

LÊ ĐÌNH LƯƠNG

**GIÁO TRÌNH
DI TRUYỀN HỌC**

Hà Nội - 2009

MỞ ĐẦU

Chúng ta đang bị cuốn hút và phần nào đang thật sự tham gia vào một cuộc cách mạng khoa học công nghệ mà thật ra đã được dự đoán trước từ những năm bảy mươi của thế kỷ XX vừa qua. Nhưng loài người vẫn bị bất ngờ vì những thành tựu nối tiếp thành tựu ngày càng dồn dập làm tăng quá nhanh tốc độ của cuộc cách mạng đó - cách mạng công nghệ sinh học hiện đại mà hạt nhân là di truyền học - lĩnh vực khoa học đã đạt tới một tầm cao lý thuyết đủ để trở thành lực lượng sản xuất trực tiếp, phục vụ hết sức hiệu quả cho các nhu cầu hàng ngày của thực tiễn xã hội, không phải chỉ của các nước phát triển mà cho cả các nước nghèo như nước ta. Điều đặc biệt và rất đặc thù ở đây là cuộc cách mạng này đã và đang tạo ra những cơ hội lớn và rất hiếm có để cho các nước nghèo có thể rút ngắn khoảng cách với các nước giàu trong việc phát triển khoa học này, nếu chúng ta biết chớp lấy và tận dụng tốt những cơ hội.

Các khái niệm về “gen”, “ADN”, mới ngày nào còn mang nặng tính hàn lâm, lý thuyết xa vời, thì giờ đây các sản phẩm từ chúng đã bày bán trên các sạp ở chợ trên khắp các châu lục. Việc thao tác ADN, các kỹ thuật di truyền phân tử - một hệ thống các phương pháp kết hợp những thành tựu khoa học lớn được giải thưởng Nobel, giờ đây đang được sử dụng phổ biến trên khắp thế giới, trong đó có nhiều phòng thí nghiệm ở nước ta, để phục vụ những nhu cầu thực tiễn đa dạng.

Ngày 12-2-2001 toàn bộ trình tự hệ gen người đã được xác định và công bố, mở ra một thời kỳ phát triển mới của di truyền học trên đối tượng con người. Đây lại là một cơ hội nữa cho các nước còn nghèo vì được thừa hưởng một kho thông tin khổng lồ và vô giá mà không phải trả tiền.

Giáo trình này được biên soạn nhằm cung cấp những kiến thức cơ bản nhất về di truyền học, tương đương với chương trình phổ thông trung học của môn học.

Trung tâm của di truyền học là gen

Từ cuối thập kỷ 60 của thế kỷ XX xuất hiện một công nghệ cao bắt nguồn từ di truyền học hiện đại gọi là kỹ thuật di truyền. Nó đã phát triển với tốc độ thần kỳ và đạt tới điểm mà hiện nay trong nhiều phòng thí nghiệm trên toàn thế giới kỹ thuật này đã trở thành công việc hàng ngày: từ tách chiết đoạn ADN đặc thù của một hệ gen của một sinh vật đến việc xác định trình tự bazơ và đánh giá chức năng của nó. **Điều đặc biệt hấp dẫn là các nhà khoa học riêng biệt có thể áp dụng kỹ thuật mà**

không cần những thiết bị đắt tiền hoặc những nguồn tài chính lớn, nằm ngoài tầm với của cá nhân các nhà khoa học riêng lẻ.

Nguyên lý cơ bản của kỹ thuật này khá đơn giản. Cơ sở của công nghệ là thông tin di truyền mã hoá trong ADN, tồn tại ở dạng các gen. Thông tin này có thể sửa đổi theo nhiều cách khác nhau để đạt tới những mục tiêu nhất định trong nghiên cứu lý thuyết, ứng dụng và trong y học. Có 4 lĩnh vực chủ yếu sử dụng thao tác di truyền là:

- Nghiên cứu lý thuyết về cấu trúc và chức năng của gen;
- Sản xuất các protein hữu ích bằng các phương pháp mới, dựa trên chức năng của gen;
- Tạo ra các thực vật, động vật và vi sinh vật chuyển gen;
- Xét nghiệm ADN để xác định đặc trưng cá thể và chẩn đoán bệnh di truyền và bệnh nhiễm trùng...

Như vậy, trung tâm của di truyền học là GEN. Trên cơ sở hiểu biết rõ cấu trúc, chức năng và qui luật vận động của chúng con người đang ngày càng làm chủ công nghệ điều khiển sự sống của muôn loài theo hướng mong muốn.

Vậy gen là gì? Hành vi của chúng ra sao? Con người đang và sẽ làm được gì để điều khiển hành vi đó phục vụ cho lợi ích của nhân loại?

Gen được Mendel phát hiện từ thế kỷ 19

Ngay từ khi lịch sử loài người bắt đầu, người ta đã tự hỏi các đặc điểm, tính trạng đã được truyền từ thế hệ này sang thế hệ khác như thế nào. Mặc dù các con thường trông giống bố hoặc mẹ hơn, nhưng nói chung các đặc điểm thường pha trộn giữa cả bố và mẹ. Nhiều thế kỷ chọn giống động vật và thực vật đã cho thấy các tính trạng có ích như tốc độ chạy của ngựa, sức mạnh của bò, kích thước của hoa quả chịu ảnh hưởng rõ ràng của những phép lai nhất định. Tuy nhiên, lúc đó không có một phương pháp khoa học nào có thể dự đoán trước kết quả của một phép lai giữa hai cặp cha mẹ cụ thể.

Cho đến tận năm 1865 một thầy tu người Áo tên là Gregor Mendel đã phát hiện ra rằng những tính trạng riêng biệt được xác định bởi các nhân tố riêng rẽ mà sau này được gọi là gen. Chính chúng là cái được di truyền từ bố mẹ sang các con. Cách nghiên cứu nghiêm túc và chặt chẽ của ông đã biến chọn giống nông nghiệp từ một nghệ thuật trở thành khoa học. Tuy nhiên, công trình của ông lúc đó chưa được thừa nhận ngay. Vì vậy, khoa học di truyền chỉ thật sự được tái sinh với việc phát minh lại công trình của Gregor Mendel vào đầu thế kỷ 20 và 40 năm sau các nguyên lý di truyền đó đã được chứng minh và làm sáng tỏ. Di truyền học vi sinh vật đã phát triển mạnh mẽ vào những năm 40 thế kỷ trước và vai trò của ADN (thành phần hóa học của gen) đã được khẳng định. Cũng vì vậy, sự hiểu biết của con người về các cơ

ché chuyển gen giữa các vi khuẩn cũng như những kiến thức cơ bản của khoa học sự sống trở nên sáng tỏ hơn rất nhiều.

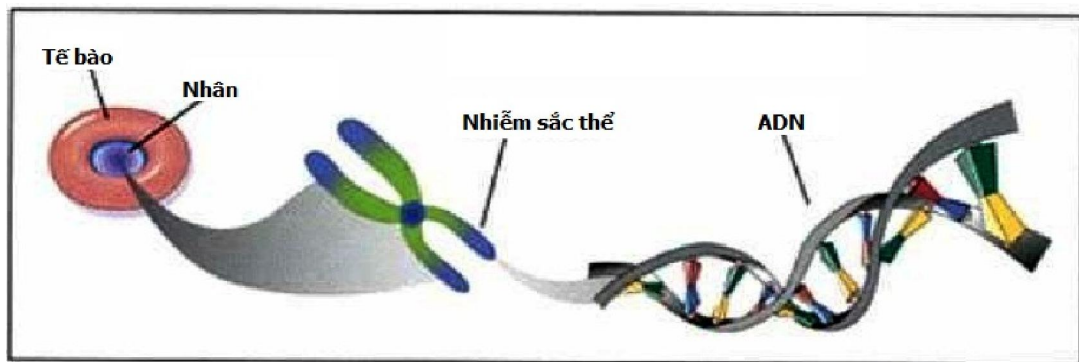
Phát minh của James Watson và Francis Crick năm 1953 về cấu trúc của ADN đã tạo ra bước đột phá cho phát triển di truyền học ở mức phân tử. Việc giải xong mã di truyền đầy đủ vào năm 1966 được coi là ngày ra đời của di truyền học hiện đại.

Từ năm 1865 đến ngày nay lịch sử phát triển của di truyền học phản ánh sự hiểu biết ngày càng tăng của con người về gen. Trước khi định nghĩa gen một cách chính xác, chúng ta hãy tạm hiểu gen là một đoạn ADN có đủ thông tin di truyền để qui định một chức năng như màu mắt của người, hình dạng của hạt đậu hay một căn bệnh nào đó.

Gen chủ yếu nằm trên các nhiễm sắc thể

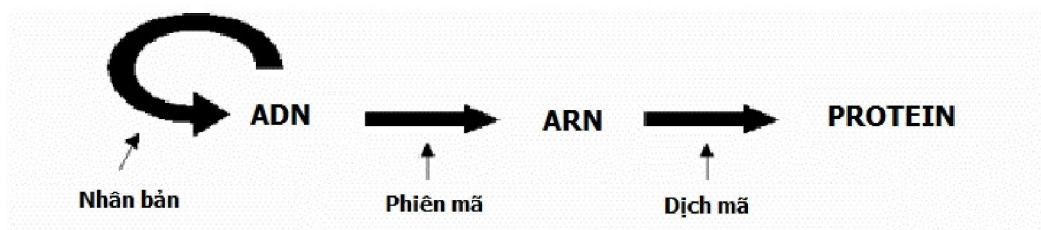
Tất cả các sinh vật đều cấu thành từ các tế bào. Các phản ứng hóa học diễn ra bên trong các tế bào gọi là quá trình trao đổi chất. Thông tin di truyền cần thiết để duy trì các tế bào tồn tại và sản sinh ra các tế bào mới được tàng trữ trong nhân tế bào các sinh vật. Thông tin di truyền được truyền từ một thế hệ sang thế hệ tiếp theo.

Nhân tế bào có chứa thông tin di truyền (ADN) là trung tâm điều khiển của tế bào. ADN được bọc gói thành các nhiễm sắc thể. Việc tổng hợp (sao chép) ADN và tổng hợp (phiên mã) ARN tương ứng diễn ra bên trong nhân. Phiên mã là bước đầu tiên trong quá trình biểu hiện của thông tin di truyền và là hoạt động trao đổi chất chủ yếu của nhân tế bào.



