



Sebenta de apoio ao estudo

Unidade curricular de Genética e Melhoramento de Plantas

Licenciatura em Agronomia 2024/2025

Lénia Rodrigues, Hélia Cardoso



Genética de Plantas

Índice

Introdução	3
Capítulo 1 – A molécula de DNA	4
Capítulo 2 – Ciclo celular	21
Capítulo 3 - Mutações	37
Capítulo 4 – Genética Mendeliana	48
Capítulo 5 – Genética quantitativa	68
Capítulo 6 – Síntese de proteínas: Transcrição e Tradução	79
Capítulo 7 – Mutações génicas	96
Capítulo 8 - Marcadores moleculares	105
Bibliografia	117

Introdução

A Genética é a ciência especializada no estudo dos genes e da hereditariedade, investigando a forma como os organismos recebem e transmitem as características biológicas de geração em geração. No contexto da agricultura, a genética é a base do melhoramento de plantas e desempenha um papel fundamental na agricultura moderna. O conhecimento dos princípios genéticos permite desenvolver cultivares mais produtivas, resistentes a pragas e doenças, e mais resilientes face às condições ambientais. A unidade curricular de *Genética e Melhoramento de Plantas* tem como objetivo proporcionar aos alunos de agronomia as ferramentas necessárias para compreender os mecanismos da hereditariedade e aplicá-los na otimização da produção agrícola.

Esta sebenta foi realizada com o objetivo de apoiar no estudo da genética aplicada às plantas, abordando desde os conceitos fundamentais, como a estrutura da molécula de DNA, a transmissão de características hereditárias e as leis de Mendel, até tópicos mais avançados, como a heritabilidade e a genética quantitativa, destacando a importância das novas biotecnologias no melhoramento vegetal.

Ao longo desta sebenta, os alunos terão acesso a explicações teóricas, exemplos práticos, aplicações da genética no contexto agronómico e curiosidades sobre diversos temas. Espera-se que esta abordagem integrada contribua para uma melhor compreensão dos princípios genéticos e para a formação de profissionais capacitados a enfrentar os desafios da agricultura do futuro.

Bom estudo!

Capítulo 1 – A molécula de DNA

O ácido desoxirribonucleico (DNA, do inglês *deoxyribonucleic acid*) é uma molécula que contém as informações essenciais para o desenvolvimento dos organismos, sendo responsável pela transmissão de todas as características genéticas no processo de reprodução dos seres vivos. É a sequência específica de bases nucleotídicas que compõem o DNA que determina a composição em aminoácidos das proteínas estruturais das células, bem como das enzimas necessárias para catalisar as reações bioquímicas.

Neste primeiro capítulo serão abordados os conceitos fundamentais acerca da estrutura e organização da molécula de DNA no núcleo celular, bem como da informação que a mesma contém.

1.1. Os nucleótidos como unidades básicas de ácidos nucleicos

Há dois tipos de ácidos nucleicos: DNA (ácido desoxirribonucleico) e RNA (ácido ribonucleico). Ambos são constituídos por nucleótidos unidos em sequências - polinucleótidos - formando longas cadeias. O nucleótido de DNA chama-se desoxirribonucleótido e o de RNA ribonucleótido, sendo a diferença ao nível do açúcar.

Um nucleótido é composto por três unidades, as quais se encontram covalentemente ligadas (figura 1):

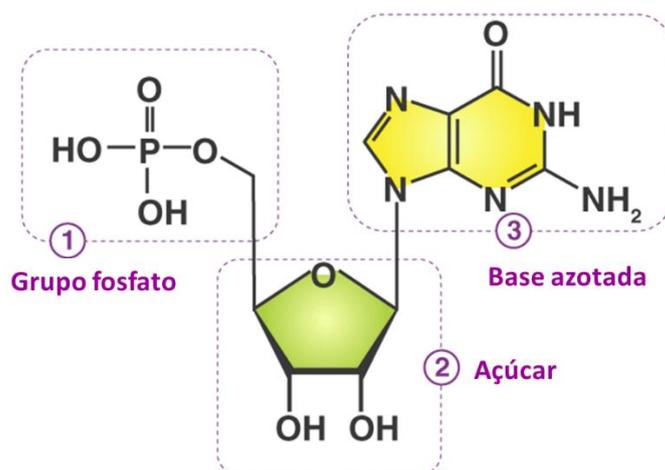


Figura 1- Estrutura química de um nucleótido (Adaptado de <https://byjus.com/neet/nucleotide/>).

1. **Grupo fosfato** – o grupo fosfato encontra-se associado ao açúcar através do carbono 5' do grupo -OH. Os nucleótidos podem conter até três grupos fosfato.
2. **Açúcar** – molécula formada por 5 carbonos sendo, por isso, classificada como pentose. A presença ou ausência de um grupo -OH ligado ao segundo carbono determina se a pentose é designada de **Ribose** (presente no RNA) ou **Desoxirribose** (presente no DNA), respetivamente.
3. **Base azotada** – Purinas ou pirimidinas (tabela 1). A molécula de DNA pode conter as bases Adenina, Guanina, Citosina ou timina. Na molécula de RNA, a base Timina é substituída por Uracilo.

Os precursores dos nucleótidos denominam-se **nucleósidos**. Os nucleósidos apresentam a mesma estrutura que um nucleótido, com exceção do grupo fosfato. Assim, podemos dizer que um nucleótido sem o grupo fosfato é denominado nucleósido (figura 2).

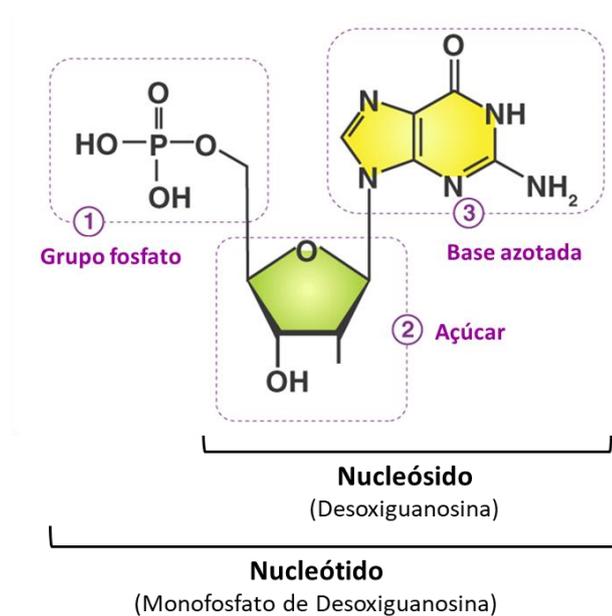
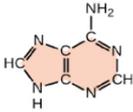
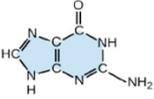
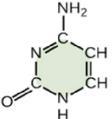
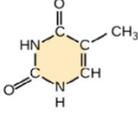
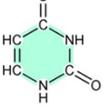


Figura 2- Diferenças ao nível da estrutura química de um nucleótido e nucleósido através do exemplo do desoxinucleótido monofosfato de desoxiguanosina (Adaptado de <https://byjus.com/neet/nucleotide/>).

Além de serem os precursores dos nucleótidos, os nucleósidos têm um papel fundamental em vários processos biológicos, desempenhando funções reguladoras, como por exemplo a adenosina que atua como molécula sinalizadora na síntese de hormonas vegetais, ou envolvidos na biossíntese de compostos secundários das plantas, como os alcaloides e flavonoides.

Tabela I – Classificação, estrutura química e características das bases azotadas que entram na composição do DNA e RNA.

Base azotada	Representação	Estrutura	Fórmula química	Características	
Purinas	Adenina	A		$C_5H_5N_5$	Duplo anel de átomos de carbono com 5 átomos de azoto na sua estrutura
	Guanina	G		$C_5H_5N_5O$	
Pirimidinas	Citosina	C		$C_4H_4N_3O$	Anel simples de átomos de carbono com 2 ou 3 átomos de azoto na sua estrutura
	Timina	T		$C_5H_6N_2O_2$	
	Uracilo	U		$C_4H_4N_2O_2$	

1.2. Formação da cadeia polimérica

A descoberta da estrutura da molécula de DNA foi um marco fundamental na história da biologia molecular. Em 1938, William Astbury obteve imagens cristalográficas do DNA que evidenciavam uma elevada densidade da molécula e uma estrutura helicoidal, determinando também o intervalo entre a repetição das suas subunidades. No entanto, a correlação entre a composição química e a estrutura não era óbvia, e não parecia ter um significado funcional. Em 1952, a confirmação definitiva que o DNA era o material hereditário foi obtida por Alfred Hershey e Martha Chase. E foi em 1953 que a molécula de DNA foi identificada pela primeira vez por James Watson e Francis Crick. A

descoberta foi publicada no artigo "*Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid*", na revista *Nature*, em abril de 1953.

Watson e Crick conseguiram decifrar a estrutura do DNA com base em diversos dados experimentais de outros investigadores: Erwin Chargaff demonstrou que havia diferenças entre os organismos na composição das bases, sugerindo os primeiros indícios da diversidade no DNA, e os estudos de difração de raios X realizados por Rosalind Franklin e Maurice Wilkins forneceram informações importantes sobre a estrutura do DNA. A partir dessas informações e das suas próprias análises, Watson e Crick propuseram o modelo da dupla hélice para a estrutura do DNA.

De acordo com o modelo proposto, os nucleótidos são unidos por ligações covalentes (ligações fosfodiéster) entre o grupo fosfato de um nucleótido e o grupo OH-3' do açúcar do nucleótido seguinte, conferindo uma grande resistência química à molécula e possibilitando a continuidade da sequência de nucleótidos (Figura 3). Cada cadeia polinucleotídica apresenta na extremidade 5' livre um grupo fosfato e na extremidade 3' livre um grupo hidroxilo, tendo-se convencionado escrever as sequências nucleotídicas no sentido 5'-3'.

O facto de conter dois grupos químicos nas suas extremidades (3'-OH e 5'-P), confere polaridade à cadeia polinucleotídica. Em condições fisiológicas normais, a pH7, os grupos fosfato são carregados negativamente, conferindo carga negativa à molécula de DNA.

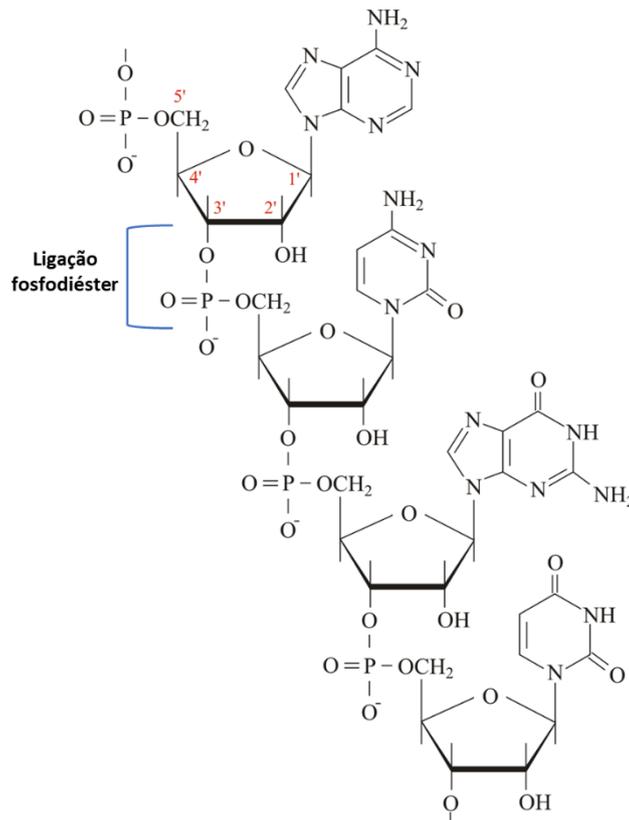


Figura 3 - Estrutura química de uma cadeia simples de DNA. Adaptado do livro *Engenharia genética: princípios e aplicações* (2001).

Como foi referido anteriormente, Chargaff demonstrou que o total das purinas era sempre igual ao total das pirimidinas, ou seja, a quantidade de Adeninas (A) era igual à de Timinas (T), e a quantidade de Guaninas (G) era igual à de Citosinas (C). Watson e Crick, em 1953, associando os dados de Chargaff, propuseram o modelo de dupla hélice para a estrutura do DNA, no qual a molécula é formada por duas cadeias polinucleotídicas enroladas uma à volta da outra, formando uma dupla hélice. O “esqueleto” de açúcar-fosfato fica posicionado para o exterior da dupla hélice e as bases azotadas (apolares) de cada cadeia estão orientadas para o interior, ligando-se às bases da outra cadeia por ligações de hidrogénio. Os emparelhamentos ocorrem entre A e T, com duas ligações de hidrogénio, e entre G e C, com três ligações de hidrogénio, determinando uma rigorosa complementaridade de bases (Figura 4). As duas cadeias polinucleotídicas de uma molécula de DNA não são, portanto, idênticas, mas sim complementares. Esta organização contribui para a estabilidade da cadeia.

Os pares de bases estão afastados 0,34nm na dupla hélice, sendo que cada volta completa da hélice (3,4nm) corresponde a 10 pares de nucleótidos. As cadeias não estão espaçadas uniformemente ao longo da dupla hélice, formando cavidades de diferentes tamanhos: cavidade maior (*major groove*) e cavidade menor (*minor groove*).

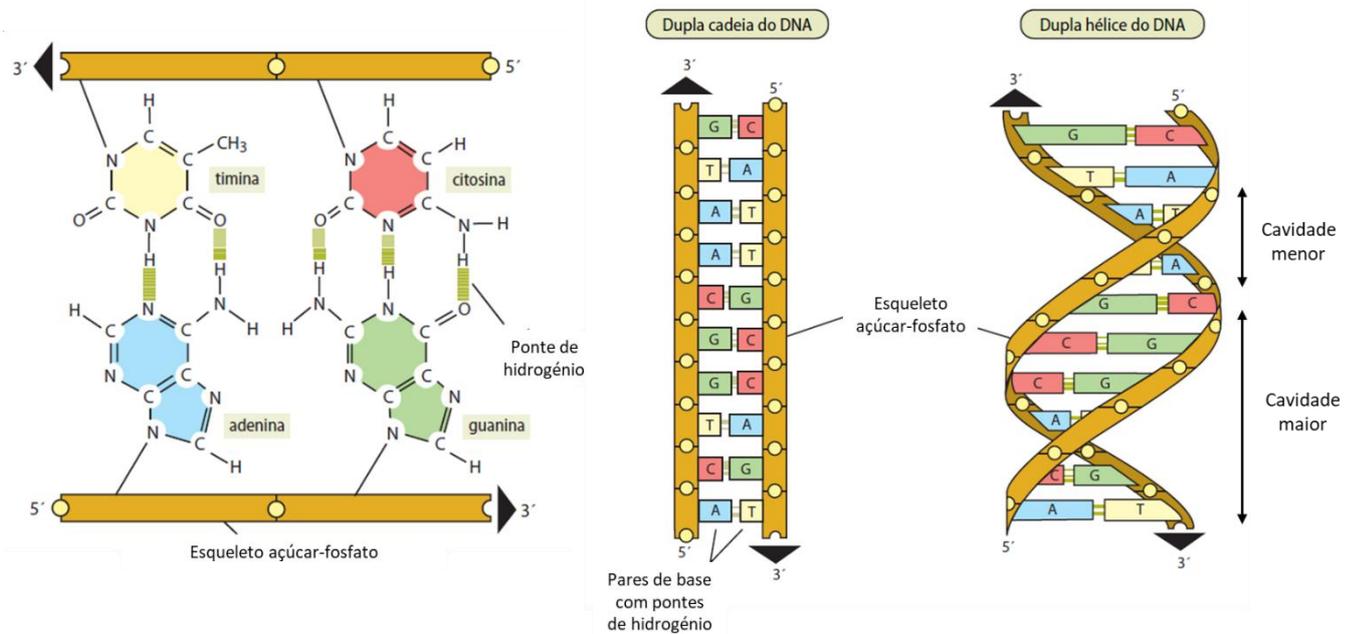


Figura 4 – Modelo para estrutura da molécula de DNA. [Adaptado de Ribeiro, M. (2009). *Genética molecular*]

1.3. Estabilidade e formas alternativas da dupla hélice

O DNA é uma molécula quimicamente estável. A estabilidade da dupla hélice é mantida devido às ligações por pontes de hidrogénio entre os pares de bases, interações entre os pares de bases empilhadas por forças de Van-der Waals e interações hidrofóbicas com o meio envolvente. Uma vez que as cadeias são unidas por pontes de hidrogénio, ligações químicas fracas, é relativamente fácil separar as duas cadeias pela ação de vários agentes, entre eles a temperatura, que rompe as ligações entre as bases. O processo de separação das duas cadeias de DNA denomina-se **desnaturação** e a temperatura necessária (T_m) para se obter desnaturação é mais elevada num DNA com maior proporção de G/C (três ligações de hidrogénio) do que num com maior proporção de A/T (duas ligações de hidrogénio). As cadeias separadas podem voltar a juntar-se

através de uma descida gradual de temperatura, num processo chamado de **renaturação**, respeitando a complementaridade de bases.

A conformação da dupla hélice do DNA depende da concentração de sais e da natureza das moléculas que interagem com o material genético. A estrutura encontrada normalmente nas condições fisiológicas estáveis é chamada de DNA B, caracterizada pela repetição regular de *minor* e *major grooves* com enrolamento helicoidal para a direita. Contudo, são possíveis outras conformações de DNA: DNA A, que existe no estado desidratado, com 11-12 nucleótidos por volta e não tendo qualquer função biológica; e DNA Z, que apresenta uma estrutura em ziguezague com 12 nucleótidos por volta (Figura 5).

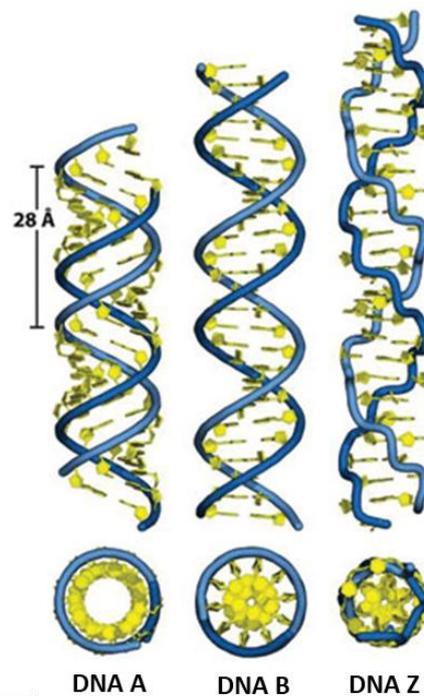


Figura 5 – Representação das três conformações possíveis de DNA: A, B e Z.
Adaptado de Nelson & Cox (2014).

1.4. A informação contida no DNA

A informação do DNA está contida na sequência em que as bases estão dispostas. O DNA contém a informação necessária para obtermos todas as proteínas que participam na totalidade das reações bioquímicas da célula. Além disso, o DNA contém ainda informação que, apesar de não conduzir à formação de proteínas, é necessária para a síntese de ácidos nucleicos.

Em 1860, Gregor Mendel introduziu o conceito de gene como unidades herdáveis discretas que dão origem a características físicas observáveis. O termo 'gene' foi utilizado pela primeira vez em 1910 pelo botânico e geneticista dinamarquês Wilhem Ludvig Johannsen, embora os estudos iniciais utilizassem o termo gene como uma entidade abstrata, uma vez que as bases químicas eram ainda desconhecidas. Só mais tarde, com a elucidação do modelo de DNA, a descoberta dos mRNA e do código genético, o gene passou a ser descrito como um fragmento de DNA que codificava para uma proteína.

Embora as definições possam sofrer pequenas variações, atualmente, um gene é definido como um segmento da molécula de DNA responsável por determinadas características herdadas geneticamente. Corresponde à unidade fundamental, física e funcional da hereditariedade. Nos organismos eucariotas, a sequência do gene que codifica para uma determinada proteína não se encontra organizada de forma contínua. As sequências codificantes que são transcritas no mRNA e, posteriormente, traduzidas numa proteína, denominam-se **exões**, sendo as mesmas separadas por **intrões**, regiões não codificantes que são removidas durante o processo de maturação do RNA. A região do gene que compreende a sequência de exões e intrões é denominada de **Open Reading Frame (ORF)**. Uma ORF tem início com o codão **START** (triplete ATG que codifica para a metionina) localizado no 1º exão e termina com um codão **STOP** (TAA, TAG, TGA – UAA, UAG, UGA). Os codões START e STOP marcam o início e fim da tradução, respetivamente. A montante da região codificante, ou seja, a 5', o gene contém o **promotor**, que regula a expressão desse gene e define como, quando e quantas vezes o mesmo se vai expressar; e a jusante, ou seja a 3', o gene apresenta o **terminador** que indica o final da transcrição. Em ambas as extremidades, um gene apresenta ainda

sequências reguladoras - **5' e 3'- UTR** (*untranslated regions*) - uma região que é transcrita, mas não é traduzida (Figura 6).

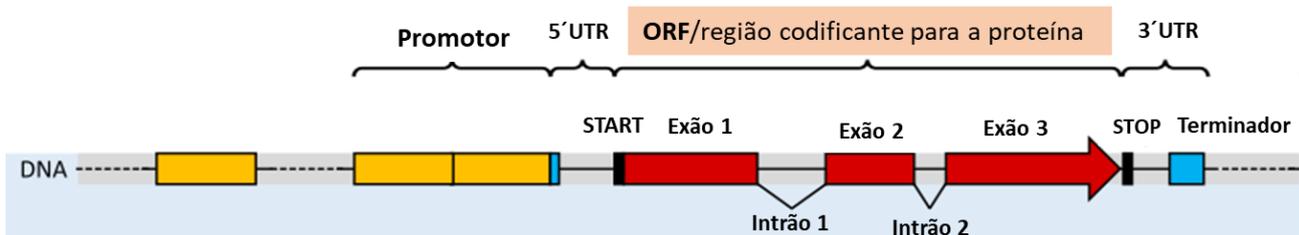


Figura 6 – Estrutura de um gene. Adaptado de [efaidnbmnnnibpcajpcglclefindmkaj/https://edisciplinas.usp.br/pluginfile.php/6579346/mod_resource/content/1/Gene_Est_Fun_Parte_2_2021v2.pdf](https://edisciplinas.usp.br/pluginfile.php/6579346/mod_resource/content/1/Gene_Est_Fun_Parte_2_2021v2.pdf)

1.5. Complexidade dos organismos e tamanho do genoma

O **genoma** diz respeito à totalidade do DNA de um organismo. Trata-se do conjunto do material genético, incluindo o genoma nuclear, mitocondrial e plastidial (no caso de células vegetais).

Por norma, o tamanho do DNA é referido em termos de pares de bases (pb), sendo a unidade comum 1kb (1000 pb) ou 1 Mb (1000000 pb). Uma molécula de 1 kb de DNA da forma B tem 340 nm de comprimento. No entanto, genomas de diferentes organismos apresentam moléculas de DNA com diferentes dimensões. O genoma de *E.coli*, por exemplo, tem aproximadamente 4×10^6 pb, enquanto que o genoma do homem apresenta aproximadamente 3 bilhões pb. Entre as plantas, o maior genoma é o da *Paris japonica*, natural do Japão, que tem 149 bilhões de pares de bases e é 50 vezes maior do que o humano.

Sabias que...

O genoma do milho tem mais genes do que o de um ser humano. O milho tem 32 mil genes e 2,3 mil milhões de pares de bases na cadeia de DNA, enquanto o homem tem entre 20-25 mil genes e 2,9 mil milhões de pares de bases. Relativamente ao número de cromossomas, o milho tem 10, enquanto que os seres humanos têm 23.

Os resultados foram publicados em 2009 na revista *Science*.



Science
The B73 Maize Genome: Complexity, Diversity, and Dynamics
Patrick S. Schnable *et al.*
Science 326, 1112 (2009).
DOI: 10.1126/science.1178534

Há 50 anos, pensava-se que organismos mais complexos tinham maior quantidade de DNA no genoma, visto que, segundo o pensamento da época, quanto mais complexa fosse uma espécie, mais DNA era necessário para armazenar toda a informação que seria necessária para originar um fenótipo mais complexo. No entanto, os avanços na ciência mostraram que essa correlação positiva entre o tamanho do genoma e a complexidade da espécie não é correta. Ao estudar a complexidade dos genomas em diferentes espécies, a comunidade científica descobriu que grande parte desse DNA não era traduzido, ou seja, não tinha informação para sintetizar uma proteína. Nesse sentido, percebeu-se que genomas muito grandes possuíam uma enorme quantidade de sequências repetitivas no DNA, as quais não codificam para proteínas. Atualmente sabe-se que essas sequências não são DNA lixo, pois podem desempenhar outras funções ao nível da regulação da expressão gênica ou da organização dos cromossomas na célula. Essas sequências, características de todos os genomas, podem ter diversas origens. O genoma humano, por exemplo, incorporou ao longo do tempo DNA viral de infecções antigas.

1.6. Organização do DNA no núcleo celular

Nos seres eucariotas, a maior parte do DNA está compartimentado no núcleo da célula, organizado em estruturas denominadas **cromossomas**. Nos cromossomas, o DNA encontra-se associado a proteínas básicas, as histonas, dando origem à cromatina.

Durante o ciclo celular, a cromatina é organizada no núcleo em diferentes níveis de condensação, atingindo o máximo de condensação na divisão celular (numa fase denominada metafase), onde os cromossomas podem ser visualizados ao microscópio. A cromatina mais condensada denomina-se **heterocromatina**, na qual os genes estão inativos e não ocorre transcrição. Por outro lado, a cromatina num estado menos condensado denomina-se **eucromatina**, na qual os genes se encontram ativos, ocorrendo a transcrição.

O nível mais simples de organização da cromatina consiste no enrolamento de pequenas regiões do DNA (147 pb de comprimento) em torno de um complexo proteico formado por oito histonas (duas moléculas de cada uma das histonas H2A, H2B,

H3 e H4). Este octâmero associado ao DNA, encontra-se ainda associado à histona H1. Esta estrutura é designada por **nucleossoma** (1º nível de condensação), a unidade fundamental da estrutura da cromatina. A cadeia de nucleossomas sofre um grau de compactação maior, originando uma fibra mais curta e mais larga, designada por solenóide ou zigzag (2º nível de condensação). Por sua vez, o solenóide associa-se a um conjunto de proteínas formando uma estrutura em Alças (3º nível de condensação). O nível mais elevado de condensação corresponde a um cromossoma em metafase (fase da mitose onde os cromossomas apresentam o mais elevado grau de condensação) (Figura 7).

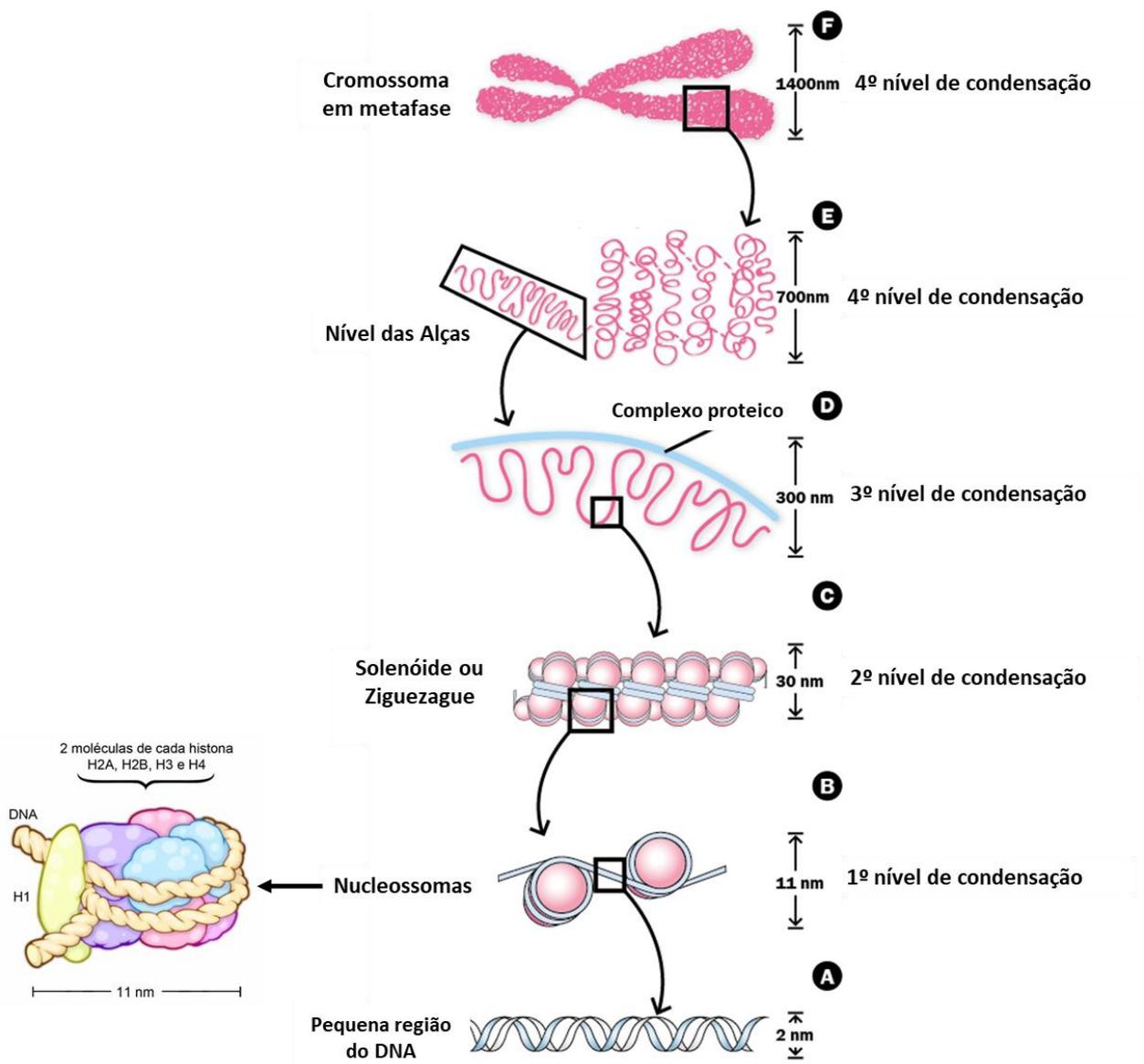


Figura 7- Diferentes níveis de compactação da cromatina.

