong cả quần thể ( $H = 2pq$ ) và trong số các thể bình thường ( $H' = 2q/1+q$ ) đều xấp xỉ bằng 0,019. Như vậy khoảng 2% số người bình thường là có mang mầm bệnh PKU. Các kết quả này thật đáng ngạc nhiên: bằng cách nào các thể dị hợp về allele lặn lại phổ biến đến như vậy, trong khi tần số bệnh thực tế là quá thấp!

Đến đây ta có thể khẳng định rằng: Nếu như ai đó có ý tưởng muốn loại bỏ một allele lặn hiếm gây bệnh nào đó ra khỏi quần thể hòng “cải thiện chủng tộc” chẳng hạn, quả là không tưởng! Thật vậy, nếu gọi  $t$  là số thế hệ cần thiết để biến đổi tần số allele ban đầu là  $q_0$  xuống còn  $q_t$  ở thế hệ thứ  $t$ , ta có  $t = 1/q_t - 1/q_0$ . Giả sử  $q_0 = 0,01$ , muốn giảm xuống còn 0,001 phải cần tới 900 thế hệ; tương tự, để giảm tần số xuống còn 0,0001 phải cần đến 9.900 thế hệ. Thử tưởng tượng ở người một thế hệ trung bình là 30 năm, thời gian ấy lớn đến dường nào ( $9.900 \times 30 = 297.000$  năm)!

### 2.3. Khảo sát trạng thái cân bằng của quần thể

Từ nguyên lý H-W và các hệ quả rút ra được ở trên cho phép ta vận dụng để xác định xem cấu trúc di truyền của một quần thể có ở trạng thái

cân bằng H-W hay không.

Dưới đây chỉ lược trình vài phương pháp tổng quát đối với một quần thể ngẫu phối (Hoàng Trọng Phán 2001), với các giả thiết và ký hiệu đã được đề cập. Trước tiên, cần nắm vững nguyên tắc này trong suy luận: Theo nguyên lý H-W, các tần số kiểu gene ở đời con được xác định nhờ tần số allele ở bố mẹ chúng. Nếu quần thể ở trạng thái cân bằng, tần số các allele sẽ như nhau ở cả hai thế hệ, vì vậy tần số allele quan sát được ở đời con có thể dùng y như thể nó là tần số allele đời bố mẹ để tính các tần số kiểu gene kỳ vọng theo nguyên lý H-W. Như vậy, về nguyên tắc, một quần thể được coi là ở trạng thái cân bằng nếu như nó thỏa mãn một trong những khả năng sau đây; ngược lại, quần thể không ở trạng thái cân bằng.

(1) Các tần số kiểu gene quan sát được ( $P$ ,  $H$  và  $Q$ ) phải xấp xỉ bằng các tần số kỳ vọng tương ứng ( $p^2$ ,  $2pq$  và  $q^2$ ), nghĩa là thành phần di truyền của quần thể phải thỏa mãn công thức H-W.

Về mặt số lượng, quần thể được coi là ở trạng thái cân bằng nếu như có sự phù hợp sát sao giữa các con số quan sát và kỳ vọng đối với mỗi kiểu gene, nghĩa là:  $N_{11} \approx p^2N$ ;  $N_{12} \approx 2pqN$ ; và  $N_{22} \approx q^2N$ .

(2) Tần số thể dị hợp quan sát phải xấp xỉ bằng tần số kỳ vọng ( $H \approx 2pq$ ), nghĩa là:  $p \cdot q \approx \frac{1}{2}H$  hay  $P \cdot Q \approx (\frac{1}{2}H)^2$

(3) Tần số của mỗi kiểu gene quan sát được giữa hai thế hệ liên tiếp là tương đương nhau. Nếu ta gọi tần số của các kiểu gene  $A_1A_1$ ,  $A_1A_2$  và  $A_2A_2$  tương ứng ở thế hệ thứ nhất là  $P_1$ ,  $H_1$  và  $Q_1$  và ở thế hệ thứ hai là  $P_2$ ,  $H_2$  và  $Q_2$ , lúc đó:  $P_1 \approx P_2$ ;  $H_1 \approx H_2$ ; và  $Q_1 \approx Q_2$ .

(4) Đối với trường hợp khảo sát cân bằng H-W hoặc giao phối ngẫu nhiên dựa trên tần số giao phối hoặc số lượng cặp giao phối của các kiểu giao phối khác nhau, ta có thể so sánh như sau:

Kiểu giao phối	Tần số		Số lượng	
	Quan sát	$\approx$ Kỳ vọng	Quan sát	$\approx$ Kỳ vọng
$A_1A_1 \times A_1A_1$	$P^2$	$p^2 \cdot p^2$	$P^2 \cdot N/2$	$p^2 \cdot p^2 \cdot N/2$
$A_1A_1 \times A_1A_2$	$2PH$	$2(p^2)(2pq)$	$2P \cdot H \cdot N/2$	$2(p^2)(2pq)N/2$
$A_1A_1 \times A_2A_2$	$2PQ$	$2(p^2)(q^2)$	$2P \cdot Q \cdot N/2$	$2(p^2)(q^2)N/2$
$A_1A_2 \times A_1A_2$	$H^2$	$(2pq)(2pq)$	$H^2 \cdot N/2$	$(2pq)(2pq)N/2$
$A_1A_2 \times A_2A_2$	$2QH$	$2(2pq)(q^2)$	$2Q \cdot H \cdot N/2$	$2(2pq)(q^2)N/2$
$A_2A_2 \times A_2A_2$	$Q^2$	$q^2 \cdot q^2$	$Q^2 \cdot N/2$	$q^2 \cdot q^2 \cdot N/2$
Tổng	1	1	$N/2$	$N/2$

(5) Phương pháp “Khi-bình phương” (Chi-square method)

Khi so sánh giữa các số liệu quan sát và kỳ vọng thường có thể có sự sai lệch không đáng kể hoặc đáng kể. Vì ranh giới phân định giữa chúng là không rõ ràng khiến ta khó mà khẳng định quần thể ở trạng thái cân bằng hoặc không. Trong trường hợp đó, ta phải sử dụng phương pháp  $\chi^2$  (xem chương 1).

*Ví dụ:* Để khảo sát trạng thái cân bằng H-W, ta hãy xét quần thể người Mỹ da trắng gốc Âu đã cho ở bảng 12.1. Từ số người mang các nhóm máu M, MN và N tương ứng là 1.787; 3.039; và 1.303 (với  $N = 6.129$ ), ta tính được tần số của các allele M và N là p và q như sau:

$$p = 1.787 + 1/2(3.039) = 0,539$$

$$\text{và } q = 1 - p = 0,461.$$

Từ đây tính được tần số kỳ vọng của các kiểu gene:

$$\text{MM } p^2 = (0,539)^2 = 0,292$$

$$\text{MN } 2pq = 2(0,539)(0,461) = 0,497$$

$$\text{NN } q^2 = (0,461)^2 = 0,211$$

Và số cá thể kỳ vọng của chúng:

$$\text{MM } p^2 \times N = 0,292 \times 6.129 = 1.787,2$$

$$\text{MN } 2pq \times N = 0,497 \times 6.129 = 3.044,9$$

$$\text{NN } q^2 \times N = 0,211 \times 6.129 = 1.296,9$$

So sánh các số liệu quan sát và kỳ vọng về từng kiểu gene ta thấy có sự phù hợp sít sao, chứng tỏ quần thể ở trạng thái cân bằng H-W.

Thật vậy, nếu kiểm tra bằng trắc nghiệm  $\chi^2$ , ta có:

$$\chi^2 = \frac{(1787 - 1787,2)^2}{1787,2} + \frac{(3039 - 3044,9)^2}{3044,9} + \frac{(1303 - 1296,9)^2}{1296,9} = 0,04$$

Tra bảng phân phối  $\chi^2$  ứng với  $P = 0,05$  và 1 bậc tự do ta tìm được trị số  $\chi^2$  bằng 3,84. Vì trị số thực tế là rất nhỏ so với trị số lý thuyết, chứng tỏ giữa các số liệu quan sát và kỳ vọng hầu như trùng khớp nhau hoàn toàn; nghĩa là, quần thể ở trạng thái cân bằng H-W.

### III. Mở rộng nguyên lý Hardy-Weinberg

Sau đây ta lần lượt xét xem khả năng áp dụng nguyên lý H-W vào ba trường hợp: một gene có nhiều hơn hai allele (đa allele), một gene có hai allele với tần số khác nhau ở hai giới tính, và gene liên kết với giới tính.

#### 1. Đa allele (multiple alleles)

Với quần thể ngẫu phối như đã nói ở trước, ở đây ta chỉ thay giả thiết

một locus A có ba allele:  $A_1, A_2$  và  $A_3$  với tần số tương ứng là  $p_1, p_2$  và  $p_3$  ( $p_1 + p_2 + p_3 = 1$ ). Khi đó trong quần thể có tất cả sáu kiểu gene với số lượng cá thể tương ứng như sau :

Kiểu gene :	$A_1A_1$	$A_2A_2$	$A_3A_3$	$A_1A_2$	$A_1A_3$	$A_2A_3$	Tổng
Số lượng :	$N_{11}$	$N_{22}$	$N_{33}$	$N_{12}$	$N_{13}$	$N_{23}$	$N$

Theo nguyên tắc, ta tính được các tần số allele:

$$p_1 = N_{11} + \frac{1}{2} (N_{12} + N_{13})$$

$$p_2 = N_{22} + \frac{1}{2} (N_{12} + N_{23})$$

$$p_3 = N_{33} + \frac{1}{2} (N_{13} + N_{23})$$

Bằng cách lập bảng tổ hợp ngẫu nhiên của các giao tử và tần số của chúng, hoặc bằng cách khai triển bình phương của một tam thức ta tính được các tần số cân bằng H-W chỉ sau một thế hệ ngẫu phối như sau:

$$(p_1 + p_2 + p_3)^2 = p_1^2 + p_2^2 + p_3^2 + 2p_1p_2 + 2p_1p_3 + 2p_2p_3 = 1$$

Tổng quát, một locus có  $n$  allele sẽ có tất cả  $n(n + 1)/2$  kiểu gene, trong đó gồm  $n$  kiểu đồng hợp và  $n(n - 1)/2$  kiểu dị hợp. Tần của một allele bất

kỳ ( $p_i$ ) được tính theo công thức: 
$$p_i = p_{ii} + \frac{1}{2} \sum_{i \neq j=1}^n P_{ij}$$

trong đó  $p_{ii}$  - tần số kiểu gene đồng hợp và  $p_{ij}$  - tần số kiểu gene dị hợp.