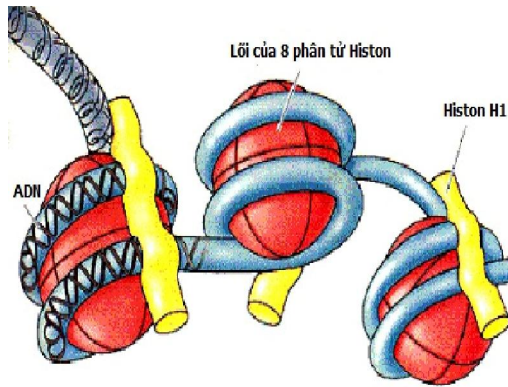
Hình 1.

Mỗi gen, tức mỗi đơn vị thông tin di truyền là một đoạn ADN mà thông tin của nó được thể hiện trong một ngôn ngữ chỉ bao gồm 4 chữ cái, mỗi chữ là một bazơ nitơ. Thông tin lưu trữ trong các sợi ADN sau đó thể hiện trong trình tự của một loại polymer sinh học khác, đó là protein.



Hình 2.

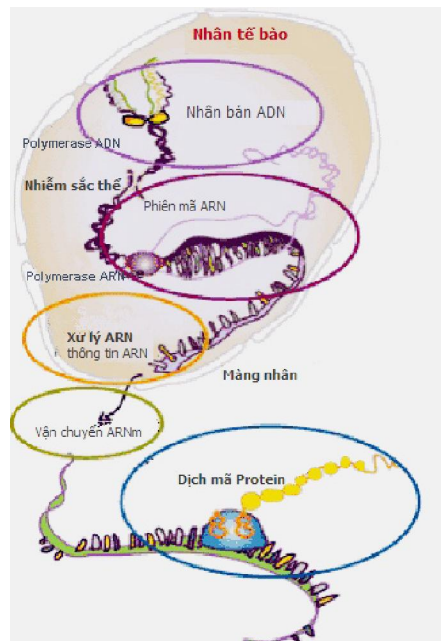
Công trình tế bào học cuối những năm 1800 đã chứng minh rằng mỗi cơ thể sống có một bộ nhiễm sắc thể đặc thù nằm trong mỗi tế bào. Cùng thời gian đó các nghiên cứu hóa sinh học đã cho thấy rằng vật liệu cấu tạo nên các nhiễm sắc thể bao gồm ADN và protein.



Hình 3.

Công trình của Avery và Hershey, trong những năm 1940 và 1950, đã chứng minh rằng ADN là phân tử di truyền.

Các công trình hoàn thành vào những năm 1960 và 1970 đã cho thấy, mỗi nhiễm sắc thể là một cấu trúc chỉ gồm một sợi ADN liên tục và rất dài. Ở sinh vật bậc cao các protein cấu trúc, mà một số trong đó là histon, đóng vai trò giá đỡ giúp cho ADN có cấu trúc nhiễm sắc thể chặt chẽ.

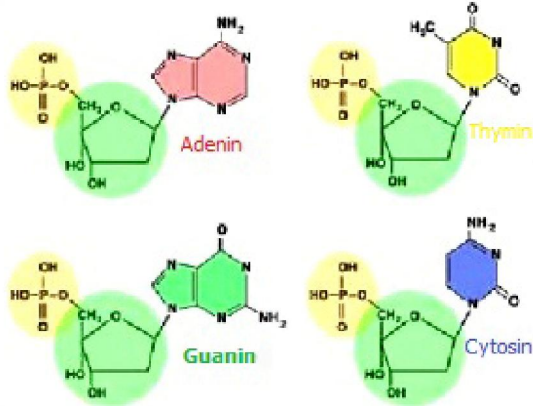


Hình 4.

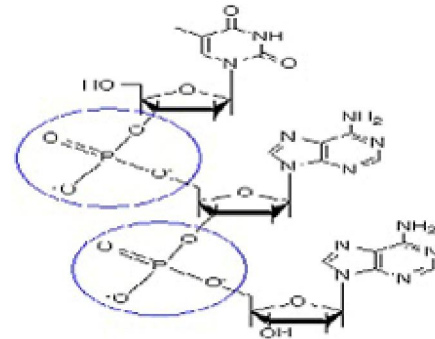
Gen có thành phần là ADN hoặc ARN

Cấu trúc của ADN (Axit DeoxyriboNucleic) gồm 4 đơn vị vật liệu gọi là các nucleotid. Mỗi nucleotid gồm đường deoxyribose, nhóm phosphat, và một trong 4 bazơ nitơ: adenin (A), thymin (T), guanin (G), và cytosin (C). Các nhóm phosphat và đường của các nucleotid kế cận nối với nhau tạo nên một polymer mạch dài. Trong khi đó các bazơ nitơ A kết với T và G kết với C tạo cho phân tử ADN hình ảnh một chuỗi xoắn kép giống như cái thang xoắn trong nhà ở. Điều này được khẳng định trong công trình năm 1953 của James Watson và Francis Crick tại phòng thí nghiệm Cavendish ở Cambridge, nước Anh.

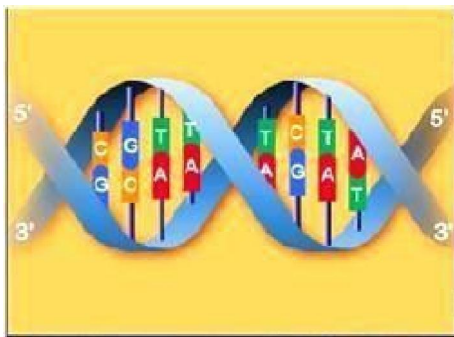
Như vậy các cặp bazơ có tác dụng gắn chuỗi xoắn kép với nhau. Nơi "bắt đầu" của sợi phân tử ADN được ký hiệu là 5'. Còn đuôi của phân tử ADN được ký hiệu là đầu 3'. Thuật ngữ 5' và 3' chỉ vị trí của bazơ tương ứng với phân tử đường trên bộ khung ADN nơi các mối liên kết phosphodiester kết nối giữa nguyên tử carbon 3' và carbon 5' của đường deoxyribose (trong ADN) hoặc đường ribose (trong ARN).



Hình 5.



Hình 6.

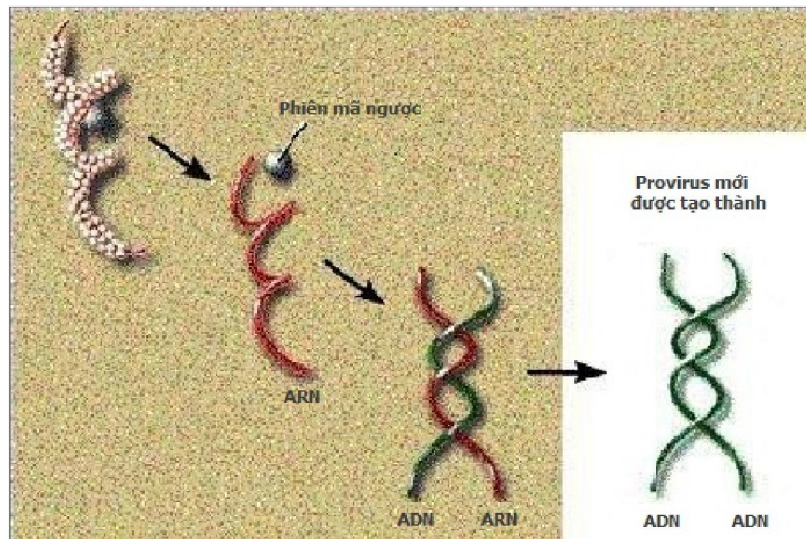


Hình 7. Hai sợi ADN trong chuỗi xoắn kép có định hướng ngược nhau

Mỗi nhiễm sắc thể chỉ bao gồm một phân tử ADN đơn lẻ. ADN của chúng ta chứa tới hơn 3 tỷ cặp bazơ - một con số khổng lồ. Tất cả lượng thông tin này phải được tổ chức theo cách để có thể bọc gói bên trong nhân tế bào bé nhỏ. Để thực hiện được việc đó, ADN phải nằm trong tổ hợp với histon để tạo thành chromatin. Histon là những protein đặc biệt để phân tử ADN có thể cuộn quanh và trở nên cô đặc hơn. Sau đó chromatin tự cuộn và tạo thành dạng nhiễm sắc thể.

Khi một tế bào phân chia thành hai tế bào con thì ADN, tất cả 46 nhiễm sắc thể, chẳng hạn, ở người, phải được sao chép (tổng hợp). Tính đặc hiệu trong việc kết cặp giữa A/T và C/G là rất quan trọng để tổng hợp sợi ADN mới giống hệt sợi ADN cha mẹ. Vì mỗi sợi này được dùng làm khuôn để tổng hợp ADN theo nguyên tắc bổ sung giữa hai sợi – sợi cũ làm khuôn và sợi mới được tổng hợp.

Nếu có trục trặc xảy ra trong quá trình sao chép ADN thì nó sẽ làm hỏng chức năng của gen. Chẳng hạn, nếu một bazơ sai được xen vào trong quá trình sao chép (một đột biến) và sai sót này xảy ra ở đoạn giữa của một gen quan trọng, thì nó có thể dẫn đến một protein bị bất hoạt. Rất may, chúng ta đã có được trong quá trình tiến hóa những cơ chế khác nhau để đảm bảo rằng những đột biến như thế được phát hiện, được sửa chữa và không bị nhân lên. Tuy nhiên, đôi khi những cơ chế này bị hỏng và đột biến không được sửa chữa vẫn cứ xảy ra. Trong trường hợp đó quá trình chuyển hóa hoặc một cấu trúc sẽ bị hủy hoại dẫn đến bệnh tật.

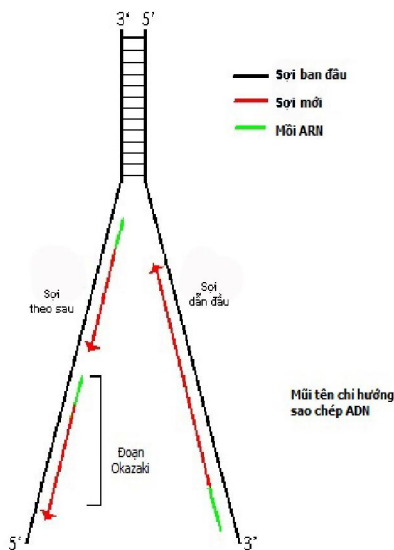


Hình 8.

Một số virus lưu trữ thông tin di truyền trong ARN, mặc dù đã có lúc người ta tin rằng ADN là phân tử di truyền duy nhất.