Adaptado de Lodish *et al.*, 2005.

Apesar do DNA se encontrar maioritariamente no núcleo da célula, outros organitos celulares contêm o seu próprio material genético, como é o caso das mitocôndrias e cloroplastos (células vegetais). Contrariamente ao genoma nuclear, que é linear, o DNA das mitocôndrias e cloroplastos apresenta as formas linear ou circular. Em ambos os casos apresentam genes necessários à realização dos processos metabólicos que decorrem nos organitos, como a respiração celular e fotossíntese.

O DNA mitocondrial (mtDNA) é semelhante ao nuclear, constituído por dupla cadeia. Embora seja comum dizer-se que o mtDNA é circular, nas plantas ele pode assumir formas complexas e dinâmicas, incluindo moléculas circulares, lineares ou formas ramificadas. Outra característica do mtDNA é o facto de não ter histonas, estando associado a outras proteínas nucleares numa estrutura denominada 'nucleóide'. Atualmente, sabe-se que os tipos de proteínas são variáveis dependendo da espécie. Embora consiga codificar muitas proteínas, o mtDNA não consegue produzir um número suficiente para funcionar de forma independente, sendo necessário importar proteínas do núcleo. Dada a sua proximidade com a cadeia respiratória, onde são produzidas espécies reativas de oxigénio, e o facto de não possuir a maquinaria enzimática de correção de erros aquando da replicação, o mtDNA é mais suscetível a acumular mutações do que o DNA nuclear.

Sabias que...

O mtDNA de células vegetais é muito maior que o de células animais.

O mtDNA humano tem um tamanho de, aproximadamente, 17kbp. Comparativamente, o menor mtDNA vegetal pertence à espécie *Marchantia polymorpha* e tem 186,6kbp.



O genoma cloroplastidial pode ser circular ou linear e é constituído por 120-180 kbp na maioria das plantas superiores. É constituído por genes relacionados com processos de obtenção de energia (ATP) pela fotossíntese, codificando para proteínas

associadas ao metabolismo fotossintético. Tal como nas mitocôndrias, parte das proteínas do cloroplasto são codificadas no núcleo, sendo transportadas após a tradução para o interior do cloroplasto.

1.7. Estrutura dos cromossomas

No núcleo dos seres eucariotas o DNA está organizado em **cromossomas**, estruturas celulares compostas por DNA e proteínas, dando origem à cromatina. Tal como foi referido anteriormente, os cromossomas podem ter diferentes níveis de condensação dependendo da fase do ciclo celular em que a célula se encontra.

O conjunto de cromossomas de uma determinada espécie denomina-se **cariótipo**, apresentando como características o número, tamanho e posição do centrómero nos cromossomas.

O cromossoma condensado apresenta uma região central denominada **centrómero**, uma constrição primária de posição variável que divide o cromossoma em dois e tem um papel fundamental no movimento dos cromossomas durante a mitose e meiose. O centrómero permite dividir o cromossoma em braços, os quais podem ter tamanhos diferentes dependendo da posição do mesmo. No centrómero localiza-se ainda o **cinetocoro**, uma estrutura formada por proteínas onde os microtúbulos do fuso acromático se ligam durante a divisão celular. Na extremidade dos cromossomas existem os **telómeros**, complexos proteicos especializados que desempenham funções importantes na proteção e estabilidade do cromossoma, estando presentes em todos os cromossomas dos eucariotas. Na maior parte dos organismos, os telómeros são compostos por longas repetições de sequências de DNA simples (5'-TTAGGG-3'), designadas repetições terminais (Figura 8).

Em cada cariótipo há pelo menos um par de cromossomas com uma constrição secundária em posições específicas denominada **Região Organizadora do Nucléolo (NOR)** (Figura 8), sendo que é à sua volta que o nucléolo se reconstitui na telófase, contendo esta região várias cópias dos genes que codificam o rRNA. Após a constrição NOR, os cromossomas podem apresentar **Satélites**, zonas de dimensão variável que constituem uma região bem definida no extremo do braço do cromossoma, contendo assim o respetivo telómero. Os satélites também estão associados à formação

do nucléolo (formado por rRNA e proteínas) no final da divisão celular, durante a telófase. Os satélites contêm múltiplas cópias repetidas dos genes que codificam o rRNA.

Antes da fase S da interfase (ciclo celular), cada cromossoma é constituído apenas por um cromatídeo. Aquando da fase S, que corresponde à replicação do DNA, o número de cromatídeos duplica e cada cromossoma é constituído por dois cromatídeos (cromatídeos irmãos) unidos pelo centrómero.

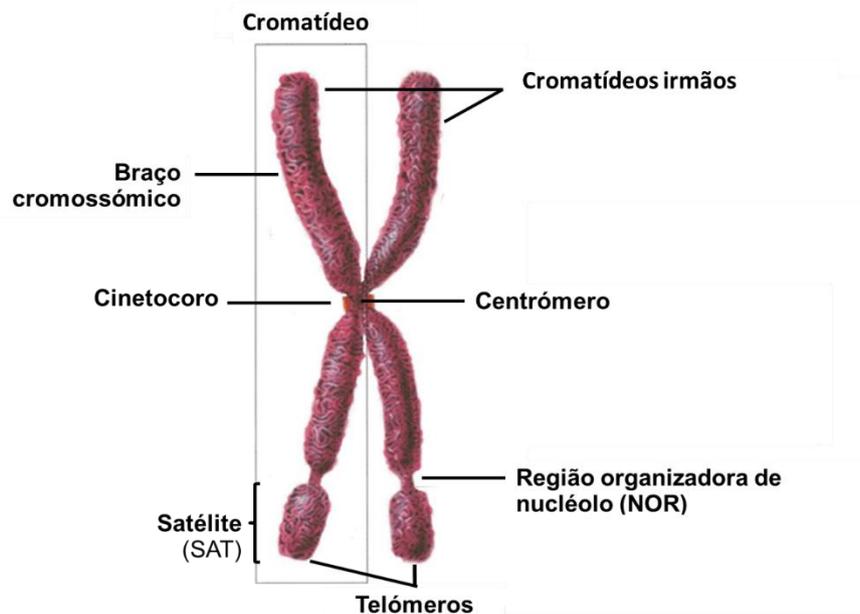


Figura 8 – Estrutura de um cromossoma.

Adaptado de <https://www.colegioweb.com.br/nucleo/a-forma-dos-cromossomos.html>

Na maioria dos organismos multicelulares, cada célula possui dois conjuntos de cromossomas equivalentes, ou seja, em cada célula os cromossomas são iguais dois a dois no que se refere ao tamanho e à forma. Os dois representantes de cada par de cromossomas denominam-se **cromossomas homólogos**. Células que possuem os dois pares de cromossomas homólogos são **células diploides (2n)**, enquanto células que possuem apenas um cromossoma do par de homólogos são denominadas **haploides (n)**. No caso da espécie humana, por exemplo, todas as células somáticas são diploides e possuem dois conjuntos de 23 cromossomas (22 pares de autossomas + 1 par de cromossomas sexuais), perfazendo um total de 46. Já as células reprodutivas (gâmetas), são células haploides e possuem apenas um conjunto de cromossomas, neste caso 23.

Os cromossomas de cada par de homólogos, além de serem morfológicamente iguais, possuem os mesmos genes, distribuídos na mesma sequência ao longo do seu comprimento. Os genes que ocupam a mesma posição (o mesmo *locus*) em cromossomos homólogos e estão envolvidos na determinação de uma mesma característica denominam-se genes alelos (Figura 9). Assim, para a mesma característica, os alelos são formas alternativas do mesmo gene que ocupam o mesmo *locus* em cromossomas homólogos. Os alelos podem possuir a mesma informação ou informação diferente para a mesma característica.

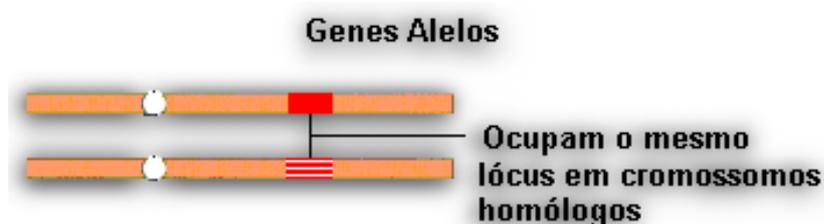


Figura 9- Genes alelos em cromossomas homólogos.

Adaptado de <https://www.infoescola.com/genetica/alelos/>

Quanto à posição do centrómero, os cromossomas podem ser classificados em **metacêntricos** (Figura 10A), na medida em que o centrómero se apresenta ao centro, sendo os braços de comprimentos aproximadamente iguais; **submetacêntricos** (Figura 10B), nos quais o centrómero está afastado do centro e os braços apresentam comprimentos nitidamente diferentes (p- braço mais curto, q- braço mais longo); **acrocêntricos** (Figura 10C), onde o centrómero está localizado bastante próximo de uma extremidade; **telocêntricos** (Figura 10D), nos quais o centrómero está localizado numa das extremidades. Neste último caso, quando o centrómero ocupa uma das extremidades não havendo mais genes acima dele, o cromossoma denomina-se **acêntrico**.

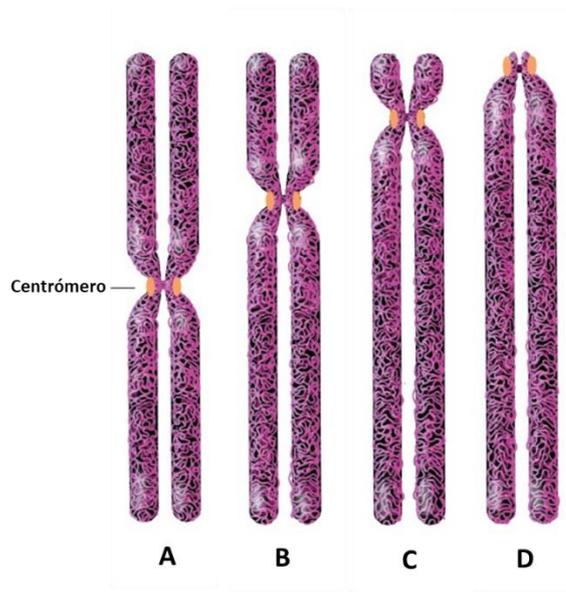


Figura 10 – Diferentes posições do centrômero no cromossoma.
Adaptado de <https://thinkbio.wordpress.com/2012/01/02/o-nucleo-celular/>

A organização estrutural dos cromossomas foi posta em evidência através de técnicas citológicas que permitem corar os cromossomas mitóticos. O bandeamento de cromossomas consiste na utilização de diferentes corantes que conduzem ao aparecimento de múltiplas bandas coradas de acordo com a condensação dos cromossomas. O padrão de bandas, específico de uma determinada espécie em termos de posição e tamanho, é evidenciado devido ao diferente grau de condensação da cromatina e à presença de cromómeros, regiões em que a cromatina é mais irregular, com engrossamentos com aspeto de granulações. Um dos métodos de coloração mais comum é o Bandeamento G, no qual o corante Giemsa, específico para os grupos fosfato do DNA, cora diferencialmente as regiões ricas em Guanina e Citosina da eucromatina (bandas claras) e regiões ricas em Adenina e Timina da heterocromatina (bandas escuras). Esta e outras técnicas de bandeamento são muito importantes na identificação e caracterização dos cromossomas, sendo utilizadas para emparelhar os cromossomas homólogos e constituir o cariótipo de um organismo, identificar modificações nos cromossomas e detetar doenças associadas a alterações cromossómicas.

Capítulo 2 – Ciclo celular

A perpetuação das espécies ocorre através da reprodução. Em organismos mais simples a reprodução pode ser uma simples duplicação, ou envolver mecanismos mais complexos de reprodução sexuada, como ocorre nos animais e nas plantas superiores. Em todos os casos, a reprodução envolve a transmissão exata da informação genética contida no DNA. Para que ocorra a transferência da informação genética de células-mãe para células-filhas, é necessário um processo de replicação fiel do DNA.

O Ciclo celular corresponde aos processos que ocorrem na célula desde o momento em que a mesma surge até ao seu processo de divisão celular, o qual dará origem a duas células. Durante esse período, a célula passa por diversos processos, como crescimento celular, multiplicação de seu material genético e divisão celular. O tempo de duração do ciclo celular varia entre os diferentes tipos de célula e é dividido em duas fases: interfase e mitose/meiose (Figura 11). A replicação do DNA acontece num momento específico do ciclo celular denominado fase S, pertencente à interfase.

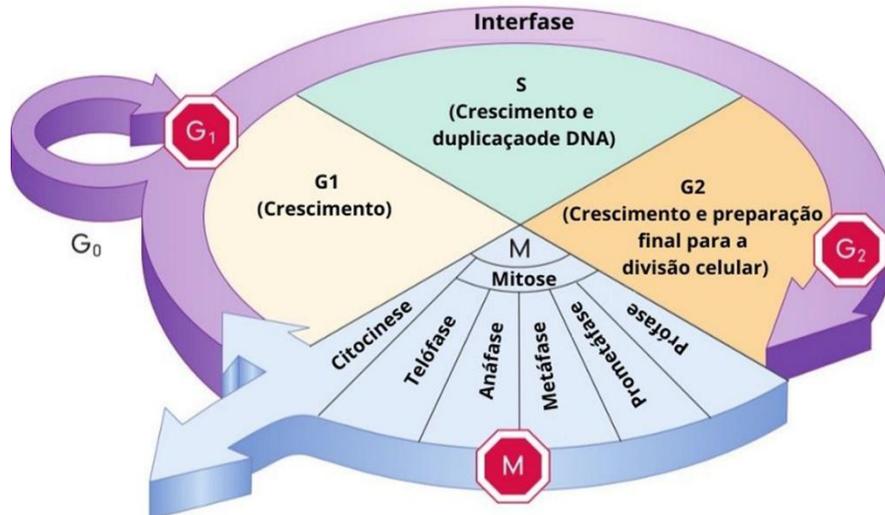


Figura 11 – Representação das diferentes fases do ciclo celular: Interfase (G1, S e G2) e Mitose.

<https://www.biologianet.com/biologia-celular/ciclo-celular.htm>

A **interfase** é a fase mais longa do ciclo celular, sendo um período de crescimento, onde a célula tem uma intensa atividade metabólica. A interfase encontra-se dividida em três períodos: G1, S e G2. O período **G1** (do inglês *gap*, intervalo) é o que precede a

duplicação do DNA, onde se verifica um crescimento celular significativo e síntese de proteínas necessárias para a fase de duplicação do DNA. Neste período os cromossomas apresentam apenas um cromátídeo. O período **S** (do inglês *synthesis*, síntese) é aquele em que ocorre a duplicação do DNA. O período **G2**, o mais curto, é o que acontece após a duplicação do DNA (os cromossomas apresentam dois cromátídeos) e onde a célula se prepara para a divisão celular.

2.1. Período S – replicação do DNA nuclear

Antes do início de uma nova divisão celular, os organismos têm de duplicar fielmente o seu DNA. O processo, rápido e preciso, é levado a cabo por um complexo multienzimático (tabela II).

De uma forma geral, a replicação do DNA inicia-se pelo desenrolar da hélice e separação das cadeias por enzimas específicas. A separação das cadeias dá origem a uma estrutura especializada de DNA, a bolha ou garfo de replicação (do inglês *replication fork*), que avança ao longo da dupla hélice e onde ocorre a adição de nucleótidos. O processo de replicação do DNA encontra-se descrito em pormenor nas seguintes etapas e representado na figura 12:

1. A replicação do DNA inicia-se pelo desenrolar da dupla hélice, executado pelas enzimas **topoisomerasas**, e separação da dupla cadeia pelas **helicases**. As helicases reconhecem as origens de replicação (*ori*), dispersas por todo o genoma, nas quais abrem e distorcem a molécula de DNA, quebrando as ligações de hidrogénio que mantêm as duas cadeias unidas, dando origem à bolha de replicação. As helicases movem-se sobre a molécula de DNA, separando a cadeia.

Com esta estrutura estabelecida entra em ação um complexo enzimático – REPLISSOMA - que garante a síntese de duas novas cadeias filhas. Para evitar que as cadeias se liguem novamente, as chamadas **proteínas SSB** (do inglês *single stranded DNA-binding proteins*) ligam-se ao DNA de cadeia simples, estabilizando-o e deixando as bases livres para a associação dos nucleotídeos.

2. A enzima **primase** sintetiza oligonucleótidos de vida curta (**primer de RNA**) que se ligam a cada cadeia - nenhuma DNA polimerase conhecida é capaz de iniciar a síntese de polinucleótidos.
3. Depois de o *primer* estar no lugar, a **DNA polimerase** envolve cada cadeia, e liga novos nucleótidos complementares às bases livres.
4. A DNA polimerase apenas consegue adicionar nucleótidos à posição 3' da cadeia original; uma das cadeias é lida na direção 3' – 5', e a nova cadeia é formada no sentido 5' – 3' (cadeia principal ou **leading strand**). A outra cadeia (cadeia secundária ou **lagging strand**) é antiparalela, e está na direção 5' – 3', no entanto a DNA polimerase não consegue adicionar nucleótidos à posição 5'; assim, os primers de RNA são continuamente adicionados e a síntese de DNA ocorrem em fragmentos (**fragmentos de Okazaki**), na direção 5' – 3', tal como a *leading strand*.
5. A enzima **DNA ligase** liga os fragmentos de Okazaki para formar uma cadeia única.

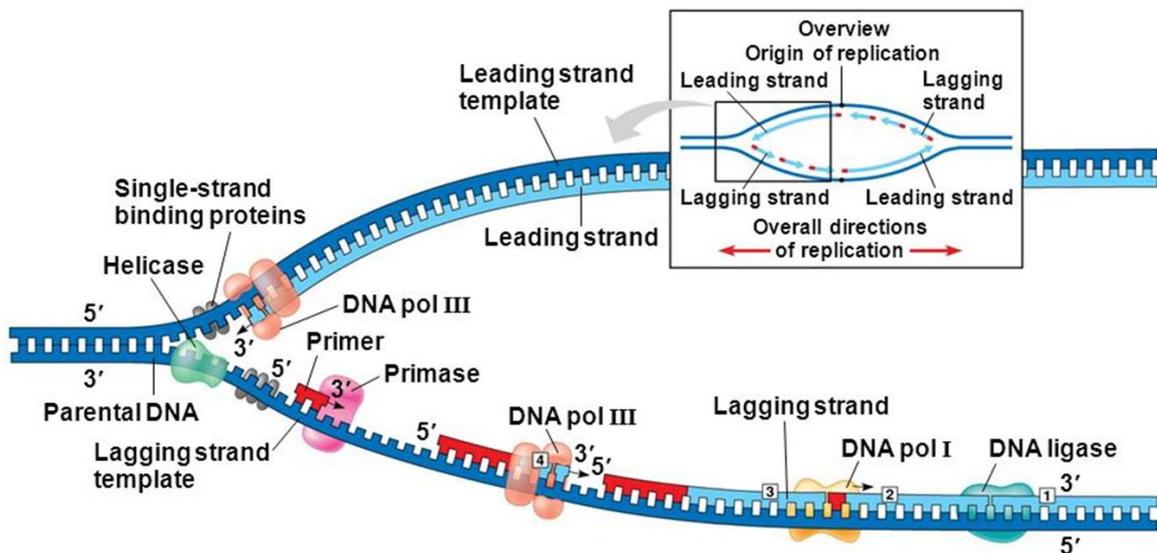


Figura 12 – Forquilha de replicação com complexo enzimático envolvido na replicação do DNA, com representação das cadeias *leading* e *lagging*.

Adaptado de <https://edisciplinas.usp.br/pluginfile.php>

A replicação do DNA dura até que as sucessivas frentes de síntese de DNA se reúnam. Nesta fase os cromossomas apresentam dois cromatídeos e transita-se para a fase G2 da interfase.

A replicação do DNA apresenta como características o facto de ser **semiconservativa, bidirecional e semidescontínua**.

A replicação é **semi-conservativa** uma vez que cada uma das cadeias simples originais fica emparelhada com uma nova cadeia, servindo de molde (do inglês *template*).

A partir de uma única origem de replicação, a replicação do DNA ocorre de forma **bidirecional**, pois são formados dois garfos de replicação a partir de cada uma das cadeias originais de DNA.

A enzima DNA polimerase catalisa a adição de nucleótidos à extremidade 3' de uma pequena cadeia polinucleotídica, ou seja, a nova cadeia cresce sempre a partir da extremidade 5' em direção à extremidade 3' pela adição de nucleótidos a 3' do *primer*, o qual se encontra emparelhado com a cadeia molde. Visto que as duas cadeias originais que constituem os moldes são antiparalelas, é necessário que a síntese aconteça de forma contínua na cadeia principal (*leading strand*), e de forma descontínua na cadeia secundária (*lagging strand*). Assim, a síntese da *leading strand* progride continuamente a partir de um *primer* na direção 5' → 3', correspondendo à direção do movimento do garfo de replicação crescente. Por outro lado, como a síntese da *lagging strand* também deve ocorrer na direção 5' → 3', a cópia da cadeia original tem de ocorrer na direção oposta à do crescimento do garfo. Este processo é conseguido pelo emparelhamento de *primers* com a cadeia original, sendo que, a partir de cada um destes *primers* ocorre a extensão de fragmentos da nova cadeia na direção 5' → 3' em segmentos descontínuos, denominados fragmentos de Okasaki (Figura 13).

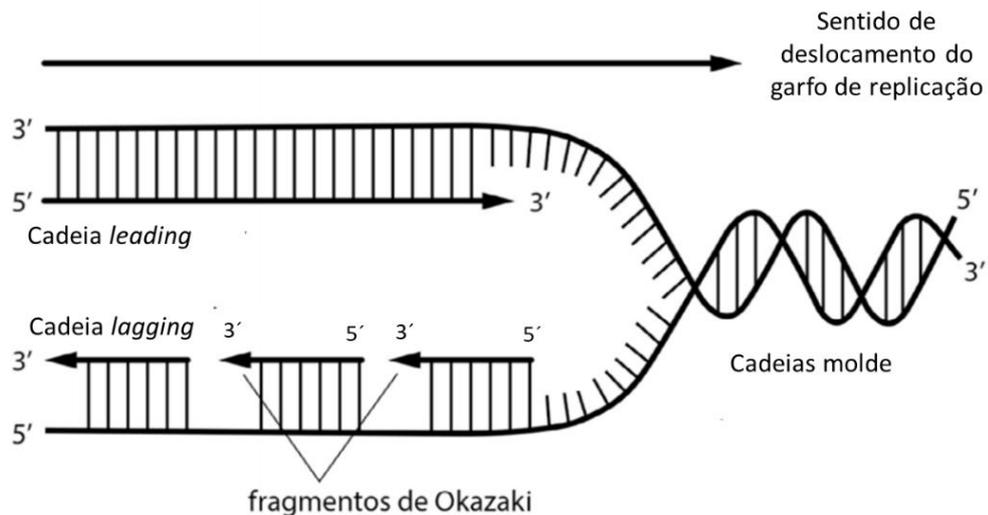


Figura 13- Estrutura do garfo de replicação com representação da síntese da cadeia *leading* e *lagging*.

Adaptado de <https://nadispersa.medium.com/replica%C3%A7%C3%A3o-do-dna-a6f7b8a8f20f>

Tabela II – Enzimas envolvidas no processo de replicação do DNA em eucariotas.

Enzima	Função
Helicases	Desenrolam a dupla hélice e separam as cadeias de DNA de modo a expor as bases para emparelhamento.
Proteínas SSB	Ligam-se ao DNA de cadeia simples, estabilizando-o.
Primase	Responsável pela síntese do <i>primer</i> de RNA (sequência de ~10 nt), complementar à cadeia molde, de onde a enzima iniciará a síntese da nova cadeia. Contrariamente à DNA polimerase não necessita ter previamente um grupo 3'OH para iniciar a síntese.
DNA Polimerase	Sintetizam uma nova cadeia de DNA, catalisando a adição de nucleótidos à extremidade 3'OH livre da nova cadeia em crescimento, de acordo com a sequência complementar. Apresenta atividade de exonuclease, permitindo-lhe corrigir erros decorridos durante a replicação.
DNA ligase	Promove a ligação fosfodiéster entre a extremidade 3' de um fragmento de <i>Okazaki</i> e a extremidade 5' do fragmento subsequente.
Topoisomerases	Exercem um papel importante auxiliando a atividade das helicases, pois desenrola a cadeia, diminuindo a tensão à medida que as helicases avançam, facilitando assim o trabalho dessas enzimas na abertura da dupla cadeia.

2.2. Regulação do ciclo celular – pontos de verificação

O ciclo celular apresenta mecanismos de controle que regulam os seus processos, como a síntese de proteínas ou a divisão celular. Esses mecanismos são de extrema importância para evitar a proliferação descontrolada das células. Na maioria das células eucarióticas, esses mecanismos atuam nos chamados **pontos de verificação** (Figura 11). Existem três pontos principais:

- **Ponto de verificação G1:** ocorre no final da fase G1, antes da fase S, e verifica se a célula teve um crescimento suficiente, se o DNA se mantém íntegro e se foram sintetizadas as proteínas adequadas/suficientes para a replicação do DNA. Quando as condições não são adequadas, este ponto de verificação impede a continuação do ciclo celular.

- **Ponto de verificação G2:** ocorre no final da fase G2 e verifica a ocorrência de erros que possam ter ocorrido durante a replicação do DNA. Caso todas as condições sejam as adequadas, a célula avança para a fase mitótica.

- **Ponto de verificação M:** ocorre na transição entre a metáfase e anáfase, verificando a ligação dos cromossomos aos microtúbulos do fuso mitótico.

Em cada ponto de verificação, o ciclo é regulado pelos chamados “reguladores do ciclo celular”, onde se destacam as proteínas ciclinas e enzimas quinases dependentes de ciclinas (Cdks). As proteínas ciclinas podem ser de quatro tipos, consoante o ponto de verificação em que participam: ciclina G1, ciclina G1/S, ciclina G2 e ciclina M. Para fazer com que o ciclo celular avance, uma ciclina deve ativar ou desativar proteínas alvo dentro da célula. As ciclinas desencadeiam os eventos do ciclo celular associando-se a uma família de enzimas chamada **Cdks**, ou seja, enzimas quinases com capacidade para fosforilar (adição de um grupo fosfato) proteínas alvo específicas, tornando-as ativas e permitindo o avanço no ciclo celular. A figura 14 representa, esquematicamente, a ativação de enzimas Cdks pelas proteínas ciclinas em dois pontos diferentes do ciclo celular: antes da entrada na fase S (figura 14A) e antes da entrada na fase de mitose (figura 14B).

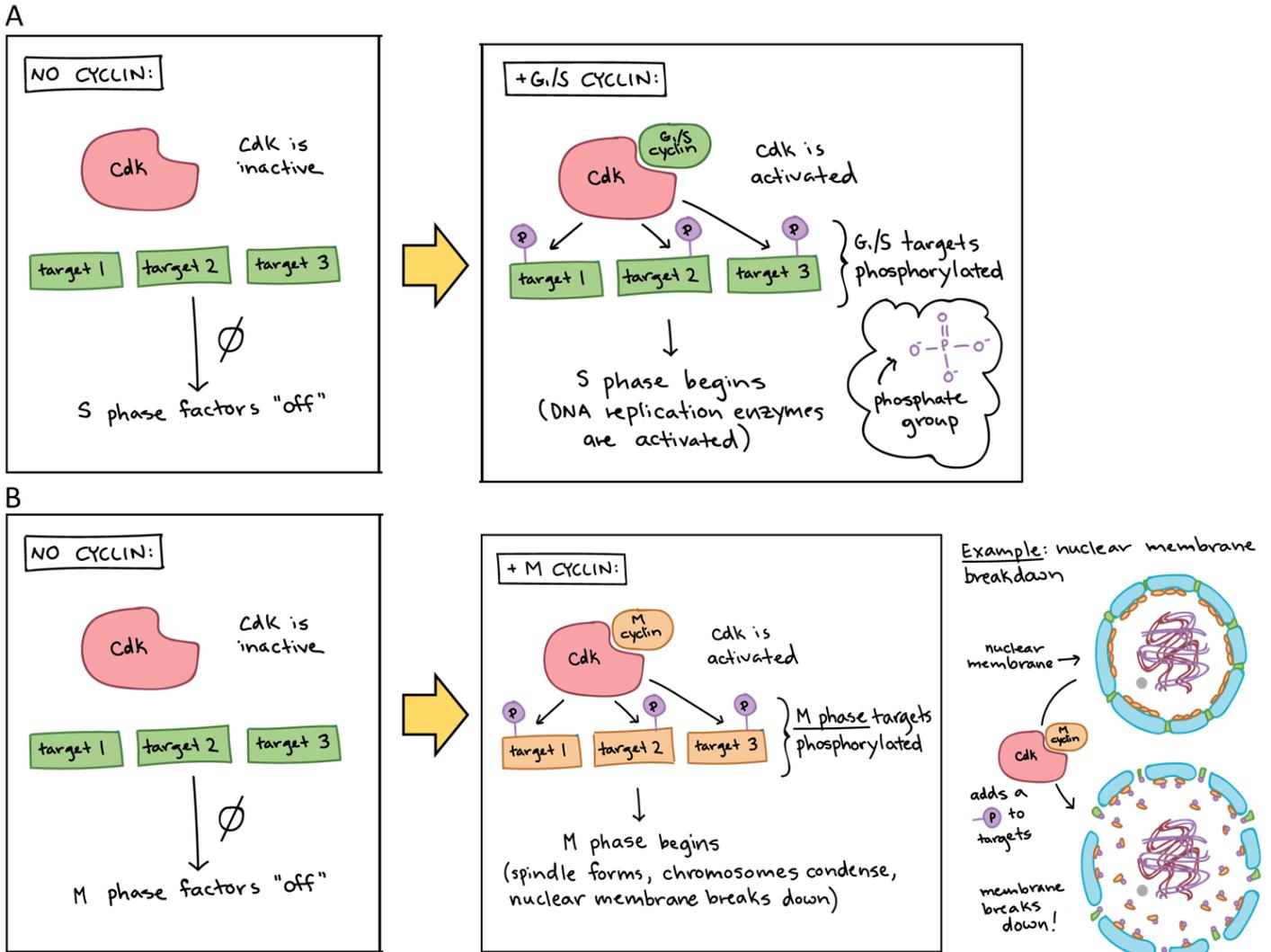


Figura 14 – Representação esquemática da interação entre proteínas ciclinas e enzimas Cdk, com consequente fosforilação de proteínas alvo, antes da fase S (A) e antes da fase mitótica (B).