*Vi dụ:* Thông thường hệ nhóm máu ABO được lấy ví dụ cho ba allele. Vì các allele  $I^A$  và  $I^B$  là đồng trội và allele  $I^O$  là lặn, nên trong quần thể người bất kỳ nào cũng sẽ có bốn nhóm máu A, B, AB và O ứng với sáu kiểu gene. Để tính các tần số allele trong trường hợp này ta phải giả định quần thể ở trạng thái cân bằng. Đặt tần số của các allele  $I^A, I^B$  và  $I^O$  lần lượt là  $p, q$  và  $r$  ( $p + q + r = 1$ ). Khi đó ta tính được tần số H-W của các nhóm máu chính là các tần số quan sát được (bảng 12.5).

Phương pháp tính các tần số allele như sau: Trước tiên, tần số allele  $I^O$  ( $r$ ) bằng các căn bậc hai của tần số nhóm máu O ( $r^2$ ). Tần số của hai allele còn lại,  $p$  và  $q$ , được tính bằng cách kết hợp tần số H-W của một nhóm máu A hoặc B với nhóm máu O theo một trong hai phương pháp sau:

Phương pháp 1	Phương pháp 2
<p>Ta có <math>f(A+O) = p^2 + 2pr + r^2 = (p + r)^2</math></p> <p><math>\Leftrightarrow p+r = \sqrt{f(A+O)}</math></p> <p><math>\Rightarrow p = \sqrt{f(A+O)} - r</math></p> <p>Tương tự, ta có :</p>	<p>Vì <math>p + q + r = 1 \Rightarrow q + r = 1 - p</math></p> <p>Bình phương 2 vế ta được:</p> <p><math>(1 - p)^2 = (q + r)^2 = f(B + O)</math></p> <p><math>\Leftrightarrow 1 - p = \sqrt{f(B + O)}</math></p>

$q = \sqrt{f(B+O)} - r$	$\Rightarrow p = 1 - \sqrt{f(B+O)}$
	Tương tự, ta có: $q = 1 - \sqrt{f(A+O)}$

Một cách tương đối, ta có thể tính  $p$  hoặc  $q$  rồi suy ra cái còn lại dựa vào tổng  $p + q + r = 1$ . Tuy nhiên, nếu tính cẩn thận cả ba tần số theo một trong hai phương pháp trên ta sẽ biết được trị số thực của chúng. Khi đó tổng các tần số allele tính được sẽ không đúng bằng đơn vị một cách chính xác. Điều này được lý giải là do tỷ lệ các kiểu gene trong mẫu không phải là các tỷ lệ H-W chính xác và hơn nữa, nhóm máu AB đã không được sử dụng trong tính toán. Vì vậy, khi kiểu hình không được sử dụng đến (ở đây là nhóm máu AB) mà có tần số cao hơn thì sự mất mát thông tin sẽ nghiêm trọng hơn, và phải cần đến một phương pháp chính xác hơn.

**Bảng 12.5** Tương quan giữa các nhóm máu, kiểu gene và tần số của chúng

Nhóm máu	Kiểu gene	Tần số	
		Kỳ vọng	Quan sát
A	$I^A I^A + I^A I^O$	$p^2 + 2pr$	0,41716
B	$I^B I^B + I^B I^O$	$q^2 + 2qr$	0,08560
O	$I^O I^O$	$r^2$	0,46684
AB	$I^A I^B$	$2pq$	0,03040
Tổng		1	1,0

Bây giờ ta hãy xét một mẫu nghiên cứu trên 190.177 phi công vương quốc Anh (UK) gồm 79.334 A, 16.279 B, 88.782 O, và 5.782 AB (Race và Sanger, 1954; dẫn theo Falconer 1989). Tương quan giữa các nhóm máu, kiểu gene và các tần số của chúng được trình bày ở bảng 12.5.

Áp dụng hai phương pháp trên ta tính được các tần số allele như sau:

Allele	Tần số	
	Phương pháp 1	Phương pháp 2
$I^A$	0,2569	0,2567
$I^B$	0,0600	0,0598
$I^O$	0,6833	0,6833
Tổng	1,0002	0,9998

## 2. Tần số allele sai biệt giữa hai giới tính

Trên thực tế, các tần số allele nhiễm sắc thể thường ở hai giới tính có thể khác nhau. Chẳng hạn, trong chăn nuôi gia súc - gia cầm tùy theo mục tiêu kinh tế là lấy sữa, thịt hoặc trứng...mà tương quan số lượng cá thể đực-cái sẽ khác nhau. Khi đó việc áp dụng nguyên lý H-W sẽ như thế nào? Để xét quần thể này, ta sử dụng ký hiệu và giả thiết sau :

Allele	Tần số	
	Giới đực	Giới cái
A <sub>1</sub>	p'	p''
A <sub>2</sub>	q'	q''
Tổng	1	1

Bằng cách lập bảng tổ hợp của các giao tử, ta xác định được cấu trúc di truyền của quần thể sau một thế hệ ngẫu phối:

$$(p'A_1 : q'A_2)(p''A_1 : q''A_2) = p'p''A_1A_1 : (p'q'' + p''q') A_1A_2 : q'q''A_2A_2$$

Rõ ràng là nó không thỏa mãn công thức H-W. Bây giờ đến lượt tần số các allele của quần thể này là như sau:

$$f(A_1) = p'p'' + \frac{1}{2}(p'q'' + p''q')$$

Thay giá trị  $q'' = 1 - p''$ , ta có:

$$f(A_1) = \frac{1}{2}(p' + p'')$$

Tương tự:  $f(A_2) = \frac{1}{2}(q' + q'')$

Đặt  $f(A_1) = p$  và  $f(A_2) = q$ , khi đó cấu trúc di truyền quần thể ở thế hệ tiếp theo sẽ thỏa mãn công thức H-W:  $p^2 A_1A_1 : 2pqA_1A_2 : q^2 A_2A_2$ .

Điều đó chứng tỏ rằng, nếu như các tần số allele (autosome) khởi đầu là khác nhau ở hai giới, thì chúng sẽ được san bằng chỉ sau một thế hệ ngẫu phối và quần thể đạt trạng thái cân bằng sau hai thế hệ.

*Ví dụ:* Một quần thể khởi đầu có tần số các allele A và a ở hai giới như sau:  $p' = 0,8$ ;  $q' = 0,2$ ;  $p'' = 0,4$ ; và  $q'' = 0,6$ . Nếu như ngẫu phối xảy ra, thì ở thế hệ thứ nhất có tần số các kiểu gene là:  $0,32AA : 0,56Aa : 0,12aa$ .

Và tần số cân bằng của mỗi allele lúc đó như sau:

$$p = \frac{1}{2}(0,8 + 0,4) = 0,32 + \frac{1}{2}(0,56) = 0,6$$

$$q = \frac{1}{2}(0,2 + 0,6) = 0,12 + \frac{1}{2}(0,56) = 0,4$$

Ở thế hệ thứ hai, quần thể đạt cân bằng với các tần số H-W là:

$$0,36AA : 0,48Aa : 0,16aa$$



### 3. Các gene liên kết trên X

Trong trường hợp các gene liên kết với giới tính, tình hình trở nên phức tạp hơn rất nhiều. Ở giới đồng giao tử, mối quan hệ giữa tần số allele và tần số kiểu gene tương tự như một gene autosome, nhưng ở giới dị giao tử chỉ có hai kiểu gene và mỗi cá thể chỉ mang một allele. Để cho tiện, ta coi giới dị giao tử là giới đực. Bây giờ ta xét hai allele  $A_1$  và  $A_2$  với tần số tương ứng là  $p$  và  $q$ , và đặt các tần số kiểu gene như sau:

Kiểu gene:	$A_1A_1$	$A_1A_2$	$A_2A_2$	$A_1$	$A_2$
Tần số :	$P$	$H$	$Q$	$R$	$S$

Theo nguyên tắc, ta xác định được tần số của một allele (ví dụ  $A_1$ ):

- ở giới cái ( $p_c$ ):  $p_c = P + \frac{1}{2}H$
- ở giới đực ( $p_d$ ):  $p_d = R$
- chung cả quần thể ( $\bar{p}$ ):  $\bar{p} = \frac{2}{3}p_c + \frac{1}{3}p_d$

*Lưu ý:* Mỗi con cái có hai nhiễm sắc thể X và mỗi con đực chỉ có một X; vì tỉ lệ đực : cái trên nguyên tắc là 1:1, cho nên  $\frac{2}{3}$  các gene liên kết giới tính trong quần thể là thuộc về giới cái và  $\frac{1}{3}$  thuộc về giới đực. Vì vậy, tần số của các allele  $A_1$  trong cả quần thể là:  $\bar{p} = \frac{2}{3}p_c + \frac{1}{3}p_d$ .