Mặc dù vậy, các virus này vẫn tạo ra protein đúng theo cách như ở các sinh vật bậc cao. Trong quá trình gây nhiễm (virus xâm nhập vào tế bào chủ), đầu tiên ARN virus phiên mã ngược trở lại ADN, sau đó chuyển sang ARN rồi sang protein phù hợp với sơ đồ đã được chấp nhận. Quá trình chuyển đổi từ ARN sang ADN gọi là phiên mã ngược, và các virus sử dụng cơ chế này gọi là retro-virus. Một polymerase chuyên dụng, gọi là transcriptase ngược, sử dụng ARN như sợi khuôn để tổng hợp phân tử ADN mạch kép bổ sung như nêu ở hình trên.

Gen có thể tự sinh ra gen



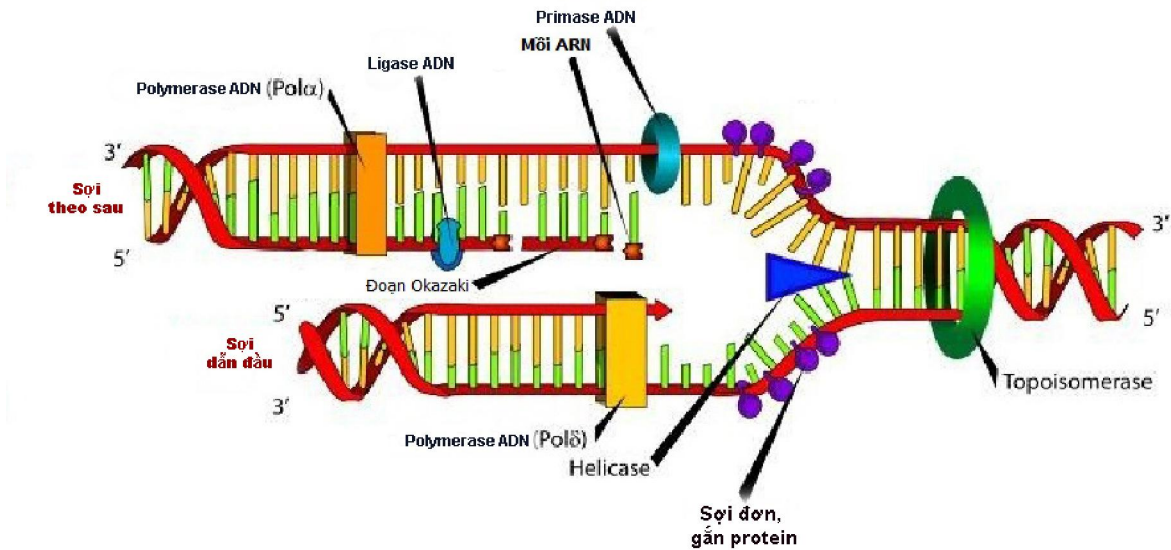
Hình 9.

Vì các gen có thành phần cấu tạo là ADN nên chúng có thể tự tạo ra chúng khi ADN sao chép. Tính đặc hiệu trong quá trình kết cặp bazơ giữa A/T và C/G giúp giải thích vì sao ADN được sao chép trước khi tế bào phân chia. Các enzym dẫn xoắn ADN bằng cách bẻ gãy các mối liên kết hydro giữa các cặp bazơ. Các bazơ không kết cặp giờ đây có thể tự do bám vào các nucleotid khác với các bazơ bổ trợ. Enzym primase bắt đầu quá trình bằng cách tổng hợp các đoạn mồi ARN ngắn bổ trợ với ADN chưa kết cặp. ADN polymerase bây giờ đính các nucleotid ADN vào một đầu của sợi bổ trợ đang được kéo dài ra do tổng hợp.

Quá trình sao chép được thực hiện liên tục trên một sợi gọi là *sợi dẫn đầu*. Vì quá trình tổng hợp bao giờ cũng bắt đầu tại đầu 3' của sợi khuôn và tiến dần về phía

đầu 5' của sợi khuôn này, nên sợi ADN mới chỉ được sao chép liên tục trên một sợi khuôn. Trên hình nó ở phía bên phải. Còn trên sợi khuôn bên trái hình, quá trình tổng hợp diễn ra với những đoạn ngắn riêng biệt có tên Okazaki. Sợi này gọi là sợi *theo sau*. Sau đó enzym có tên là ligase sẽ nối các đoạn Okazaki với nhau, tạo nên sợi ADN mới hoàn chỉnh.

Bây giờ ta có 2 phân tử ADN, một phân tử là sợi gốc ban đầu và một sợi mới có trình tự các nucleotid bổ trợ (bổ sung) với nó. Hai sợi giống hệt nhau về trình tự sắp xếp các nucleotid.



Hình 10.

Ngôn ngữ của gen đơn giản nhưng chứa đầy thông tin

Thông tin di truyền giống như một ngôn ngữ. Chúng ta sử dụng các chữ cái trong bảng alphabet để tạo ra các từ, sau đó nối các từ lại với nhau để tạo ra các câu, các đoạn văn và các cuốn sách. Trong ngôn ngữ ADN:

Bảng alphabet chỉ có 4 chữ cái A, T, G và C.

Mỗi chữ cái là một hợp chất hóa học gọi là bazơ hoặc nucleotid. Bốn chữ cái này được dùng để tạo ra các từ di truyền gọi là codon.

Khác ngôn ngữ bình thường, tất cả các từ di truyền chỉ gồm 3 chữ cái.

Các từ này kết hợp với nhau tạo nên các câu gọi là gen có chức năng xác định trình tự các axit amin trong polypeptid.

Tại cuối mỗi câu có một từ đặc biệt dùng làm dấu chấm câu gọi là codon dừng. Tất cả các câu nối lại với nhau tạo thành cuốn sách chứa đựng toàn bộ thông tin di truyền về một sinh vật gọi là hệ gen.

Chúng ta hãy làm một phép so sánh giữa tiếng Việt và Ngôn ngữ di truyền:

So sánh các ngôn ngữ	
Tiếng Việt	Ngôn ngữ Di truyền
<ul style="list-style-type: none"> - Ta có 24 chữ cái để tạo nên các từ - Các từ có chiều dài khác nhau - Chúng ta ghép các từ để tạo nên câu - Mỗi câu kết thúc bằng một dấu chấm - Các câu hợp lại tạo ra Sách 	<ul style="list-style-type: none"> - ADN có 4 phân tử để tạo nên các bộ ba - Các bộ ba có chiều dài bằng nhau - Các bộ ba phối hợp tạo nên các gen - Mỗi gen kết thúc bằng một bộ ba vô nghĩa - Tất cả các gen hợp lại tạo nên hệ gen

Bazơ thứ 2 trong codon

		U	C	A	G	
Bazơ thứ nhất trong codon	U	Phe	Ser	Tyr	Cys	U
		Phe	Ser	Tyr	Cys	C
		Leu	Ser	STOP	STOP	A
		Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U	
	Leu	Pro	His	Arg	C	
	Leu	Pro	Gln	Arg	A	
	Leu	Pro	Gln	Arg	G	
A	Ile	Thr	Asn	Ser	U	
	Ile	Thr	Asn	Ser	C	
	Ile	Thr	Lys	Arg	A	
	Met	Thr	Lys	Arg	G	
G	Val	Ala	Asp	Gly	U	
	Val	Ala	Asp	Gly	C	
	Val	Ala	Glu	Gly	A	
	Val	Ala	Glu	Gly	G	

Bazơ thứ 3 trong codon

Hình 11.

Ngôn ngữ di truyền của ADN cung cấp thông tin cần thiết cho việc sản xuất các protein để thực hiện các quá trình trao đổi chất đảm bảo cho sự sống của sinh vật. Các protein lại có ngôn ngữ riêng của chúng với bảng "alphabet" có 20 "chữ cái". Các chữ cái này là các axit amin.

ARN được dùng để "dịch" ngôn ngữ di truyền sang ngôn ngữ protein. Nó lấy thông tin từ gen trên sợi ADN để tạo ra các protein cần thiết cho sự sống.

Đọc theo gen (cũng là đọc theo ADN) các thông tin dùng cho các axit amin được lưu trữ trong các từ gồm 3 chữ cái gọi là codon. Mỗi codon chỉ định một axit amin cụ thể. Bằng cách đọc bộ codon này các protein đặc thù được sinh ra từ mã di truyền. Các codon trên ADN mã hóa cho axit amin đặc thù. Có 20 axit amin tìm thấy trong các protein tự nhiên.

Bên dưới là một đoạn trong ngôn ngữ của gen:

CCG ACG TCC GAA GAG TGA CCG ACG TCC GAA GAG TGA CCG ACG TCC
GAA GAG TGA CCG ACG TCC GAA GAG TGA CCG ACG TCC GAA GAG TGA CCG ACG
TCC GAA GAG TGA CCG ACG TCC GAA GAG TGA CCG ACG TCC GAA GAG TGA CCG
ACG TCC GAA GAG TGA CCG ACG TCC GAA GAG TGA CCG ACG TCC GAA GAG TGA
CCG ACG TCC GAA GAG TGA CCG ACG TCC GAA GAG TGA CCG ACG TCC GAA GAG
GAA GAG TGA CCG ACG TCC GAA GAG TGA CCG ACG TCC GAA GAG TGA CCG ACG
TCC GAA GAG TGA CCG

Gen bị biến đổi trở thành đột biến

ADN của hai cá thể thuộc cùng một loài có độ giống nhau rất cao - chỉ khác nhau một nucleotid trên 1000. Một đột biến, trong trường hợp đơn giản nhất, là sự thay đổi trong một đoạn ADN. Thay đổi này có thể hoặc không thể dẫn đến sự thay đổi trong protein mà gen đó mã hóa. Thay đổi không ảnh hưởng đến trình tự protein được gọi là hiện tượng đa hình và là một phần của sự biến dị bình thường trong hệ gen người. Tuy nhiên, biến đổi trong ADN thường làm hỏng chức năng của gen mà chúng ta gọi là những biểu hiện lâm sàng. Protein bị biến đổi bắt nguồn từ đột biến có thể phá vỡ cách thức hoạt động của gen và dẫn đến bệnh tật. Các đột biến này biểu hiện như thế nào phụ thuộc vào khả năng di truyền riêng biệt của từng cá thể và vào sự tương tác với môi trường.

Ngoài ra, biến đổi này có thể truyền hoặc không truyền lại cho các thế hệ kế tiếp. Nếu bệnh ung thư không mang tính gia đình, là đột biến xảy ra tại các tế bào soma tách biệt thì nó sẽ không truyền lại cho các thế hệ tiếp theo. Chỉ những đột biến xảy ra trong ADN của giao tử (tinh trùng hoặc tế bào trứng) sẽ truyền lại cho thế hệ sau. Nếu đột biến truyền lại cho thế hệ con thì chúng sẽ mang đột biến đó trong tất cả các tế bào của cơ thể.