Adaptado de <https://pt.khanacademy.org/science/biology/cellular-molecular-biology/stem-cells-and-cancer/a/cell-cycle-regulators>

Durante o processo de replicação do DNA, mecanismos de detecção e reparação de erros asseguram a (quase) perfeita fidelidade da replicação do DNA. Alguns erros durante a replicação são reparados imediatamente (*proofreading*) devido à atividade de exonuclease da DNA polimerase (removendo, por exemplo, um nucleótido incorretamente adicionado), enquanto outros são reparados depois da replicação (*mismatch repair, MMR*).

2.3. Divisão celular – mitose e meiose

Após o período de crescimento celular e preparação para a divisão celular, que ocorre durante a interfase, a célula divide-se numa fase denominada **mitose**. A mitose é fundamental para o desenvolvimento, crescimento e reposição de células perdidas. É um processo fisiológico que permite que uma célula somática (diploide, $2n$), ou seja, com o número total de cromossomas que caracteriza a espécie, origine duas células-filhas com informação genética idêntica. É uma das etapas mais curtas do ciclo celular, tendo uma duração aproximada de 30 minutos.

A mitose é dividida em quatro fases: **profase**, **metafase**, **anafase** e **telofase**. As diferentes fases encontram-se esquematizadas na figura 15.

Profase – Corresponde à primeira fase da mitose, ocorrendo a espiralização e condensação dos cromossomas, que se encontravam no seu máximo grau de despiralização na interfase. À medida que a profase prossegue, os cromossomas, ainda que formando um emaranhado onde é impossível individualizar cada um, vão-se tornando mais evidentes, por serem cada vez mais curtos e grossos. Em muitos casos, começa a ser possível ver que cada cromossoma se diferencia longitudinalmente em dois filamentos, ou seja, dois cromatídeos (resultantes da replicação do DNA na interfase). Simultaneamente, o nucléolo ou nucléolos, que eram bem evidentes no núcleo interfásico e início da profase, vão-se dissipando e acabam por desaparecer.

Nas células animais, e nas de algumas plantas inferiores, existem, nos pólos do fuso acromático, os centríolos, estruturas constituídas por dois cilindros de nove tripletos microtubulares cada, dispostos em posição ortogonal. Os centríolos têm um ciclo paralelo ao da mitose, replicando-se na interfase. Durante a prófase, cada um dos dois novos centríolos migra para posições opostas da célula, formando os pólos do fuso acromático ou fuso mitótico.

No final da profase ocorre a rutura do invólucro nuclear e a mistura entre o citoplasma e o nucleoplasma.

Os cromossomas (com dois cromatídeos irmãos) ligam-se pelos seus centrómeros (cinetócoro) às fibras do fuso acromático e movem-se de forma a disporem-se no plano designado por **placa equatorial**, em posição equidistante em relação aos

pólos do fuso, agora completo. O período que decorre desde a rutura do invólucro nuclear até à disposição dos cromossomas com os seus centrómeros na placa equatorial designa-se por **prometáfase**.

Metáfase – Corresponde à fase da mitose na qual os cromossomas se dispõem com os seus centrómeros na placa equatorial, encontrando-se no seu mais elevado grau de condensação. Os braços dos cromossomas não mostram qualquer orientação especial e dispõem-se, ou no plano da placa equatorial ou para um ou outro lado desta. Os cromossomas mostram bem evidente a sua diferenciação longitudinal em dois cromátídeos, podendo estes estar ainda enrolados um sobre o outro.

Anáfase – Quando a posição de equilíbrio é atingida na metáfase, inicia-se a separação dos centrómeros e o movimento de cada um dos cromátídeos (agora designados cromossomas filhos) para os pólos do fuso acromático. A anáfase é o período que decorre desde o início da separação dos centrómeros até ao momento em que os cromossomas filhos atingem os pólos do fuso acromático. É nesta fase que se efetua a divisão do material cromossómico em duas metades, de forma a distribuir por cada núcleo resultante da divisão exatamente um exemplar do código genético da célula original.

Telófase – Inicia-se quando os cromossomas filhos atingem os pólos do fuso e pode considerar-se como o processo inverso da profase. Os cromossomas despiralizam-se, deixando de ser visível a sua individualização. Reaparecem os nucleolos que tinham desaparecido na profase e volta a formar-se o invólucro nuclear. Os núcleos filhos retomam, assim, o seu aspeto da interfase.

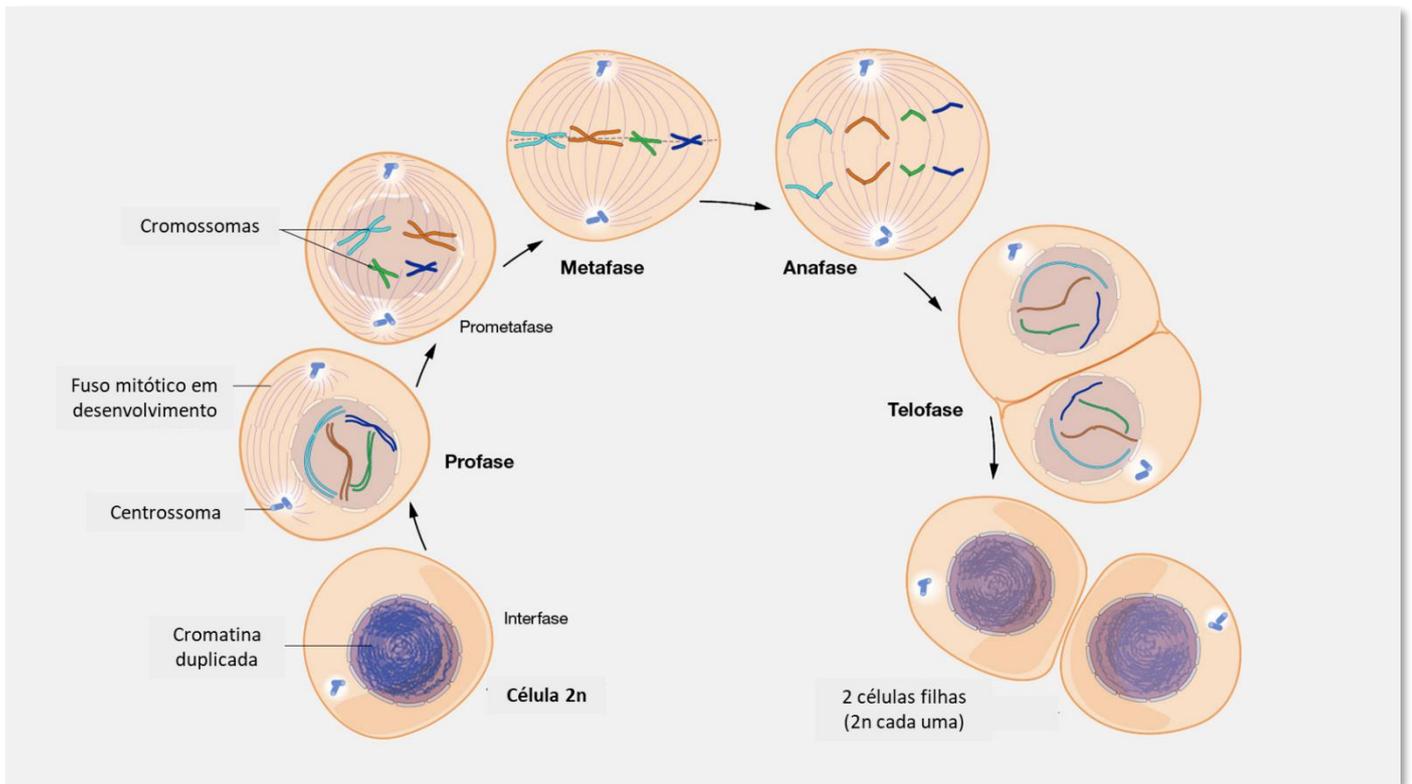


Figura 15 – Esquema representativo das diferentes fases da mitose.

Adaptado de <https://www.genome.gov/es/genetics-glossary/Mitosis>

A mitose é o processo pelo qual a célula divide o seu núcleo em dois (**cariocinese**), de forma que cada núcleo filho receba um código genético idêntico. Após a mitose, a célula divide-se em duas num processo denominado **citocinese**.

O processo de citocinese difere entre células animais e vegetais. Nas células animais, a citocinese consiste na clivagem da célula por um estrangulamento da zona equatorial (Figura 16A). A clivagem inicia-se nas células animais logo após a telófase. Na zona equatorial da célula começa a dar-se um estrangulamento, com o aparecimento do sulco de clivagem, que progride até à completa separação da célula em duas. Por seu lado, nas células vegetais, a citocinese consiste na formação de um septo que divide a célula em duas. Nas células vegetais, após ou durante a telófase, verifica-se o aparecimento, no plano onde estava a placa equatorial, de uma série de vesículas que coalescem, formando o fragmoplasto, que cresce centrifugamente (do centro da célula para a periferia), formando-se, assim, as membranas que dividirão a célula em duas (Figura 16B). Sobre cada uma dessas membranas deposita-se o material (pectina) para formar as respetivas paredes celulares.

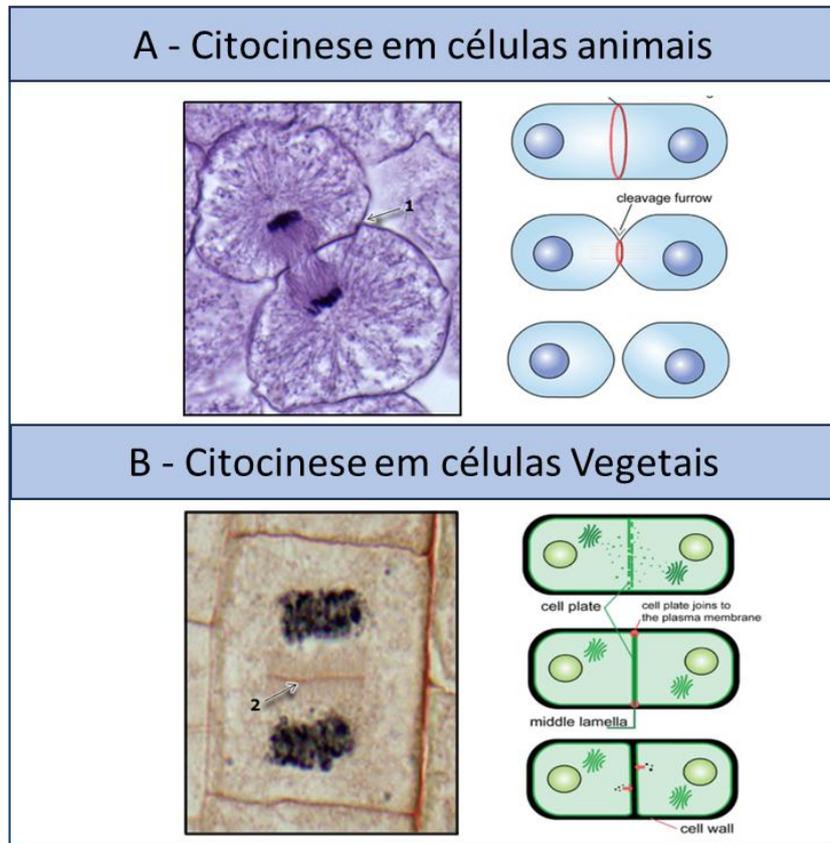


Figura 16 – Citocinese em células animais (A) com formação da zona de estrangulamento (representado na imagem por 1); citocinese em células vegetais (B) com formação do fragmoplasto (representado na imagem por 2).

Adaptado de <https://pt.slideshare.net/slideshow/31-knp/29023547>

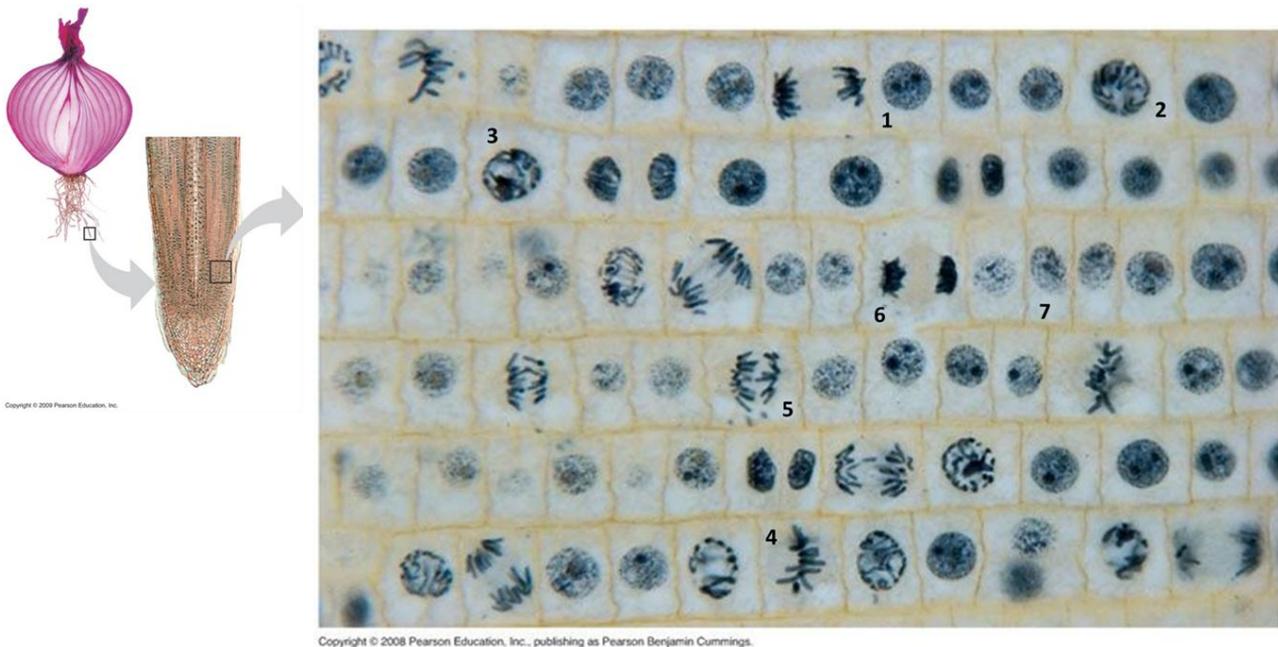


Figura 17 – Exemplo de diferentes fases do ciclo celular em células da extremidade da raiz da cebola: 1- interfase; 2,3- profase; 4- metáfase; 5- anafase; 6- telófase; 7- citocinese.

Adaptado de <https://bcphysics180.wordpress.com/2015/01/12/day-61-onion-root-tip-lab/>

A **meiose** é uma forma de divisão nuclear muito importante entre os organismos com reprodução sexuada, resultando na formação de células reprodutoras designadas gâmetas nos animais e esporos nos vegetais. No final do processo meiótico, formam-se quatro células-filhas, cada uma com metade do número de cromossomas (haploide) originalmente presentes na célula-mãe (diploide).

A meiose é o conjunto de **duas divisões celulares** consecutivas e complementares: **meiose I** – divisão reducional ou heterotípica, na qual se verifica a separação dos cromossomas homólogos, reduzindo para metade o número de cromossomas de cada célula ($2n \rightarrow n$); **meiose II** – divisão equacional ou homotípica, ocorrendo a separação dos cromatídeos irmãos, mantendo o número de cromossomas ($n \rightarrow n$).

A meiose I e II são divididas em quatro fases cada um, com os mesmos nomes das fases da mitose. As diferentes fases da meiose I e II encontram-se esquematizadas na figura 18.

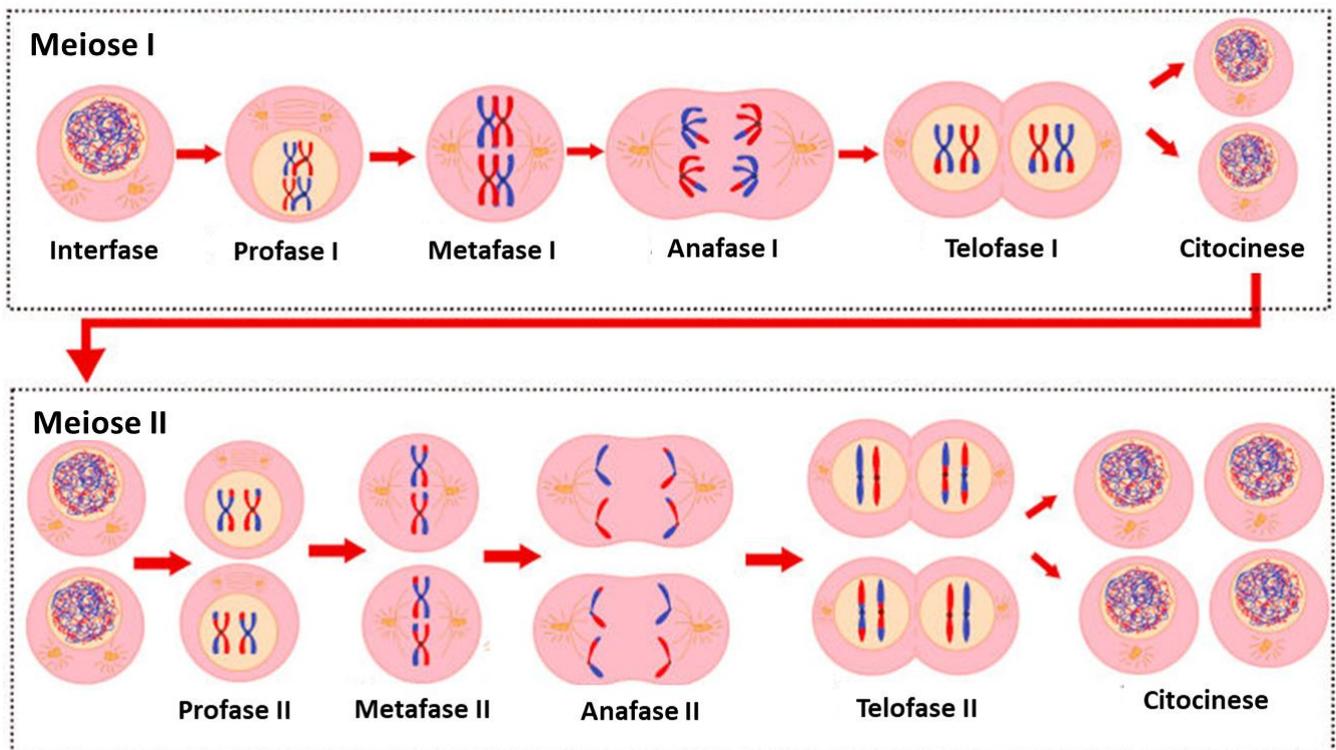


Figura 18 - Esquema representativo das diferentes fases da meiose I e II.

Adaptado de <https://pt.vecteezy.com/arte-vetorial/6065312-diagrama-de-meiose-divisao-celula-e-o-processo-celulas-passam-para-dividir-ilustracao-vetorial>.

A **profase I** é uma fase longa, ocorrendo depois da interfase. Devido à sua complexidade, a profase I divide-se nas seguintes fases: **Leptóteno, Zigóteno, Paquíteno, Diplóteno, Diacinese** (figura 19).

- a) **Leptóteno** - Início da condensação dos cromossomas, tornando-se visíveis ao microscópio.
- b) **Zigóteno** – À medida que condensam, os cromossomas homólogos colocam-se lado a lado, emparelhando perfeitamente ao longo do seu comprimento, formando sinapses.
- c) **Paquíteno** - Completam-se as sinapses e entre os cromossomas homólogos ocorre o fenómeno de *crossing-over*, processo definido como a troca de segmentos entre cromátídeos não irmãos de cromossomas homólogos, formando-se *tétradas cromatídicas* ou *bivalentes*. Os pontos onde os cromossomas se tocam e se realiza a troca de material genético denominam-se **pontos de quiasma**. O resultado de um evento de *crossing over* são dois cromossomas recombinantes com uma combinação génica diferente da encontrada nos progenitores, contribuindo assim para o **aumento da diversidade genética**.
- d) **Diplóteno** - Os cromossomas homólogos começam a separar-se, mas continuam ligados pelos pontos de quiasma.
- e) **Diacinese** – Última fase da prófase I. Os cromossomas tornam-se mais condensados, forçando os pontos de quiasma a deslocarem-se para os telómeros, desfazendo-se progressivamente. Ao movimento dos quiasmas para os telómeros chama-se **terminalização**. O invólucro nuclear desorganiza-se, os nucléolos desintegram-se e inicia-se a diferenciação do **fuso acromático**.

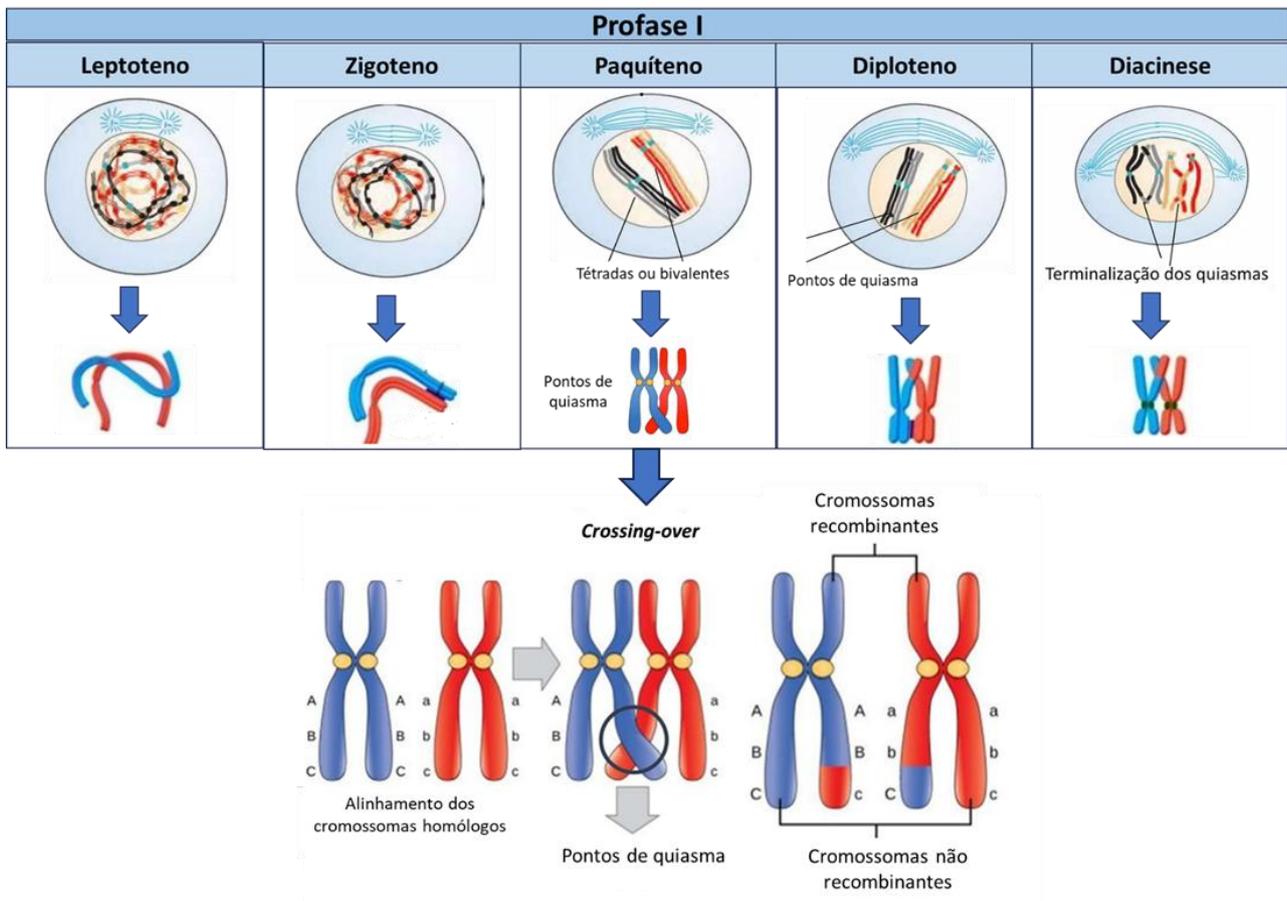


Figura 19 – Representação das diferentes fases da Profase I, com destaque para o fenómeno de *crossing-over* que ocorre na fase Paquíteno.

Adaptado de <https://thinkbio.wordpress.com/2012/01/04/ciclo-celular/>

Na **metáfase I** os cromossomas bivalentes ligam-se aos microtúbulos do fuso acromático e ficam alinhados na **placa equatorial**. Um microtúbulo de um dos pólos liga-se ao cinetocoro (conjunto de proteínas que existe no centrómero) de um cromossoma homólogo e o microtúbulo do outro pólo liga-se ao outro membro do par de homólogos.

A **anafase I** é marcada pela separação dos cromossomas homólogos (cada um constituído por dois cromatídeos), seguindo cada um para pólos opostos da célula, num processo denominado **segregação independente dos homólogos**. Os cromatídeos irmãos permanecem unidos pelo centrómero e migram juntos.

Na **telófase I**, os cromossomas, distribuídos em dois grupos, cada um em cada pólo da célula, descondensam-se, o invólucro nuclear reorganiza-se e os nucléolos

reaparecem. Surgem, assim, dois núcleos novos, cada um com metade do número de cromossomas presentes no núcleo original.

Após a primeira divisão meiótica, as duas células filhas separam-se num processo designado **citocinese**. As células em seguida avançam para a meiose II sem duplicar o seu DNA.

Na **meiose II**, as duas células resultantes da divisão I entram ao mesmo tempo na **profase II**. A profase II é um estado muito rápido, caracterizado pela desintegração do invólucro nuclear, pela condensação dos cromossomas e pelo início da diferenciação do fuso acromático da metáfase II.

Na **metáfase II** os cromossomas, constituídos cada um por dois cromatídeos, dispõem-se com os centrómeros no plano equatorial do respetivo fuso acromático.

Na **anafase II**, o centrómero de cada cromossoma divide-se, proporcionando a separação dos dois cromatídeos, que migram para os pólos opostos da célula.

Na **telofase II** os cromossomas tornam-se menos condensados, o invólucro nuclear reorganiza-se e os nucléolos reaparecem, completando-se assim a segunda divisão meiótica. De seguida, o citoplasma divide-se através de citocinese, dando origem a quatro células filhas haploides, com informação genética diferente da célula mãe. Estas células filhas originam as células reprodutoras.

Capítulo 3- Mutações

O DNA é uma molécula muito estável, que se replica com elevada precisão, no entanto, apesar da sua estabilidade, é suscetível de sofrer mutações.

As mutações podem ser consideradas a diferentes níveis, podendo envolver alterações ao nível dos cromossomas ou apenas alterações de pequenas sequências de DNA num ou mais genes, sendo classificadas de **mutações cromossómicas** ou **mutações génicas**, respetivamente.

Uma mutação apenas será visível se implicar uma alteração fenotípica. O efeito de uma mutação deve ser considerado com referência a um fenótipo contra o qual o mutante pode ser comparado, que geralmente é o fenótipo do tipo selvagem, ou seja, o fenótipo mais comum em populações naturais do organismo.

As mutações podem ou não ser hereditárias caso aconteçam em células germinativas ou células somáticas, respetivamente. As mutações somáticas podem acontecer em qualquer divisão celular, desde a primeira divisão do zigoto, até às divisões celulares que substituem as células de um indivíduo adulto. As mutações somáticas não são transmitidas para a geração seguinte, no entanto, a célula com a mutação irá transmitir a alteração para todas as suas células descendentes.

As mutações podem ocorrer de forma espontânea ou ser induzidas por agentes mutagénicos, como por exemplo radiação ultravioleta, raios X ou agentes químicos.

3.1. **Mutações cromossómicas**

Tal como o nome indica, mutações cromossómicas implicam alterações ao nível dos cromossomas, podendo afetar o seu número, e são por isso classificadas de mutações cromossómicas **numéricas**, ou a sua estrutura, sendo classificadas de mutações cromossómicas **estruturais**. As mutações cromossómicas conduzem a alterações do cariótipo de um organismo.

3.1.1. **Mutações cromossómicas numéricas**

Mutações cromossómicas numéricas causam alterações no número de cromossomas (ploidia) de uma célula ou indivíduo, podendo ser resultado da adição de múltiplos do conjunto único de cromossomas – **euploidia** – ou por adição ou perda de um ou mais cromossomas – **aneuploidia**.

Os cariótipos euploides apresentam um número de cromossomas igual a um múltiplo do seu conjunto básico, podendo ser considerados normais ou anomalias dependendo da espécie. Por exemplo, células animais apresentam células em diferentes estados de ploidia: a maioria das células são diploides, os gâmetas são considerados haploides e as células hepáticas são poliploides. Nestes casos, estas euploidias são consideradas normais ou geneticamente programadas. Por outro lado, euploidias não programadas são consideradas anomalias.

As **euploidias** alteram o conjunto dos cromossomas, podendo ser resultado da não disjunção de cromossomas homólogos ou de não ocorrência de citocinese. Podem ainda ocorrer de forma espontânea ou induzida, utilizando-se neste último caso inibidores mitóticos, como a colchicina, um dos inibidores mitóticos mais utilizados em plantas. A colchicina é um alcaloide extraído da planta *Colchicum autumnale*, utilizado como fármaco em diversas aplicações. Na qualidade de agente anti-mitótico, a colchicina atua sobre as fibras do fuso acromático durante a divisão celular, impedindo a sua formação ou promovendo a sua fragmentação, impedindo assim a separação dos cromossomos na anafase. Como consequência, as células iniciam o ciclo celular seguinte com a quantidade de DNA duplicado.

De acordo com o número de conjuntos cromossómicos básicos apresentados, os cariótipos podem ser sub-classificados em:

- a) **Monoploide** (n) – organismos que apresentam um único conjunto básico de cromossomas;
- b) **Diploide** ($2n$) – organismos possuem dois conjuntos de cromossomas;
- c) **Poliploides** ($3n, 4n\dots$) – indivíduos cujo cariótipo apresenta mais de dois conjuntos cromossómicos (exemplo: $3n$ – triploide; $4n$ – tetraploide)

Sabias que...