Rõ ràng là các tần số allele ở hai phần đực và cái là khác nhau, do đó quần thể không ở trạng thái cân bằng. Trong khi tần số allele trong cả quần thể không thay đổi qua các thế hệ, nhưng sự phân phối các allele giữa hai giới có sự dao động khi quần thể tiến dần đến sự cân bằng. Điều này được chứng minh như sau. Theo quy luật liên kết gene trên X, các con đực nhận các gene liên kết giới tính chỉ từ các cơ thể mẹ, vì vậy  $p_d$  ở thế hệ con bằng với  $p_c$  ở thế hệ trước; các con cái nhận các gene liên kết giới tính đồng đều từ cả hai bố mẹ, vì vậy  $p_c$  ở thế hệ con bằng trung bình cộng của  $p_d$  và  $p_c$  ở thế hệ trước. Nếu dùng dấu phẩy trên đầu để chỉ tần số allele thế hệ con, ta có:

$$p'_d = p_c$$

$$p'_c = \frac{1}{2}(p_c + p_d)$$

Từ đây xác định được mức chênh lệch hay là hiệu số giữa các tần số allele của hai giới:  $p'_c - p'_d = \frac{1}{2}(p_d + p_c) - p_c = -\frac{1}{2}(p_c - p_d)$

Nghĩa là, hiệu số của các tần số allele giữa hai giới ở thế hệ con bằng một nửa hiệu số của các tần số allele giữa hai giới ở thế hệ bố mẹ của nó, nhưng ngược dấu. Như vậy, sự phân bố các allele giữa hai giới có sự giao động theo quy luật sau: Cứ sau một thế hệ, mức chênh lệch đó giảm đi một nửa và như thế quần thể tiến dần đến trạng thái cân bằng cho đến khi các tần số gene ở hai giới là cân bằng nhau, nghĩa là  $p_c = p_d = \bar{p}$ .

*Ví dụ:* Theo kết quả một mẫu nghiên cứu trên mèo ở Luân Đôn (Searle, 1949; trong Falconer 1989) cho thấy trong số 338 mèo cái có 277 con lông đen (BB), 54 con thể khảm (BO) và 7 con lông da cam (OO), và trong số 353 mèo đực có 311 đen (B) và 42 da cam (O). Tính trạng này tuân theo quy luật di truyền liên kết với giới tính như đã đề cập trước đây.

Để kiểm tra xem quần thể có ở trạng thái cân bằng hay không, trước tiên ta hãy xem liệu có bằng chứng nào về sự giao phối ngẫu nhiên? Phép thử đầu tiên là xem tần số allele ở hai giới có giống nhau không. Tính toán cụ thể cho thấy các tần số gene ở hai giới khác nhau không đáng kể.

$$\text{- Ở giới cái: } f(B) = p_c = (2 \times 277) + 54 / (2 \times 338) = 0,8994$$

$$f(O) = q_c = (2 \times 7) + 54 / (2 \times 338) = 0,1006$$

$$\text{- Ở giới đực: } p_d = 311 / 353 = 0,881$$

Từ tần số các allele ở giới cái, ta tính được số cá thể kỳ vọng của mỗi kiểu gene ở giới này như sau:

Số cá thể	Kiểu gene			Tổng
	BB	BO	OO	
Quan sát	277	54	7	338
Kỳ vọng	273,2	61,2	3,4	338
$\chi^2_{(1)} = 4,6$	P = 0,04			

Kết quả cho thấy các số liệu quan sát không phù hợp lắm với số kỳ vọng mà chủ yếu là các số liệu thấp (kiểu BO và OO). Nếu vậy thì sự không nhất quán đó có thể là do giao phối ngẫu nhiên, nhưng cũng có thể do thị hiếu của con người thiên về các màu sắc đã làm sai lệch mẫu, không đại diện được cho quần thể. Qua sự phân tích này cùng với sự sai khác chút ít về tần số gene giữa hai giới đã nói ở trên, chúng ta chẳng có lý do gì để nghĩ rằng quần thể này không ở trạng thái cân bằng.

#### IV. Nội phối (*inbreeding*)

Bên cạnh các quần thể giao phối ngẫu nhiên đã xét, còn có các ngoại lệ đối với giả định này ở một số loài, chẳng hạn như các thực vật hoặc động vật không xương sống tự thụ tinh. Nội phối là sự *giao phối không ngẫu nhiên* (nonrandom mating) xảy ra giữa các cá thể có quan hệ họ hàng gần; nó có tầm quan trọng đặc biệt ở người, bởi vì nhiều người mang các bệnh di truyền lặn sinh ra từ sự kết hôn họ hàng. Hơn nữa, nội phối được sử dụng trong chọn giống thực vật và cả động vật để tạo ra các dòng mang những đặc tính mong muốn nào đó.

Sự giao phối không ngẫu nhiên đối với các kiểu gene xảy ra trong các

quần thể, trong đó các cá thể giao phối hoặc là có quan hệ họ hàng gần hơn hoặc là ít gần hơn so với kỳ vọng giao phối ngẫu nhiên từ quần thể. Kết quả của hai kiểu giao phối này được gọi tương ứng là *nội phối* (inbreeding) và *ngoại phối* (outbreeding). Mặc dù cả hai kiểu giao phối này tự nó không làm thay đổi tần số allele, nhưng đều làm thay đổi tần số các kiểu gene (ở mọi locus trong bộ gene).

\* Trong một quần thể nội phối, tần số của các thể đồng hợp tăng lên và tần số của các thể dị hợp giảm xuống so với các tỉ lệ giao phối ngẫu nhiên.

\* Trong một quần thể ngoại phối, tình hình xảy ra ngược lại, với tần số các thể dị hợp tăng và tần số các thể đồng hợp giảm so với các tỉ lệ giao phối ngẫu nhiên.

### 1. Tự thụ tinh (self-fertilization)

Kiểu cực đoan nhất của nội phối là sự tự thụ tinh, trong đó hạt phấn và noãn (hay tinh trùng và trứng) được sinh ra trên cùng một cá thể. Hình thức sinh sản này phổ biến ở một số nhóm thực vật. Trong trường hợp tự thụ tinh hoàn toàn, một quần thể được phân thành nhiều dòng nội phối mau chóng trở nên đồng hợp cao độ. Đó là trường hợp các dòng đậu thuần chủng được sử dụng bởi Mendel.

**Bảng 12.6 Sự biến đổi tần số kiểu gene trong một quần thể khởi đầu với chỉ những thể dị hợp tự thụ tinh hoàn toàn qua nhiều thế hệ**

Thế hệ	Tần số kiểu gene			Tần số allele	Hệ số nội phối
	AA	Aa	aa	a	(F)
0	0	1	0	0,5	0
1	1/4	1/2	1/4	0,5	1/2
2	3/8	1/4	3/8	0,5	3/4
...					
n	$\frac{1-1/2^n}{2}$	$1/2^n$	$\frac{1-1/2^n}{2}$	0,5	$1-1/2^n$
$\infty$	1/2	0	1/2	0,5	1

Để minh họa cho điều này, ta giả sử ở thế hệ bố mẹ có ba kiểu gene AA, Aa và aa với tần số tương ứng là P, H và Q. Khi sự tự thụ tinh là hoàn toàn, các kiểu gene AA và aa sinh ra đời con tương ứng toàn là AA và aa; còn kiểu gene Aa theo quy luật phân ly Mendel sẽ cho đời con gồm một nửa là Aa và nửa kia phân đồng đều cho hai kiểu đồng hợp, AA và aa. Khi đó tần số của các kiểu gene AA, Aa và aa ở thế hệ con tương ứng là:  $(P + \frac{1}{4}H)$ ,  $(\frac{1}{2}H)$  và  $(Q + \frac{1}{4}H)$ . Như vậy, sau một thế hệ tự thụ tinh hoàn

toàn, tần số thể dị hợp giảm đi một nửa so với bố mẹ, trong khi tần số của hai kiểu đồng hợp tăng lên.

Để hiểu rõ sự nội phối làm thay đổi các tỷ lệ kiểu gene ra sao, ta hãy xét quần thể ban đầu gồm chỉ những thể dị hợp Aa khi tự thụ tinh hoàn toàn qua nhiều thế hệ (bảng 12.6). Vì cứ sau mỗi thế hệ mức dị hợp tử lại giảm đi một nửa so với thế hệ trước nó, nên ở thế hệ thứ n mức dị hợp bằng  $(\frac{1}{2})^n$  của trị số ban đầu:  $H_n = (\frac{1}{2})^n H_0$ , trong đó  $H_0$  và  $H_n$  là tỷ lệ thể dị hợp ở thế hệ ban đầu và thế hệ thứ n. Và tỷ lệ mỗi kiểu đồng hợp là  $(1 - \frac{1}{2})/2$ . Khi n tiến đến vô hạn thì thành phần dị hợp bị triệt tiêu và chỉ còn lại hai thành phần đồng hợp với tần số là  $\frac{1}{2}$ . Lúc này, tần số mỗi kiểu gene bằng tần số allele.

Một điểm quan trọng của nội phối là ở chỗ, mặc dù các tần số kiểu gene có thể bị thay đổi nhiều, nhưng các tần số allele vẫn được giữ nguyên không đổi qua các thế hệ. Bạn có thể tự chứng minh điều này?

## 2. Hệ số nội phối (inbreeding coefficient)

Để mô tả hiệu quả của nội phối lên các tần số kiểu gene nói chung, ta sử dụng phép đo gọi là *hệ số nội phối* (F). Trị số này được định nghĩa là xác suất mà hai allele tại một locus trong một cá thể là giống nhau về nguồn gốc (các allele được coi là giống nhau về nguồn gốc khi hai allele đó trong một cơ thể lưỡng bội bắt nguồn từ một allele cụ thể của tổ tiên).

*Tính chất của hệ số nội phối* (F):

- + Trị số F chạy từ 0 đến 1 (xem cột cuối ở bảng 12.6).
- +  $F = 1$  khi tất cả các kiểu gene trong quần thể là đồng hợp chứa các allele giống nhau về nguồn gốc.
- +  $F = 0$  khi không có các allele giống nhau về nguồn gốc.
- + Trong một quần thể ngẫu phối có kích thước lớn, F được coi là gần bằng 0, bởi vì bất kỳ sự nội phối nào cũng có thể xảy ra giữa các cá thể họ hàng rất xa và vì vậy sẽ có tác dụng nhỏ lên hệ số nội phối.