Các loại đột biến chủ yếu xảy ra ở mức phân tử được liệt kê và giải thích trong thí dụ ở bảng dưới:

Gen bị biến đổi là đột biến

Bình thường (Normal)

CAC ANH CHI HAY HOC NUA HOC MAI.

Nhầm nghĩa (Missense)

CAC ANH CHA HAY HOC NUA HOC MAI..

Vô nghĩa (Nonsense)

CAC ANH CHI.

Dịch khung do mất đoạn (Frameshift – deletion)

CAC NHC HIH AYH OCN UAH OCM AI.

Dịch khung do thêm đoạn (Frameshift – insertion)

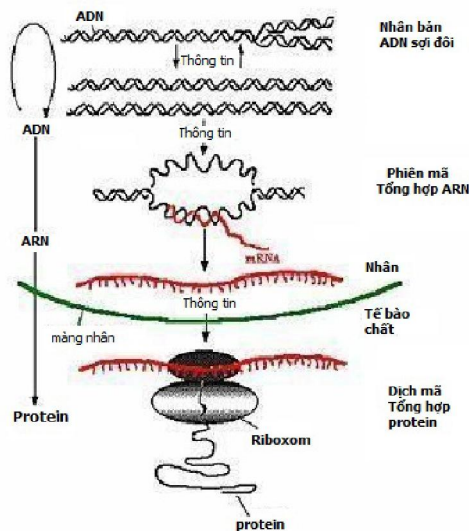
CAC ANH CHI HHA YHO CNU AHO CMA I

Ngoài các đột biến xảy ra mức phân tử, còn có các loại đột biến phát hiện được ở mức nhiễm sắc thể như đảo đoạn, chuyển đoạn, mất đoạn, đột biến tăng giảm số lượng nhiễm sắc thể. Tất cả đều dẫn tới những hậu quả ở mức phân tử ADN và ngày nay đều có thể phát hiện được dễ dàng bằng các chỉ thị phân tử, kể cả trước sinh và sau sinh. Hậu quả của những đột biến này thường hết sức nghiêm trọng. Chẳng hạn, một số dạng ung thư máu là hậu quả của các đột biến chuyển đoạn nhiễm sắc thể.

Các đột biến hóa sinh, trong đó có đột biến khuyết dưỡng, được sử dụng phổ biến trong các nghiên cứu cơ bản và nghiên cứu ứng dụng trên đối tượng vi sinh vật.

Một loại đột biến có cơ chế tác dụng khá đặc biệt gọi là đột biến ức chế, khi chúng không ảnh hưởng trực tiếp đến gen nơi chúng xảy ra mà tác động lên sự biểu hiện của các gen khác.

Con đường từ gen đến tính trạng



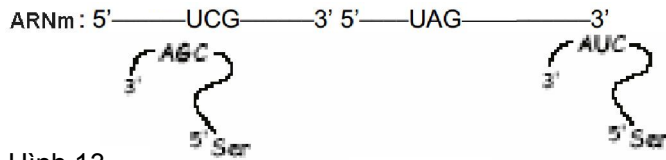
Thuyết trung tâm của sinh học phân tử

Phần tổng kết dưới đây về những quá trình tham gia vào việc chuyển thông tin từ ADN sang các protein dùng để xây dựng cơ thể chúng ta. Đôi khi quá trình này được gọi là Thuyết trung tâm của di truyền học.

- Sao chép là quá trình ADN tự copy mình để chuyển sang tế bào mới trong khi phân chia (nhân lên) của tế bào.

- Phiên mã là quá trình trình tự ADN (các nucleotid) trong gen được sử dụng để tạo ra sợi mRNA (ARN thông tin) giống nó, sau đó sợi này được dùng để điều khiển việc tổng hợp protein.

- Dịch mã là quá trình mà ở đó trình tự của mRNA được dùng để hướng dẫn việc xây dựng trình tự axit amin trong protein.



Hình 13.

Sự cố xảy ra ở bất cứ quá trình nào trong số kể trên đều có thể dẫn đến sự phá vỡ chức năng bình thường của gen và dẫn tới bệnh tật.

Bên dưới là danh sách 20 axit amin và công thức hóa học của chúng do GS. Douglas J. Burks xây dựng và trao tặng.

$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ (\text{CH}_2)_3 \\ \\ \text{NH} \\ \\ \text{C}=\text{NH}_2 \\ \\ \text{NH}_2 \end{array}$ <p>Arginin (Arg / R)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{C}=\text{O} \\ \\ \text{NH}_2 \end{array}$ <p>Glutamin (Gln / Q)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{C}_6\text{H}_5 \end{array}$ <p>Phenylalanin (Phe / F)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{C}_6\text{H}_4 \\ \\ \text{OH} \end{array}$ <p>Tyrosin (Tyr / Y)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{C}_8\text{H}_6\text{N}_2 \end{array}$ <p>Tryptophan (Trp / W)</p>
$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ (\text{CH}_2)_4 \\ \\ \text{NH}_2 \end{array}$ <p>Lysin (Lys / K)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{H} \end{array}$ <p>Glycin (Gly / G)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_3 \end{array}$ <p>Alanin (Ala / A)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{C}_4\text{H}_3\text{N}_2 \end{array}$ <p>Histidin (His / H)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{OH} \end{array}$ <p>Serin (Ser / S)</p>
$\begin{array}{c} \text{H}_2\text{C} \\ \\ \text{C} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \end{array}$ <p>Prolin (Pro / P)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{COOH} \end{array}$ <p>Axit Glutamic (Glu / E)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{COOH} \end{array}$ <p>Axit Aspartic (Asp / D)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_3 \end{array}$ <p>Threonin (Thr / T)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{SH} \end{array}$ <p>Cystein (Cys / C)</p>
$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{S} \\ \\ \text{CH}_3 \end{array}$ <p>Methionin (Met / M)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{CH} \\ \\ \text{CH}_3 \quad \text{CH}_3 \end{array}$ <p>Leucin (Leu / L)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH}_2 \\ \\ \text{C}=\text{O} \\ \\ \text{NH}_2 \end{array}$ <p>Asparagin (Asn / N)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{HC} - \text{CH}_3 \\ \\ \text{CH}_2 \\ \\ \text{CH}_3 \end{array}$ <p>Isoleucin (Ile / I)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} \begin{array}{l} \diagup \text{O} \\ \diagdown \text{O} \end{array} \\ \\ \text{CH} \\ \\ \text{CH}_3 \quad \text{CH}_3 \end{array}$ <p>Valin (Val / V)</p>

Hình 14.

Các protein có vai trò vô cùng to lớn và đa dạng trong cơ thể. Chúng có trách nhiệm vận chuyển, lưu giữ và xây dựng khung cấu trúc của các tế bào. Chúng tạo ra kháng thể, bộ máy enzym để xúc tác các phản ứng hóa sinh học cần thiết cho các hoạt động trao đổi chất. Cuối cùng, các protein là thành phần quan trọng trong nhiều hormon, và các protein co bóp có nhiệm vụ co bóp các cơ và sự vận động của tế bào.

Thí dụ về một số protein quen biết là hemoglobin, collagen, hormon thyroid, insulin, và myosin. Bệnh tật thường là sự biểu hiện một chức năng không phù hợp của protein do ảnh hưởng của di truyền, của môi trường, hoặc cả hai.

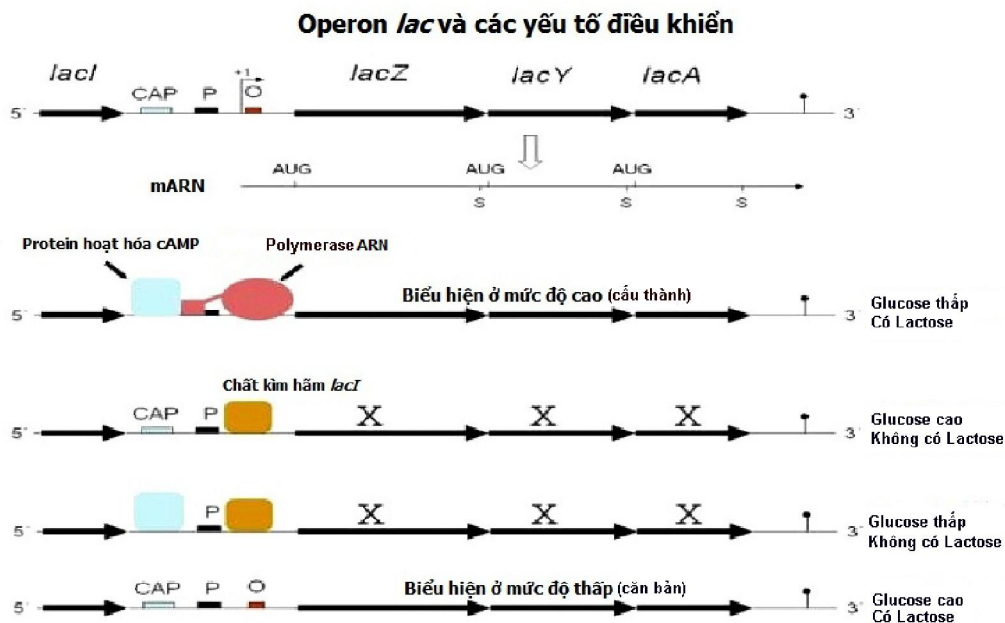
Gen có thể bật và tắt

Qua cơ chế phân bào mà chúng ta sẽ tìm hiểu ở phần sau của tài liệu này, hầu như tất cả các tế bào của một cơ thể đều mang cùng một bộ gen giống nhau. Nếu như tất cả các gen đó cùng hoạt động với mức độ như nhau trong suốt thời gian tồn tại của cơ thể, từ khi hình thành hợp tử trong bụng mẹ đến khi sinh ra và mất đi, thì hình dạng, mới chỉ nói về hình dạng, của tất cả chúng ta đều giống nhau và giống như một quả bóng (hình cầu). Nhưng thực tế đã không diễn ra như vậy. Những gen khác nhau ở những mô và tế bào khác nhau có thể hoạt động chức năng với cường độ khác nhau trong những thời gian khác nhau. Nói cách khác, các gen không sản xuất ra tất cả protein của chúng vào mọi lúc, việc sản xuất protein chịu một sự điều khiển nào đó. Các nhà nghiên cứu Pháp lần đầu tiên đã làm sáng tỏ cơ chế điều hòa hoạt động gen trên đối tượng vi khuẩn gọi là cơ chế biểu hiện gen phân hóa (differential gene expression).