A poliploidia é muito comum em plantas, mas rara em animais.



Na natureza, um aumento do número do conjunto de cromossomas foi um fator muito importante para o aparecimento de novas espécies de plantas.



A descodificação do mecanismo que conduz à obtenção destas alterações cromossómicas (não disjunção cromossómica), associado às vantagens que plantas poliploides apresentam em termos agronómicos, conduziu ao desenvolvimento de protocolos que permitem a duplicação cromossómica e regeneração de plantas euploides. Os organismos monoploides, por exemplo, apresentam grande interesse ao nível do melhoramento de plantas uma vez que permitem o desenvolvimento de linhas diploides completamente homocigotas - **duplos haploides** - um recurso essencial em genética e melhoramento por permitir o desenvolvimento de novas cultivares, localização de genes com interesse agronómico ou o desenvolvimento de marcadores para características de interesse, aumentando assim a eficiência do melhoramento de plantas. A obtenção de duplos haploides tem interesse acrescido no estudo de características associadas a genes recessivos, cujo estudo é impossibilitado em organismos heterocigotas.

Células haploides ocorrem naturalmente na fase gametofítica das plantas superiores. A obtenção de plantas duplo haploides a partir de monoploides inclui métodos *in vitro*, como a ginogénese (cultura *in vitro* de ovário e flor) e androgénese (cultura *in vitro* de anteras e micrósporos), dependendo da espécie em estudo. Na cultura *in vitro* de micrósporos (um dos métodos mais utilizados), o micrósporo (grão de pólen antes da divisão mitótica, n) é colocado em meio de cultura apropriado para o desenvolvimento de uma planta monoploides (n). As plantas obtidas (monoploides) são selecionadas tendo em conta as características de interesse e submetidas a tratamento com inibidores mitóticos (ex. colchicina) para duplicação do número de cromossomas,

obtendo-se assim um duplo haploide ($2n$). As plantas duplos haploides podem ser utilizadas em programa de melhoramento convencional (por exemplo, hibridação).

A técnica de obtenção de duplo haploides já foi aplicada a mais de 250 espécies, tais como na planta do tabaco (*Nicotiana tabacum*), colza (*Brassica napus*), trigo (*Triticum aestivum*), milho (*Zea mays*), aveia (*Avena sativa*), batata (*Solanum tuberosum*), couve-flor (*Brassica oleracea*), entre outras.

Plantas duplo haploides têm sido muito utilizadas ao nível do melhoramento de plantas, apresentando como vantagens uma rápida produção de linhagens totalmente homozigóticas, contribuindo para a purificação de linhagens e aumento da eficiência da seleção.

Dependendo da composição do genoma, os organismos poliploides podem ser classificados em **autopoliploides** ou **alopoliploides**. Organismos **autopoliploides** possuem múltiplas cópias do conjunto único de cromossomas (x) de um mesmo genoma (conjuntos de cromossomas da mesma espécie, totalmente homólogos). Apresentam grande interesse ao nível do melhoramento de plantas no desenvolvimento de novas cultivares, visto que as mesmas são, geralmente, plantas com maior vitalidade e robustez, com maior produtividade e conseqüente rentabilidade. A obtenção de poliploides pode ocorrer espontaneamente na natureza, sendo exemplos a alfafa, o amendoim, a batata e o café, com genomas tetraploides e as bananas, variedade cultivada, com genoma triploide, ou de forma induzida, como acontece na melancia ou citrinos, com genoma triploide, resultantes do cruzamento de um tetraploide (progenitor feminino) com um diploide (progenitor masculino). A obtenção de organismos triploides tem grande interesse ao nível do melhoramento, visto que os mesmos são caracterizados por apresentarem um crescimento mais vigoroso e maior produtividade, maior tolerância a stresses bióticos e abióticos e frutos grandes com poucas ou nenhuma sementes (apirenes).

Certos organismos são compostos por, no mínimo, 2 genomas diferentes, - **alopoliploides** - resultado da hibridação de duas espécies diferentes, mas com uma proximidade genética elevada (hibridação interespecífica). O aloploiploide resultante possui os conjuntos cromossómicos derivados de duas ou mais espécies (cromossomas homeólogos, ou seja, cromossomas parcialmente homólogos). São comuns na natureza,

estando na base da evolução de diferentes espécies. Tomando como exemplo o caso do trigo mole (*Triticum aestivum*): os 42 cromossomas correspondem à fórmula hexaplóide $6x=42$ agrupados por três cariótipos que se representam pelas letras A, B e D e provêm de *T.durum* (A e B), outra espécie alopoliplóide ($2n=4x=28$) e *T.tauschii* (D), diplóide ($2n=2x=14$). O *Triticum aestivum* é proveniente da hibridação entre um alopoliplóide AABB e um diplóide DD, resultando num zigoto triplóide ABD que, por sua vez, se desenvolveu no hexaplóide AABBDD por endoduplicação (figura 20). Durante a meiose os cromossomas A B D de *T.aestivum* portam-se entre si como cromossomas não-homólogos: os cromossomas A emparelham e segregam apenas com cromossomas A, os B apenas com cromossomas B e os D apenas com cromossomas D – **cromossomas homeólogos**.

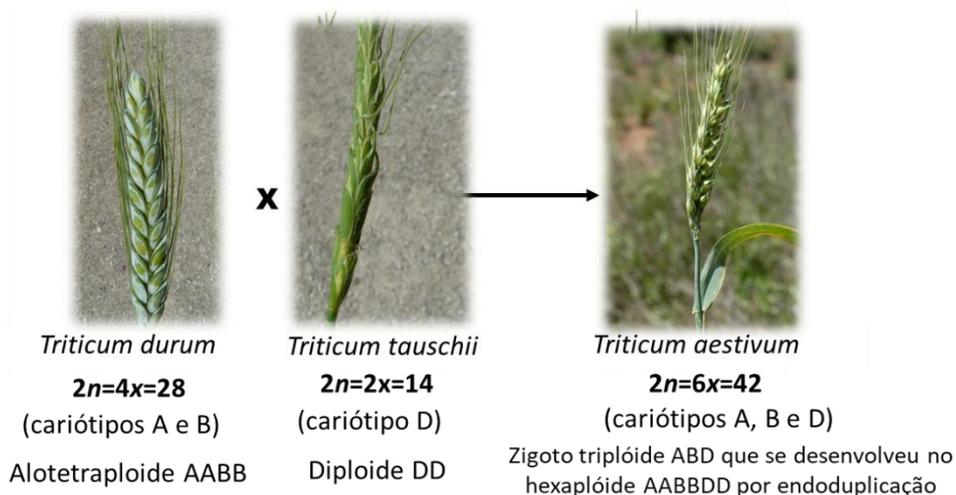


Figura 20 – Esquema, com representação do cariótipo, da hibridação entre *T.durum* e *T.tauschii* originando *T.aestivum* (adaptado de manual de genética, Professor Paulo de Oliveira).

Ao contrário das euploidias, em que se verificam constituições cromossómicas com múltiplos inteiros do número básico n da espécie, nas **aneuploidias** observam-se alterações de apenas um ou poucos cromossomas a mais ou a menos do que o normal da espécie. As Aneuploidias podem ocorrer devido a: i) perda do centrómero, levando a que o cromossoma seja perdido durante o processo de mitose ou meiose; ii) movimento retardado dos cromossomas durante a anáfase; iii) assinapse, ou seja, o não emparelhamento dos cromossomas homólogos; iv) não-disjunção dos cromossomas

homólogos ou cromátides irmãs durante a meiose ou mitose. Este evento leva a que alguns gâmetas ou células somáticas contenham um cromossoma extra ao passo que outros terão um cromossoma a menos. As aneuploidias podem ser classificadas em **hipossomias** ou **hiperssomias**, consoante haja perda ou ganho de cromossomas, respetivamente. No grupo das hipossomias, diz-se que há uma **monossomia** quando falta um único cromossoma num par de homólogos ($2n-1$) ou uma **nulissomia**, quando se verifica a perda de ambos os cromossomas de um par de homólogos ($2n-2$). No caso das hiperssomias, estas classificam-se em **trissomias** quando se verifica o ganho de um único cromossoma e um par de homólogos apresenta três cromossomas ($2n+1$), ou **tetrassomias** quando ocorre o ganho de dois cromossomas homólogos, existindo quatro cópias de um determinado cromossoma ($2n+2$). As aneuploidias podem ocorrer nos autossomas ou nos cromossomas sexuais.

3.1.2. Mutações cromossómicas estruturais

Quando as alterações nos cromossomas ocorrem ao nível da sua estrutura, estamos perante **mutações cromossómicas estruturais**. Estas mutações resultam de quebras na dupla cadeia do DNA, seguidas de ligação dos fragmentos, dando origem a alterações na estrutura dos cromossomas. Estas mutações podem ocorrer de forma natural durante o *crossing-over*, devido ao emparelhamento desigual dos cromossomas homólogos durante a prófase I da meiose, ou induzidas, devido à ação de agentes mutagénicos, como os raios X e raios UVC ou UVB que quebram ligações químicas entre os nucleótidos.

As alterações cromossómicas estruturais podem ter efeito *balanced* ou *unbalanced*, dependendo do conteúdo da informação genética final que resulta dos rearranjos após as quebras das ligações. Se a quebra nos cromossomas resultar em perda ou acréscimo de determinado fragmento, isso irá resultar num desequilíbrio da dosagem génica e estamos perante um efeito *unbalanced*. Por outro lado, se a quebra nos cromossomas homólogos resultar numa troca de fragmentos entre os mesmos, não haverá perda da informação genética causada pelo novo arranjo estrutural. Neste caso, estamos perante um efeito *balanced*. As alterações cromossómicas estruturais

unbalanced podem ser do tipo **duplicação ou deleção**, enquanto as alterações *balanced* podem ser do tipo **inversão** ou **translocação**.

a) **Duplicação**

Ocorre quando um fragmento se encontra duplicado no genoma. As sequências duplicadas podem variar em tamanho, localização no genoma e orientação. Tomamos como exemplo a seguinte sequência normal: ABCDE●FG. Quando as duplicações estão numa posição adjacente entre si, estamos perante uma duplicação em *Tandem* (ABCBCDE●FG) ou *Tandem* reversa (ABCCBDE●FG). Por outro lado, quando as duplicações ocorrem em diferentes segmentos génicos, a mesma pode ser classificada como *Displaced* (deslocada - ABCDE●FBCG) ou *Displaced Reverse* (ABCDE●FCBG).

Estas alterações podem ser resultado de um emparelhamento desigual dos cromossomas homólogos durante a prófase I da meiose. Quando isso acontece, o *crossing-over* origina cromátídeos recombinantes com a estrutura alterada devido à inserção de outro fragmento.

As alterações fenotípicas provocadas pelas duplicações devem-se à introdução de cópias adicionais e conseqüente desequilíbrio nas quantidades de produtos génicos, o que se traduz em alterações ao nível das proteínas. No entanto, os eventos de duplicação são comuns nos eucariotas e foram a fonte de origem de novos genes que adquiriam funções distintas do gene que lhe deu origem.

b) **Deleção**

As deleções são caracterizadas pela perda de um segmento cromossómico. Tomando como exemplo a sequência normal ABCDE●G, uma deleção poderá originar a sequência ABF●G. Uma grande deleção é facilmente identificada porque o **cromossoma fica mais curto**, obrigando o seu homólogo a fazer uma alça durante a profase I para permitir que regiões homólogas se alinhem.

O efeito ao nível do fenótipo depende do(s) gene(s) que se encontra(m) localizado(s) na região que sofreu a deleção. Em indivíduos homozigóticos a maioria destas deleções são letais porque correspondem à perda total dos genes localizados na região que sofreu a deleção. Em indivíduos heterozigóticos, estes podem ser portadores de deficiências

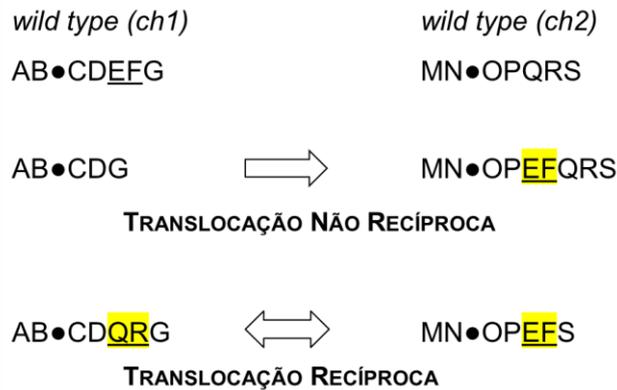
graves devido a i) um desequilíbrio nos produtos génicos, ii) ao fenómeno de pseudodominância, em que se verifica a expressão de genes recessivos ou mutações recessivas localizadas no segmento não-deletado, ou iii) à deleção de genes haploinsuficientes, ou seja, genes que requerem a presença de duas cópias para produzir o fenótipo normal do indivíduo (tipo selvagem).

c) Inversão

Nas mutações causadas por inversão, não existe perda (deleção) nem ganho (duplicação) de informação génica. Para que ocorra inversão deve haver, obrigatoriamente, dois pontos de quebra no cromossoma, ocorrendo uma inversão de 180° do segmento cromossómico. A ordem dos genes no cromossoma é alterada. As inversões podem ser classificadas em pericêntricas, quando a região invertida no cromossoma envolve o centrómero, ou paracêntricas quando a região invertida no cromossoma não envolve o centrómero, ou seja, compreende apenas a região de um braço do cromossoma. Uma inversão pode ser identificada durante a profase I porque os cromossomas homólogos formam uma alça para poderem alinhar, a alça de inversão. Por norma, as inversões alteram a regulação da expressão do gene, traduzindo-se em alterações fenotípicas.

d) Translocação

Nas translocações ocorre o movimento de material genético entre cromossomas não homólogos ou dentro do mesmo cromossoma. Quando a troca de fragmentos cromossómicos ocorre entre cromossomas não homólogos sem haver perda de informação genética, esta translocação denomina-se translocação recíproca, pois há troca de partes cromossómicas entre os não homólogos. Por outro lado, se um fragmento é translocado para outro cromossoma, sem haver troca entre eles, a translocação denomina-se não recíproca.



Das translocações recíprocas faz parte um tipo de translocação muito comum envolvendo cromossomas acrocêntricos, a translocação Robertsoniana, devendo o nome ao geneticista norte-americano W. R. B. **Robertson**, o primeiro a observar este tipo de mutações em gafanhotos, no ano de 1916. Nestas translocações, dois cromossomas acrocêntricos fundem-se próximos da região do centrómero com perda dos seus braços curtos. Esta translocação ocorre nos cinco pares de cromossomas humanos (13, 14, 15, 21 e 22) que são acrocêntricos.

A maioria das translocações recíprocas não afetam significativamente o fenótipo do indivíduo (não são letais) porque não há acréscimo nem diminuição de informação génica – ***balanced translocations***. Um exemplo típico de translocação que pode surgir na descendência é a síndrome de Down (trissomia 21), onde a translocação no cromossoma 14 contará como o terceiro homólogo do par 21.

As translocações Robertsonianas têm aplicabilidade ao nível do melhoramento de plantas. Existem muitos estudos na literatura acerca de translocações que envolvem diversas espécies, como o trigo (*Triticum aestivum* L.), centeio (*Secale cereale* L.), cevada (*Hordeum vulgare* L.), entre outras. Translocações heterozigotas induzidas são muito utilizadas nos programas de melhoramento de trigo com a finalidade de transferência de genes de resistência a fatores bióticos e abióticos, bem como genes que favorecem o rendimento de grãos. O centeio é uma espécie passível de translocar segmentos cromossómicos para o trigo e, pelo fato de possuir muitos genes de resistência a stresses, tais como genes de resistência à ferrugem da folha e do colmo e genes que favorecem o rendimento de grãos, são muito utilizado como dadores em programas de melhoramento. Estes segmentos são transferidos a partir das translocações, pois os

cromossomos do trigo tendem a quebrar simultaneamente, resultando na incidência de segmentos pequenos de cromossomos do centeio incorporados no genoma do trigo.

Sabias que...



O centeio possui muitos genes de interesse agronômico, dos quais se destacam o *Lr26*, *Yr19*, *Pm8* e o *Sr31*, os quais conferem resistência à ferrugem da folha, à ferrugem amarela, ao oídio e à ferrugem do colmo, respectivamente, além de genes que favorecem características agronômicas.



Capítulo 4 – Genética Mendeliana

Certamente que todos já ouvimos comentários sobre como somos parecidos com os nossos pais ou irmãos em determinadas características. De facto, ao longo da história da humanidade sempre houve muita curiosidade sobre a herança biológica, ou seja, sobre como os filhos herdam as características dos pais. Para encontrar as respostas a estas questões, precisamos recorrer à genética. A **genética** é a área da biologia que estuda o DNA, o material hereditário, e a forma como o mesmo se transmite ao longo de gerações, de como se expressa e como pode ser alterado.

O cientista **Johann Gregor Mendel** foi o primeiro a descobrir os mecanismos básicos da hereditariedade sendo considerado, por isso, o pai da genética. Foi entre 1856 e 1854 que Mendel desenvolveu as suas experiências com plantas, publicando os seus resultados em 1865 e 1866. No entanto, apesar da importância dos seus estudos, os seus resultados não tiveram o impacto e reconhecimento merecido pela comunidade científica. Foi apenas mais tarde, em 1915 com a descoberta do cromossoma por Thomas Morgan, que a genética Mendeliana se tornou a essência da genética clássica.

Segundo a teoria de Mendel, os caracteres são determinados por unidades discretas que são herdadas intactas através de gerações. Apesar de não ter definido a palavra, Mendel introduziu assim o conceito de *gene*. Para comprovar a sua teoria, Mendel realizou os seus trabalhos científicos com variedades de ervilha da espécie *Pisum sativum*. As principais razões que o levaram a optar por realizar os estudos com esta espécie foram a sua facilidade de cultivo, a presença de características distintas e facilmente diferenciáveis e o ciclo de vida curto e elevada descendência, o que permite obter várias gerações em pouco tempo. As ervilhas são plantas hermafroditas, já que numa mesma flor há órgãos reprodutores masculinos e femininos. Os óvulos de uma flor são fecundados, na maioria dos casos, pelos seus próprios grãos de pólen, ocorrendo o que chamamos de autofecundação.

Antes de iniciar cada cruzamento, Mendel certificava-se que estava a trabalhar com linhas puras, ou seja, plantas idênticas entre si numa determinada característica, produzindo descendência com o mesmo fenótipo quando cruzadas entre si. Para realizar cruzamentos com plantas que naturalmente se autofecundavam, Mendel recorreu à polinização artificial, realizando o que chamamos de fecundação cruzada. Nesta técnica, antes do pólen estar maduro e do estigma estar pronto para o receber, Mendel abria a

flor e cortava os estames contendo os grãos de pólen. O estigma dessa flor era assim fecundado com grãos de pólen trazidos de outra planta com as características desejadas.

Para os cruzamentos, Mendel selecionou variedades geneticamente puras para a geração parental (geração P), características grandemente dependentes do genótipo e com efeito ambiental residual, e variedades que apresentavam apenas dois fenótipos possíveis. Foram sete as características estudadas por Mendel e encontram-se resumidas na seguinte tabela:

	Características	Fenótipo contrastante	
1	Cor da flor	Branca	Lilás
2	Posição da flor	Axial	Terminal
3	Cor da semente	Amarela	Verde
4	Forma da semente	Lisa	Rugosa
5	Forma da vagem	Inflada	Estreita
6	Cor da Vagem	Amarela	Verde
7	Altura da planta	Alta	Anã

Mendel teve ainda o cuidado de analisar apenas uma característica de cada vez (monohibridismo). Assim, ao cruzar plantas altas com plantas baixas, por exemplo, ele considerava apenas a altura das plantas, não tendo em consideração características como a cor da semente ou a posição das flores no caule.

Mendel iniciou as suas experiências de monohibridismo realizando fecundação cruzada de plantas consideradas linhas puras com características contrastantes, cruzando, por exemplo, uma linhagem pura de plantas com flor branca com uma linhagem pura de plantas com flor lilás. Essa geração é chamada 'parental' (geração P), os seus descendentes formavam a primeira geração, ou geração F1, e os descendentes da F1 eram cultivados e autofecundados, formando a segunda geração, ou F2.

Em relação aos resultados dos cruzamentos, Mendel observou que as plantas que formavam a geração F1 apresentavam características físicas iguais apenas a uma das plantas da linhagem parental. Para perceber o que acontecia à outra característica, Mendel cultivou as sementes de F1 e permitiu que as mesmas sofressem autofecundação, formando a geração F2. Foi nessa geração que Mendel verificou que a

outra característica não havia desaparecido. Com base nesse resultado, Mendel denominou a característica que desaparece na F1 como **recessiva**, voltando a aparecer na F2, e a que não desaparece na F1 como **dominante**. Mendel observou ainda que na F2 a proporção de plantas com a característica dominante era de 75% e a de plantas com a característica recessiva era de 25%, ou seja, a relação era de 3 plantas com a característica dominante para 1 planta com a recessiva (3:1).

Os resultados constantes e concordantes obtidos levaram Mendel a formular a 1ª lei de Mendel ou lei da segregação, segundo a qual, nos organismos diploides, a **hereditariedade de uma característica fenotípica** é determinada por **dois fatores herdáveis** (genes alelos de um *locus*) que são **segregados** (separam-se) quando são formados os gâmetas, sendo que metade dos gâmetas tem um membro do par (gene alelo) e a outra metade dos gâmetas tem o outro membro (o outro gene alelo).

Atualmente sabemos que um gene pode existir em diferentes estados, os **alelos**. Mendel desconhecia este termo e referia-se aos alelos como fatores. Assim, o gene que codifica para a característica textura da semente, ocorre em dois estados, ou seja, apresenta dois alelos, o liso (R) e rugoso (r).

Segundo o **princípio da segregação**, a **hereditariedade de características individuais é determinada por dois genes alelos segregados nos gâmetas**. Cada indivíduo diploide tem dois alelos que codificam para uma característica. Quando um indivíduo apresenta, pelo menos num par de cromossomas homólogos, um par de genes constituído por diferentes alelos (exemplo *Tt*), este é considerado **heterozigótico** (híbrido). Por outro lado, se um indivíduo apresenta, num par de cromossomas homólogos, genes alelos iguais (exemplo *TT* ou *tt*), é considerado **homozigótico** (linha pura).

Na terminologia atual, um gene alelo corresponde a uma das diferentes formas de um gene que pode existir num determinado *locus* (*loci* – plural). *Locus* corresponde à posição do gene alelo no cromossoma.

Os **genes alelos** localizam-se no mesmo *locus* (mesma posição nos homólogos) e codificam para um carácter herdável (por exemplo, forma da semente), embora possam determinar diferentes fenótipos (lisa/rugosa). O conjunto de alelos de um indivíduo

define o genótipo (RR , Rr , rr). A figura 21 representa o *loci* dos alelos que codificam para o fenótipo forma da semente.

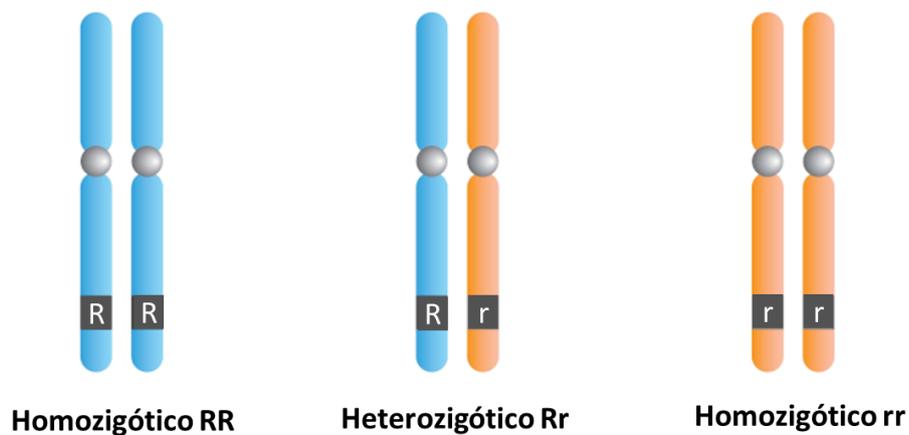


Figura 21 - Localização, nos cromossomas homólogos, dos alelos que codificam para a forma da semente (lisa ou rugosa). RR- semente lisa; Rr- semente lisa; rr- semente rugosa.

Adaptado de <https://mundoeducacao.uol.com.br/biologia/homozigoto-heterozigoto.htm>

Ao nível molecular R e r são idênticos na maior parte da sequência de DNA. Os dois alelos (R e r) terão tido origem no mesmo gene básico que por acumulação de mutações deu origem a duas variantes distintas. O alelo R (dominante) é composto por 3500 nucleótidos, enquanto o r (recessivo) apresenta, além dos 3500 nucleótidos originais, 800 nucleótidos adicionais.

Neste exemplo da forma da semente, o gene dominante (R) é aquele que define o fenótipo em genótipos heterozigotas (sementes lisas – RR ou Rr) e o gene recessivo (r) é o que define o fenótipo em genótipos apenas homozigotas (sementes rugosas – rr).

O comportamento dos cromossomas e dos alelos durante a meiose (divisão celular que origina a formação dos gametas) permite-nos compreender os resultados de Mendel. Na figura 22 podemos observar o comportamento dos alelos (representados pelas letras R e r) responsáveis pela determinação da forma da semente. Na anafase I, os cromossomas homólogos separam-se e, conseqüentemente, os alelos de um determinado gene também. Na meiose II, durante a anafase II, são os cromatídeos irmãos que se separam. Assim, num indivíduo com o genótipo Rr , metade dos gametas tem o alelo R e a outra metade o r .

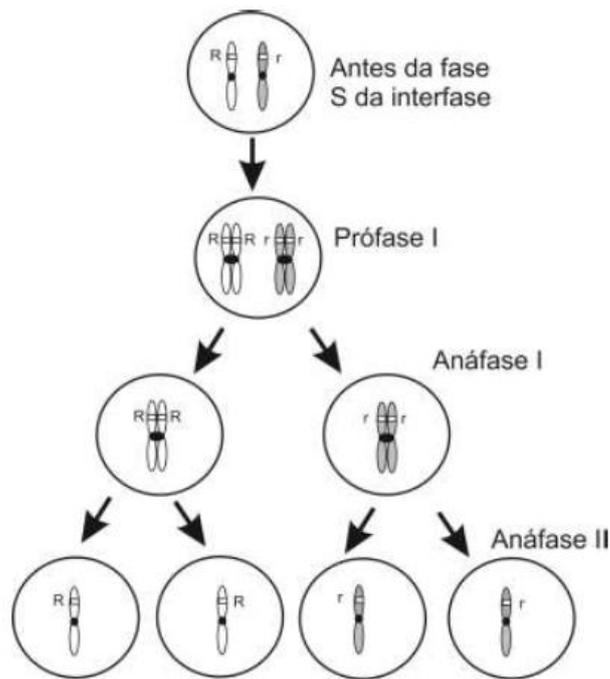
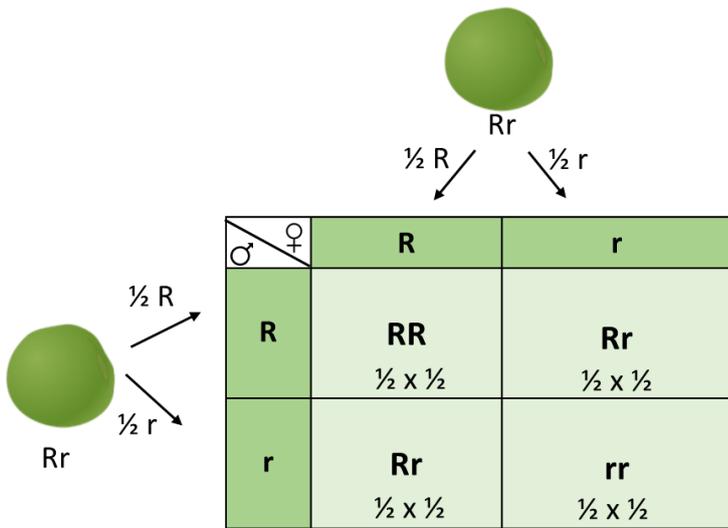


Figura 22 - Separação dos cromossomos e dos alelos responsáveis pela forma da semente durante a meiose. Adaptado de Souza *et al.*, 2015.

Uma forma mais simples de entender e representar os cruzamentos de Mendel é através da utilização do **quadro de Punnett**. Desenvolvido por Reginald C. Punnett (1917), permite prever as proporções genotípicas e fenotípicas de um cruzamento genético. Trata-se de um diagrama onde estão representados os gâmetas dos progenitores e os possíveis genótipos, sendo utilizado para prever os resultados de um determinado cruzamento.