Giả sử rằng quần thể gồm ba kiểu gene AA, Aa và aa được phân tách thành một tỷ lệ nội phối (F) và một tỷ lệ ngẫu phối (1 - F). Trong quần thể nội phối, tần số của AA, Aa, và aa tương ứng là p, 0, và q. Đây là tỷ lệ của các dòng được kỳ vọng đối với mỗi kiểu gene, nếu như sự tự thụ tinh hoàn toàn diễn ra liên tục. Bằng cách cộng các tỷ lệ nội phối và ngẫu phối với nhau và sử dụng mối quan hệ  $q = 1 - p$ , lúc đó tần số các kiểu gene trở thành như sau (xem bảng 12.7):

$$P = p^2 + Fpq$$

$$H = 2pq - 2Fpq$$

$$Q = q^2 + Fpq$$

Trong mỗi phương trình trên, số hạng đầu là tỷ lệ H-W của các kiểu gene và số hạng sau là độ lệch so với trị số đó. Lưu ý rằng các cá thể đồng hợp, ví dụ AA, có thể hoặc là do hai allele giống nhau về nguồn gốc, nghĩa là bắt nguồn từ cùng một allele tổ tiên (số hạng  $Fpq$ ) hoặc là do hai allele giống nhau về loại sinh ra qua ngẫu phối (số hạng  $p^2$ ). Độ lớn của hệ số nội phối phản ánh độ lệch của các kiểu gene so với các tỷ lệ H-W; nghĩa là, lúc  $F = 0$  thì các hợp tử đạt tỷ lệ H-W, và khi  $F > 0$  do có nội phối, thì xảy ra sự giảm thiểu các thể dị hợp và đôi thừa các thể đồng hợp.

**Bảng 12.7 Tần số của các kiểu gene khác nhau khi trong quần thể xảy ra cả nội phối lẫn ngẫu phối**

Kiểu gene	Nội phối (F)	Ngẫu phối (1 - F)	Tổng
AA	$Fp$	$(1 - F)p^2$	$Fp + (1 - F)p^2 = p^2 + Fpq$
Aa	–	$(1 - F)2pq$	$(1 - F)2pq = 2pq - 2Fpq$
aa	$Fq$	$(1 - F)q^2$	$Fq + (1 - F)q^2 = q^2 + Fpq$
	F	1 - F	1 1

### 3. Tính toán hệ số nội phối

Có hai cách ước tính hệ số nội phối, đó là dựa vào các tần số kiểu gene hoặc là dựa vào các phả hệ.

• Với phương pháp thứ nhất, ta ước tính hệ số nội phối trong một quần thể tự nhiên bằng cách sử dụng biểu thức về tần số các thể dị hợp đã cho ở trên. Qua đó ta có thể tìm ra biểu thức cho F như sau:

$$H = 2pq - 2Fpq = (1 - F)2pq$$

$$1 - F = H/2pq$$

$$\text{Suy ra } F = 1 - (H/2pq)$$

Từ phương trình trên cho thấy hệ số nội phối (F) là một hàm của tỷ số giữa mức dị hợp tử quan sát được (H) và mức dị hợp tử kỳ vọng ( $2pq$ ). Trường hợp có nội phối, H nhỏ hơn  $2pq$ , vì vậy  $F > 0$ . Nếu như không có thể dị hợp nào cả ( $H = 0$ ), thì hệ số nội phối bằng 1.

Nhiều loài thực vật có hệ thống giao phối bao gồm cả tự thụ phấn và giao phấn tự do với các cá thể khác. Nếu như tỷ lệ tự thụ phấn cao, thì hầu như tất cả các cá thể trong quần thể là các thể đồng hợp. Ví dụ, một quần thể thực vật gồm ba kiểu gene AA, Aa và aa với các tần số tương ứng là  $P = 0,70$ ,  $H = 0,04$  và  $Q = 0,26$ . Ta có thể ước tính hệ số nội phối như sau :

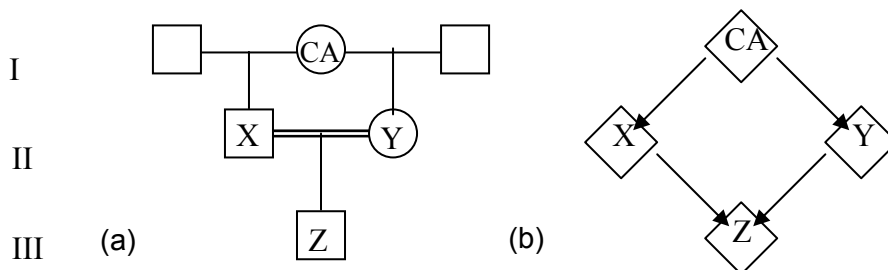
Trước tiên, tính được các tần số allele A và a (p và q) :

$$p = 0,70 + \frac{1}{2} (0,04) = 0,72 \quad \text{và} \quad q = 1 - p = 0,28$$

$$\text{Vây hệ số nội phối } F = 1 - (0,04/2 \times 0,72 \times 0,28) = 0,901$$

Trị số F ở đây rất cao, gợi ý rằng hầu hết quần thể này sinh sản bằng tự thụ phấn và chỉ một số rất nhỏ là tạp giao.

• Phương pháp thứ hai để thu nhận hệ số nội phối cho đời con là từ một phả hệ trong đó có xảy ra sự *giao phối cận huyết* (consanguineous mating). Trong trường hợp này ta sử dụng một phả hệ để tính xác suất của các tổ hợp chứa các allele giống nhau về nguồn gốc ở đời con. Ví dụ, ta hãy tính hệ số nội phối cho một đời con của hai *anh chị em bán đồng huyết* (half-sibs), tức các cá thể sinh ra từ cùng một bố (hoặc mẹ). Hình 12.1a cho phả hệ về kiểu giao phối này, trong đó X và Y là hai anh em có cùng mẹ nhưng khác cha. Người mẹ của X và Y được biểu thị là *tổ tiên chung* (CA = common ancestor). Còn hai người cha không góp phần vào hệ số nội phối được biểu diễn bằng các hình vuông trắng. Ở hình 12.1b, cùng một phả hệ như thế nhưng biểu diễn theo một cách khác, bỏ qua các ký hiệu cha mẹ còn các dấu quả trám biểu thị cho tất cả các cá thể, vì giới tính không quan trọng trong việc xác định hệ số nội phối ở đây. Các mũi tên trên hình vẽ chỉ hướng truyền từ bố mẹ đến con cái.



**Hình 12.1** Phả hệ minh họa sự kết hôn giữa hai anh em bán đồng huyết, X và Y. (a) với tất cả các cá thể; (b) không có bố. Ở đây CA = tổ tiên chung, và đường kẻ đôi chỉ sự giao phối cận huyết.

Giả sử người mẹ (CA) có kiểu gene là Aa. Để tính hệ số nội phối, ta cần phải biết xác suất mà đứa cháu của bà, Z, có kiểu gene AA hoặc aa, là giống nhau về nguồn gốc đối với một trong hai allele của bà. Trước tiên ta xét Z là AA, chỉ có thể xảy ra nếu như mỗi bên X và Y đều đóng góp vào Z một giao tử chứa A. Xác suất của allele A trong X là xác suất mà một allele A đến từ CA, hay  $\frac{1}{2}$ . Vì xác suất truyền đạt allele A từ X sang Z cũng là  $\frac{1}{2}$ , nên xác suất kết hợp của hai sự kiện này là  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$  (qui tắc nhân xác suất). Tương tự, xác suất để Z nhận được allele A từ Y là  $\frac{1}{4}$ . Vì vậy xác suất của một đứa con AA nhận được allele A từ mỗi bên X và Y

là  $\frac{1}{4} \times \frac{1}{4} = 1/16$  hay 0,0625. Bằng phương pháp này ta tính được xác suất của một đứa con có kiểu gene aa là 1/16. Như vậy xác suất toàn bộ các tổ hợp có chứa các allele giống nhau về nguồn gốc ở Z lúc đó là  $1/16 + 1/16 = 1/8$  hay 0,125 (qui tắc cộng xác suất).

Để đơn giản, trong tính toán hệ số nội phối từ một phả hệ người ta đã đề xuất một phương pháp gọi là kỹ thuật đếm chuỗi (*chain-counting technique*). Một chuỗi đối với một tổ tiên chung cho trước bắt đầu với một bố mẹ của cá thể nội phối, ngược trở lên phả hệ cho đến tổ tiên chung, và trở lại với bố mẹ đó. Ví dụ, từ hình 12.1 ta lập được chuỗi đơn giản X-CA-Y. Số cá thể trong chuỗi (n) được dùng để tính hệ số nội phối trong công thức sau đây:  $F = (1/2)^n$ . Với ví dụ trên, hệ số nội phối là  $(1/2)^3 = 0,125$ .

## V. Các nhân tố tác động lên thành phần di truyền quần thể

Trong tự nhiên, thành phần di truyền của các quần thể nói chung là không ổn định như nguyên lý Hardy-Weinberg đã vạch ra, mà luôn luôn bị biến động dưới ảnh hưởng của nhiều nhân tố khác nhau. Lần đầu tiên, vấn đề này được Charles Darwin nêu lên trong tác phẩm *Nguồn gốc các loài bằng con đường chọn lọc tự nhiên* (On the Origin of Species by Means of Natural Selection) năm 1859. Theo quan niệm hiện nay, có bốn nhân tố cơ bản làm thay đổi tần số các allele của các quần thể và xác định quá trình tiến hóa, đó là: đột biến, biến động di truyền ngẫu nhiên, di-nhập cư và chọn lọc tự nhiên. Về chi tiết, có thể xem thêm trong *Bài giảng Di truyền học quần thể* (Hoàng Trọng Phán 2003); ở đây chúng ta chỉ tìm hiểu sơ lược vai trò và hiệu quả của mỗi nhân tố.

### 1. Đột biến

*Đột biến* (mutation) có nhiều loại khác nhau như đã được trình bày ở các chương 3 và 8; ở đây chúng ta chỉ đề cập đến vai trò, tính chất và áp lực của các đột biến gene tự phát đối với quá trình tiến hóa và chọn lọc. Phần lớn đột biến mới xuất hiện có hại cho cơ thể

Đột biến là nguồn cung cấp chủ yếu các biến dị di truyền mới trong một quần thể-loài, vì vậy nó được xem là một quá trình quan trọng đặc biệt trong di truyền học quần thể. Nói chung, phần lớn đột biến mới xuất hiện là có hại, một số đột biến là trung tính và chỉ một số ít là có lợi cho bản thân sinh vật. Theo thuyết trung tính (Kimura 1983), đại đa số các biến đổi tiến hóa ở cấp độ phân tử được gây nên không phải bằng chọn lọc Darwin mà bằng sự cố định ngẫu nhiên các thể đột biến trung tính hoặc hầu như trung tính về mặt chọn lọc; áp lực đột biến và biến động di truyền ngẫu nhiên chiếm ưu thế trong sự biến đổi tiến hóa ở cấp độ phân tử.