Thí nghiệm đó như sau: Khi có đường lactose trong môi trường thì vi khuẩn *E. coli* mở toàn bộ nhóm gen để chuyển hóa đường. Các nhà nghiên cứu phát hiện thấy lactose đã loại bỏ một nhân tố ức chế khỏi ADN. Sự loại bỏ nhân tố ức chế đã bật điện cho việc sản xuất của gen.

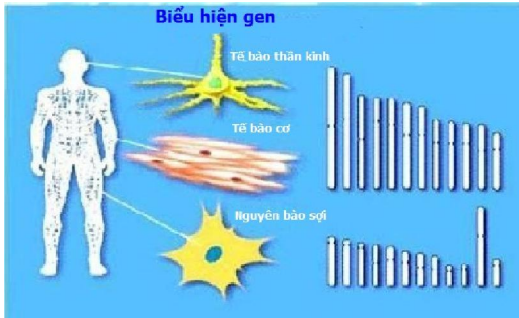
Gen tạo ra nhân tố ức chế gọi là gen điều hòa. Phát minh này đã làm thay đổi quan niệm trước đó về sự phát triển của sinh vật bậc cao. Các tế bào không những có kế hoạch di truyền dùng cho các protein cấu trúc trong ADN của chúng mà còn có cả chương trình điều hòa di truyền để biểu hiện các kế hoạch nói trên.

Các chi tiết của chương trình điều hòa được nghiên cứu trên đoạn lac operon - đơn vị điều hòa của gen. Sơ đồ nguyên lý hoạt động của operon trình bày ở dưới.



Hình 15.

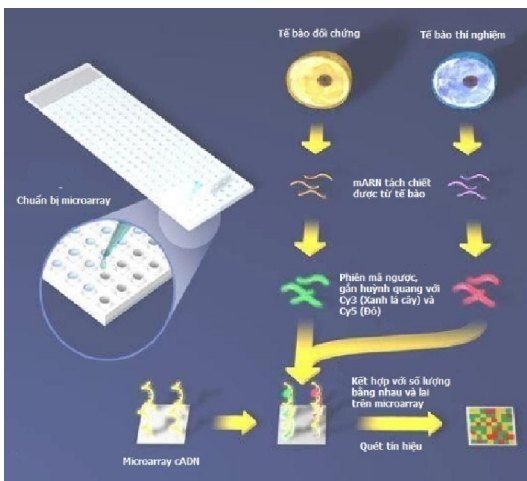
Các gen khác nhau hoạt động trong các tế bào khác nhau



Hình 16.

Tất cả các tế bào trong cơ thể đều mang một bộ gen giống nhau chứa thông tin di truyền đầy đủ. Nhưng chỉ khoảng 20% gen biểu hiện ở mỗi thời điểm cụ thể. Các protein khác nhau được tạo ra ở các tế bào khác nhau phù hợp với chức năng của tế bào. Sự biểu hiện của gen được kiểm soát và điều hòa chặt chẽ.

Phần lớn sinh vật có cơ thể bao gồm nhiều loại tế bào khác nhau được chuyên hóa để thực hiện những chức năng khác nhau. Chúng được gọi là các tế bào biệt hóa, khác với các **tế bào gốc**. Tế bào gan chẳng hạn, không có các nhiệm vụ hóa sinh như các tế bào thần kinh. Vậy là mỗi tế bào của cơ thể có cùng một cấu trúc di truyền, nhưng những loại tế bào khác nhau làm cách nào để có những cấu trúc khác nhau và chức năng hóa sinh khác nhau như thế? Vì chức năng hóa sinh được xác định bởi các enzym (protein) đặc hiệu, nên những bộ gen khác nhau phải được bật và tắt ở các loại tế bào khác nhau. Đó là cách các tế bào phân hóa.



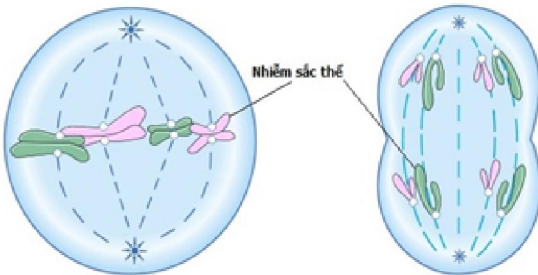
Hình 17.

Quan niệm này về sự biểu hiện đặc thù tế bào của gen đã được chứng minh bằng các thí nghiệm lai - những thí nghiệm có thể xác định các mARN duy nhất có trong mỗi loại tế bào. Gần đây, các phép thử ADN và các chip gen đã tạo điều kiện sàng lọc nhanh chóng tất cả hoạt động gen của một sinh vật. Do vậy, sự biểu hiện đồng thời của các gen khi phản ứng với các nhân tố bên ngoài có thể được sử dụng và xét nghiệm như chỉ ra trên hình bên trên do GS. Douglas J. Burks trao tặng.

Các gen vận động cùng với nhiễm sắc thể theo qui luật nguyên phân và giảm phân

Tuyệt đại đa số các gen ở sinh vật nhân chuẩn nằm trên các nhiễm sắc thể. Vì vậy, nhiễm sắc thể vận chuyển đi đâu thì tất cả các gen định vị trên nó vận chuyển đến đó. Như vậy, sự vận động của các gen phụ thuộc vào cách các nhiễm sắc thể phân bố trong quá trình **nguyên phân** và **giảm phân**.

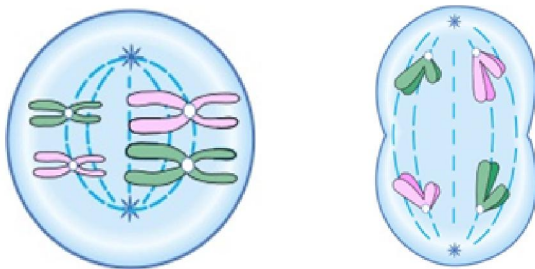
Nguyên phân sinh ra các tế bào giống hệt nhau về mặt di truyền; trong khi các sản phẩm giảm phân lại khác biệt về mặt di truyền vì sự phân ly độc lập.



Hình 18. **Nguyên phân**

tương đồng tách nhau ra để chuyển về hai tế bào con sau đó với xác suất 50:50. Điều đó xảy ra với tất cả các cặp nhiễm sắc thể tương đồng có trong tế bào đang giảm phân. Do vậy, từ mỗi tế bào ban đầu, sau giảm phân tạo ra hai tế bào có cấu trúc di truyền khác nhau vì chúng có hai tổ hợp khác nhau các nhiễm sắc thể, đồng nghĩa với hai bộ gen khác. Chúng ta biết, mỗi tế bào lưỡng bội chứa hai bộ nhiễm sắc thể. Một của mẹ và một bộ từ bố. Một cặp nhiễm sắc thể tương đồng bao gồm một chiếc của mẹ và một chiếc của bố.

Nguyên phân là quá trình mà ở đó các nhiễm sắc thể trong nhân của tế bào nhân chuẩn được chia thành 2 phần giống hệt nhau về mặt di truyền. Do vậy có thể gọi nguyên phân là photocopy vì từ một tế bào ban đầu, sau nguyên phân hình thành hai tế bào có cấu trúc di truyền giống hệt nhau. Còn trong quá trình giảm phân, thông tin di truyền được tổ hợp lại theo cách: Mỗi cặp nhiễm sắc thể



Hình 19. **Giảm phân**

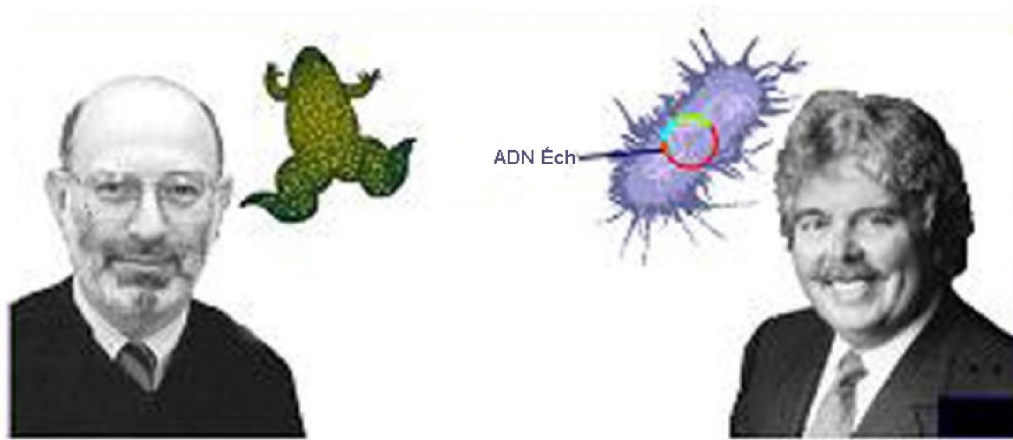
độc lập của các nhiễm sắc thể cha và mẹ. Do vậy, các giao tử do giảm phân sinh ra sẽ có những tổ hợp alen khác nhau vì chúng là kết quả của cả hai hiện tượng tái tổ hợp và sắp xếp độc lập.

Chúng chính là các đơn vị di truyền Mendel, cho thấy sự phân ly và tổ hợp độc lập. Các nhiễm sắc thể tương đồng mang những gen giống nhau nhưng có thể có những **alen** khác nhau.

Các gen nằm trên các nhiễm sắc thể khác nhau cũng chịu sự chi phối của qui luật tổ hợp

Các gen có thể chuyển qua lại giữa các loài

Vì tính phổ biến của mã di truyền mà enzym polymerase của một sinh vật có thể dùng để phiên mã chính xác gen của một sinh vật khác. Thí dụ, những loài vi khuẩn khác nhau có thể nhận được các gen kháng chất kháng sinh bằng cách trao đổi các đoạn ADN nhỏ gọi là plasmid. Vào đầu những năm 1970, các nhà nghiên cứu ở California đã dùng kiểu trao đổi này để chuyển một phân tử ADN tái tổ hợp từ một loài sang loài khác. Vào đầu những năm 1980, các nhà khoa học khác đã cải tiến kỹ thuật và đưa một gen của người vào vi khuẩn *E. coli* để tạo ra insulin tái tổ hợp của người và hormon sinh trưởng.



Hình 20.

Stanley Cohen (bên trái) và Herbert Boyer (bên phải) đã thực hiện thí nghiệm đầu tiên về kỹ thuật di truyền vào năm 1973. Họ đã chứng minh rằng gen của ếch có thể chuyển sang các tế bào vi khuẩn *E. coli* và biểu hiện ở chúng.