Frequência genotípica:

RR = $\frac{1}{4}$ ou 25%

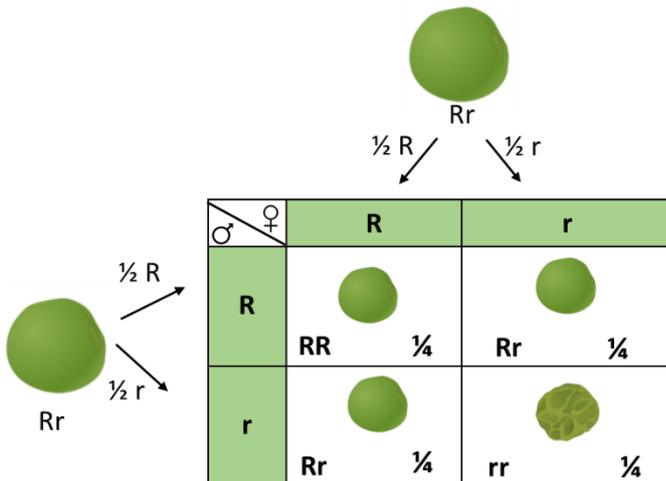
Rr = $\frac{2}{4}$ ou 50%

rr = $\frac{1}{4}$ ou 25%

$\frac{1}{4} : \frac{1}{2} : \frac{1}{4}$ ou 25%:50%:25%

Proporção genotípica:

1(RR):2(Rr):1(rr)



Frequência fenotípica:

RR + Rr = $\frac{1}{4} + \frac{2}{4} = \frac{3}{4}$ ou 75%

rr = $\frac{1}{4}$ ou 25%

$\frac{3}{4} : \frac{1}{4}$ ou 75%:25%

Proporção fenotípica:

RR + Rr: 1+2=3

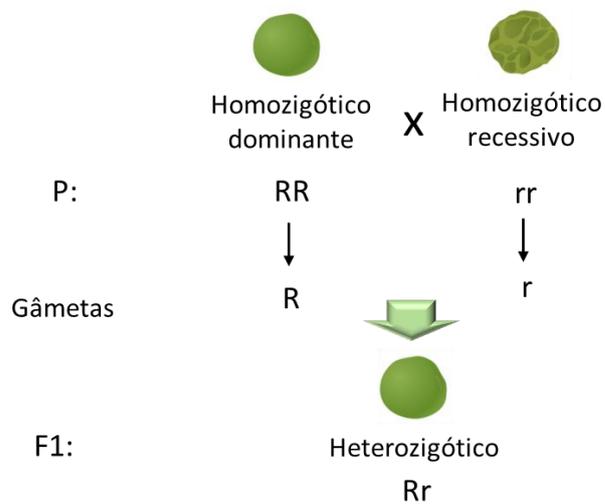
rr: 1

3  : 1 

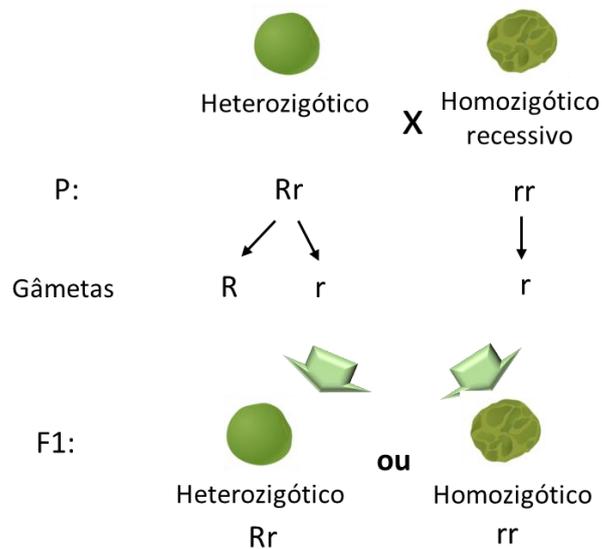
4.1. Cruzamento teste

Uma das estratégias para determinar o genótipo (heterozigótico ou homozigótico) de um indivíduo que apresenta o fenótipo dominante consiste em realizar um cruzamento teste (ou *Test cross*). O cruzamento teste é um cruzamento entre um indivíduo com o fenótipo dominante, para o qual queremos saber o genótipo, com um indivíduo de fenótipo recessivo e que terá o genótipo homozigótico recessivo.

Se o indivíduo testado for homozigótico, originará apenas um tipo de descendência, pois produz apenas um tipo de gâmetas:



No entanto, se o indivíduo testado for heterozigótico, originará dois tipos de descendentes: 50% com o fenótipo dominante e 50% com o fenótipo recessivo, como pode ser observado no seguinte esquema:



Até aqui temos estudado apenas cruzamentos entre linhas puras para uma única característica (monohíbridismo). No entanto, Mendel foi mais longe nos seus ensaios, realizando cruzamentos entre linhas puras para duas características (dihíbridismo). Continuando com o exemplo da forma da semente, acrescentamos agora a característica cor da semente. Neste caso, Mendel estabeleceu uma linhagem pura de forma de semente lisa (**RR**) e cor do revestimento da semente amarela (**AA**) e outra linhagem pura de forma da semente rugosa (**rr**) e cor do revestimento da semente verde (**aa**). A partir do cruzamento destas linhagens, ou seja, **RRAA** x **rraa**, Mendel obteve uma F1 com genótipo **RrAa** (heterozigótico duplo). A geração F2, resultante da autofecundação da F1, apresenta uma relação de 9/16 para sementes com ambas as características dominantes (R-A-), 3/16 para uma característica dominante e outra recessiva (R-aa ou rrA-) e 1/16 para ambas as características recessivas (rraa). Estes resultados encontram-se esquematizados na figura 23.

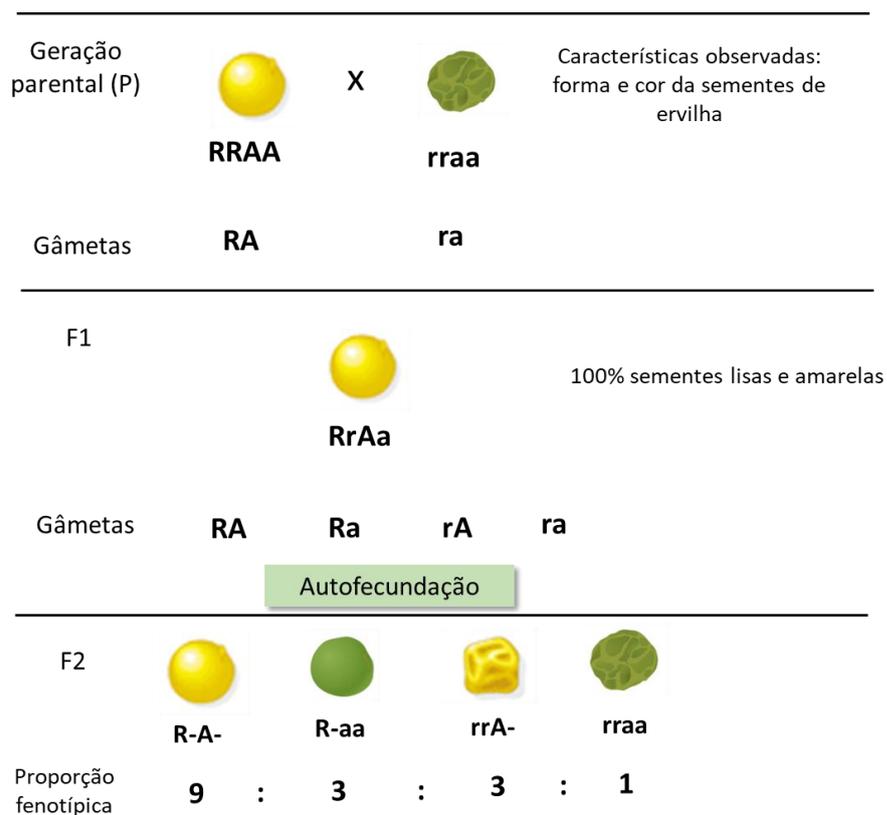


Figura 23 – Esquema dos cruzamentos dihíbridos de Mendel considerando as características forma e cor da semente de ervilha.

O resultado obtido na geração F2, resultante da autofecundação da F1, pode ser visualizado num quadro de Punnett, como o seguinte:

		Gâmetas				
		RA	Ra	rA	ra	
Gâmetas	♀ ♂	RA	RRAA (Lisa/amarela)	RRAa (Lisa/amarela)	RrAA (Lisa/amarela)	RrAa (Lisa/amarela)
	♀	Ra	RRAa (Lisa/amarela)	RRaa (Lisa/verde)	RrAa (Lisa/amarela)	RRaa (Lisa/verde)
	♀	rA	RrAA (Lisa/amarela)	RrAa (Lisa/amarela)	rrAA (Rugosa/amarela)	rrAa (Rugosa/amarela)
	♀	ra	RrAa (Lisa/amarela)	RRaa (Lisa/verde)	rrAa (Rugosa/amarela)	rraa (Rugosa/verde)

Os resultados obtidos nestes cruzamentos conduziram à elaboração da **2ª lei de Mendel** ou **princípio da segregação independente**, que afirma que durante a formação dos gâmetas, na meiose, os alelos localizados em diferentes pares de cromossomas homólogos segregam-se de forma independente. Tomando novamente como exemplo a forma e cor da semente, a separação dos alelos *R* e *r* ocorre de forma independente da separação dos outros dois alelos, *A* e *a*. A figura 24 esquematiza a 2ª lei de Mendel.

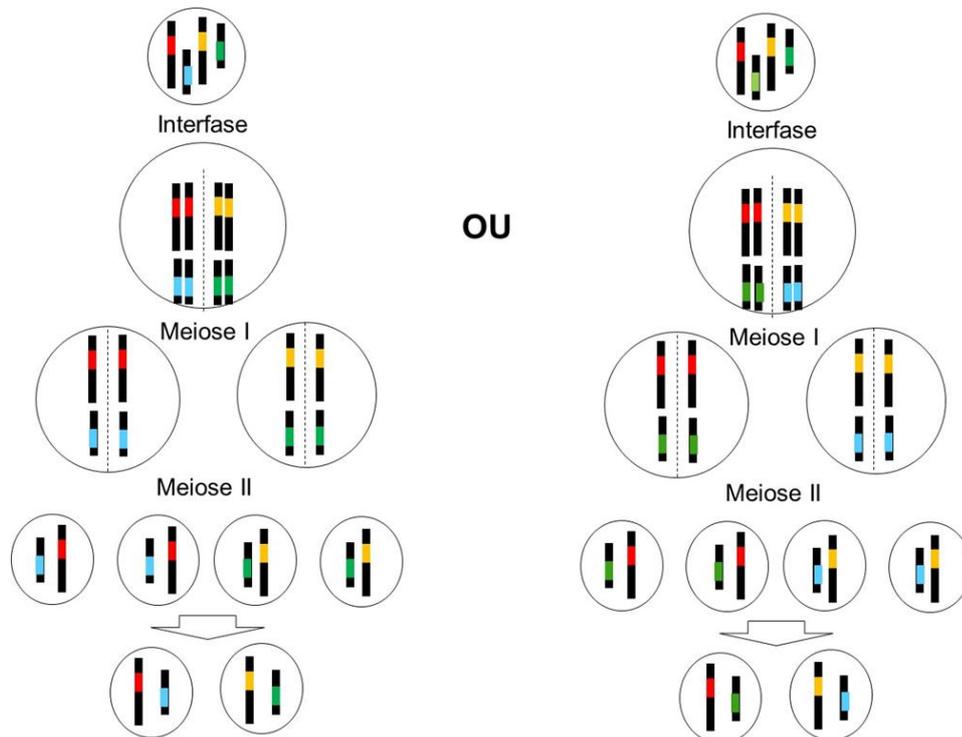


Figura 24- Esquema representativo da 2ª lei de Mendel. As diferentes cores representam diferentes alelos em diferentes *loci*.

No entanto, a 2ª lei de Mendel não é universal e só se aplica a genes que estão localizados em cromossomas separados ou, se localizados no mesmo cromossoma, se encontrarem em *loci* muito afastados. Genes localizados no mesmo cromossoma segregam juntos durante a anafase I da meiose e ficarão localizados no mesmo gâmeta. Apenas com eventos de *crossing over* poderá ocorrer essa separação.

Quando se trata de um cruzamento com mais de dois genes com dois alelos, podemos visualizar os resultados desse cruzamento de uma forma mais simples, recorrendo a um diagrama ramificado. Vejamos o exemplo para três genes com dois alelos e dominância completa nos três genes (***AaBbDd***) (Figura 25).

Cruzamento: ***AaBbDd*** x ***AaBbDd***

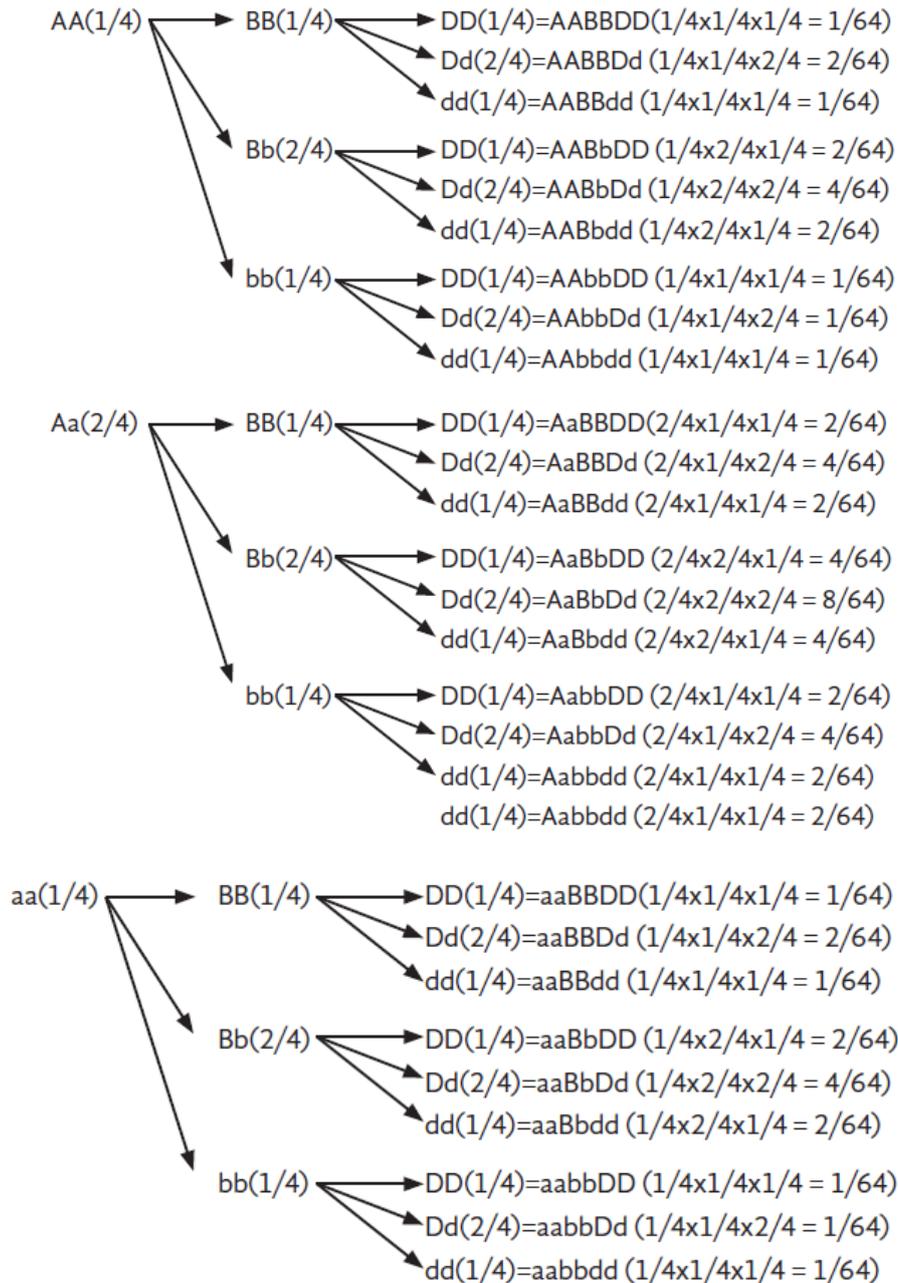


Figura 25 – Possíveis resultados para os genótipos, representados num diagrama ramificado, resultantes do cruzamento *AaBbDd* x *AaBbDd*. Adaptado de Souza et al., 2015.

Os princípios básicos do modelo de hereditariedade proposto por Mendel conseguem explicar como as características são herdadas em diversos organismos, mantendo-se como a base do conhecimento sobre herança genética. No entanto, as proporções fenotípicas descritas por Mendel nem sempre se verificam, observando-se alguns desvios das frequências.

Estes desvios das frequências devem-se ao acaso? Qual a probabilidade de ocorrer um desvio? Para dar resposta a estas perguntas realizamos o teste de χ^2 . Este teste baseia-se na diferença entre os valores esperados e observados, determinando a probabilidade da sua ocorrência. Calcula-se através da seguinte fórmula:

$$\chi^2 = \sum \frac{(\text{observado} - \text{esperado})^2}{\text{esperado}}$$

Tendo em consideração o seguinte exemplo:



Genótipo Aa

x



Genótipo Aa

Valores observados	Valores esperados	Desvio do esperado
30 (roxas) :20 (rosas)	37,5 (roxas) :12,5 (rosas)	Roxas: 7,5 Rosas: - 7,5
Total observado: 50 flores		

Aplicando a fórmula para o cálculo do teste de χ^2 para as flores roxas e rosas, obtemos um valor de $\chi^2=6,0$. Para associarmos o valor obtido através do teste de χ^2 a uma probabilidade, precisamos calcular o número de graus de liberdade, um conceito estatístico que indica quantas observações (fenótipos) independentes existem ao realizar o teste do χ^2 , e é obtido através da fórmula $n-1$ (n representa o número de fenótipos diferentes esperados). No exemplo das flores roxas e rosas, o número de graus de liberdade será 1. Consultando a seguinte tabela estatística, conseguimos associar o número de graus de liberdade ao valor obtido com o teste χ^2 e determinar o nível de significância:

		Nível de Significância:									
Graus de Liberdade:		99,5%	99%	97,5%	95%	90%	10%	5%	2,5%	1%	0,5%
		0,995	0,99	0,975	0,95	0,90	0,10	0,05	0,025	0,01	0,005
n-1	1	-	-	0,001	0,004	0,016	2,706	3,841	5,024	6,635	7,879
	2	0,010	0,020	0,051	0,103	0,211	4,605	5,991	7,378	9,210	10,597
	3	0,072	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,345	12,838
	4	0,207	0,297	0,484	0,711	1,064	7,779	9,488	11,143	13,277	14,860
	5	0,412	0,554	0,831	1,145	1,610	9,236	11,071	12,833	15,086	16,750
	6	0,676	0,872	1,237	1,635	2,204	10,645	12,592	14,449	16,812	18,548
	7	0,989	1,239	1,690	2,167	2,833	12,017	14,067	16,013	18,475	20,278
	8	1,344	1,646	2,180	2,733	3,490	13,362	15,507	17,535	20,090	21,955
	9	1,735	2,088	2,700	3,325	4,168	14,684	16,919	19,023	21,666	23,589
	10	2,156	2,558	3,247	3,940	4,865	15,987	18,307	20,483	23,209	25,188
	11	2,603	3,053	3,816	4,575	5,578	17,275	19,675	21,920	24,725	26,757
	12	3,074	3,571	4,404	5,226	6,304	18,549	21,026	23,337	26,217	28,299
	13	3,565	4,107	5,009	5,892	7,042	19,812	22,362	24,736	27,688	29,819
	14	4,075	4,660	5,629	6,571	7,790	21,064	23,685	26,119	29,141	31,319
	15	4,601	5,229	6,262	7,261	8,547	22,307	24,996	27,488	30,578	32,801

Através desta tabela, verificamos que há uma $P < 2,5\%$ de que o desvio que observamos entre os números esperado e observado de flores roxas e rosas possa ser devido ao acaso. Existindo uma *diferença significativa* indica que haverá **outras razões** que expliquem a diferença entre os valores esperados e observados **que não o acaso**.

Que outros fatores contribuem para a variabilidade não explicável pelo acaso e que alteram as proporções descritas por Mendel na F2?

1. Hereditariedade citoplasmática

A hereditariedade do DNA extra-nuclear, presente no citoplasma de mitocôndrias e cloroplastos, não se rege pelas leis de Mendel. Os genes presentes no DNA mitocondrial e cloroplastidial não se transmitem pelos gametas masculinos, sendo a hereditariedade citoplasmática mediada exclusivamente pelo gametófito feminino.

Se uma mutação der origem a um novo alelo no cloroplasto, a sua segregação ocorrerá aleatoriamente devido às sucessivas divisões da mitose e meiose, não havendo nenhuma lei que permita fazer previsões.

No início do século XX, Carl Correns, um botânico alemão, realizou várias experiências genéticas utilizando a planta maravilha (*Mirabilis jalapa*) (Figura 26).



Figura 26 – Planta Maravilha (*Mirabilis jalapa*).

<https://www.blog-flores.pt/flores-de-exterior/maravilha-mirabilis-jalapa/>

Os cruzamentos realizados com as plantas desta espécie resultaram em três fenótipos diferentes: plantas com ramos verdes (devido à presença de cloroplastos), plantas com ramos brancos (devido à presença de cloroplastos sem clorofila) e plantas com ramos de coloração variada (devido à presença de cloroplastos com e sem clorofila), sendo estas últimas denominadas 'variegadas'. Dos cruzamentos realizados, Correns entendeu que era a cor do ramo da planta que produz a oosfera que influenciava a cor do ramo descendente. Com efeito, ramos maternos puramente brancos ou verdes, produziam descendentes puramente brancos e verdes, respetivamente, enquanto ramos maternos variegados poderiam produzir os três tipos de cor na descendência, sem qualquer proporção previsível. Este resultado permitiu compreender que era a informação genética presente no citoplasma do gâmeta feminino que determinava a cor dos descendentes. No caso específico desta espécie, alguns cloroplastos presentes no citoplasma contêm mutações que os incapacita para a produção de clorofila e, consequentemente, da cor verde, apresentando a cor branca. Durante as muitas divisões celulares, algumas células terminarão com um conjunto de cloroplastos normais (com cor verde), outras receberão um conjunto de cloroplastos com mutação (apresentando manchas brancas) e outras, ainda, terão uma mistura de cloroplastos normais e mutados, apresentando manchas verdes e brancas.

2. Diferentes tipos de dominância

Quando ocorre dominância de um alelo sobre outro, ou seja, quando o alelo dominante mascara por completo o efeito do alelo recessivo, os indivíduos heterozigóticos (Aa) têm o mesmo fenótipo que os homozigóticos para o alelo dominante (AA), existindo apenas dois fenótipos possíveis, o associado ao homozigótico dominante ou heterozigótico e o fenótipo associado ao homozigótico recessivo (aa). Assim, o cruzamento entre dois indivíduos heterozigóticos (Aa x Aa) resultará em 75% dos indivíduos com fenótipo dominante e 25% com fenótipo recessivo, na proporção fenotípica de 3:1. Já a proporção genotípica traduz-se em 1 (AA) :2 (Aa) :1 (aa). Neste caso, em a geração F2 apresenta uma proporção genotípica diferente da fenotípica, estamos perante um tipo de **dominância completa**.

No caso de ausência de dominância entre os alelos de um gene, cada genótipo corresponde a um fenótipo e as proporções fenotípicas e genotípicas são iguais. Por exemplo, nas plantas boca-de-leão (*Antirrhinum majus*), o cruzamento entre uma planta homozigótica de flores brancas ($F^B F^B$) e uma planta homozigótica de flores vermelhas ($F^V F^V$) originará descendência com flores cor-de-rosa ($F^B F^V$). Este tipo de relação entre alelos, com um fenótipo heterozigótico intermediário entre os dois fenótipos homozigóticos, é chamado de **dominância incompleta**. Ainda neste exemplo, a autofecundação de um planta heterozigótica ($F^B F^V$), originará uma proporção genotípica de $1 F^B F^B : 2 F^B F^V : 1 F^V F^V$, que corresponde à proporção fenotípica de 1 planta com flores brancas : 2 plantas com flores cor-de-rosa : 1 planta com flores brancas (Figura 27).

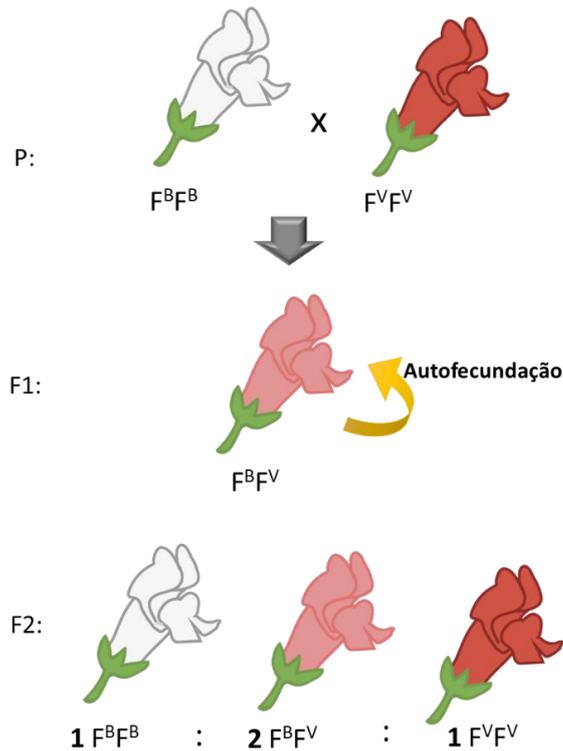


Figura 27 – Exemplo de dominância incompleta na hereditariedade da cor da flor da planta boca-de-leão (adaptado de <https://pt-pt.khanacademy.org/science/biology/classical-genetics>).

Quando dois alelos são expressos integralmente e em simultâneo num organismo heterozigótico, conduzindo ao aparecimento de um terceiro fenótipo (genótipo híbrido), estamos perante um caso de **codominância** (Figura 28).

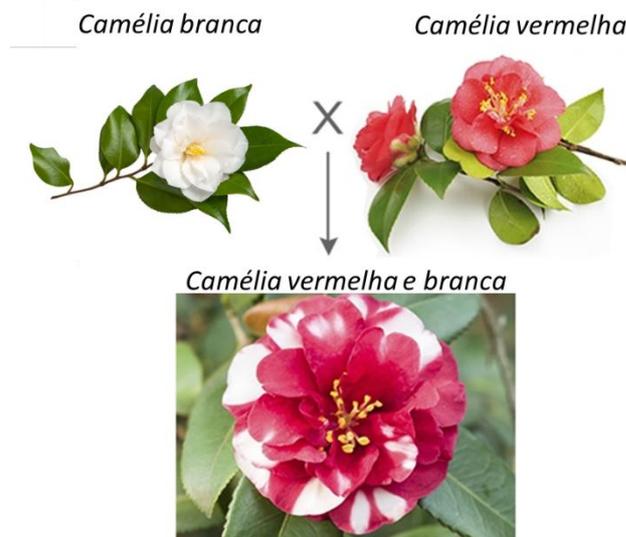


Figura 28 – Exemplo de codominância resultante do cruzamento entre uma camélia branca e vermelha (*Camelia japónica*).

Por que razão não existe sempre dominância completa? Por que razão existem casos de dominância incompleta e codominância? A dominância não é uma regra absoluta, mas sim uma consequência do funcionamento bioquímico dos genes e proteínas. No caso da dominância completa, o alelo dominante mascara completamente a expressão do alelo recessivo. Isto acontece porque, na maioria das vezes, o alelo dominante codifica uma proteína funcional, enquanto o recessivo pode ser uma versão não funcional do gene. A consequência fenotípica da presença de ambos os alelos recessivos (homozigótico recessivo) depende do papel da proteína no organismo. Caso a proteína desempenhe um papel fundamental para a sobrevivência do organismo, a presença de ambos os alelos recessivos pode ser letal. No caso da codominância, ambos os alelos são expressos ao mesmo tempo, sem que um mascare o outro. Neste tipo de dominância, cada alelo codifica uma proteína funcional diferente, mas ambas são produzidas. Algumas características são ainda determinadas pela quantidade de proteína produzida. Na dominância incompleta, por exemplo, um alelo pode produzir menos proteína do que o outro, resultando num fenótipo intermediário.

3. Genes pleiotrópicos

A pleiotropia é o fenómeno no qual um único par de alelos codifica em simultâneo para diferentes características. Um dos exemplos mais comuns é o das cebolas vermelhas e brancas. Neste caso, a cor das cebolas é determinada pelo mesmo gene que determina a produção de uma substância fungicida. Cebolas vermelhas apresentam o genótipo recessivo e resistência a um determinado fungo parasita. Já as cebolas brancas apresentam o genótipo dominante e suscetibilidade ao fungo. Neste caso, o alelo recessivo é responsável pela cor vermelha e ao mesmo tempo pela produção de uma substância fungicida.

4. Interações entre *loci*

Muitas características fenotípicas resultam da ação de vários genes. O aparecimento de diferentes fenótipos resulta, muitas vezes, da **interação entre genes não alélicos**, localizados em *loci* diferentes. Um exemplo de uma característica poligénica

é a cor do grão do milho, característica controlada por **três loci** (três genes cada um com dois alelos: AA, BB e CC) que codificam para a cor branca e vermelha. A cor do grão pode variar de branco a vermelho-escuro dependendo da concentração de pigmento, sendo que a concentração do pigmento aumenta com o número de genes dominantes presentes. Assim, um genótipo aabbcc originará um fenótipo cuja cor do grão é muito clara, enquanto um genótipo AABBCC originará um fenótipo cuja cor do grão é vermelho-escuro. Por outro lado, um genótipo AaBbCc dará origem a um grão de cor rosa-escuro.

Quando a interação génica se define por uma hierarquia entre os dois *loci*, ou seja, quando um gene atua inibindo a ação do outro para uma determinada característica, estamos perante casos de **epistasia**. Dois ou mais genes alelos interferem **suprimindo ou mascarando** a expressão do fenótipo de outro par de alelos localizado num *locus* distinto. O gene que suprime ou mascara o gene localizado noutra *locus* impedindo a expressão do seu fenótipo, define-se como epistático, sendo o segundo (o que fica 'mascarado') o hipostático. O fenótipo expresso corresponde ao do gene epistático. A epistasia pode ser classificada como dominante ou recessiva, consoante o efeito epistático resulte da presença de um gene dominante ou de um genótipo homocigótico recessivo, respetivamente. Por norma, os genes envolvidos na epistasia encontram-se envolvidos na mesma via bioquímica, codificando para enzimas diferentes que contribuem para o aparecimento de determinado fenótipo. Nessa via bioquímica, os alelos epistáticos geralmente atuam antes dos genes dos quais vão inibir a expressão (hipostáticos).

5. Epigenética

As diferentes manifestações de uma mesma característica são muitas vezes atribuídas a variações na molécula de DNA. No entanto, a expressão fenotípica é também muito suscetível às condições ambientais. As modificações que ocorrem no genoma em resposta ao ambiente e a estados patológicos é objeto de estudo da epigenética. O termo 'epigenética' foi definido em 1942, por Conrad Waddington, como o *ramo da biologia que estuda as interações casuais entre genes e seus produtos que trazem o fenótipo ao ser*. Nesse sentido, o termo epigenética refere-se ao estudo das vias moleculares que alteram a expressão dos genes para resultarem

num determinado fenótipo. Atualmente, a **epigenética** é definida como ***o estudo das alterações na função do gene que podem ser herdadas por mitose ou meiose e que não envolvem mudança na sequência de nucleotídeos do DNA***. Modificações epigenéticas dizem respeito a modificações bioquímicas nos nucleótidos da molécula de DNA ou noutros elementos que compõem a cromatina.