Để xét xem hiệu quả của đột biến lên sự biến đổi di truyền trong một

quần thể, ta xét hai allele A (kiểu đại) và a (gây hại) với tần số ban đầu tương ứng là p và q; gọi u là tỷ lệ đột biến thuận từ A thành a cho một giao tử mỗi thế hệ, và v là tỷ lệ đột biến nghịch từ a thành A. Các allele A do đột biến thuận thành a đã làm tăng tần số của allele a lên một lượng là up, trong khi đó tần số allele a do đột biến nghịch có thể bị giảm đi một lượng là vq. Như vậy, nhìn toàn cục thì sau mỗi thế hệ sự biến đổi trong tần số của allele a ( $\Delta q$ ) do đột biến là:

$$\Delta q = up - vq$$

Trị số dương cực đại cho sự biến đổi này là u, khi p = 1 và q = 0 (nghĩa là tất cả các allele đều là kiểu đại). Trị số âm cực đại là v, khi p = 0 và q = 1. Tuy nhiên do các tỷ lệ đột biến u và v nói chung là nhỏ, nên sự biến đổi được kỳ vọng do đột biến cũng rất nhỏ. Chẳng hạn, nếu ta cho u =  $10^{-5}$ , v =  $10^{-6}$  và q = 0,0, lúc đó:

$$\Delta q = (0,00001)(1,0) - (0,000001)(0,0) = 0,00001$$

Mặc dù đột biến chỉ gây một hiệu quả nhỏ trong tần số allele ở mỗi thế hệ, nhưng nó lại có tầm quan trọng căn bản trong việc xác định mức độ gây ra các bệnh di truyền hiếm. Trên thực tế, sự cân bằng giữa đột biến (làm tăng tần số của allele bệnh) và chọn lọc (làm giảm tần số của allele bệnh) có thể lý giải mức độ quan sát được của các bệnh như bạch tạng chẳng hạn. Ngoài ra, đột biến cùng với sự biến động di truyền ngẫu nhiên cho phép giải thích hợp lý cho số lượng lớn các biến đổi phân tử quan sát được gần đây ở nhiều loài (xem Kimura 1983).

## 2. Biến động di truyền ngẫu nhiên

*Biến động di truyền ngẫu nhiên* (genetic random drift), hay nói gọn là biến động di truyền, đó là những sự biến đổi ngẫu nhiên vô hướng về tần số allele trong tất cả các quần thể, nhưng đặc biệt là ở các quần thể nhỏ.

Biến động di truyền là một quá trình thuần túy ngẫu nhiên, mang tính xác suất và tỷ lệ nghịch với kích thước quần thể. Trong một quần thể lớn, thông thường biến động di truyền chỉ gây ra một sự thay đổi ngẫu nhiên nhỏ. Nhưng trong các quần thể nhỏ, nó có thể gây ra những sự biến động lớn về tần số allele qua các thế hệ khác nhau. Chính hiện tượng này là nguyên nhân tạo nên sự khác biệt đa dạng về mặt di truyền ở các quần thể nhỏ và dần dần dẫn tới sự cách ly sinh sản trong quá trình tiến hóa của loài. Và sự biến động di truyền có thể là nguyên nhân làm cho mức dị hợp tử thấp quan sát được ở một số loài có nguy cơ bị diệt vong.

## 3. Dòng gene hay sự di nhập cư

*Di-nhập cư* (migration) hay *dòng gene* (gene flow) là sự di chuyển của các cá thể từ một quần thể này sang một quần thể khác, kéo theo việc đưa



vào các allele nhập cư mới thông qua sự giao phối và sinh sản sau đó.

Như vậy, dòng gene không làm thay đổi các tần số allele của cả loài, nhưng có thể làm biến đổi cục bộ (tần số allele so với nguyên lý H-W) khi tần số các allele của những cá thể di cư tới là khác với của các cá thể cư trú tại chỗ. Giả sử các cá thể từ các quần thể xung quanh di cư tới một quần thể địa phương ta nghiên cứu với tốc độ  $m$  mỗi thế hệ và giao phối với các cá thể cư trú ở đó. Và cũng giả sử rằng tần số allele A của quần thể nguồn gene nhập cư là  $P$  và của quần thể nghiên cứu là  $p_0$ . Khi đó, ở thế hệ thứ nhất sau nhập cư, tần số allele A của quần thể nghiên cứu sẽ là:

$$p_1 = (1 - m)p_0 + mP = p_0 - m(p_0 - P)$$

Sự biến đổi  $\Delta p$  về tần số allele sau một thế hệ là:  $\Delta p = p_1 - p_0$

Thay trị số  $p_1$  thu được ở trên, ta có:  $\Delta p = -m(p_0 - P)$

Điều đó có nghĩa là, tỷ lệ các cá thể di cư càng lớn và sự chênh lệch giữa hai tần số allele càng lớn, thì đại lượng  $\Delta p$  càng lớn. Lưu ý rằng  $\Delta p = 0$  chỉ khi hoặc  $m = 0$  hoặc  $(p_0 - P) = 0$ . Như vậy, trừ phi sự di cư dừng lại ( $m = 0$ ) còn thì tần số allele sẽ tiếp tục biến đổi cho đến khi nó trở nên giống nhau giữa quần thể địa phương và quần thể phụ cận ( $p_0 - P = 0$ ).

Sau thế hệ thứ nhất, hiệu số về tần số allele giữa hai quần thể trên là:

$$\begin{aligned} p_1 - P &= p_0 - m(p_0 - P) - P \\ &= (1 - m)(p_0 - P) \end{aligned}$$

Tương tự, sau thế hệ thứ hai, hiệu số về tần số allele đó sẽ là:

$$p_2 - P = (1 - m)^2 \cdot (p_0 - P)$$

Và sau  $n$  thế hệ di cư, ta có:

$$p_n - P = (1 - m)^n \cdot (p_0 - P)$$

Công thức này cho phép ta tính toán hiệu quả của  $n$  thế hệ di cư ở tốc độ  $m$  nào đó, nếu như biết được tần số các allele khởi đầu ( $p_0$  và  $P$ ):

$$p_n = (1 - m)^n \cdot (p_0 - P) + P$$

Hoặc nếu như biết được tần số các allele khởi đầu ( $p_0$  và  $P$ ), tần số allele của quần thể nghiên cứu tại một thời điểm nào đó ( $p_n$ ) cũng như số thế hệ  $n$ , ta có thể tính được tốc độ dòng gene ( $m$ ).

*Ví dụ:* Ở Mỹ (USA), những người có nguồn gốc hỗn chủng da trắng Capca (Caucasian) và da đen Châu Phi (African) được coi là thuộc quần thể người da đen. Sự pha tạp về chủng tộc có thể xem như là một quá trình của dòng gene từ quần thể Capca sang quần thể da đen. Tần số của allele  $R^0$  ở locus xác định các nhóm máu rhesus là  $P = 0,028$  ở các quần thể Capca nước Mỹ. Trong số các quần thể Châu Phi mà từ đó các tổ tiên của

người Mỹ da đen di cư đến, tần số allele  $R^o$  là 0,630. Tổ tiên Châu Phi của những người Mỹ da đen đã đến nước Mỹ cách đây khoảng 300 năm hay khoảng 10 thế hệ; nghĩa là  $n = 10$ . Tần số allele  $R^o$  trong số những người Mỹ hiện giờ là  $p_n = 0,446$ .

Bằng cách biến đổi lại phương trình trên, ta có:

$$(1 - m)^n = (p_n - P) : (p_o - P)$$

Thay các trị số đã cho, ta có:

$$(1 - m)^{10} = (0,446 - 0,028) : (0,630 - 0,028) = 0,694$$

$$1 - m = \sqrt[10]{0,694} = 0,964$$

Suy ra:  $m = 0,036$

Như vậy, dòng gene chuyển từ những người Mỹ Capca vào trong quần thể người Mỹ da đen đã xảy ra ở tốc độ tương đương với trị số trung bình là 3,6% mỗi thế hệ. Những tính toán tương tự bằng cách sử dụng các tần số allele ở nhiều locus khác cho các kết quả hơi khác một chút. Hơn nữa, mức độ pha tạp hỗn chủng có thể khác nhau ở các vùng khác nhau của nước Mỹ; nhưng rõ ràng là sự trao đổi gene đáng kể đã xảy ra (Ayala và Kiger 1980, tr.644; Hartl *et al* 1988, tr.214).

#### 4. Chọn lọc tự nhiên

Cả ba quá trình gây biến đổi tần số gene đã xét trên đây đều có một điểm chung là không một quá trình nào định hướng đối với sự *thích nghi* (adaptation). Theo nhận định của một số tác giả (Mayer 1974; Ayala và Kiger 1980), các quá trình này là ngẫu nhiên đối với sự thích nghi, do đó tự thân chúng sẽ phá hoại tổ chức và các đặc tính thích nghi của sinh vật. Chỉ có *chọn lọc tự nhiên* (natural selection) mới là quá trình thúc đẩy sự thích nghi và hạn chế các hiệu quả phá hoại tổ chức của các quá trình khác. Trong ý nghĩa đó, chọn lọc tự nhiên là quá trình tiến hóa khốc liệt nhất, bởi vì chỉ có nó mới có thể giải thích được bản chất thích nghi, *tính đa dạng* (diversity) và có tổ chức cao của các sinh vật.

Ý tưởng về chọn lọc tự nhiên như là quá trình nền tảng, là động lực của sự biến đổi tiến hóa do Charles Darwin và Alfred Russel Wallace độc lập đưa ra năm 1858. Lý luận tiến hóa bằng chọn lọc tự nhiên đã được phát triển đầy đủ, với chứng cứ ủng hộ xác đáng, trong cuốn *Nguồn gốc các loài* (The Origin of Species) do Darwin xuất bản năm 1859.