Công nghệ ADN tái tổ hợp - tức Kỹ thuật di truyền - hạt nhân của công nghệ sinh học hiện đại, đã giúp con người “nhìn” được vào bên trong cách thức hoạt động các chi tiết của gen. Trong các trường hợp không thể kiểm tra chức năng gen trên mô hình động vật thì đầu tiên gen có thể biểu hiện trên vi khuẩn hoặc trên tế bào động vật nuôi cấy. Tương tự, kiểu hình của đột biến gen và hiệu quả của thuốc và các tác nhân khác có thể được kiểm nghiệm bằng các hệ thống tái tổ hợp. Sự truyền gen này có thể xảy ra trong tự nhiên thông qua hiện tượng biến nạp. Điều đó chứng tỏ rằng điều mà các nhà di truyền học có thể thực hiện được quan trọng hơn rất nhiều so với người ta vốn nghĩ trước đây.

Kỹ thuật thao tác gen cũng như truyền gen sẽ được mô tả chi tiết ở các phần sau.

Hệ gen là bộ đầy đủ các gen của một sinh vật

Trong di truyền học kinh điển hệ gen của một sinh vật lưỡng bội nhân chuẩn, chẳng hạn con người, là một bộ đầy đủ các nhiễm sắc thể hoặc các gen có trong một giao tử; có nghĩa là một tế bào soma bình thường có chứa hai hệ gen đầy đủ. Các sinh vật đơn bội bao gồm vi khuẩn, cổ sinh vật, virus hay ty thể, tức một tế bào chỉ có chứa một hệ gen riêng lẻ, thì thường hệ gen đó nằm trong một phân tử ADN hay ARN (ở một số virus) mạch vòng hay mạch thẳng. Trong sinh học phân tử hiện đại hệ gen của một sinh vật là thông tin di truyền được mã hóa trong ADN hoặc ARN.

Hệ gen bao gồm cả các gen và đoạn ADN không mã hóa protein. Thuật ngữ này được giáo sư Hans Winkler đặt ra vào năm 1920, khi ông làm việc tại Đại học tổng hợp Hamburg, Đức.

Chính xác hơn, hệ gen của một sinh vật là trình tự di truyền đầy đủ của một bộ nhiễm sắc thể; thí dụ, một trong hai bộ nhiễm sắc thể có trong tế bào soma ở một sinh vật lưỡng bội. Thuật ngữ hệ gen có thể áp dụng một cách đặc thù để chỉ một bộ đầy đủ các ADN trong nhân (thí dụ "hệ gen nhân") nhưng cũng có thể áp dụng cho ADN nằm trong các cơ quan tử của tế bào như hệ gen ty thể, hệ gen Lạp thể. Ngoài ra, thuật ngữ hệ gen còn được dùng cho các phần tử di truyền không phải nhiễm sắc thể như virus, plasmid và các yếu tố di truyền vận động (transposable elements). Khi người ta nói rằng hệ gen của một loài sinh sản hữu tính nào đó được giải trình tự, thì điều đó có nghĩa là đã xác định được trình tự của một bộ nhiễm sắc thể thường và các nhiễm sắc thể giới tính.

Cả số lượng cặp bazơ và số gen đều rất khác nhau ở các loài khác nhau. Hiện nay, số lượng gen lớn nhất là vào khoảng 60.000 gen ở động vật nguyên sinh (protozoan), gần gấp 3 lần số gen trong hệ gen người.

Cần chú ý rằng hệ gen không phản ánh tính đa dạng di truyền hay đa hình di truyền của một loài.





Bảng dưới so sánh kích thước của các hệ gen ở những sinh vật khác nhau:



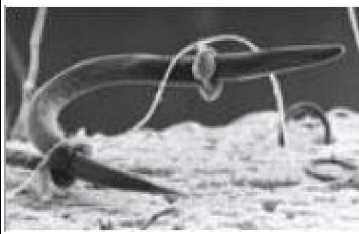

Sinh vật	Kích thước hệ gen (cặp bazơ)	Nguồn tài liệu
Virus, Bacteriophage MS2	3,569	First sequenced RNA-Genome [3]
Virus, SV40	5,224	[4]
Virus, Phage -X174;	5,386	First sequenced DNA-Genome [5]
Virus, Phage	48,502	

Bacterium, Haemophilus influenzae	1,830,000	First genome of living organism, July 1995 [6]
Bacterium, Carsonella ruddii	160,000	Smallest non-viral genome.[7]
Bacterium, Buchnera aphidicola	600,000	
Bacterium, Wigglesworthia glossinidia	700,000	
Bacterium, Escherichia coli	4,600,000	[8]
Amoeba, Amoeba dubia	670,000,000,000	Largest known genome.[9]
Plant, Arabidopsis thaliana	157,000,000	First plant genome sequenced, Dec 2000.[10]
Plant, Genlisea margaretae	63,400,000	Smallest recorded flowering plant genome, 2006.[10]
Plant, Fritillaria assyrica	130,000,000,000	
Plant, Populus trichocarpa	480,000,000	First tree genome, Sept 2006
moss, Physcomitrella patens	480,000,000	First genome of a bryophyte, January 2008 [11]
Yeast, Saccharomyces cerevisiae	12,100,000	[12]
Fungus, Aspergillus nidulans	30,000,000	
Nematode, Caenorhabditis elegans	98,000,000	First multicellular animal genome, December 1998 [13]
Insect, Drosophila melanogaster aka Fruit Fly	130,000,000	[14]
Insect, Bombyx mori aka Silk Moth	530,000,000	
Insect, Apis mellifera aka Honey Bee	1,770,000,000	
Fish, Tetraodon nigroviridis, type of Puffer fish	385,000,000	Smallest vertebrate genome known
Mammal, Homo sapiens	3,200,000,000	
Fish, Protopterus aethiopicus aka Marbled lungfish	130,000,000,000	Largest vertebrate genome known

Các sinh vật khác nhau có chứa nhiều gen chung

Tất cả các sinh vật đều lưu trữ thông tin di truyền trong cùng những phân tử như nhau, ADN hoặc ARN. Mã di truyền chung ghi trên các phân tử này là bằng chứng tất yếu cho thấy tất cả các sinh vật đều có một tổ tiên chung. Quá trình tiến hóa của các dạng sống bậc cao đòi hỏi sự phát triển của những gen mới để hỗ trợ những chương trình khác nhau của cơ thể và những kiểu dinh dưỡng khác nhau. Mặc dù vậy, các sinh vật phức tạp vẫn duy trì nhiều gen điều khiển các chức năng chuyển hóa nòng cốt được thực hiện từ quá khứ sơ khai của chúng.

	Các gen chung của các sinh vật với con người	% gen chung với con người
	<p>Khi Chimpanzee, <i>Pantroglodytes</i>, 30 000 genes Chimpanzees have about the same number of genes as humans. But then why can't they speak? The difference could be in a single gene, FOXP2, whichin the chimpanzee is missing certain sections.</p>	98%
	<p>Chuột, <i>Mus musculus</i>, 30000 genes Thanks to mice, researchers have been able to identify genes linked to skeletal development, obesity and Parkinson's disease, to name but a few.</p>	90%
	<p>Cá Vàng, (Zebra Fish, <i>Danio rerio</i>), 30000 genes 85% of the genes in these little fish are the same as yours. Researchers use them to study the role of genes linked to blood disease such as anemia falciforme and heart disease.</p>	85%
	<p>Ruồi dấm, <i>Drosophila melanogaster</i>, 13600 genes For the past 100 years, the fruit fly has been (Image: David M.Phillips, Visuals Unlimited,Inc.)</p>	36%

	<p>Cây Arabidopsis thaliana, 25000 genes. This little plant, from the mustard family, is used as a model for the study of all flowering plants. Scientists use its genes to study hepatolenticular degeneration, a disease causing copper to accumulate in the human liver. (Image: Wally Eberhart, Visuals Unlimited, Inc.)</p>	<p>26%</p>
	<p>Nấm men, Saccharomyces cerevisiae, 6275 genes You have certain genes in common with this organism that is used to make bread, beer and wine. Scientists use yeast to study the metabolism of sugars, the cell division process, and diseases such as cancer. (Image: Kessel & Shih, Visuals Unlimited, Inc.)</p>	<p>23%</p>
	<p>Giun đũa, Caenorhabditis elegans, 19000 genes Just like you, this worm possesses muscles, a nervous system, intestines and sexual organs. That is why the roundworm is used to study genes linked to aging, to neurological diseases such as Alzheimer's, to cancer and to kidney disease.</p>	<p>21%</p>
	<p>Vi khuẩn, Escherichia coli, 4800 genes The E. coli bacterium inhabits your intestines. Researchers study it to learn about basic cell functions, such as transcription and translation. (Image: Fred Hossler, Visuals Unlimited, Inc.)</p>	<p>7%</p>

Các gen được duy trì qua quá trình tiến hóa của sinh vật; tuy nhiên chúng có thể bị biến đổi hoặc được lấy từ các sinh vật khác. Các vi khuẩn có thể trao đổi với nhau các plasmid kháng thuốc kháng sinh, các virus có thể đính gen của chúng vào vật chủ. Một số gen của động vật có vú cũng có thể bị virus lấy đi và sau đó đính vào vật chủ khác của chúng.

Như vậy, tất cả các sinh vật đều có những gen cổ xưa bắt nguồn từ buổi ban đầu khi loài người có các gen chung với các sinh vật. Nếu con người có nhiều gen chung với các loài khác như vậy thì cái gì tạo nên sự khác biệt ở con người? Cái gì đã biến con người thành một sinh linh phức hợp biết học, biết nói, biết nghĩ và cảm nhận? Và cái gì đã làm cho người ta khác biệt nhau?

Chúng ta có nhiều gen chung với con chuột và con giun nhiều hơn ta tưởng! Bất chấp ngoại hình khác biệt, chúng ta có số gen chung với các loài khác một cách đáng ngạc nhiên (xem bảng trên). Mặc dù không phải các gen này có chung các nucleotid trong cùng một trình tự sắp xếp như nhau, nhưng chức năng của chúng có độ tương đồng đủ để so sánh với nhau. Chắc rằng chúng bắt nguồn từ một tổ tiên chung sống cách đây khoảng 3,5 tỷ năm. Các nhà khoa học cho rằng quá trình tiến hóa hệ gen tổ tiên này trở thành cơ sở cho mỗi loài mà ngày nay chúng ta quan sát thấy.