Os mecanismos epigenéticos mais estudados são a metilação do DNA e modificação das histonas. A metilação do DNA consiste na adição de um radical metil na posição 5' de uma molécula de citosina na sequência do DNA, geralmente seguida por uma guanina. As regiões metiladas tendem a ser menos transcricionalmente ativas, tendo impacto sobre o fenótipo. A ausência de metilação do DNA proporciona uma cromatina com uma estrutura mais aberta, facilitando o início da transcrição.

As histonas são proteínas nucleares que se ligam ao DNA, formando a cromatina. O mecanismo epigenético consiste em modificações nas caudas das histonas, as quais estão associadas a uma maior ou menor compactação da cromatina. Quando ocorre uma maior compactação, os sítios de ligação dos fatores de transcrição tornam-se inacessíveis, dificultando o processo de síntese proteica.

Capítulo 5 – Genética quantitativa

Em programas de melhoramento, os caracteres submetidos a seleção podem ser classificados como qualitativos ou quantitativos dependendo da natureza das características observadas. **Caracteres qualitativos**, como por exemplo a forma e cor da semente ou a cor ou posição da flor, dizem respeito a características que podem ser categorizadas em classes discretas e distintas. Apesar de serem controlados por poucos genes, o genótipo tem um grande efeito sobre os caracteres qualitativos, sendo os mesmos pouco influenciados pelo ambiente. O melhoramento de características qualitativas, na maioria das vezes, envolve a seleção de indivíduos que apresentam o fenótipo desejado, seguindo padrões mendelianos simples, como dominância/recessividade. Por outro lado, **caracteres quantitativos**, como a altura da planta, produção de grãos ou comprimento da espiga, referem-se a características com variação contínua dentro de uma população, ou seja, tudo o que pode ser medido e ter qualquer valor numa escala de números reais. Na maioria das populações, a variabilidade fenotípica dos caracteres quantitativos segue uma **distribuição normal** (ou curva de Gauss), representado graficamente por um histograma. Essa distribuição ocorre devido à soma de dois fatores principais:

1. **Efeitos genéticos múltiplos:** este tipo de caracteres são influenciados por vários genes (poligenia), cada um contribuindo com um pequeno efeito.
2. **Influência ambiental:** O ambiente afeta cada indivíduo de forma ligeiramente diferente, suavizando os extremos e contribuindo para uma distribuição contínua.

A maioria dos caracteres com grande importância econômica e agrícola são do tipo quantitativo. O melhoramento de características quantitativas envolve a seleção da geração parental com base em valores genéticos ou outros índices de seleção determinados através de análises estatísticas. Parâmetros estatísticos populacionais como média, variância e desvio padrão, são também muito utilizados para obter informações relevantes para o melhoramento.

O ramo da genética que estuda a variação de uma determinada característica quantitativa, numa dada população, denomina-se **genética quantitativa**. Este ramo da

genética procura determinar a proporção da variação fenotípica influenciada pelo genótipo e pelo ambiente, traduzindo-se na seguinte expressão:

$$\text{Fenótipo (P)} = \text{fatores genéticos (G)} + \text{fatores ambientais (E)}$$

Tal como referido anteriormente, caracteres quantitativos são extremamente influenciados pelo ambiente, sendo que o mesmo genótipo pode produzir diferentes fenótipos dependendo das condições ambientais onde se desenvolve o organismo. A plasticidade fenotípica refere-se à capacidade de um único genótipo expressar diferentes fenótipos - formas alternativas de morfologia, estado fisiológico ou comportamental - em relação às características variáveis do ambiente. Como a plasticidade fenotípica está intimamente ligada à capacidade de adaptação de um organismo perante constrangimentos ambientais impostos durante o seu desenvolvimento, diferentes genótipos podem convergir em fenótipos similares, originando fenocópias. O termo fenocópias diz respeito a fenótipos semelhantes ao de organismos mutantes, mas sem a presença dos genes mutantes respetivos, sendo o resultado de condições ambientais específicas. Em moscas da fruta, por exemplo, uma mutação autossómica recessiva dá origem a moscas com olhos muito pequenos. O mesmo fenótipo pode ser observado quando larvas de moscas normais (sem mutação) são expostas a metaborato de sódio.

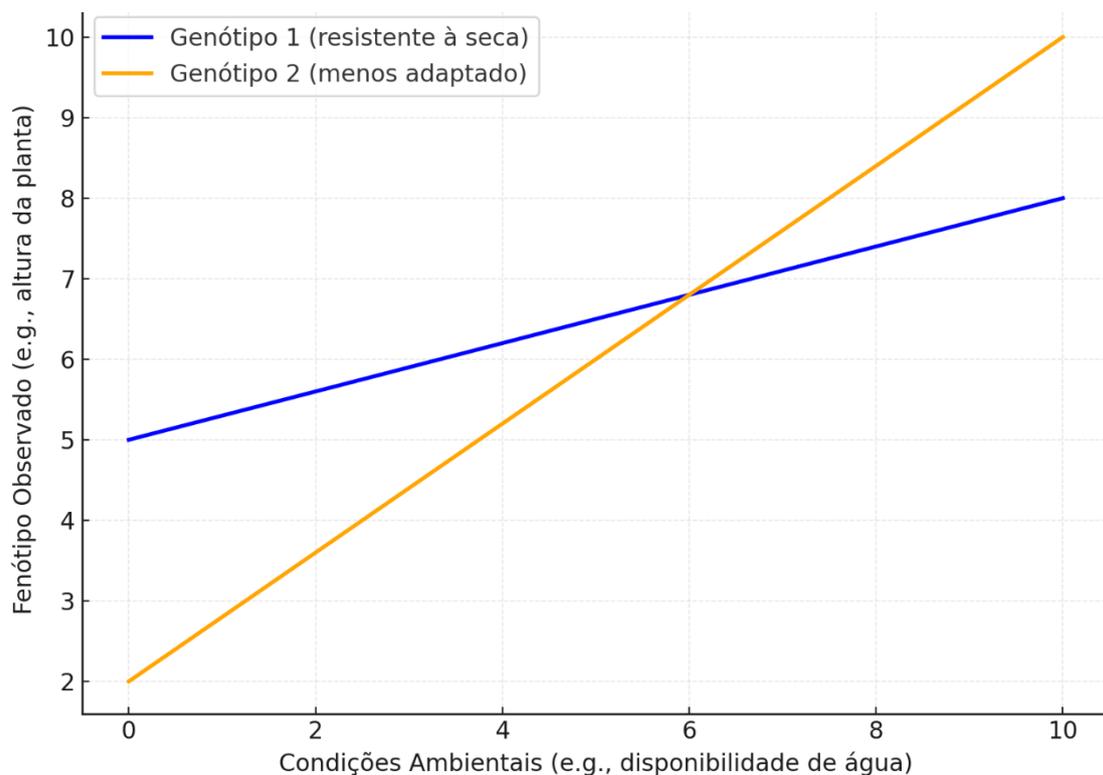
A capacidade de um genótipo produzir diferentes fenótipos em resposta às condições ambientais denomina-se **norma de reação**. A norma de reação de um dado genótipo corresponde a um gráfico de respostas fenotípicas ao longo de um gradiente ambiental, normalmente representado por uma linha ou curva, refletindo o efeito do ambiente na variação fenotípica.

A norma de reação pode ser entendida como o conjunto de respostas fenotípicas possíveis de um genótipo ao longo de uma gama de condições ambientais. Ela demonstra a **plasticidade fenotípica**, ou seja, a capacidade de um organismo ajustar a sua expressão fenotípica de acordo com mudanças no ambiente.

Por exemplo:

- Uma planta com um genótipo específico pode apresentar variações na altura dependendo da disponibilidade de recursos como água, luz ou nutrientes.
- Da mesma forma, a produção de biomassa, o teor de proteínas em sementes ou a resistência a doenças em plantas podem variar conforme as condições ambientais.

A norma de reação é frequentemente representada por gráficos, onde o eixo X corresponde às condições ambientais (ex.: níveis de nutrientes, temperaturas) e o eixo Y corresponde ao fenótipo observado (ex.: altura, produção de sementes). Por exemplo, uma planta resistente à seca pode manter uma altura relativamente constante (norma de reação plana) em diferentes condições de disponibilidade de água. Por outro lado, outra planta, menos adaptada, pode mostrar uma redução acentuada na altura em ambientes de seca (norma de reação inclinada).



Neste exemplo gráfico, o genótipo 1 (resistente à seca) mostra uma resposta moderada ao ambiente, com um fenótipo relativamente estável em diferentes condições

ambientais. O genótipo 2 (menos resistente), apresenta uma resposta fenotípica mais acentuada às condições ambientais, indicando maior sensibilidade.

A análise da norma de reação permite-nos compreender como um único genótipo pode produzir uma gama de fenótipos ao longo de diferentes condições ambientais, destacando a plasticidade fenotípica. No entanto, nem sempre a expressão de um fenótipo está garantida, mesmo em condições favoráveis. É aqui que entra o conceito de **penetrância génica**.

A Penetrância génica diz respeito à percentagem de indivíduos com determinado genótipo que expressa o fenótipo correspondente. No entanto, algumas vezes, o genótipo não dá origem ao fenótipo esperado. Quando tal acontece estamos perante um caso de Penetrância incompleta. Uma das causas possíveis para a ocorrência de penetrância incompleta está relacionada com as condições ambientais, na medida em que as mesmas podem inibir a expressão do fenótipo de alguns indivíduos, originando uma segunda classe fenotípica. Por exemplo, uma planta com um genótipo associado à resistência à seca pode expressar essa característica apenas num ambiente moderadamente seco, enquanto num ambiente extremamente seco essa resistência pode não se manifestar.

Apesar de indivíduos da mesma espécie compartilharem genes muito semelhantes e se desenvolverem também num ambiente muito semelhante, os mesmos exibem diferenças ao nível do fenótipo. Estas diferenças são explicadas pelo efeito do acaso ou ruído de desenvolvimento (do inglês *Developmental noise*), e é determinado por fatores como a expressão de genes e sinalização celular. O ruído de desenvolvimento pode ajudar os indivíduos a adquirir a capacidade de se adaptar ao ambiente e contribuir para os seus padrões únicos de desenvolvimento. A maioria dos estudos de ruído de desenvolvimento são focados em animais, no entanto, o diferente desenvolvimento de raízes laterais em plantas da mesma espécie ou a diferente capacidade de germinação de sementes são exemplos deste fenómeno.

Ao analisar a variabilidade fenotípica em caracteres complexos, como resistência a doenças, altura ou produtividade das plantas, torna-se evidente que essas características raramente são controladas por um único gene. Em vez disso, são influenciadas por múltiplos genes, cada um contribuindo para o fenótipo final. Esses genes, ou regiões do genoma associadas à variação fenotípica de um caráter quantitativo, são chamados de **Quantitative Trait Loci (QTLs)**. Os QTLs desempenham um papel fundamental na interpretação da norma de reação e da penetrância gênica, ajudando a explicar como diferentes genes, em conjunto com o ambiente, ajustam a expressão fenotípica em populações naturais ou cultivadas. Por exemplo, a norma de reação para a altura de uma planta pode ser atribuída à interação de múltiplos QTLs com fatores ambientais, como a disponibilidade de nutrientes ou água. Além deste exemplo, a penetrância de um genótipo resistente a uma doença pode depender da presença de QTLs específicos que regulam essa resistência em condições ambientais variadas.

No contexto prático do melhoramento de plantas, os QTLs permitem identificar genes que respondem a fatores ambientais, por exemplo.

5.1. Heritabilidade

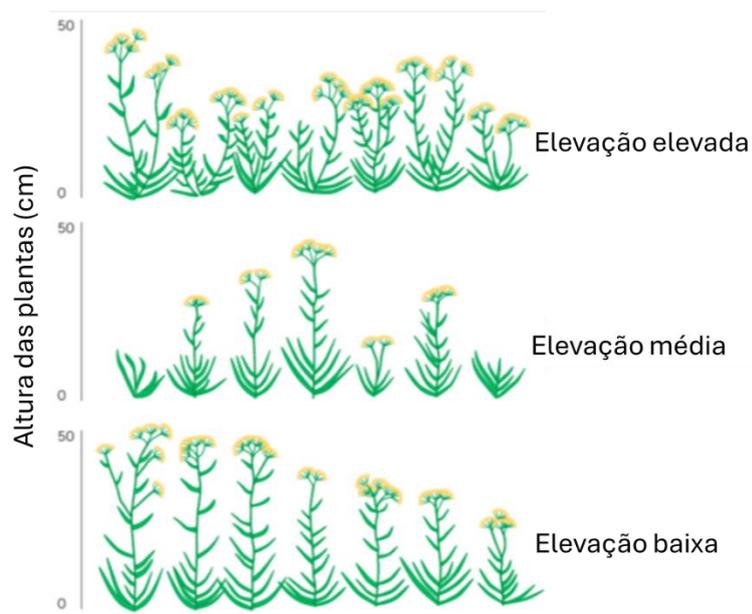
Até agora percebemos que o fenótipo é condicionado pelo genótipo, pelo ambiente e pelo efeito do acaso. No entanto, se num programa de melhoramento pretendermos determinar se a variação observável numa característica fenotípica numa determinada população é transmissível às gerações seguintes, utilizamos o coeficiente genético denominado heritabilidade. A **heritabilidade** define-se como a proporção da variação fenotípica que é atribuída a fatores genéticos, e por isso hereditários, independentemente do ambiente.

A heritabilidade é um conceito muito importante ao nível do melhoramento genético, visto que a mesma está relacionada com a transmissão de características genéticas de uma geração para a geração seguinte. Determinados parâmetros com elevado impacto agronómico, como o rendimento, a resistência a doenças e parâmetros relacionados com a qualidade, apresentam na maioria das vezes, baixa heritabilidade, o

que significa uma grande influência ambiental na sua expressão. Nesses casos, o impacto do melhoramento genético será reduzido.

A heritabilidade pode ser definida em sentido lato (H^2) ou sentido estrito (h^2). A heritabilidade em sentido lato reflete todas as contribuições genéticas, incluindo a aditiva, dominante e epistática, que determinam a variância fenotípica entre indivíduos numa população. Por seu lado, a heritabilidade em sentido estrito estuda a proporção do fenótipo que é transmissível da geração parental para os descendentes, ou seja, indica-nos quanto é que a variação fenotípica visível numa população é causada por efeitos genéticos, de modo que possa ser transmitida dos progenitores à geração seguinte.

Tendo em consideração o seguinte exemplo:



Neste exemplo, todas as plantas têm o mesmo genótipo, mas crescem de forma diferente consoante a elevação (alta, média e baixa), sugerindo que **fatores ambientais** (como temperatura, oxigênio, disponibilidade de nutrientes) influenciam o crescimento de forma significativa. Neste caso, a heritabilidade é baixa. Se as plantas mantivessem um padrão de altura similar entre cada nível de elevação, apesar das condições ambientais, isso indicaria uma maior heritabilidade.

No melhoramento de plantas, h^2 é utilizada para prever alterações no fenótipo em resposta à seleção (**R**). A resposta à seleção corresponde à evolução da média fenotípica da geração parental para a geração de descendentes (geração *offspring*) e traduz-se na seguinte fórmula: **$R = h^2S$** , onde h^2 corresponde à heritabilidade em sentido estrito e S corresponde ao diferencial de seleção direcionada, ou seja, a diferença entre a média do fenótipo dos **indivíduos selecionados** para os cruzamentos para produção da geração seguinte e a média do fenótipo observado na população original **antes da seleção**.

A resposta à seleção pode ser avaliada tanto numa geração quanto em várias. Se quisermos calcular a resposta à seleção ao longo de várias gerações, devemos atualizar a fórmula para: **$R = t h^2S$** , onde t representa o número de gerações. Neste caso, o S pode ser calculado como a **média do diferencial de seleção** ao longo das gerações, visto que há fatores que podem afetar a expressão fenotípica e ter impacto ao nível do S, como por exemplo efeitos ambientais.

Num programa de melhoramento, se decidirmos seguir uma linha de melhoramento focado numa característica utilizando uma ou várias gerações, a heritabilidade é muito importante para determinar quão eficaz será a seleção artificial para essa característica, sendo estimada utilizando a expressão **$h^2 = R/S$** . Os valores de **h^2** podem variar entre 0 e 1, sendo o valor mínimo indicativo de que a variância fenotípica é unicamente devida à variância ambiental e, neste caso, a seleção será ineficaz, pois a variação fenotípica é exclusivamente devida a fatores ambientais, pelo que a característica em causa não será transmitida à geração seguinte. Pelo contrário, valores de **$h^2=1$** significa que a contribuição genética é o único fator responsável pelas diferenças fenotípicas observáveis entre indivíduos. No geral, se uma característica tiver uma heritabilidade alta, as mudanças na população podem ocorrer mais rapidamente através da seleção artificial, uma vez que a variação genética é uma parte significativa da variação total. Valores de **h^2** entre 0 e 0,25 são indicativos de uma heritabilidade baixa; entre 0,25 e 0,5 são indicativos de uma heritabilidade média e entre 0,5 e 1 indicam uma heritabilidade alta.

5.2. Genética de populações

O ramo da genética que estuda a variação ao nível do genótipo dentro e entre as populações de uma espécie denomina-se **genética de populações**. Trata-se do estudo das frequências genéticas e genotípicas nas populações e das forças evolutivas capazes de as alterar ao longo das gerações. A genética de populações estuda como os genes e alelos são distribuídos e como suas frequências mudam ao longo do tempo, descrevendo as variações genéticas presentes em grupos de indivíduos da mesma espécie. Quando o *pool* génico de uma população se mantém inalterado ao longo de gerações sucessivas, dizemos que a população está em **equilíbrio genético**. Sempre que existam alterações no *pool* génico, diz-se que a população está em evolução.

A **frequência genotípica** de uma população diz respeito à proporção ou percentagem de indivíduos que pertencem a cada genótipo. Por seu lado, a **frequência alélica** (genética ou génica) diz respeito à proporção dos diferentes alelos em cada *locus*. A figura 29 exemplifica estes conceitos essenciais em genética populacional.

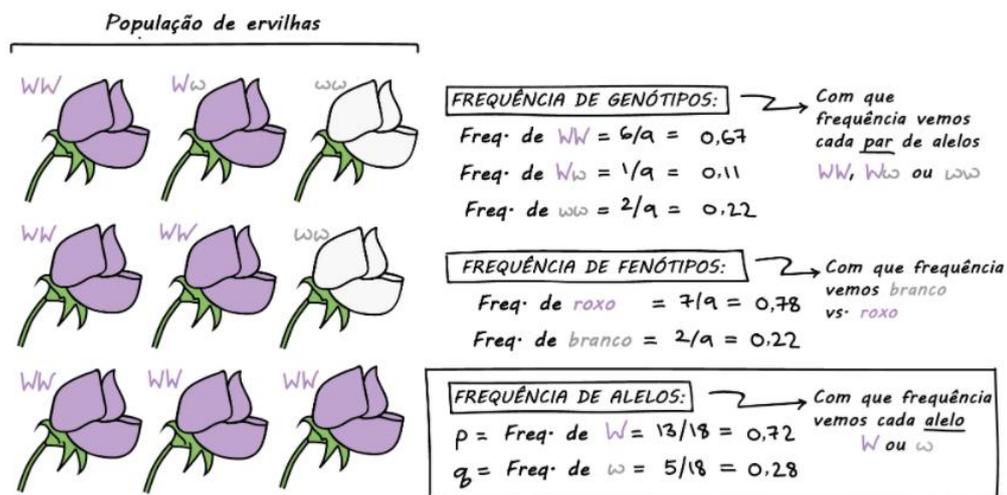


Figura 29 - Diagrama representativo do cálculo da frequência genotípica, fenotípica e alélica de uma população. (Adaptado de <https://pt.khanacademy.org/science/ap-biology/natural-selection/hardy-weinberg-equilibrium/a/allele-frequency-the-gene-pool>)

As frequências genotípicas são derivadas das frequências alélicas, sendo que para o cálculo das mesmas devemos assumir que a população se encontra em **Equilíbrio de Hardy-Weinberg**. Segundo o teorema de Hardy-Weinberg, em populações infinitamente grandes, em que os cruzamentos se processem ao acaso e não se verifique a pressão de

fatores evolutivos (seleção natural, migração, mutação, etc.) não haverá alterações no *pool* génico, pelo que as frequências génicas e genotípicas permanecerão constantes de geração em geração (em equilíbrio genético). Ou seja, **se uma população estiver em equilíbrio HW, as frequências dos alelos e dos genótipos não mudam ao longo das gerações.**

• **Frequências genotípicas:** a frequência de cada genótipo é calculada por:

$$f(AA) = \frac{N_{AA}}{N}$$

$$f(Aa) = \frac{N_{Aa}}{N}$$

$$f(aa) = \frac{N_{aa}}{N}$$

- N_{AA} = número de indivíduos com genótipo AA
- N_{Aa} = número de indivíduos com genótipo Aa
- N_{aa} = número de indivíduos com genótipo aa
- N = total de indivíduos na população

Onde **$f(AA) + f(Aa) + f(aa) = 1$**

• **Frequências alélicas:** a frequência do alelo A é representada por p e a frequência do alelo a é representada por q. Então, a frequência de cada alelo é calculada por:

$$p = \frac{2N_{AA} + N_{Aa}}{2N}$$

$$q = \frac{2N_{aa} + N_{Aa}}{2N}$$

- N_{AA} = número de indivíduos com genótipo AA
- N_{Aa} = número de indivíduos com genótipo Aa
- N_{aa} = número de indivíduos com genótipo aa
- N = total de indivíduos na população

A soma das frequências alélicas sempre será: **$p+q=1$**

Se a população seguir a **lei de Hardy-Weinberg**, as frequências genotípicas podem ser previstas a partir das frequências alélicas. O equilíbrio de Hardy-Weinberg representa-se pela fórmula **$p^2+2pq+q^2=1$** , onde *p* é a frequência do alelo dominante (por exemplo, alelo A), *q* é a frequência do alelo recessivo (por exemplo, alelo a), p^2 é a frequência genotípica de homozigotos dominantes (por exemplo, AA), $2pq$ é a frequência genotípica de heterozigotos (por exemplo, Aa), q^2 é a frequência genotípica de homozigotos recessivos (por exemplo, aa).

Uma população **só estará em equilíbrio HW** se atender a estas condições:

1. **A população tem de ser grande** (evita deriva genética).
2. **Não há seleção natural** (todos os genótipos têm a mesma chance de sobreviver e se reproduzir).
3. **Não há mutações** (as frequências alélicas não se alteram por mutação).
4. **Não há migração** (nenhum alelo entra ou sai da população).
5. **O acasalamento é aleatório** (não há preferência por certos genótipos).

Se **qualquer uma dessas condições for violada**, a população **não estará em equilíbrio HW**, e as frequências genotípicas precisarão ser determinadas diretamente dos dados observados.

Por norma, as populações não estão em equilíbrio Hardy-Weinberg, visto que as frequências de alguns dos seus genes tendem em mudar. Sempre que se verificar uma mudança nas frequências alélicas na população ao longo de gerações, podemos dizer que a população evoluiu. Os mecanismos de evolução compreendem a ocorrência de mutações, introduzindo um novo genótipo na população, acasalamentos aleatórios, migração, deriva genética e seleção natural. A trajetória evolutiva de um determinado gene pode ser o resultado da ação simultânea de diferentes mecanismos de evolução.

Capítulo 6 – Síntese de proteínas: Transcrição e Tradução

As proteínas são polímeros fundamentais para a estrutura, função, diferenciação e desenvolvimento de um organismo. O conhecimento acerca do processo de síntese proteica foi um dos maiores desafios da biologia e um marco fundamental na história da biologia molecular. Os primeiros estudos que contribuíram para elucidar o processo de síntese de proteínas datam de 1954, quando o físico George Gamow sugeriu que as proteínas deveriam ser sintetizadas diretamente no DNA. No entanto, estudos preliminares de Brachet (1944) e Caspersson (1947) já sugeriam que existia uma molécula intermediária entre o DNA e a proteína, o RNA. A descoberta do papel do RNA mensageiro foi fundamental na compreensão deste processo. Estas considerações resultaram na hipótese conhecida como o *Dogma Central da Biologia*, segundo o qual o fluxo da informação genética presente no DNA flui obrigatoriamente desde o DNA até ao RNA, e deste para a proteína.

A sequência de nucleótidos presentes no DNA (gene) codifica a sequência de aminoácidos presentes na proteína através de duas etapas: primeiro, a informação contida na cadeia de DNA é **transcrita** numa cadeia simples de RNA; segundo, a informação presente no RNA é **traduzida** numa sequência de aminoácidos de acordo com o código genético.

O Ácido Ribonucleico (**RNA**, do inglês *Ribonucleic Acid*), é uma molécula fundamental no processo global de síntese proteica. À semelhança do DNA, o RNA é constituído por nucleótidos ligados por **ligações fosfodiéster**. O RNA possui um grupo hidroxilo no átomo de carbono 2' do açúcar ribose (enquanto o DNA tem apenas um átomo de hidrogénio), tornando-o uma molécula quimicamente instável e mais facilmente degradado. Ao contrário do DNA, o RNA apresenta-se tipicamente como uma cadeia simples, sendo a cadeia polinucleotídica composta pelas bases azotadas A (adenina), U (uracilo), C (citosina) e G (guanina). Os nucleótidos da cadeia de RNA podem estabelecer ligações entre si devido à ocorrência de regiões complementares, conduzindo ao estabelecimento de estruturas tridimensionais estáveis e importantes para cada uma das classes de RNA.

Há três classes de RNA envolvidas na síntese proteica: RNA mensageiro (**mRNA**), RNA ribossomal (**rRNA**) e RNA de transferência (**tRNA**). De uma forma simplificada, a informação é transcrita do DNA para o mRNA, atuando este como intermediário do produto final dos genes codificadores. As moléculas de mRNA, que se formam por complementaridade de bases do DNA, transportam a informação do núcleo até aos ribossomas – organito presente no citoplasma e constituído por proteínas e rRNA. O tRNA transporta os aminoácidos até ao ribossoma, onde decorrerá a síntese proteica.

No mecanismo de síntese proteica em eucariotas podemos considerar três etapas:

- 1- **Transcrição** da informação do DNA para a molécula de RNA mensageiro
- 2- **Processamento** do pré-mRNA em mRNA funcional
- 3- **Tradução** da informação presente no mRNA numa proteína

Transcrição

Por norma, apenas um gene, ou no máximo alguns genes, são transcritos em RNA ao mesmo tempo. A transcrição é um processo muito seletivo, sendo os genes transcritos apenas quando necessário. No entanto, o tamanho de um gene pode variar de organismo para organismo, ou mesmo genes diferentes que pertencem a um determinado organismo, podendo ser formados por dezenas ou milhões de pares de bases (pb).

O molde para a síntese de RNA (transcrição), como para a síntese de DNA (replicação), é uma cadeia simples da dupla hélice de DNA. No entanto, ao contrário da replicação, a transcrição ocorre **em apenas uma das duas cadeias** de nucleótidos do DNA – a **cadeia molde** ou **não codificante**. A molécula de RNA que é sintetizada é **complementar e antiparalela** à cadeia de molde de DNA. Logo, o transcrito de RNA tem a mesma orientação e sequência de nucleótidos que a cadeia não-molde, com a exceção de que o **U** no RNA substitui o **T** no DNA. A cadeia que não serve de molde, e que ocorre de 5´-3´, denomina-se **cadeia codificante** do DNA (Figura 30).

Num gene, o primeiro nucleótido transcrito é denominado +1, sendo numerados negativamente os nucleótidos a montante do primeiro nucleótido, e os que se encontram a jusante numerados positivamente.

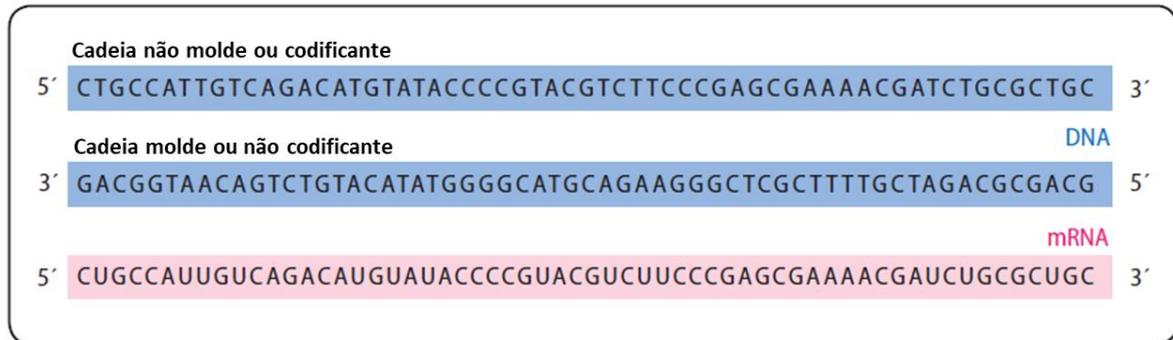


Figura 30 – DNA transcrito e RNA resultante, com representação das cadeias codificante e não codificante de DNA. Adaptado de Ribeiro, 2009.

À semelhança da replicação, a transcrição ocorre na direção **5'-3'**. Os nucleótidos de RNA [**ribonucleósidostrifosfato (NTPs)**] são adicionados um de cada vez ao **grupo 3'-OH** da molécula de RNA em crescimento. A enzima que catalisa a síntese de RNA a partir de uma cadeia de DNA é a **RNA polimerase**. Para tal, esta enzima utiliza como substrato os trifosfatos 5' de ribonucleósidos (ATP, GTP, CTP e UTP). Nas células eucarióticas existem três tipos de RNA polimerases: RNA polimerase I, II e III. A RNA polimerase I encontra-se no nucléolo e é responsável por transcrever os rRNA. A RNA polimerase II está presente no nucleoplasma e é a enzima responsável pela síntese do RNA nuclear, o qual é posteriormente processado em mRNA. Finalmente, a RNA polimerase III encontra-se igualmente presente no nucleoplasma e é responsável pela síntese de tRNA e rRNA.

Ao contrário das DNA polimerases, as RNA polimerases conseguem iniciar uma cadeia de RNA, ou seja, não necessitam de um 3'-OH livre de um nucleótido já incorporado, ou seja, não necessita de um *primer*. No entanto, não apresentam ação exonucleotídica, o que faz com que o RNA incorpore nucleótidos erradamente mais frequentemente que o DNA.