Theo Hartl *et al* (1988, 1997), trên quan điểm thuyết tổng hợp hiện đại, có thể hình dung chọn lọc tự nhiên xảy ra dựa trên ba điểm chính: (1) Ở mọi sinh vật, đời con được sinh ra nhiều hơn số sống sót và sinh sản; (2) Các cá thể khác nhau về khả năng sống sót và sinh sản và phần lớn những

khác biệt này là do kiểu gene; (3) Trong mỗi thế hệ, các kiểu gene sống sót sẽ sinh sản nhiều hơn và quyết định sự phân bố lại các kiểu gene ở thế hệ sau. Hậu quả là, các allele tăng cường sự sống sót và sinh sản sẽ gia tăng tần số từ thế hệ này sang thế hệ khác, và quần thể đó sẽ ngày càng sống sót và sinh sản tốt hơn với môi trường của nó.

Trên quan điểm đó, *chọn lọc tự nhiên được định nghĩa là sự sống sót và sinh sản biệt hóa của các kiểu gene.*

Để hiểu được các tác dụng của chọn lọc lên biến dị di truyền, ta phải xét xem *độ phù hợp tương đối* (relative fitness) của các kiểu gene khác nhau. Nó có nhiều thuật ngữ đồng nghĩa như: *độ phù hợp Darwin* (Darwinian fitness), *giá trị chọn lọc* (selective value), hay *giá trị thích nghi* (adaptive value). Tựu trung, nó có thể được định nghĩa như là khả năng tương đối của các kiểu gene khác nhau trong việc truyền lại các allele cho những thế hệ tương lai (Weaver và Hedrick 1997).

Bởi vì chọn lọc tự nhiên tác động bằng sự sinh sản biệt hóa, nên ta có thể xem nó là số đo hiệu năng sinh sản của một kiểu gene. Để cho tiện, các nhà di truyền học thường đặt định trị số độ phù hợp ( $w$ ) bằng 1 cho kiểu gene có hiệu năng sinh sản cao nhất. Một đơn vị đo liên quan là *hệ số chọn lọc* (selection coefficient), được ký hiệu bằng  $s$ , và được định nghĩa là  $s = 1 - w$ . Hệ số chọn lọc đo mức độ giảm bớt độ phù hợp của một kiểu gene. Giả sử mỗi thế hệ các kiểu gene AA và Aa đều sinh được 100 con, còn thể đồng hợp lặn sinh được 80 con; nếu ta coi độ phù hợp của các cá thể mang allele trội là 1, thì độ phù hợp của các thể đồng hợp lặn là 0,8. Hiệu số của các trị số độ phù hợp này chính là hệ số chọn lọc ( $s$ ), và trong trường hợp này  $s = 1 - 0,8 = 0,2$ . Nếu như các kiểu gene có khả năng sống sót và sinh sản như nhau thì  $s = 0$ ; nếu một kiểu gene nào đó gây chết hoặc làm bất thụ hoàn toàn thì  $s = 1$ .

#### 4.1. Chọn lọc và đột biến

Chọn lọc có xu hướng đào thải các allele có hại ra khỏi quần thể, trong khi đột biến có thể tạo ra các allele có hại mới. Quần thể sẽ giữ nguyên trạng thái phân bố các kiểu gene nếu như tần số đột biến mới xuất hiện vừa đúng bằng tần số allele bị chọn lọc đào thải. Sau đây ta thử xét sự cân bằng này trong trường hợp "Chọn lọc chống lại các đồng hợp tử lặn".

Giả sử A là allele bình thường và a là allele có hại với tần số tương ứng của chúng là  $p$  và  $q$ . Khi đó độ phù hợp hay giá trị thích nghi của các kiểu gene AA, Aa và aa tương ứng là 1: 1: 1- $s$ . Trong trường hợp này tốc độ đào thải allele a khỏi quần thể bởi chọn lọc sẽ là  $sq^2$ . Nếu cho rằng tốc độ đột biến thuận (A  $\rightarrow$  a) là  $u$ , thì tốc độ xuất hiện allele a mới trong quần thể là  $up$ . Vì  $p \approx 1$  (do tần số a rất thấp) nên có thể coi  $up \approx u$ . Với cơ chế

ngẫu phối, quần thể sẽ ở trạng thái cân bằng khi tốc độ xuất hiện đột biến mới bằng tốc độ đào thải, nghĩa là  $u = sq^2$ , hay khi tần số allele lặn trong quần thể ở mức  $q = \sqrt{u/s}$ . Tương tự, đối với allele trội,  $u = sp$  hay  $p = u/s$ .

*Ví dụ:* Tần số mắc bệnh PKU ở trẻ sơ sinh là khoảng 4 trên 100.000; do đó  $q^2 = 4 \times 10^{-5}$ . Hiệu quả sinh sản của các bệnh nhân không được chữa trị là zero, hay  $s = 1$ . Khi đó  $u = sq^2 = 4 \times 10^{-5}$ .

Tần số allele này trong các quần thể người là  $q = \sqrt{4 \times 10^{-5}} = 6,3 \times 10^{-3}$  và tần số của các thể dị hợp là:  $2pq \approx 2q = 2(6,3 \times 10^{-3}) = 1,26 \times 10^{-2}$

Điều đó có nghĩa là, trong 100 người có khoảng 1,3 người mang allele đó, mặc dù có 4 trong 100.000 người mắc bệnh PKU. Tần số của allele này có mặt trong các thể dị hợp bằng một nửa của  $1,26 \times 10^{-2}$  hay  $6,3 \times 10^{-3}$ ; và tần số của allele đó ở các thể đồng hợp là  $4 \times 10^{-5}$ . Do vậy các allele PKU có mặt trong các thể dị hợp nhiều hơn  $6,3 \times 10^{-3} / 4 \times 10^{-5} = 158$  lần so với các thể đồng hợp. Như đã nói từ đầu, các allele hiếm tồn tại trong quần thể hầu hết ở các thể dị hợp.

#### 4.2. Ưu thế dị hợp tử

Ưu thế dị hợp tử hay *siêu trội* (overdominance) là trường hợp chọn lọc ưu ái ủng hộ các thể dị hợp nhiều hơn cả hai dạng đồng hợp tử. Khi đó chọn lọc không loại thải allele nào. Ở mỗi thế hệ, các thể dị hợp sẽ sinh sản mạnh cho nhiều con cháu hơn các thể đồng hợp và chọn lọc sẽ giữ lại cả hai allele cho đến khi quần thể tạo được trạng thái cân bằng, với các tần số allele không thay đổi.

Một ví dụ nổi bật về hiện tượng siêu trội trong các quần thể người là bệnh thiếu máu hồng cầu hình liềm, một bệnh phổ biến ở châu Phi và châu Á. Bệnh này có liên quan đến một dạng sốt rét do ký sinh trùng phổ biến gây ra là *Plasmodium falciparum*. Allele  $Hb^S$  gây chết trước tuổi trưởng thành ở những người đồng hợp tử  $Hb^S Hb^S$ . Tần số allele này có thể cao hơn 10% ở các vùng có sốt rét nói trên, bởi vì các thể dị hợp  $Hb^A Hb^S$  đề kháng được sự nhiễm sốt rét, trong khi các thể đồng hợp  $Hb^A Hb^A$  thì không có khả năng đó.

### Câu hỏi và Bài tập

1. (a) Định nghĩa các khái niệm sau: di truyền học quần thể, quần thể, vốn gene, tần số allele và tần số kiểu gene. (b) Đặc trưng di truyền của các quần thể giao phối và nội phối là gì?

2. Tại sao nói khái niệm căn bản và công cụ thiết yếu của di truyền học

quần thể là tần số allele chứ không phải tần số kiểu gene? Hãy thiết lập các công thức tính tần số allele khác nhau và cho ví dụ.

3. (a) Nêu nội dung và các điều kiện nghiệm đúng của nguyên lý Hardy-Weinberg. Từ đó hãy chứng minh và nêu các hệ quả và ứng dụng của nó. (b) Thế nào là trạng thái cân bằng di truyền của quần thể? Các phương pháp khảo sát trạng thái cân bằng di truyền của quần thể?

4. Nguyên lý Hardy-Weinberg được mở rộng cho các trường hợp đa allele, tần số allele sai khác giữa hai giới tính và các gene liên kết với giới tính như thế nào?

5. Hãy tính tần số các allele và khảo sát trạng thái cân bằng theo các cách khác nhau đối với hệ nhóm máu M-N ở các quần thể người sau đây (số liệu từ F.J. Ayala và J.A. Kiger 1980):

Quần thể	M	MN	N	Tổng
Thổ dân Úc	22	216	492	730
Bộ lạc da đỏ Navaho	305	52	4	361
Người Mỹ gốc Capca	1787	3039	1303	6129

6. Hãy tính tần số allele trong các trường hợp sau:

a) Các thể đồng hợp lặn nhiều gấp hai lần các thể dị hợp.

b) Các thể đồng hợp lặn nhiều gấp sáu lần các thể dị hợp.

7. Màu sắc vỏ ốc sên châu Âu được kiểm soát bởi ba allele ở một locus đơn:  $C^B$  (nâu),  $C^P$  (hồng) và  $C^Y$  (vàng). Allele màu nâu là trội so với hồng và vàng; màu hồng trội so với vàng; màu vàng là lặn hoàn toàn. Trong một quần thể ốc sên các màu sắc được phân bố như sau: 472 nâu, 462 hồng, 66 vàng. Nếu như quần thể này ở dạng cân bằng, tần số các allele sẽ ra sao?

8. Chứng mù màu do một allele lặn  $m$  liên kết trên X gây ra. Cứ mười người có một người nam mắc chứng mù màu. Hỏi: (a) Tỷ lệ đó ở nữ giới là bao nhiêu? (b) Nếu một nửa số trẻ sinh ra ở mỗi giới tính bị mù màu, thì tỷ lệ các cuộc hôn nhân gây ra hậu quả đó là bao nhiêu? (c) Nếu như tất cả trẻ sinh ra đều bình thường, thì tỷ lệ các cuộc hôn nhân đó là như thế nào?