Sự tương đồng về trình tự ADN ở gen béo phì

Người	CTGCCACTTG	CCCTGGGCCA	GTGGCCTGCA
Đười ươi	CTGCCACTTG	CCCTGGGCCA	GTGGCCTGGA
Ngựa	CTGCCCTTG	CCCCAGGCCA	GGGGTGTGGA
Bò	CTGCCCTTG	CCGCAGGTCA	GGCCCTGGA
Lợn	CTGCCCTTG	CCCCGGGCCA	GGGGCTGGA
Chó	CTGCCCTTG	CCCCAGGCCA	GGCCCTGGA
Chuột	CTGTCCTTG	CCTCAGGCCA	GTGGCCTGCA

Chính vì vậy, thành phần của nhiều gen rất giống nhau. Hình trên bảng bên trái cung cấp một thí dụ về gen béo phì obesity (ob) ở những động vật khác nhau, trình tự của chúng rất giống nhau. Bức tranh phía dưới thậm chí cho thấy các trình tự giống hệt nhau ở những sinh vật rất khác nhau, từ nấm men đến con người.

Hình 21.

Các sinh vật chia sẻ các gen tương tự bởi vì chúng có nguồn gốc từ một tổ tiên chung. Ngay cả con người và nấm men cũng có các gen tương tự nhau.



..... Glu Tyr Lys Ile Val Val Val Gly Gly Gly Gly Val Gly Lys Ser Ala Leu Thr Ile Gln Phe Ile Gln Ser Tyr Phe.. **NẤM MEN**
 Glu Tyr Lys Ile Val Val Val Gly Gly Gly Gly Val Gly Lys Ser Ala Leu Thr Ile Gln Leu Ile Gln Asn His Phe.. **NẤM NHẬY**
 Glu Tyr Lys Leu Val Val Val Gly Pro Gly Gly Val Gly Lys Ser Ala Leu Thr Ile Gln Leu Ile Gln Asn His Phe.. **RUỒI GIẤM**
 Glu Tyr Lys Leu Val Val Val Gly Ala Gly Gly Val Gly Lys Ser Ala Leu Thr Ile Gln Leu Ile Gln Asn His Phe.. **GÀ**
 Glu Tyr Lys Leu Val Val Val Gly Ala Gly Gly Val Gly Lys Ser Ala Leu Thr Ile Gln Leu Ile Gln Asn His Phe.. **NGƯỜI**

Hình 22.

Các gen có thể được cắt, ghép, sửa đổi

Sự tiến bộ trong bất kỳ lĩnh vực khoa học nào đều phụ thuộc vào những kỹ thuật và phương pháp hiện có, có khả năng mở rộng phạm vi và đạt độ chính xác cao của các thí nghiệm cần thực hiện. Trong hơn 50 năm qua điều đó đã được minh chứng một cách hết sức ngoạn mục làm nổi bật lên lĩnh vực di truyền học phân tử - nòng cốt của Công nghệ sinh học hiện đại. Lĩnh vực này phát triển hết sức nhanh chóng đạt tới đỉnh mà ngày nay trong rất nhiều phòng thí nghiệm trên toàn thế giới việc nghiên cứu gen đã trở thành thực tế sôi sục hàng ngày, bao gồm tách chiết đoạn ADN đặc thù từ hệ gen của một sinh vật, xác định trình tự các bazơ và đánh giá chức năng của nó. Điều đặc biệt ấn tượng là công nghệ này hết sức dễ dàng sử dụng đối với từng cá nhân nhà khoa học mà không cần đến những thiết bị qui mô cũng như nguồn tài chính lớn ngoài tầm với của các phòng nghiên cứu được trang bị vừa phải.

Mặc dù có rất nhiều kỹ thuật phức tạp và đa dạng liên quan, nhưng nguyên lý cơ bản của thao tác di truyền lại khá đơn giản. Quan điểm chung mà các công nghệ dựa vào là thông tin di truyền được mã hóa trong ADN và được bố trí ở dạng các gen, là nguyên liệu nguồn có thể được thao tác theo nhiều cách khác nhau nhằm đạt tới những mục tiêu cụ thể.

Sau đây là những phương pháp và công nghệ chủ yếu dùng hàng ngày tại các phòng thí nghiệm di truyền học hiện đại:

Tách chiết ADN. Bước đầu tiên cần thực hiện để có được ADN mong muốn phù hợp với mục tiêu thí nghiệm. Nguyên lý chung giống nhau, tuy có khác biệt nhỏ ở 3 loại đối tượng khác nhau: động vật, thực vật và các tế bào nhân sơ như vi khuẩn.

Các phương pháp lai phân tử. Dựa trên tính chất bổ trợ giữa 2 sợi ADN mạch kép hoặc giữa sợi ADN và ARN có thể có các phương pháp Southern blotting, Northern blotting và lai tại chỗ (including fluorescent in situ hybridization - FISH). Các kỹ thuật lai cho phép tách biệt các gen quan tâm từ một hỗn hợp gồm nhiều đoạn ADN, ARN hoặc cả hai loại.

Sửa đổi ADN bằng enzym. Dựa vào đặc tính hoạt động đặc hiệu của các enzym lên ADN người ta có thể cắt ADN tại những điểm mong muốn (enzym giới hạn), sau đó có thể nối các đoạn ADN lại với nhau bằng enzym ADN ligase, tạo nên đoạn ADN có thành phần và nguồn gốc khác nhau, thuật ngữ chuyên môn gọi là *ADN tái tổ hợp*. Ngoài ra, còn nhiều loại enzym khác có đặc thù hoạt động khác trên ADN như phân hủy ADN từ một đầu nhất định, hay từ giữa sợi ADN ra ngoài - Nói chung, ngày nay trong tay các nhà khoa học có những công cụ phong phú để tạo ra các đoạn ADN mang những chức năng cần thiết cho các nhu cầu đa dạng của con người.

Gắn vào vector. Để chuyển ADN được tạo ra bằng các phương pháp trên, thường cần gắn ADN đó vào các cấu trúc ADN mạch vòng nhỏ gọi là plasmid, cosmid hay một loại virus, tựa như đưa hàng lên xe để chuyển đi. Sau đó nhân số lượng các vector lên để sẵn sàng cho các thí nghiệm chuyển gen đến cá thể vật chủ thích hợp. Do vậy kỹ thuật này thường gọi là *tách dòng gen*.

Tách dòng gen được sử dụng trong các lĩnh vực sau: nhận dạng gen đặc hiệu, lập bản đồ hệ gen, sản xuất các protein tái tổ hợp, và tạo ra các sinh vật chuyển gen (GMO).

Phản ứng chuỗi trùng hợp (PCR). Là kỹ thuật không thể thiếu trong các phòng thí nghiệm di truyền học hiện đại. Bản chất của phương pháp là cho phép nhân lên hàng triệu lần đoạn ADN quan tâm chỉ trong vòng hơn một giờ đồng hồ. GS. James Watson, người cùng Francis Crick tìm ra cấu trúc không gian chuỗi xoắn kép của ADN đã nhận xét: “Kỹ thuật PCR đã cách mạng hóa công nghệ sinh học”. Đặc biệt đối với các nước nghèo như nước ta, kỹ thuật này đã tạo ra cơ hội ngàn năm có một để vượt đỏi các nước tiên tiến, do không cần đầu tư lớn mà vẫn làm được khoa học và công nghệ ở mức độ cao.

Phương pháp PCR do Kary Mullis phát minh vào giữa những năm 1980, dựa trên những nguyên lý tổng hợp ADN tự nhiên trong tế bào và gồm các bước sau: 1) Tách sợi ADN mạch kép thành 2 sợi đơn bằng nhiệt độ cao, để mỗi sợi đơn có thể làm khuôn để tổng hợp các sợi mới; 2) Gắn 2 đoạn mồi (các đoạn ADN ngắn có chiều dài từ 18 - 30 nucleotid) vào 2 sợi ADN khuôn ở nhiệt độ thích hợp; 3) Tổng hợp 2 sợi ADN mới bắt đầu từ vị trí 2 mồi trên đã gắn vào ở một nhiệt độ khác. Các bước này được lặp lại 20-40 lần và một đoạn ADN mới nằm giữa 2 mồi được tổng hợp.