A primeira fase da transcrição inicia-se com o reconhecimento de uma sequência específica no gene denominada **promotor**. Os promotores são segmentos do DNA que

contêm sequências especificamente reconhecidas por proteínas reguladoras (fatores de transcrição), determinando o local do DNA onde a enzima RNA polimerase se deve encaixar para iniciar a síntese de RNA, ou seja, o local de início da transcrição. A sequência do promotor (TATAAA, também conhecida como TATA box) é reconhecida **por proteínas reguladoras (fatores de transcrição)**, as quais se ligam a essa sequência. O primeiro fator de transcrição a reconhecer a sequência denomina-se TFIID, o qual se liga ao promotor na TATA box. A transcrição tem início quando a RNA polimerase II se liga à região do promotor, assim como mais fatores de transcrição gerais (figura 31).

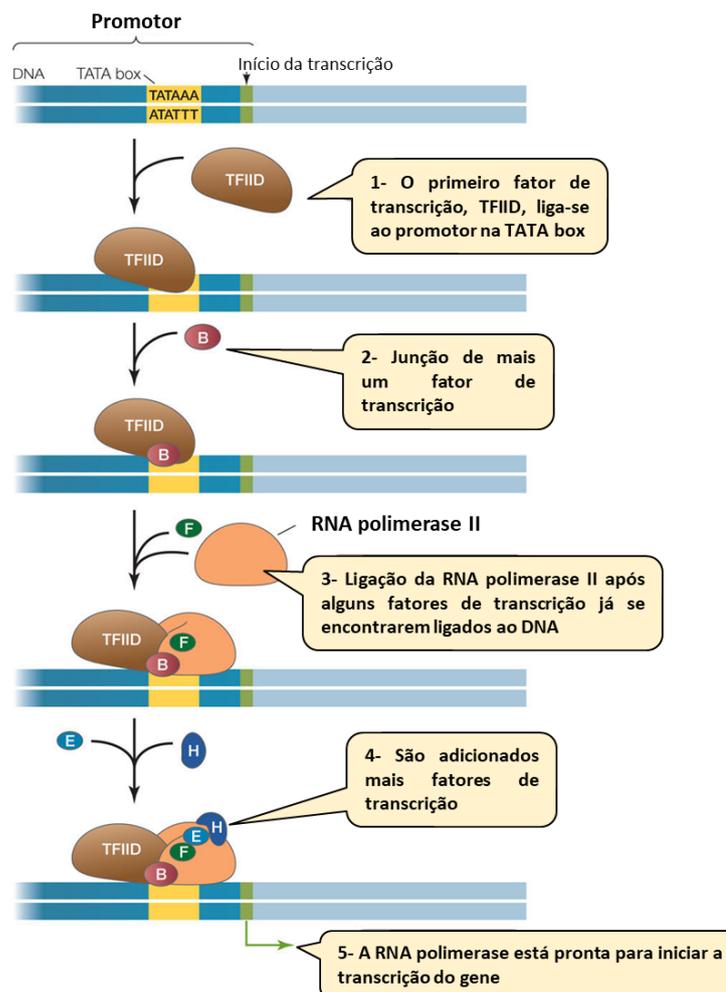


Figura 31 – Representação do complexo que determina o início da transcrição em células eucarióticas.

Adaptado de https://digfirpublished.macmillanusa.com/pol2e/pol2e_ch11_3.html

A RNA polimerase II provoca a separação das duas cadeias, sendo que, à medida que a enzima avança, as duas cadeias desenrolam-se e separam-se, permitindo que os

nucleótidos de RNA existentes no meio celular se ligem por ação da RNA polimerase II, usando como molde apenas uma das duas cadeias de DNA. Forma-se desta forma, na direção 5'-3', a molécula de **RNA pré-mensageiro** (pré-mRNA). À medida que decorre a transcrição, a molécula de DNA restabelece as ligações quebradas entre as bases complementares, fecha-se e enrola-se.

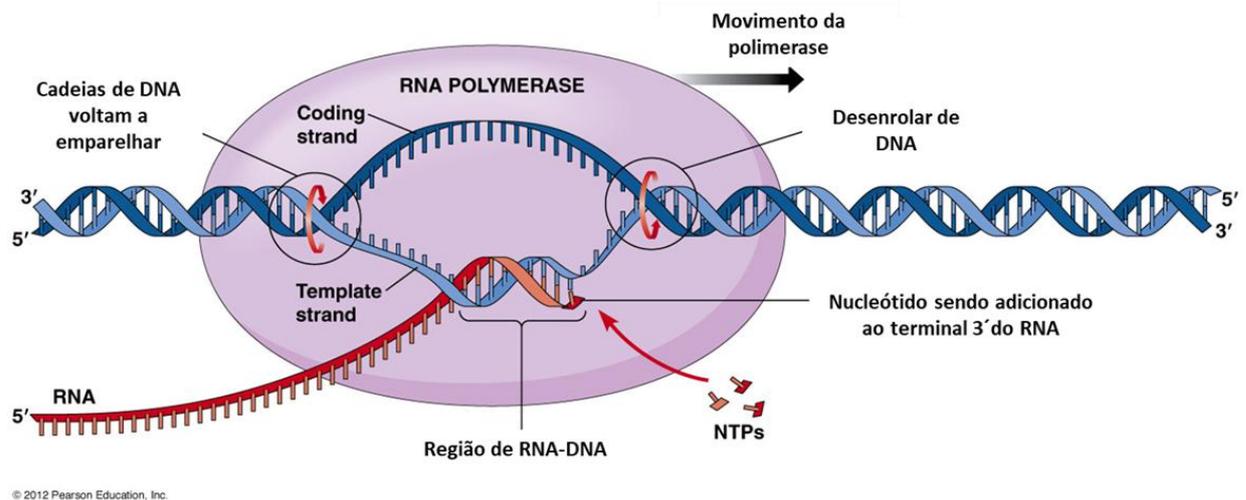


Figura 32 – Representação do processo de transcrição. Adaptado de <https://www.rubendries.com/project/transcription/>

A transcrição pára quando a RNA polimerase II encontra uma sequência de terminação ou terminador, uma sequência de cerca de 40 pb, rica em C e G, localizados a jusante da extremidade 3' da região codificante. A sequência de terminação forma uma estrutura em gancho (*hairpin*) na zona rica em G e C devido a ligações complementares que ocorrem no próprio RNA. Esta estrutura serve de sinal para separar da cadeia molde o RNA sintetizado e a RNA polimerase II.

O processo de transcrição envolve o consumo de energia, sendo a mesma proveniente da dissociação dos trifosfatos dos ribonucleótidos em monofosfatos.

Processamento

O **transcrito primário (pré-mRNA)** resultante da transcrição não é ainda o mensageiro (mRNA) a ser traduzido no citoplasma, sendo alvo de um processamento no núcleo para dar origem ao RNA maduro. Esse processamento consiste na adição de um nucleótido de 7-metilguanina invertido (*cap*) na extremidade 5' do transcrito primário,

e na adição de uma **cauda poliadenilica - poly(A) tail** - na extremidade 3' - processo de poliadenilação. A adição destas estruturas terminais nas extremidades 5' e 3' é realizada por uma enzima polimerase e tem como objetivo proteger o transcrito da atividade das exonucleases do citoplasma. O processamento finaliza com a remoção dos intrões e ligação dos exões, um processo reconhecido como **Splicing alternativo** (figura 33). Este processo é realizado por ribonucleoproteínas (snRNPs) que reconhecem as extremidades típicas dos intrões (5' - **GU** e 3' - **AG**). O *splicing* ocorre dentro de um grande complexo chamado **spliceossoma**.

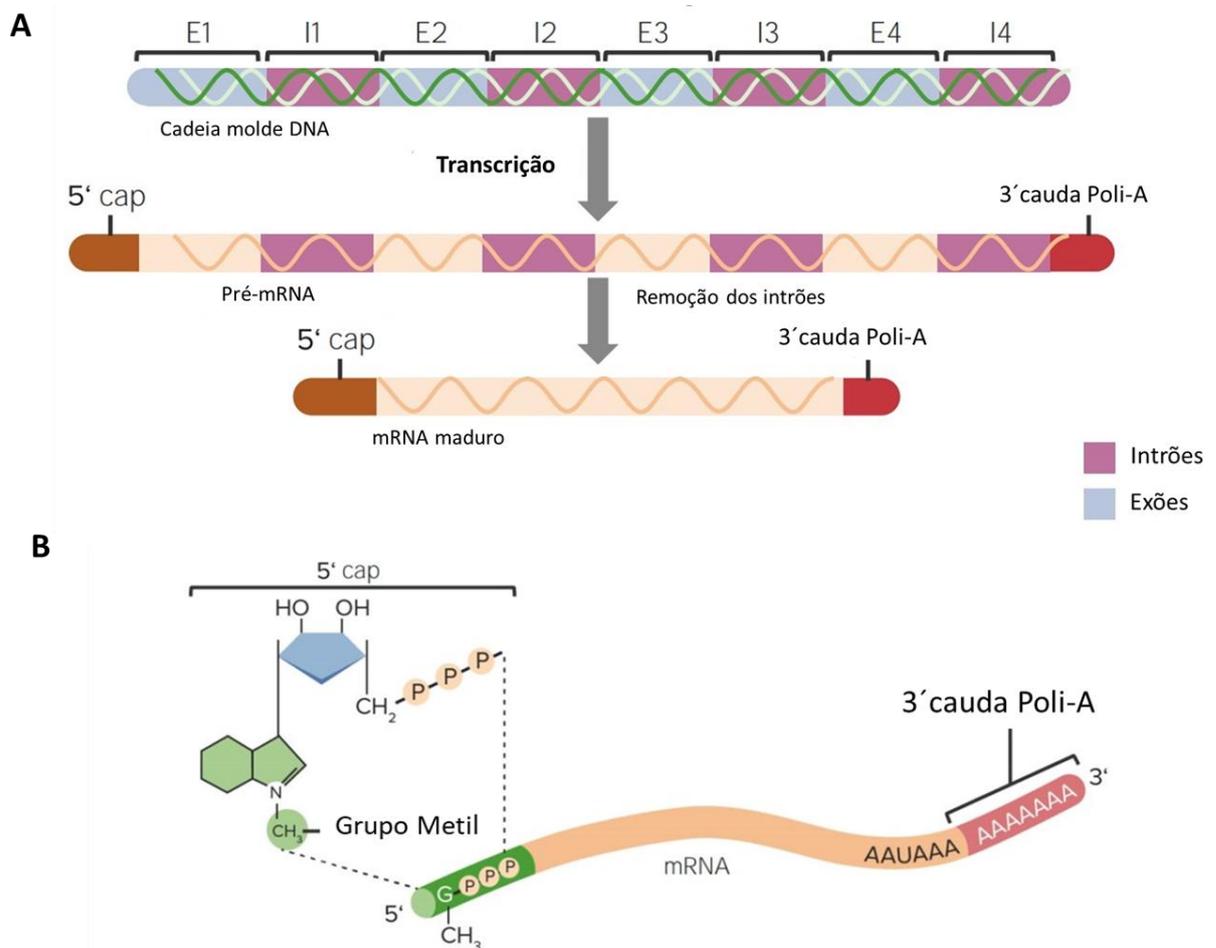


Figura 33 – A- Processamento do pré-mRNA em RNA maduro. **B-** Detalhe das extremidades 5' e 3' com a adição das estruturas *cap* e *cauda Poli-A*, respetivamente. Adaptado de <https://www.lecturio.com/pt/concepts/modificacoes-pos-transcricionais-processamento-de-arn/>

Após a remoção dos intrões, os exões podem ser unidos em várias combinações, como resultado do *splicing* alternativo, o que significa que, a partir do mesmo gene, podem obter-se diferentes proteínas em diferentes tecidos.

Tipos de RNA

Apesar de quimicamente similares, tendo em conta a sua função, são reconhecidos diversos tipos de RNA. Todos os tipos de RNA são sintetizados no núcleo, sendo que alguns são translocados para o citoplasma e participam diretamente na síntese de proteínas.

Os tipos de RNA principais que participam nas diferentes etapas da síntese proteica são:

- 1. RNA mensageiro (mRNA)** – estrutura específica de cadeia única para a síntese de uma determinada proteína e muito ativa metabolicamente. Apesar do mRNA das células eucarióticas ter estabilidade suficiente para possibilitar sua purificação, é menos estável que os outros tipos de RNA. É sintetizado e processado no núcleo, sendo de seguida translocado para o citoplasma, onde vai ser lido pelos ribossomas. Todos os mRNA citoplasmáticos contêm na extremidade 3' uma cadeia de aproximadamente 100-300 resíduos de adenina, a cauda poli-A, cuja função está relacionada com o transporte da molécula para o citoplasma, e na extremidade 5' um grupo metil ligado a um nucleótido de guanina, uma modificação denominada cap.
- 2. RNA ribossomal (rRNA)** – faz parte da estrutura dos ribossomas, desempenhando um papel fundamental na síntese de proteínas. Existem diferentes tipos de rRNA. Em eucariotas, os principais tipos de rRNA são o 18S, 5.8S, 28S (todos eles fazem parte da subunidade maior dos ribossomas) e o 5S (presente na subunidade menor). Em eucariotas, o rRNA é sintetizado no nucléolo.
No ribossoma, o rRNA define a posição de ligação do mRNA. O rRNA 18S, presente na subunidade menor do ribossoma, reconhece a sequência cap do

mRNA. Essa interação é essencial para o posicionamento correto do mRNA, permitindo que a tradução se inicie no ponto certo e prossiga de forma eficiente. O rRNA possui ainda atividade catalítica, sendo responsável pela formação da ligação peptídica entre os aminoácidos durante a síntese proteica. Essa função catalítica é atribuída ao rRNA 28S nas subunidades ribossomais de eucariotos.

- 3. RNA de transferência (tRNA)** - Molécula de RNA à qual se liga um aminoácido específico, transportando-o até ao ribossoma. O tRNA possui uma estrutura secundária característica que se assemelha a um trevo, com quatro extremidades principais. Numa das extremidades possui uma região de três nucleótidos para ligação ao mRNA chamada **anti-codão**. Na extremidade 3' possui um local de ligação ao aminoácido (esta ligação ocorre através da enzima aminoacil-tRNA sintetase). Quando dobrado, o tRNA assume uma estrutura tridimensional em forma de "L". Essa conformação é muito importante para o seu funcionamento dentro do ribossoma, permitindo que o anti-codão esteja adequadamente posicionados para interagir com o mRNA. Cada tRNA é específico para um determinado aminoácido e possui um anti-codão correspondente ao codão do mRNA que codifica esse aminoácido. Existem no mínimo 20 tipos de tRNA, um para cada aminoácido padrão.

- 4. RNA de interferência (RNAi)** – Pequenas moléculas de RNA, nomeadamente **siRNA (small interfering RNA)** e **miRNA (microRNA)**, que regulam a expressão génica, silenciando genes específicos. Este processo ocorre através da degradação do mRNA alvo ou da inibição da sua tradução, impedindo a produção da proteína correspondente.

Tradução

Na tradução, a informação contida na sequência de nucleótidos do mRNA é traduzida numa sequência de aminoácidos. É um dos processos de maior custo energético para a célula, dependente de GTP e ATP. Este processo envolve os tRNA, rRNA, os ribossomas, e o mRNA; também são necessários aminoácidos e diversas proteínas específicas.

Enquanto os tRNA e os ribossomas são reciclados pela célula, podendo ser utilizados para a síntese de diferentes proteínas, o mRNA é específico para cada proteína, sendo continuamente produzido e degradado.

O **mRNA** é translocado do núcleo para o citoplasma através dos poros do envólucro nuclear. No citoplasma, o mRNA é traduzido pelos ribossomas, produzindo-se um polipéptido pela leitura dos codões que se sucedem ao codão de iniciação (AUG), no mesmo sentido da transcrição.

Os ribossomas são constituídos por, aproximadamente, 50% de rRNA e 50% de proteínas. O ribossoma é composto por duas subunidades, uma maior e uma menor, que se juntam durante a síntese proteica. Nos eucariotas, a subunidade menor denomina-se de **40S** e a maior de **60S**, juntas formam o ribossoma **80S** (a nomenclatura "S" refere-se à unidade de Svedberg, uma medida de sedimentação que reflete o tamanho e a forma da partícula).

A subunidade menor do ribossoma é responsável por fazer a leitura do mRNA. Esta subunidade contém o sítio onde o mRNA se liga e onde ocorre o emparelhamento do codão (do mRNA) com o anti-codão (do tRNA). Cada ribossoma contém três sítios de ligação à molécula de tRNA: o **sítio A** (ou aminoacil), no qual o primeiro aminoácido é adicionado pelo aminoacil-tRNA; o **sítio P** (ou peptidil), que recebe o aminoacil-tRNA ligado ao polipéptido em crescimento; e o **sítio E** (ou de saída) onde se liga o tRNA vazio após libertar o aminoácido. Os aminoácidos ligam-se ao tRNA através da ativação pela enzima acil-tRNA sintetase. Existe uma acil-tRNA sintetase para cada aminoácido.

Nos eucariotas, os ribossomas podem estar livres no citoplasma ou ligados ao retículo endoplasmático rugoso.

Tal como acontece na transcrição, também a tradução ocorre em etapas consecutivas: **iniciação**, **alongamento** e **finalização**. Cada uma destas etapas será descrita de seguida.

Iniciação

A iniciação corresponde à primeira etapa da tradução, onde ocorre a ligação de uma molécula de RNA mensageiro à subunidade menor do ribossoma. Em eucariotas, a subunidade menor dos ribossomas, com a ajuda de fatores de iniciação, reconhece a estrutura *cap* na extremidade 5' do mRNA, à qual se liga. Em seguida, a subunidade menor move-se ao longo do mRNA até atingir o codão de iniciação – AUG. O codão de iniciação encontra-se no meio de uma sequência específica, denominada sequência de Kozak: 5'-ACCAUGG-3'.

Após o mRNA atingir o codão de iniciação, o fator de iniciação IF2 (componente proteico), ativado pelo GTP, conduz o primeiro tRNA, que transporta o aminoácido metionina, ao complexo de iniciação numa região do ribossoma denominada **sítio P**, estabelecendo-se a ligação entre o codão de iniciação e o anti-codão do tRNA. Em seguida, a subunidade maior do ribossoma liga-se e a tradução continua. Esta etapa da tradução encontra-se representada na figura 34.

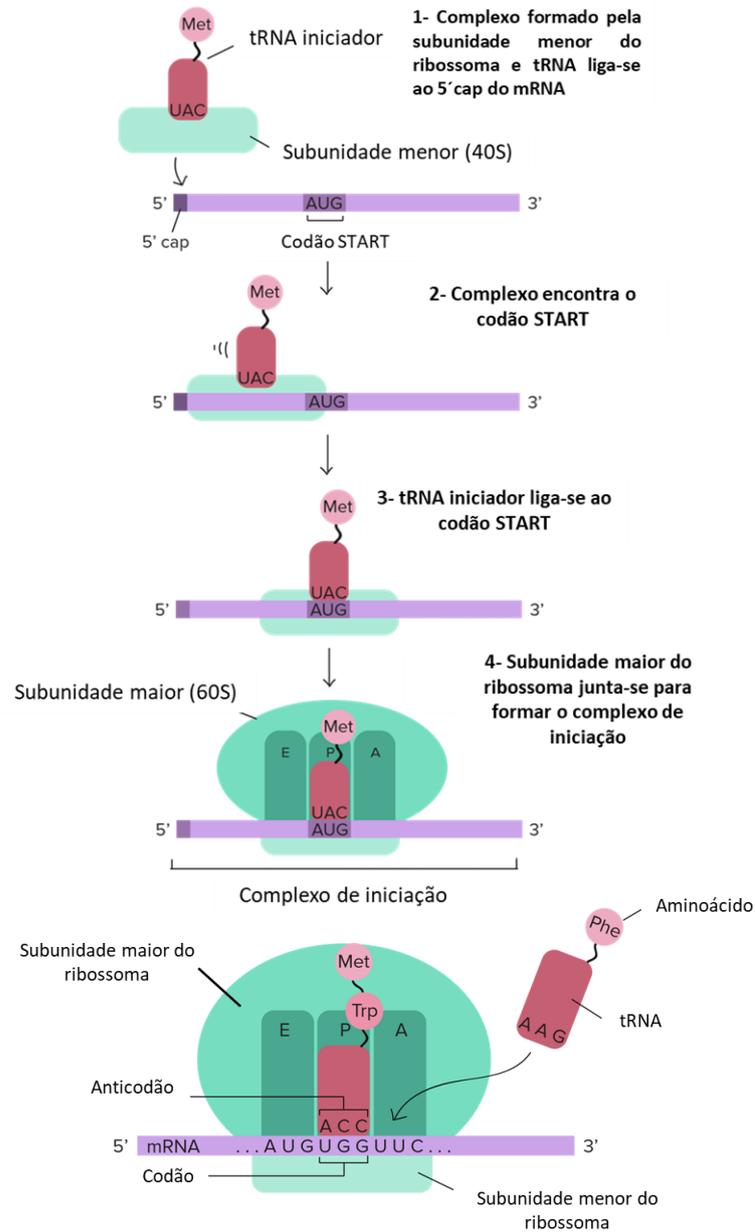


Figura 34 – Representação esquemática da sequência de acontecimentos que ocorre na primeira fase da tradução, a iniciação. Adaptado de <https://www.khanacademy.org/science/biology/gene-expression-central-dogma/translation-polypeptides/a/the-stages-of-translation>

Alongamento

Nesta etapa, os tRNA ligados aos aminoácidos, com a ajuda de fatores de alongamento EF-Tu (componente proteico), ativado pelo GTP, são transportados até ao ribossoma, sendo posicionados no sítio A pela complementaridade codão/anti-codão. Neste momento o fator de alongamento é libertado. De seguida, estabelece-se uma ligação peptídica entre os dois aminoácidos, quebrando-se a ligação entre o primeiro tRNA e o aminoácido que transportava. A ligação peptídica é catalisada pela enzima Peptidil transferase. Por seu lado, o tRNA que se encontrava no sítio A desloca-se para o sítio P, deixando o sítio A livre para receber um novo tRNA com outro aminoácido. Quando tal acontece, o tRNA que se encontrava no sítio P desloca-se para o sítio E, sendo posteriormente libertado para o citoplasma, sendo carregado com outro aminoácido.

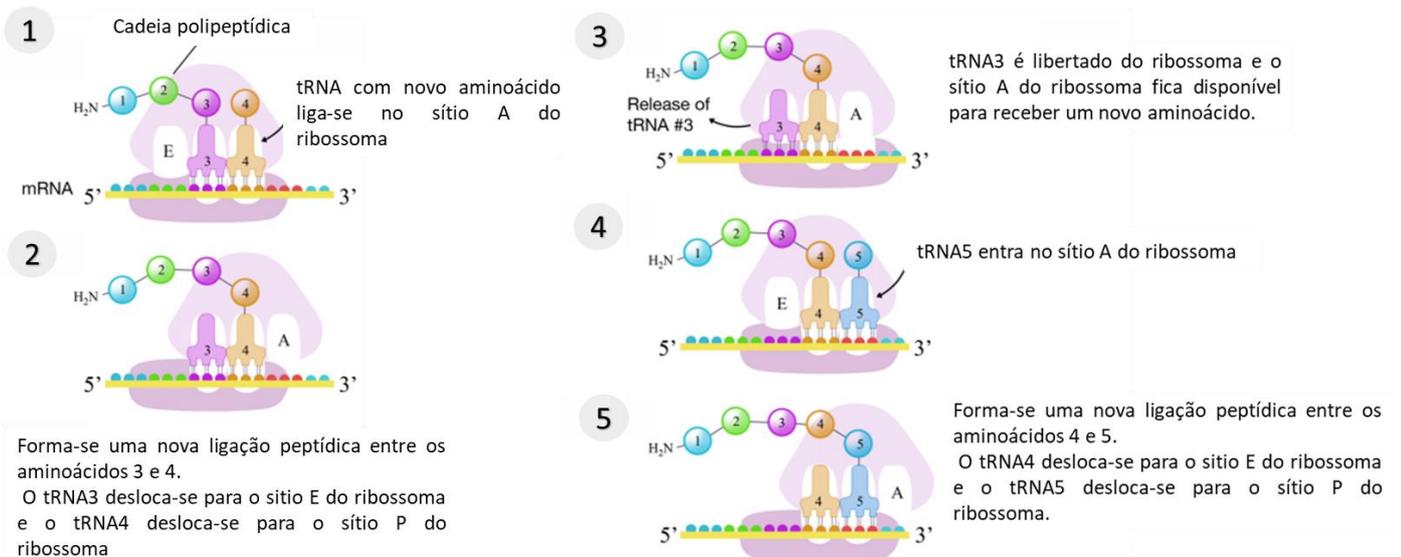


Figura 35 – Representação esquemática da sequência de acontecimentos que ocorrem na segunda fase da tradução, o alongamento. Adaptado de https://chem.libretexts.org/Courses/Fullerton_College/Introductory_Biochemistry/17%3A_Nucleic_Acid_s/17.08%3A_Flow_of_Genetic_Information/17.8.06%3A_Translation

Sabias que...



No processo de tradução, a etapa de alongamento evoluiu a uma velocidade de 15-20 aa/seg em bactérias e 5-10 aa/seg numa célula humana. Esta diferença é atribuída à maior complexidade dos ribossomas eucarióticos, bem como à regulação mais elaborada do processo de tradução.

A maior velocidade em bactérias permite uma adaptação rápida a mudanças ambientais, essencial para a sobrevivência em ambientes variáveis.



Finalização

A etapa de alongamento prossegue até que o ribossoma encontra na molécula de mRNA um codão STOP ou codão de finalização (**UAA**, **UAG** ou **UGA**). Nesse momento, o codão de finalização é reconhecido por proteínas específicas (fatores de libertação), as quais se ligam no sítio A do ribossoma, alterando a atividade da peptidil transferase (adição de H₂O em vez de aminoácido), produzindo a clivagem do péptido do último tRNA. Na fase de finalização, a cadeia polipeptídica liberta-se do tRNA, a molécula de mRNA liberta-se do ribossoma e as duas subunidades separam-se (figura 36).

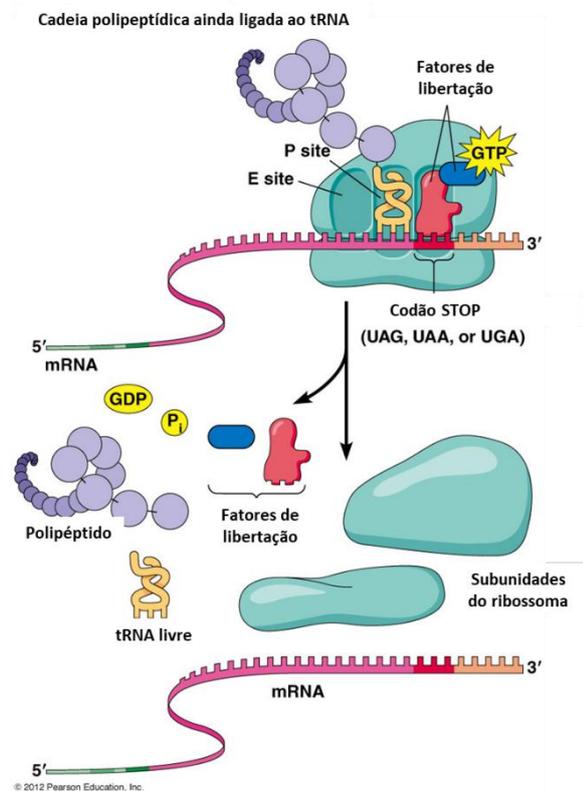


Figura 36 – Representação esquemática da última fase da tradução, a finalização. Adaptado de <https://unacademy.com/lesson/prokaryotic-translation-termination/767RPH57>

Nos organismos eucariotas, o mesmo mRNA pode ser traduzido simultaneamente por diversos ribossomas, dando origem a estruturas denominadas **polissomas**, que evoluem progressivamente ao longo da molécula de RNA (figura 37). Este processo é fundamental para aumentar a eficiência da tradução e da expressão génica, permitindo que uma única molécula de mRNA produza rapidamente muitas cópias da mesma

proteína. Isso é particularmente importante em células que necessitem aumentar rapidamente a quantidade de uma determinada proteína em resposta a estímulos ou durante certos estágios de desenvolvimento.

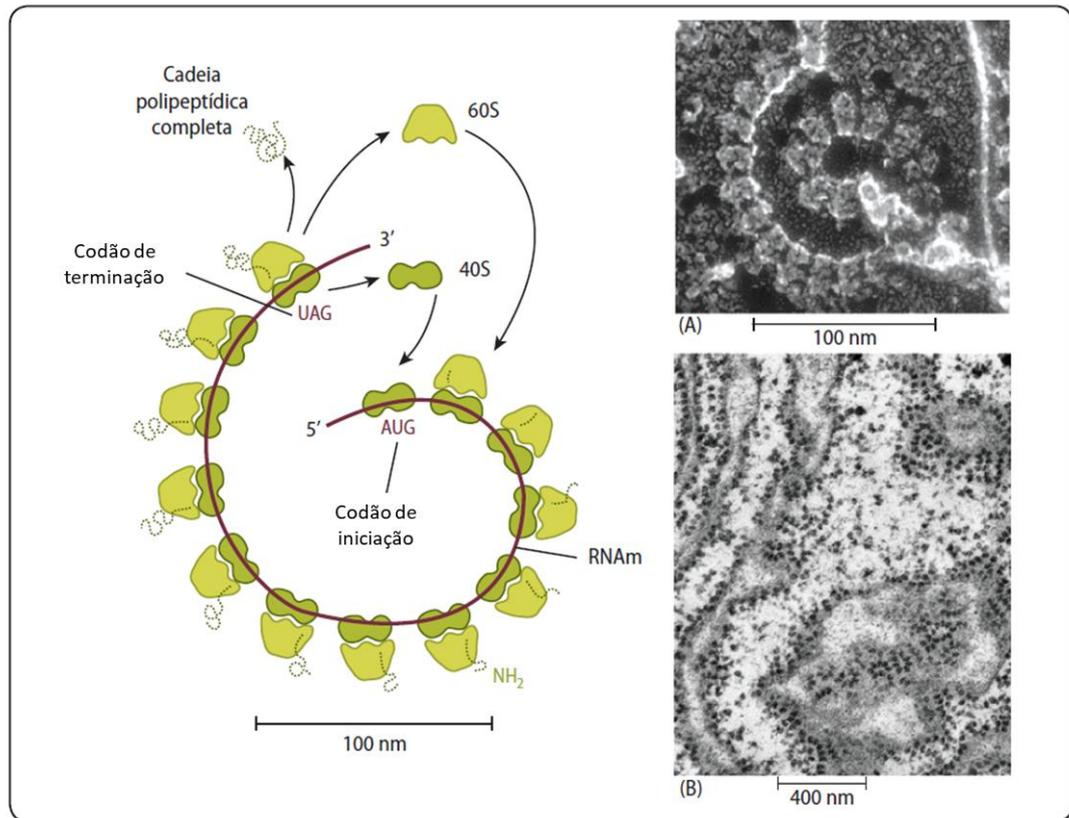


Figura 37 – Representação esquemática e microfotografia eletrônica de mRNA sendo traduzido simultaneamente por diversos ribossomos. Adaptado de Ribeiro, 2009.