9. Chứng minh rằng khi một quần thể tiến tới trạng thái cân bằng đối với một gene liên kết giới tính, thì: (a) Sự dao động về tần số allele giữa hai giới thu hẹp dần sau mỗi thế hệ theo quy luật sau:  $(p'_c - p'_d) = -\frac{1}{2}(p_c - p_d)$ , trong đó  $p_c$  và  $p_d$  là các tần số gene tương ứng của hai giới cái và đực ở đời bố mẹ, còn  $p'_c$  và  $p'_d$  là các tần số tương ứng ở đời con của nó. (b) Tần số gene của cả quần thể ( $\bar{p}$ ) là không đổi qua các thế hệ và bằng  $\frac{2}{3}p_c + \frac{1}{3}p_d$ .

10. (a) Ở một quần thể sinh vật, tần số của các thể đồng hợp trội, dị hợp và đồng hợp lặn tương ứng là 0,67; 0,06 và 0,27. Hệ số nội phối là bao nhiêu? (b) Nếu như trong một quần thể với hai allele ở một locus nhiễm sắc thể thường có  $p = 0,8$ ,  $q = 0,2$ , và tần số của các thể dị hợp là 0,20, thì hệ số nội phối là bao nhiêu?

## Tài liệu Tham khảo

### Tiếng Việt

Trần Bá Hoàn. 1988. *Học thuyết tiến hóa*. NXB Giáo Dục, Hà Nội.

Kimura M. 1983. *Thuyết tiến hóa phân tử trung tính* (Bản dịch của Hoàng Trọng Phán). NXB Thuận Hóa, Huế -1993.

Mayer E. 1974. *Quần thể, loài và tiến hóa*. (Bản dịch của Lương Ngọc Toàn, Hoàng Đức Nhuận, Nguyễn Đức Khảm và Nguyễn Văn Thảo). NXB Khoa học và Kỹ thuật, Hà Nội, 1981.

Hoàng Trọng Phán. 2001. Hệ quả của một số định luật di truyền và ứng dụng của chúng trong giảng dạy di truyền học. Trong: *Kỷ yếu Hội nghị Khoa học Khoa Sinh-KTNN*, Trường ĐHSPT Hà Nội, tr.5-8.

Hoàng Trọng Phán. 2003. *Di truyền học quần thể* (Bài giảng lưu hành nội bộ). Trường ĐHSPT Huế.

### Tiếng Anh

Ayala FJ, Kiger JA. 1980. *Modern Genetics*. Benjamin/Cummings Publishing Company, Inc, Menlo Park, CA.

Crow JF. 1986. *Basic Concepts in Population, Quantitative, and Evolutionary Genetics*. WH Freeman and Co., New York.

Crow JF. 1993. Mutation, mean fitness, and genetic load. In: *Oxford in Evolutionary Biology, Volume 9* (Futuyma D., Antonovics J. eds), pp 3-42, Oxford University Press, Inc., Oxford -New York...

Falconer DS, Mackay TFC. 1996. *Introduction to Quantitative Genetics*. 4<sup>th</sup> ed., Longman Group, Harlow.

Hartl DL, Freifelder D, Snyder LA. 1988. *Basic Genetics*. Jones and Bartlett Publishers, Inc., Boston - Portola Valley.

Hartl DL, Clark AG. 1997. *Principles of Population Genetics*. 3rd ed., Sinauer Associates Inc., Sunderland, Massachusetts. ISBN: 0-87893-306-9.

Holman DJ. 2003. *Human Population Genetics* BIO A 482 Autumn 2003). <<http://faculty.washington.edu/djholman/bioa482/>>

Ridley M. 1993. *Evolution*. Blackwell Scientific Publishing, Boston.

Standfield WD. 2002. *Genetics*, 4th ed. Schaum's Outline Series. McGraw-Hill Inc., Toronto.

Suzuki DT, Griffiths AJF, Miller JH, Lewontin RC. 1989. *An Introduction to Genetic Analysis*. 4<sup>th</sup> ed., W-H Freeman and Company, New York.

Yablokov V.A. 1986. *Population Biology: Progress and Problems on Natural Populations*. Mir Publishers, Moscow.

Weaver RF, Hedrick PW. 1997. *Genetics*. 3<sup>rd</sup> ed., McGraw-Hill Companies, Inc./ Wm.C.Brown Publishers, Dubuque, IA.

Wrights 1988. Surfaces of selective value revisited. *The American Naturalist* 131: 115-123.

### **Một số trang web**

<http://wbar.uta.edu/jands/jands.htm>

<http://science.kennesaw.edu/~rmatson/Biol%203380/3380species.html>

<http://www.users.nac.net/jmele/ECO.HW.html>

<http://nitro.biosci.arizona.edu/courses/EEB182/Lecture04/lect4.html>

# Mục lục

<b>Lời nói đầu</b>	11
<b>Mở đầu</b>	
	<b>Hoàng Trọng Phán</b>
I. Khái niệm di truyền học	13
II. Lược sử phát triển của di truyền học	13
III. Đối tượng và các lĩnh vực nghiên cứu của di truyền học	17
IV. Các phương pháp nghiên cứu của di truyền học	18
V. Các nguyên tắc nghiên cứu và phương pháp học tập di truyền học	20
VI. Di truyền học với công nghệ sinh học, tin học và các vấn đề xã hội	21
<b>Chương 1: Cơ sở của Di truyền học Mendel</b>	
	<b>Hoàng Trọng Phán</b>
I. Tiểu sử Mendel - Cha đẻ của di truyền học	24
II. Đối tượng và phương pháp thí nghiệm của Mendel	25
1. Đối tượng	25
2. Phương pháp	26
III. Lai một tính và nguyên lý phân ly	26
1. Kết quả thí nghiệm lai một tính	26
2. Giải thích và kiểm chứng nguyên lý phân ly	27
3. Nguyên lý phân ly và tính phổ biến của nó	28
IV. Lai hai tính và nguyên lý phân ly độc lập	28
1. Kết quả thí nghiệm lai hai tính	28
2. Giải thích và nội dung nguyên lý phân ly độc lập	29
V. Sự di truyền Mendel ở người	30
1. Các tính trạng lặn	31
2. Các tính trạng trội	32
VI. Lý thuyết xác suất trong dự đoán và phân tích di truyền học	33
1. Một số khái niệm và tính chất cơ bản của xác suất	33
2. Một số nguyên lý xác suất cơ bản	34
VII. Phương pháp $\chi^2$ ( <i>Chi-square method</i> ) trong đánh giá độ phù hợp giữa các số liệu quan sát và kỳ vọng	39



## **Chương 2: Mở rộng và Áp dụng của Di truyền học Mendel**

	<b>Hoàng Trọng Phán</b>	43
I. Các kiểu quan hệ giữa các gene allele đối với một tính trạng		43
1. Các kiểu trội hoàn toàn, không hoàn toàn và đồng trội		43
2. Tác động của gene gây chết ( <i>lethals</i> )		45
3. Hiện tượng đa allele ( <i>multiple allelism</i> )		46
II. Tính đa hiệu của gene ( <i>pleiotropy</i> )		47
III. Các kiểu tương tác giữa các gene không allele		48
1. Tương tác bổ trợ ( <i>complementary</i> )		49
2. Tương tác át chế ( <i>epistasis</i> )		52
3. Tương tác cộng gộp- sự di truyền đa gene và các tính trạng số lượng		54
IV. Các mối quan hệ kiểu gene - kiểu hình		59
1. Thường biến và mức phản ứng		59
2. Độ thâm nhập ( <i>penetrance</i> ) và độ biểu hiện ( <i>expressivity</i> )		60

## **Chương 3: Cơ sở Tế bào của sự Sinh sản, Di truyền và Biến dị**

	<b>Hoàng Trọng Phán</b>	67
I. Sinh sản hữu tính và tính ổn định của các bộ nhiễm sắc thể		67
II. Hình thái học nhiễm sắc thể eukaryote		70
1. Kích thước nhiễm sắc thể		70
2. Tâm động và các kiểu nhiễm sắc thể		71
3. Các kiểu băng nhiễm sắc thể ( <i>chromosomal bands</i> )		72
III. Chu kỳ tế bào và nguyên phân		74
1. Chu kỳ tế bào ( <i>cell cycle</i> )		75
2. Nguyên phân ( <i>mitosis</i> )		76
IV. Giảm phân, sự phát sinh giao tử và thụ tinh		78
1. Giảm phân ( <i>meiosis</i> )		78
2. Sự phát sinh giao tử ( <i>gametogenesis</i> )		81
3. Sự thụ tinh ( <i>fertilization</i> )		83
V. Các biến đổi của nhiễm sắc thể		84
1. Các biến đổi về cấu trúc nhiễm sắc thể		84
2. Các biến đổi về số lượng nhiễm sắc thể		91

## **Chương 4: Di truyền học Nhiễm sắc thể**

	<b>Hoàng Trọng Phán</b>	103
I. Trường phái Morgan với thuyết di truyền nhiễm sắc thể		103
1. Tầm quan trọng của ruồi giấm <i>Drosophila</i>		103
2. Thuyết di truyền nhiễm sắc thể		107
II. Sự xác định giới tính ( <i>sex determination</i> )		108
1. Sự xác định giới tính do kiểu gene ( <i>GSD</i> )		108
2. Sự xác định giới tính do môi trường ( <i>ESD</i> )		112
III. Sự di truyền liên kết với giới tính ( <i>sex-linked inheritance</i> )		113
1. Đặc điểm di truyền của các gene trên nhiễm sắc thể X và Y		113
2. Sự bất hoạt của nhiễm sắc thể X và một số vấn đề liên quan		117
3. Các tính trạng giới hạn bởi giới tính và chịu ảnh hưởng của giới tính		120
IV. Liên kết và tái tổ hợp của các gene trên một nhiễm sắc thể		121
1. Khám phá về sự trao đổi chéo ở ruồi giấm		122
2. Liên kết gene hoàn toàn (giảm phân không có trao đổi chéo)		124
V. Trao đổi chéo và lập bản đồ di truyền		125
1. Tần số tái tổ hợp		125
2. Bản đồ di truyền		127
3. Trao đổi chéo bốn sợi		131
VI. Lập bản đồ gene từ các phép lai phân tích ba điểm		134
1. Trao đổi chéo kép với việc xác định trật tự và khoảng cách các gene		134
2. Độ nhiễu và hệ số trùng hợp ( <i>coefficient of coincidence</i> )		137
VII. Lập bản đồ bằng phân tích bộ bốn ( <i>tetrad analysis</i> )		139