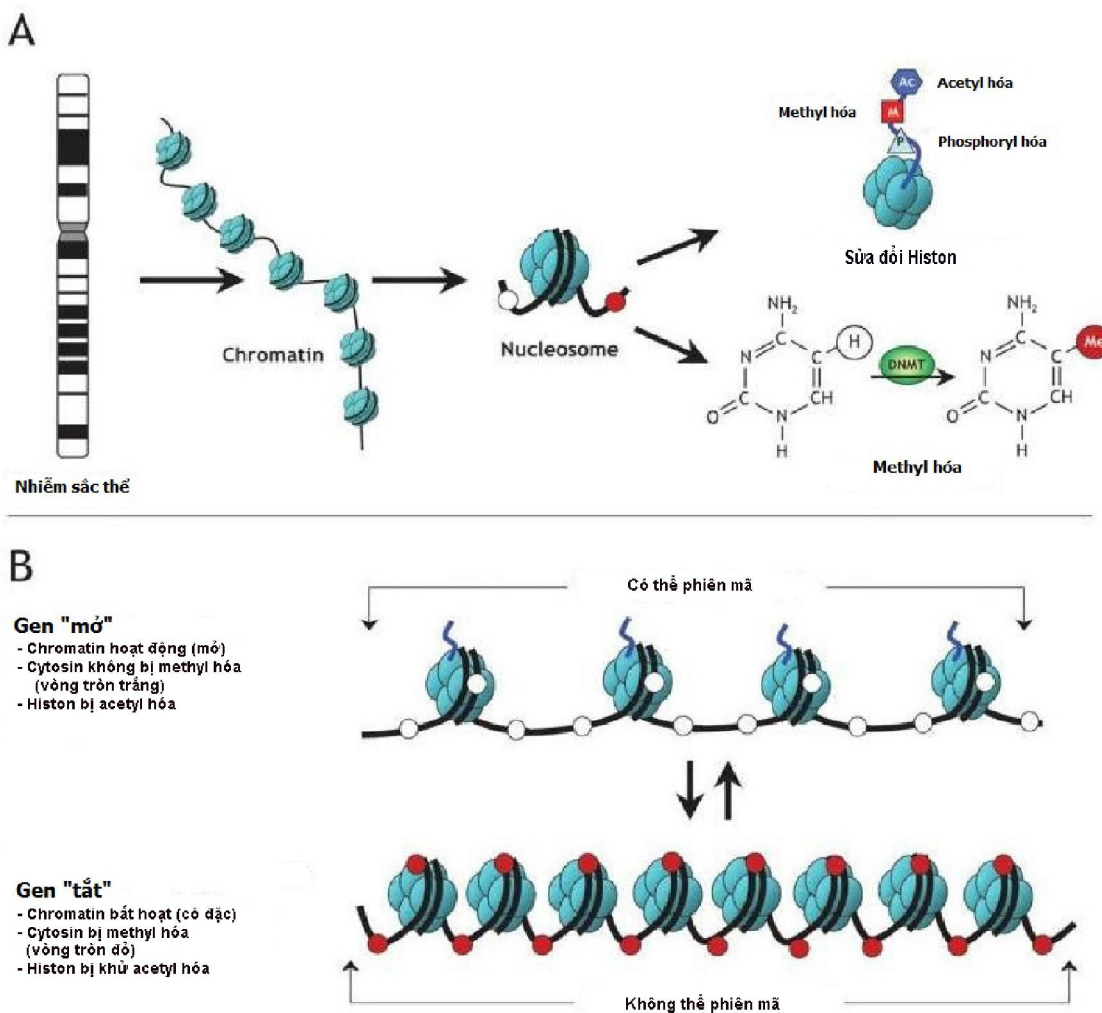
Ứng dụng của PCR

1. Chẩn đoán bệnh trong di truyền y học, vi sinh vật y học và y học phân tử.
2. Định typ HLA trong ghép tạng.
3. Phân tích ADN trong vật liệu cổ sinh.
4. Xác định đặc trưng cá thể trong hình sự và trong xét nghiệm huyết thống.
5. Chuẩn bị mẫu dò các axit nucleic.
6. Sàng lọc dòng và xây dựng bản đồ.
7. Nghiên cứu đa dạng di truyền của các loài.

Giải trình tự ADN. Một công nghệ cho phép giải trình tự trực tiếp các đoạn ADN thay vì cố gắng tìm ra trình tự các gen, trình tự đột biến ADN và các gen mới bằng các phương pháp truyền thống như phân tích RFLP, nhiễm sắc thể di chuyển hay thậm chí các thí nghiệm tải nạp, tiếp hợp ở vi khuẩn. Ngày nay giải trình tự ADN đã đạt đến giai đoạn tự động hóa và được sử dụng hàng ngày trong nhiều phòng thí nghiệm cho nhiều mục đích khác nhau.

Ngoại di truyền (Epigenetics) là một cách điều khiển hoạt động gen hiệu quả

Ngoại di truyền là lĩnh vực nghiên cứu những *thay đổi di truyền trong biểu hiện gen* mà không thay đổi trong trình tự ADN. Các nghiên cứu đã cho thấy rằng các cơ chế ngoại di truyền cung cấp thêm một tầng kiểm soát phiên mã cho phép điều hòa cách thức mà gen biểu hiện. Các cơ chế này là các hợp phần then chốt trong sự phát triển và sinh trưởng bình thường của tế bào. Các bất thường ngoại di truyền đã phát hiện thấy là những nhân tố gây ra ung thư, rối loạn di truyền và các hội chứng nhi khoa. Chúng cũng có thể là những nhân tố góp phần vào các bệnh tự miễn dịch và lão hóa. Bài này sẽ giới thiệu những nguyên lý cơ bản của các cơ chế ngoại di truyền và sự góp phần của chúng cho sức khỏe con người cũng như những hậu quả lâm sàng của các sai lệch ngoại di truyền. Bài cũng sẽ đề cập việc sử dụng con đường ngoại di truyền trong các cách tiếp cận mới đối với chẩn đoán và điều trị có định hướng thông qua phổ lâm sàng.



Hình 23.

Lĩnh vực mới này sẽ có ảnh hưởng rộng lớn đến y học, đặc biệt đối với việc nghiên cứu các biến đổi di truyền trong chức năng gen mà không biến đổi trình tự ADN. Lĩnh vực phát triển rất nhanh này đang tạo ra những cơ hội mới hết sức hấp dẫn cho chẩn đoán và điều trị các rối loạn lâm sàng phức tạp. Nguyên lý cơ bản của ngoại di truyền là hiện tượng *methyl hóa ADN* và việc *sửa đổi histon*.

(A) **Sơ đồ sửa đổi ngoại di truyền** Các sợi ADN cuộn gói xung quanh các khối bát phân histon, tạo ra các thể nhân (nucleosome), cấu trúc này được tổ chức thành chromatin - đơn vị vật liệu cấu thành nhiễm sắc thể. Quá trình sửa đổi histon đặc hiệu điểm và thuận nghịch diễn ra tại nhiều điểm thông qua acetyl hóa, methyl hóa và phosphoryl hóa. Methyl hóa ADN xảy ra tại vị trí 5 của các gốc cytosin trong một phản ứng được xúc tác bởi các enzym ADN methyltransferases (DNMTs). Đồng thời, những sửa đổi này cung cấp một dấu hiệu (signature) ngoại di truyền riêng biệt có tác dụng điều hòa cấu trúc chromatin và biểu hiện gen.

(B) **Sơ đồ thay đổi thuận nghịch** trong cấu trúc chromatin ảnh hưởng đến biểu hiện gen: các gen được biểu hiện (bật) khi chromatin mở (hoạt động), và chúng bị bất hoạt (đóng) khi chromatin bị cô đặc (câm). Trên hình các vòng trắng là các cytosin không bị methyl hóa; các vòng đen là các cytosin bị methyl hóa.

Hậu quả lâm sàng của các sai lệch ngoại di truyền và các cơ chế điều hòa chịu ảnh hưởng của ADN qua suốt thời gian sống của con người. Ngay lập tức sau thụ tinh, hệ gen của người bố nhanh chóng bị demethyl hóa ADN và sửa đổi histon. Hệ gen của mẹ thì demethyl hóa từ từ, và tất nhiên một làn sóng methyl hóa mới của phôi được khởi đầu tạo dựng một kế hoạch chi tiết cho các mô của phôi đang phát triển. Kết quả là mỗi tế bào có một mẫu hình ngoại di truyền của riêng nó được duy trì nghiêm ngặt để điều hòa sự biểu hiện gen cần thiết. Những dao động trong các mẫu hình được bố trí chặt chẽ này của quá trình methyl hóa ADN và sửa đổi histon có thể dẫn đến những rối loạn bẩm sinh và những hội chứng thai nhi đa hệ thống hoặc những con người tiềm ẩn trạng thái bệnh tật như ung thư rải rác và các rối loạn suy thoái thần kinh.

Lão hóa Cả việc tăng cường hoặc thuyên giảm trong quá trình methyl hóa ADN đều liên quan tới quá trình lão hóa. Các bằng chứng tích lũy cho đến nay cho thấy những thay đổi methyl hóa phụ thuộc tuổi tham gia vào sự phát triển các rối loạn thần kinh, tính tự miễn dịch và ung thư ở những người già. Những biến đổi methyl hóa diễn ra theo tuổi có thể bao gồm việc bất hoạt các gen liên quan ung thư. Ở một số mô mức độ cytosin bị methyl hóa tăng cao trong các tế bào đang lão hóa và quá trình demethyl hóa này có thể kích thích tính bất ổn định của nhiễm sắc thể gây ra những tái cấu trúc làm tăng nguy cơ ung thư. Trong các mô khác như ruột, việc tăng cao methyl hóa có thể là sự kiện tiềm ẩn mắc ung thư ruột kết khi tuổi cao.

Ung thư và liệu pháp ngoại di truyền Ung thư là một quá trình bao gồm nhiều bước mà ở đó những sai lệch di truyền và ngoại di truyền tích lũy lại và

chuyển một tế bào bình thường thành tế bào ung thư xâm lấn hoặc di căn. Các mẫu hình methyl hóa ADN bị biến đổi làm thay đổi sự biểu hiện của các gen liên quan đến ung thư. Quá trình methyl hóa ADN giảm sẽ hoạt hóa các gen ung thư và khởi đầu sự bất ổn định của nhiễm sắc thể, trong khi quá trình tăng methyl hóa ADN lại khởi đầu cho sự câm lặng các gen ức chế ung thư. Tỷ lệ mắc phải hiện tượng tăng siêu methyl hóa, đặc biệt trong ung thư rời rạc, dao động tùy thuộc gen liên quan và loại ung thư mà ở đó sự cố xảy ra.

Cho đến nay, các liệu pháp ngoại di truyền còn ít về số lượng nhưng một số trong đó đang được nghiên cứu trong các thử nghiệm lâm sàng hoặc đã được chấp nhận cho áp dụng điều trị những kiểu ung thư đặc thù. Một số chất đồng đẳng như azacitidin được gắn vào ADN đang sao chép, ức chế methyl hóa và tái hoạt hóa các gen đã bị câm trước đó. Azacitidin có hiệu quả trong thử nghiệm lâm sàng pha I khi điều trị hội chứng myelodysplast và bệnh bạch cầu do siêu methyl hóa gây ra. Các oligonucleotid đối nghĩa cũng cho những kết quả rất hứa hẹn trong thử nghiệm lâm sàng pha I đối với cá khối u rắn và ung thư thận. Tương tự, các phân tử nhỏ như axit valproic cũng đang được sử dụng để làm chết các tế bào ung thư. Sự phối hợp liệu pháp ngoại di truyền với liệu pháp hóa học truyền thống có thể đưa lại hiệu quả cao hơn vì chúng tái kích hoạt các gen câm trong đó có các gen ức chế ung thư, đồng thời làm cho các tế bào nhờn thuốc nhạy cảm trở lại với liệu pháp chuẩn bình thường.

Con đường phía trước Kiến thức của chúng ta về các cơ chế ngoại di truyền đã tăng lên trong 20 năm qua đã bắt đầu chuyển hóa thành những cách tiếp cận mới trong chẩn đoán phân tử và điều trị nhằm đích thông qua phổ lâm sàng. Với sự hoàn thành Dự án hệ gen người, Dự án ngoại hệ gen người đã được đề xuất sẽ tạo ra các bản đồ methyl hóa cho cả hệ gen. Bằng cách nghiên cứu cả các mô khỏe và mô bệnh các vùng hệ gen đặc thù sẽ được nhận dạng, chúng tham gia vào quá trình phát triển, vào sự biểu hiện đặc thù mô, vào tính nhạy cảm với môi trường và vào sự phát sinh bệnh. Việc sử dụng các bản đồ ngoại di truyền này sẽ dẫn đến các liệu pháp ngoại di truyền điều trị các rối loạn phức tạp xuyên suốt phổ lâm sàng.