## **Chương 5: Bản chất Hoá học và Tái bản của Vật chất Di truyền**

	<b>Hoàng Trọng Phán</b>	147
I. Bằng chứng vật chất di truyền là các nucleic acid		147
1. Các thí nghiệm biến nạp ở vi khuẩn		147
2. Tính ổn định của hàm lượng DNA ở các sinh vật bậc cao		148
3. Các thí nghiệm ở virus		149
II. Thành phần hoá học và cấu trúc của DNA		150
1. Cấu trúc của một nucleotide		150

2. Cấu trúc chuỗi polynucleotide	151
3. Thành phần hoá học và cấu trúc của chuỗi xoắn kép DNA	152
4. Sơ lược về các đặc tính hoá lý của các nucleic acid	155
III. Kích thước bộ gene và tính phức tạp về mặt tiến hoá	157
1. Sơ lược về bộ gene của các virus, prokaryote và eukaryote	157
2. Mối quan hệ giữa kích thước bộ gene và tính phức tạp về tiến hoá	159
3. Hàm lượng DNA và nghịch lý giá trị C	159
4. Về kích thước DNA các bào quan	161
IV. Tổ chức phân tử của các nhiễm sắc thể	161
1. Cấu trúc chất nhiễm sắc	161
2. Sơ lược các thành phần DNA trong bộ gene eukaryote	163
V. Tái bản DNA ( <i>DNA replication</i> )	164
1. Các nguyên tắc và đặc điểm chung của tái bản DNA	164
2. Các enzyme tham gia tái bản DNA	167
3. Cơ chế tái bản DNA	168
VI. Tái bản của các bộ gene RNA ( <i>RNA genomes</i> )	174
1. Đặc điểm tái bản của các bộ gene RNA virus	174
2. Tái bản của bộ gene RNA	174
3. Phiên mã ngược	175
<b>Chương 6: Gene và Quá trình Sinh tổng hợp Protein</b>	
	<b>Hoàng Trọng Phán</b>
I. Sự phát triển của khái niệm gene	179
1. Các quan niệm của Mendel và Morgan về gene	179
2. Giả thuyết một gene - một enzyme của Beadle và Tatum	180
3. Quan niệm của Benzer về các đơn vị cấu trúc và chức năng di truyền	180
4. Mối quan hệ gene - cistron ở các prokaryote và eukaryote	182
II. Cấu trúc và chức năng của protein	186
1. Cấu trúc protein	186
2. Chức năng protein	187
III. Mã di truyền	189
1. Bằng chứng di truyền học về mã bộ ba	190
2. Giải mã di truyền	190

3. Các đặc tính của mã di truyền	191
4. Những ngoại lệ so với mã di truyền "phổ biến"	192
5. Sự linh hoạt trong việc kết cặp anticodon-codon	193
IV. Cơ chế phiên mã ( <i>transcription</i> ) và sửa đổi sau phiên mã	194
1. Các RNA và đặc điểm chung của phiên mã	194
2. Các RNA polymerase của prokaryote và eukaryote	196
3. Các promoter ở các prokaryote và eukaryote	196
4. Các giai đoạn của quá trình phiên mã	197
5. Sự sửa đổi sau phiên mã đối với các mRNA eukaryote	198
V. Cấu trúc và chức năng của các loại RNA và ribosome	200
1. RNA thông tin ( <i>messenger RNA = mRNA</i> )	200
2. RNA vận chuyển ( <i>transfer RNA = tRNA</i> )	200
3. RNA ribosome ( <i>ribosomal RNA = rRNA</i> )	201
4. Ribosome	201
VI. Cơ chế dịch mã ( <i>translation</i> )	202
1. Hoạt hoá amino acid	202
2. Cơ chế của quá trình dịch mã (tổng hợp polypeptide)	202
<b>Chương 7: Sự Điều hoà Biểu hiện của Gene</b>	
<b>Trương Thị Bích Phượng</b>	208
I. Các nguyên lý điều hoà và mức độ kiểm soát phiên mã	202
II. Điều hoà biểu hiện gene ở prokaryote	209
1. Cấu trúc của operon	210
2. Điều hoà dương tính operon lactose	212
3. Điều hoà âm tính operon tryptophan	213
4. Phiên mã dờ ( <i>attenuation</i> )	215
III. Điều hoà biểu hiện gene ở eukaryote	217
1. Sự biến đổi DNA	218
2. Các promoter	218
3. Những trình tự tăng cường phiên mã ( <i>enhancer</i> )	219
4. Trình tự bất hoạt gene ( <i>gene silencing</i> )	220
5. Promoter chọn lọc ( <i>alternative promoter</i> )	220
6. Splicing chọn lọc	220

## **Chương 8: Đột biến Gene, Tái tổ hợp và các Yếu tố Di truyền Vận động**

**Trương Thị Bích Phượng** 223

I. Đột biến gene	223
1. Các kiểu đột biến gene	223
2. Các tác nhân gây đột biến gene	228
3. Cơ chế phân tử của các đột biến gene	228
II. Sửa chữa và bảo vệ DNA	232
III. Các yếu tố di truyền vận động ( <i>transposable genetic elements</i> )	237
1. Các yếu tố di truyền vận động ở prokaryote	237
2. Các yếu tố di truyền vận động ở eukaryote	239

## **Chương 9: Sự Di truyền Tế bào chất**

**Trương Thị Bích Phượng** 245

I. Sự di truyền tế bào chất	245
1. Sự di truyền của các gene lập thể	245
2. Sự di truyền của các gene ty thể	246
3. Hiệu quả dòng mẹ lên chiều xoắn vô ốc	249
II. Lập bản đồ ở ty thể và lập thể	251
1. Lập bản đồ gene của DNA lập thể	251
2. Lập bản đồ gene của DNA ty thể	253
III. Di truyền học phân tử các bào quan	254
1. Các bộ gene lập thể (cpDNA)	254
2. Các bộ gene ty thể (mtDNA)	255

## **Chương 10: Đại cương về Công nghệ DNA Tái tổ hợp**

**Hoàng Trọng Phán** 258

I. Các công cụ chính của kỹ thuật tạo dòng DNA tái tổ hợp	258
1. Các enzyme giới hạn	258
2. Các vector thông dụng trong kỹ thuật di truyền	260
3. Thiết lập phân tử DNA tái tổ hợp <i>in vitro</i>	261
II. Tạo dòng gene hay DNA tái tổ hợp	263
1. Nguyên tắc chung	263
2. Quy trình tạo dòng gene tái tổ hợp	264
3. Tổng hợp và tạo dòng cDNA	267

III. Các phương pháp biểu hiện các gene được tạo dòng	267
IV. Ứng dụng của công nghệ DNA tái tổ hợp	269
1. Công nghệ DNA tái tổ hợp với việc nghiên cứu bộ gene	269
2. Công nghệ DNA tái tổ hợp với y-dược học	271
3. Kỹ thuật di truyền với các sinh vật biến đổi gene	274

### **Chương 11: Di truyền học Người**

	<b>Trần Quốc Dung</b>	278
I. Các phương pháp nghiên cứu di truyền học người		278
1. Phương pháp phân tích phả hệ ( <i>genealogy analysis</i> )		278
2. Phương pháp nghiên cứu trẻ sinh đôi		279
3. Phương pháp di truyền tế bào học người		279
4. Phương pháp nghiên cứu quần thể		279
5. Các kỹ thuật sinh học phân tử		280
II. Các phương pháp lập bản đồ di truyền người		280
1. Phân tích liên kết ( <i>linkage analysis</i> )		280
2. Các phương pháp lập bản đồ vật lý ( <i>physical mapping</i> )		280
III. Nhiễm sắc thể Y và chất nhiễm sắc giới tính của người		283
1. Nhiễm sắc thể Y của người		283
2. Chất nhiễm sắc thể giới tính của người		284
IV. Sự di truyền các gene trội-lặn trên nhiễm sắc thể thường và nhiễm sắc thể giới tính		285
1. Sự di truyền các gene trội-lặn trên nhiễm sắc thể thường		285
2. Sự di truyền các gene trội-lặn trên nhiễm sắc thể giới tính		286
V. Di truyền y học		288
1. Các bệnh di truyền do rối loạn chuyển hoá và các bệnh nhiễm sắc thể		288
2. Cơ sở di truyền ung thư		291
VI. Tư vấn di truyền y học		292

### **Chương 12: Di truyền học Quần thể**

	<b>Hoàng Trọng Phán</b>	296
I. Các khái niệm cơ bản của di truyền học quần thể		296
1. Quần thể ( <i>population</i> )		296
2. Các hệ thống giao phối		296

3. Vốn gene ( <i>gene pool</i> )	297
4. Tần số kiểu gene và tần số allele	297
II. Nguyên lý Hardy-Weinberg và trạng thái cân bằng của quần thể	299
1. Nguyên lý Hardy-Weinberg	299
2. Những ứng dụng của nguyên lý Hardy-Weinberg	302
III. Mở rộng nguyên lý Hardy-Weinberg	305
1. Đa allele ( <i>multiple alleles</i> )	305
2. Tần số allele sai biệt giữa hai giới tính	308
3. Các gene liên kết trên X	309
IV. Nội phối ( <i>inbreeding</i> )	310
1. Tự thụ tinh ( <i>self-fertilization</i> )	311
2. Hệ số nội phối ( <i>inbreeding coefficient</i> )	312
3. Tính toán hệ số nội phối	313
V. Các nhân tố tác động lên thành phần di truyền quần thể	315
1. Đột biến	315
2. Biến động di truyền ngẫu nhiên	316
3. Dòng gene hay sự di nhập cư	316
4. Chọn lọc tự nhiên	318