Diferenças entre procariontos e eucariontos

Existem diferenças significativas no processo de transcrição e tradução em procariontos e eucariontos, estando as mesmas relacionadas com a localização, estrutura do mRNA, início da tradução, velocidade, subunidades ribossomais, modificações após a transcrição e complexidade das proteínas. As principais diferenças encontram-se resumidas na tabela III.

Tabela III – principais diferenças entre procariotas e eucariotas nos processos de transcrição e tradução.

Característica	Procariotas	Eucariotas
Localização	Transcrição e tradução ocorrem no citoplasma.	Transcrição ocorre no núcleo e tradução no citoplasma.
Estrutura do mRNA e início da tradução	Um único mRNA pode codificar várias proteínas diferentes. O mRNA possui vários sítios de ligação ao ribossoma (sequência de Shine-Dalgarno).	Cada mRNA codifica uma única proteína. O ribossoma reconhece o codão de iniciação presente na sequência de Kozak.
Velocidade da tradução	Tradução é mais rápida.	Tradução é mais lenta.
Subunidades ribossomais	Ribossomas 70S, compostos por uma subunidade menor 30S e uma subunidade maior 50S.	Ribossomas 80S, compostos por uma subunidade menor 40S e uma subunidade maior 60S.
Processamento do mRNA	O mRNA não passa por modificações significativas após a transcrição e é rapidamente traduzido.	O mRNA sofre várias modificações após a transcrição, incluindo adição de cap 5', <i>splicing</i> , e adição de uma cauda poli-A na extremidade 3'.
Processamento proteico	As proteínas recém-sintetizadas geralmente não passam por modificações pós-traducionais complexas.	As proteínas passam por modificações pós-traducionais, como fosforilação, glicosilação, acetilação e clivagem proteolítica.

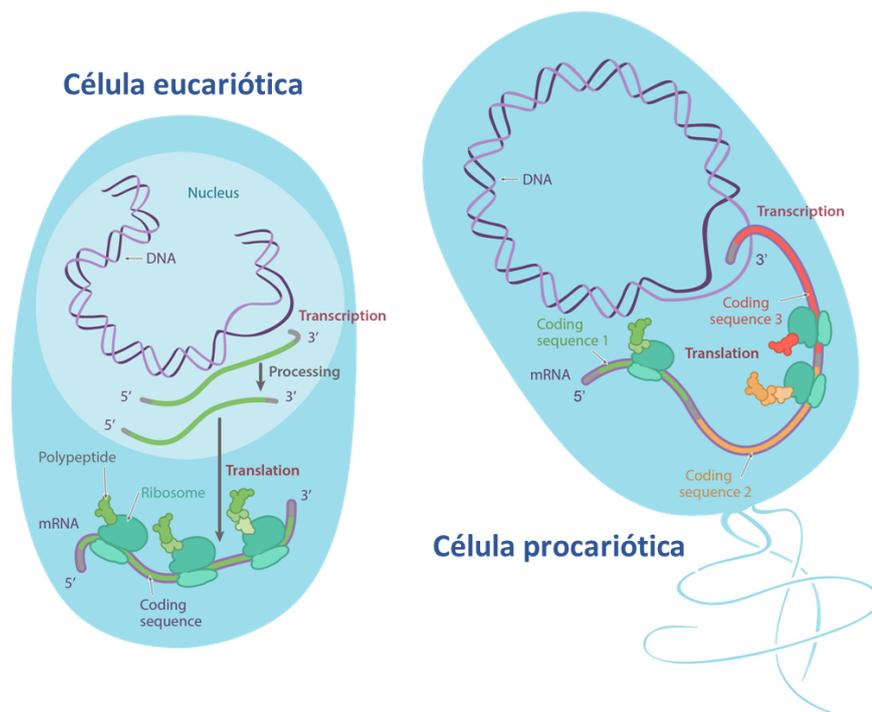


Figura 38 – Representação esquemática das diferenças no processo de síntese proteica entre uma célula eucariótica e procariótica. Adaptado de <https://www.khanacademy.org/science/biology/gen-expression-central-dogma/translation-polypeptides/a/the-stages-of-translation>.

O código genético

A conversão dos nucleótidos do mRNA em aminoácidos durante a tradução é mediada pelo código genético. Os 20 aminoácidos são codificados por combinações dos quatro nucleótidos diferentes, reunidos três a três, formando 64 combinações possíveis, ou seja, 64 codões possíveis (figura 39).

Há 61 codões que codificam os 20 aminoácidos e 3 codões (codões de terminação ou codões *stop*) que não codificam nenhum aminoácido. O codão de iniciação (AUG) codifica para o aminoácido metionina.

		Segunda Letra								
		U		C		A		G		
Primeira Letra	U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
		UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys	C
		UUA	Leu	UCA	Ser	UAA	STOP	UGA	STOP	A
		UUG	Leu	UCG	Ser	UAG	STOP	UGG	Try	G
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U	
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg	C	
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	A	
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg	G	
A	AUU	Iso	ACU	Thr	AAU	Asn	AGU	Ser	U	
	AUC	Iso	ACC	Thr	AAC	Asn	AGC	Ser	C	
	AUA	Iso	ACA	Thr	AAA	Lys	AGA	Arg	A	
	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg	G	
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U	
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C	
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly	A	
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly	G	

©BIOINNOVA
inovabiologia.com

Figura 39 – O código genético. 61 codões (três letras a negrito na imagem) codificam para os 20 aminoácidos. O codão de iniciação é o AUG (a verde na figura), que codifica para a metionina. Os três codões de terminação (ou STOP) não codificam aminoácidos (a vermelho na figura).

Diz-se que o código genético é **redundante/degenerado**, pois existem diferentes codões que codificam o mesmo aminoácido (com exceção da metionina e triptofano). Além disso, o código genético **não é ambíguo**, pois a cada codão corresponde um só aminoácido. O código genético é ainda considerado **universal**, pois é igual na maioria dos organismos.

Capítulo 7 – Mutações génicas

Apesar do DNA ser uma molécula estável e com um processo de replicação eficiente, algumas sequências de nucleótidos podem ser alteradas, de forma espontânea ou induzida, dando origem a **mutações génicas**. Estas mutações ocorrem num número reduzido de nucleótidos da molécula de DNA associado a um determinado gene, resultando no aparecimento de um novo alelo.

Estas mutações podem ocorrer em qualquer uma das componentes do gene, no entanto, quando ocorrem na região codificante (CDS – *coding sequence*), o seu significado biológico pode ter um impacto maior (dependendo do tipo de mutação).

As mutações génicas referem-se a qualquer alteração na sequência de nucleotídeos de um gene e podem envolver mudanças em um ou mais nucleotídeos. A nível molecular, estas mutações classificam-se em **mutações pontuais** ou **rearranjos**. Quando se dá a **alteração de um nucleótido apenas**, trata-se de uma mutação pontual, afetando apenas um par nucleotídico. Por outro lado, os rearranjos envolvem alterações que afetam **vários pares de bases**. Ao nível das mutações génicas, as mutações pontuais são, geralmente, as mais comuns, ocorrendo com mais frequência durante a replicação do DNA ou devido a danos causados por agentes mutagénicos.

No caso das mutações pontuais por **substituição** (também denominadas de substituição de base única), como o próprio nome indica, ocorre a substituição de um nucleótido por outro, podendo ter efeito, ou não, no fenótipo. Tendo em conta as consequências ao nível da função da respetiva proteína, as mutações pontuais por substituição classificam-se em três categorias: **mutação silenciosa** (ou sinónima), **mutação missense** e **mutação nonsense**:

- **Mutação silenciosa** – substituição de um par de bases por outro, conduzindo à formação de um códon diferente, mas que continua a codificar o aminoácido original (figura 39 B). Não se observa qualquer tipo de alteração na função da proteína nem ao nível do fenótipo. Constituem grande parte das mutações pontuais detetadas.

- **Mutação missense** – substituição de um par de bases por outro, conduzindo à formação de um códon que codifica um aminoácido diferente do original. Se o novo aminoácido codificado for quimicamente semelhante ao original, a função da proteína não sofre alterações significativas e trata-se de uma **mutação missense neutra ou**

conservadora (figura 40 C). Por outro lado, se o aminoácido original for substituído por outro quimicamente diferente, a função da proteína sofre alterações e trata-se de uma mutação missense não conservadora (figura 40 D).

Sabias que...



Nas uvas, a ausência de sementes (apirenia) pode ocorrer naturalmente devido a dois processos fisiológicos distintos: **partenocarpia** (os frutos desenvolvem-se completamente sem fertilização e nunca formam sementes) ou **estenoespermocarpia** (ocorre fertilização inicial dos óvulos, mas o embrião e as sementes não se desenvolvem completamente).

As condições de partenocarpia e estenoespermocarpia geralmente envolvem mutações *missense* no gene *VviAGL11* que regula o desenvolvimento do fruto e das sementes.



Estenoespermocarpia

Partenocarpia

- **Mutação nonsense** – substituição de um codão, que codifica para um aminoácido, por um codão STOP, conduzindo à terminação prematura da cadeia polipeptídica. Por norma, a proteína não é funcional (figura 40 E).

Sabias que...



O tamanho do grão do arroz é determinado por uma mutação *nonsense* no gene *GRAIN SIZE 3 (GS3)*

O gene **GS3** desempenha um papel crítico na regulação do crescimento e desenvolvimento dos grãos de arroz, atuando como um regulador negativo do crescimento do grão. Mutações no gene **GS3** podem resultar em uma perda ou alteração de sua função regulatória. Por exemplo, mutações *nonsense* que inativam ou reduzem a função do GS3 resultam em grãos mais longos, porque o efeito limitador sobre o crescimento é atenuado.



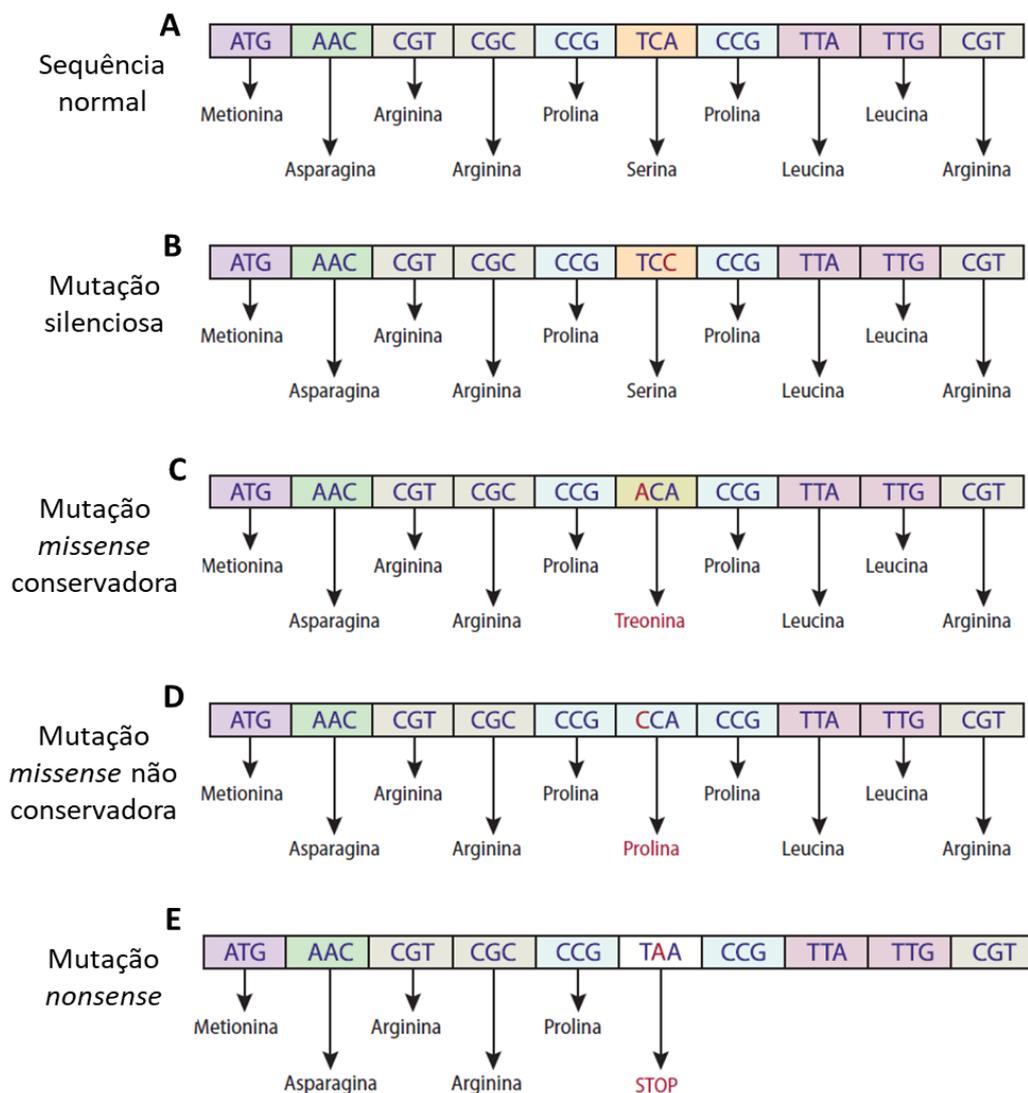
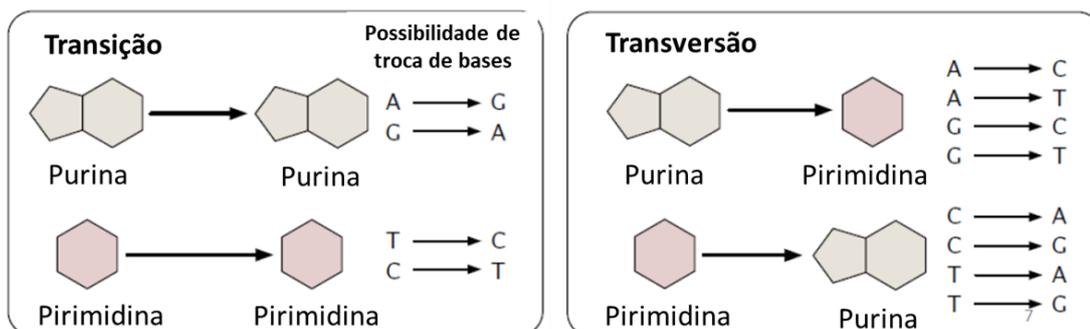


Figura 40 – Classificação das mutações pontuais por substituição de acordo com o seu efeito na tradução. Adaptado de Ribeiro, 2009.

As mutações por substituição podem ocorrer por transição ou transversão. No caso das **transições**, uma purina é substituída por outra purina diferente ou, em alternativa, uma pirimidina é substituída por outra pirimidina diferente. Nas **transversões**, uma purina é substituída por uma pirimidina ou uma pirimidina é substituída por uma purina.



As mutações gênicas por **inserção** ou **deleção** correspondem à adição ou a remoção, respectivamente, de um ou mais nucleótidos durante a transcrição do DNA para mRNA. Estas mutações caracterizam-se por provocar alterações na *Reading frame* (“grelha de leitura”), obtendo-se uma sequência de aminoácidos diferente da original, uma vez que o ribossoma identifica codões de forma desfasada a partir do local onde ocorreu a alteração (a 3’ da mutação). Geralmente originam-se proteínas não funcionais. Estes tipos de mutações classificam-se em **Mutações frameshift**.

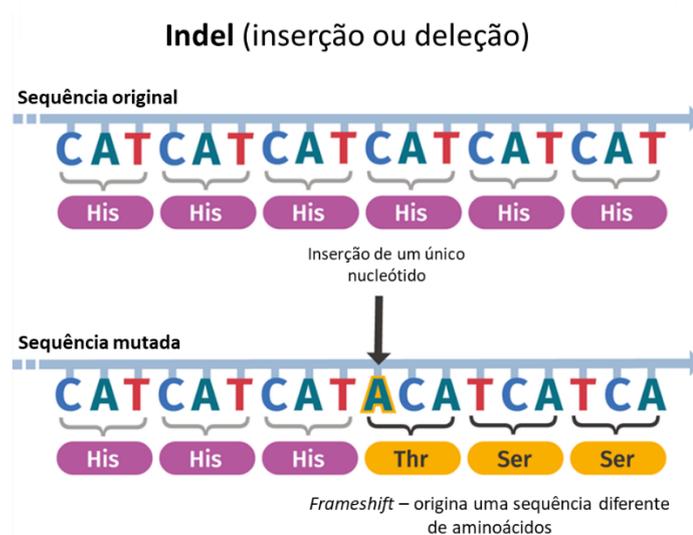


Figura 41 – Representação de uma mutação *frameshift*. Adaptado de [https://sigmaearth.com/pt/variantes-e-mutações do DNA genômico](https://sigmaearth.com/pt/variantes-e-mutações-do-dna-genômico).

No entanto, nem todas as inserções e deleções conduzem a *frameshifts*. Uma vez que os codões são compostos por três nucleótidos, inserções e deleções consistindo de múltiplos de três nucleótidos não vão alterar a *reading frame*, embora a adição ou remoção de um ou mais aminoácidos possa afetar o fenótipo. Estas mutações são chamadas de **inserções ou deleções em frame**.

Quando uma mutação ocorre em sequências codificantes do DNA, estas podem ter efeitos diretos e significativos na função da proteína resultante. No entanto, se as mutações ocorrerem em sequências reguladoras ou outras não codificantes, as consequências ao nível da proteína têm menor impacto, alterando normalmente a quantidade em que as proteínas são produzidas e não a sua função.

Tendo em conta o efeito ao nível da função da proteína, as mutações podem ser classificadas em **loss-of-function** (perda de função) e **gain-of-function** (ganho de função). As mutações *loss-of-function*, tal como o nome indica, causam a redução ou perda completa da função normal de uma proteína. As mutações *loss-of-function* são frequentemente recessivas, e os indivíduos diploides devem ser homocigóticos para a mutação para poderem exibir os efeitos da perda da proteína funcional. No caso das mutações *gain-of-function*, dão origem a uma nova função ou aumentam a atividade normal de uma proteína. Estas mutações são frequentemente dominantes na sua expressão.

Nomenclatura para a descrição das mutações genéticas

A nomenclatura para a descrição das mutações genéticas segue diretrizes estabelecidas para garantir clareza e consistência na comunicação científica. A nomenclatura padrão usada atualmente é a recomendada pela *Human Genome Variation Society* (HGVS). Para evitar erros na interpretação acerca da alteração na sequência, a descrição é precedida por uma letra indicando o tipo de sequência onde ocorreu a mutação:

"g." para sequência genómica.

"c." para sequência de DNA codificante (cDNA), ou seja, a sequência do mRNA correspondente ao gene.

"p." para sequência da proteína resultante da mutação.

"r." para sequência de RNA.

"m." para sequência do DNA mitocondrial.

Exemplos:

1. **Substituição de aminoácidos**

p.Gly12Val - Indica que na posição 12 da proteína, a glicina (Gly) foi substituída por valina (Val).

2. **Deleções de aminoácidos**

p.Lys76del – Indica que o aminoácido lisina (Lys) foi eliminado na posição 76.

3. Inserções de aminoácidos

p.Arg76_Gly77insSer – Indica que entre a arginina (Arg) na posição 76 e a glicina (Gly) na posição 77, foi inserida uma serina (Ser).

4. Mutação frameshift

p.Gly12fs - Indica uma mudança na grelha de leitura a partir do aminoácido 12, resultando em uma proteína com uma sequência completamente alterada após essa posição.

5. Substituição de nucleótidos

c.76A>T - Significa que na posição 76 do cDNA, uma adenina (A) foi substituída por uma timina (T).

6. Deleções de nucleótidos

c.76_78del - Indica que as bases nas posições 76 a 78 foram eliminadas.

7. Inserções de nucleótidos

c.76_77insT - Indica que entre as posições 76 e 77, uma timina (T) foi inserida.

8. Indels (Inserção e Deleção combinadas)

c.76_77delinsTT - Significa que uma deleção entre as posições 76 e 77 foi seguida por uma inserção de duas timinas (TT).

Origem das mutações génicas

Dependendo da sua origem, as mutações podem ser classificadas como **espontâneas** ou **induzidas**. Mutações espontâneas resultam de alterações moleculares que ocorrem espontaneamente no DNA, sem uma causa aparente. Pelo contrário, as mutações induzidas resultam da exposição a agentes mutagénicos exógenos.

Mutações génicas espontâneas – as mutações espontâneas podem ser o resultado de erros durante a replicação do DNA, podendo ocorrer substituição, inserção ou deleção de nucleótidos, ou o resultado de alterações químicas espontâneas no DNA, onde se destacam a depurinação e desaminação. No primeiro caso, ocorre uma quebra da ligação entre uma base purina (A ou G) e o açúcar desoxirribose, resultando na perda da base e produzindo um sítio apurínico. Nestes casos, normalmente é incorporada uma

adenina na cadeia de DNA recém-sintetizada, oposta ao sítio apurínico, conduzindo a um erro incorporado, sendo o mesmo transformado num erro de replicação no próximo ciclo de replicação. A desaminação corresponde à perda de um grupo amina (NH₂) numa base. Esta modificação pode alterar as propriedades de emparelhamento de uma base. Por exemplo, a desaminação da citosina produz o uracilo, que emparelha com a adenina durante a replicação. Após outro ciclo de replicação, a adenina emparelha com a timina, criando um par T::A em vez do par original C::G.

Mutações gênicas induzidas – as mutações podem resultar da exposição a agentes mutagênicos exógenos. Os agentes mutagênicos podem ser classificados em: a) agentes físicos - radiações de alta energia capazes de penetrar os tecidos e provocar lesões no DNA, como radiações ionizantes (raios X e os raios γ); b) agentes químicos - substâncias que reagem com o DNA e provocam modificações químicas nas suas bases (exemplo, inibidores mitóticos); c) agentes biológicos - ferramentas de engenharia genética, como CRISPR/Cas9.

Em que células pode ocorrer uma mutação?

Nos organismos multicelulares, as mutações podem ocorrer nas células somáticas ou germinativas.

Células somáticas são todas as células do organismo que não estão envolvidas na produção de gametas. No caso particular das plantas, isso inclui a maior parte dos tecidos, como folhas, caules, raízes e flores (exceto as células que darão origem aos gametas). As mutações somáticas podem ocorrer em qualquer célula somática do organismo e geralmente afetam apenas a célula onde a mutação ocorreu e as células descendentes dessa célula por mitose, dando origem a um aglomerado de células diferentes (setor mutante). No caso das plantas, em termos fenotípicos, o setor mutante só poderá ser observado se contrastar visualmente com o aspeto normal das células tipo selvagem (Figura 42). As mutações somáticas não são transmitidas para a descendência, sendo os seus efeitos limitados ao organismo individual onde a mutação ocorreu. No entanto, em plantas, partes mutantes podem ser propagadas vegetativamente, perpetuando a mutação nos novos indivíduos clonados.



Figura 42 – Exemplos de mutações somáticas. A - mutação somática que alterou a expressão de genes responsáveis pela cor das pétalas numa parte da flor. As células com a mutação expressam uma cor diferente das células não mutadas, resultando numa flor bicolor; B - mutação somática que ocorreu durante o desenvolvimento inicial do fruto, levando à fusão parcial de dois frutos adjacentes. Tal pode ter sido causado por uma alteração na divisão celular ou no desenvolvimento do tecido meristemático, onde as células são inicialmente pluripotentes; C – mutação somática que afetou a pigmentação da casca do fruto apenas numa parte, resultando numa fruta com duas cores diferentes. A mutação pode ter ocorrido numa célula precursora do fruto durante seu desenvolvimento inicial, herdando as células essa característica.

As células germinativas são aquelas que dão origem aos gametas. Nas plantas, as células germinativas estão localizadas nas estruturas reprodutivas, como os estames e o ovário. Mutações que ocorrem em células germinativas podem ser transmitidas para a descendência, sendo a mutação herdada pelas gerações seguintes. A poliploidia é uma mutação em que o número total de cromossomas de um organismo apresenta múltiplas cópias nas células germinativas, podendo ocorrer durante a formação dos gametas devido a erros na segregação dos cromossomas durante a meiose. Plantas poliploides, normalmente, apresentam características desejáveis, como frutos maiores ou maior vigor. A banana e o trigo são exemplos de espécies que sofreram este tipo de mutação.

Tanto as mutações somáticas como as germinativas têm sido exploradas ao nível do melhoramento de plantas. As mutações somáticas podem ser usadas para originar variedades ornamentais únicas, enquanto as mutações germinativas são essenciais para o desenvolvimento de novas cultivares de plantas agrícolas.

Capítulo 8- Marcadores moleculares

No contexto do melhoramento de plantas, a seleção de características de interesse agronômico com base em marcadores genéticos teve início nos anos 80, com os trabalhos desenvolvidos por Tanksley (Tanksley et al., 1983) em tomate (*Solanum lycopersicum*). Nesse estudo, os investigadores demonstraram como os marcadores moleculares podiam ser utilizados para mapear características genéticas em espécies vegetais, estabelecendo uma base para a identificação de *loci* associados a características agronômicas importantes.

A utilização de marcadores moleculares no melhoramento de plantas apresenta várias vantagens comparativamente à tradicional observação das características fenotípicas, também denominados marcadores morfológicos. Entre as vantagens destaca-se a **maior precisão**, visto que os marcadores morfológicos são influenciados por fatores ambientais; **redução do tempo necessário para selecionar as características de interesse**, uma vez que os marcadores moleculares permitem a seleção de plantas em estágios iniciais (como na fase de plântula), reduzindo significativamente o tempo necessário para o desenvolvimento de novas variedades; **identificação de características complexas**, como o rendimento do grão ou a qualidade do fruto, que são frequentemente controladas por múltiplos genes (QTLs) e influenciadas por fatores ambientais.

Marcador molecular corresponde a uma variação (polimorfismo) no genoma da planta, passível de ser visualizada recorrendo a técnicas de biologia molecular (PCR – reação de polimerização em cadeia, enzimas de restrição, eletroforese, sequenciação), diretamente associado a uma variação fenotípica. Os marcadores moleculares podem ser utilizados para identificar diferentes espécies ou indivíduos portadores de determinada característica fenotípica.

Polimorfismo é o termo usado para descrever a ocorrência frequente e simultânea de **variação genética** dentro de uma população de mais de um genótipo. Os polimorfismos são observáveis quando um gene, com um efeito distinto e reconhecível, ocorre numa população a uma taxa que é muito frequente para ser apenas atribuída ao acaso ou a uma mutação. Estas variantes genéticas devem estar presentes numa população com uma frequência de pelo menos 1%. Os **SNPs** (do inglês *Single Nucleotide Polymorphisms*) são o tipo de variação genética mais comum entre os indivíduos de uma

espécie. Um SNP representa uma diferença de um único nucleótido no DNA (A, T, C ou G) numa posição específica do genoma, sendo mantido de geração em geração. Os SNPs são extremamente comuns no genoma, podendo ocorrer em regiões codificantes ou não codificantes, e apresentam uma estabilidade muito elevada ao longo de gerações.

Além de substituições de nucleótidos, na base dos polimorfismos podem ainda estar inserções ou deleções (InDels), duplicações, translocações ou inversões.

Características desejáveis de um marcador molecular:

- **Elevado grau de polimorfismo:** o marcador deve ser polimórfico, ou seja, deve apresentar variações entre os indivíduos para permitir a distinção entre diferentes alelos ou genótipos.
- **Heritabilidade codominante:** marcadores codominantes permitem a distinção entre indivíduos homocigotos e heterocigotos, proporcionando informações completas sobre o genótipo.
- **Elevada ocorrência no genoma:** Idealmente, um marcador não deve ser tão comum a ponto de ocorrer em grande frequência no genoma, porque isso poderia limitar sua utilidade. Marcadores muito comuns podem não ser informativos o suficiente para distinguir entre diferentes genótipos ou para associar com precisão características fenotípicas específicas.
- **Distribuição aleatória no genoma:** o marcador deve estar distribuído aleatoriamente ao longo do genoma, permitindo a cobertura de todas as regiões do genoma (codificantes e não-codificantes), o que é essencial para um mapeamento genético abrangente e para a identificação de associações entre marcadores e características fenotípicas.
- **Comportamento neutro seletivo:** o marcador deverá ser neutro a variações nas condições ambientais, não deve estar sujeito à seleção natural, para que possa refletir com precisão a variação genética.

- **De fácil acesso/disponibilidade e rápida análise:** o processo de detecção do marcador deve ser relativamente simples, rápido e económico, preferencialmente sem a necessidade de técnicas complexas ou equipamento especializado.

- **Elevada reprodutibilidade:** a técnica usada para identificar o marcador deve ser altamente reprodutível, permitindo que os mesmos resultados sejam obtidos em diferentes ensaios.

Classificação dos Marcadores genéticos

A classificação dos marcadores genéticos pode ser feita tendo em conta as técnicas utilizadas no seu estudo, tendo em conta a heritabilidade, a natureza do marcador ou pela variação genética que eles detetam. Cada tipo de marcador tem as suas vantagens e limitações, sendo que a escolha é feita tendo em conta os objetivos específicos do estudo ou do programa de melhoramento genético.

1. Classificação tendo em conta as técnicas base:

Tendo em conta a tecnologia utilizada, os marcadores genéticos classificam-se em dois grandes grupos: **marcadores morfológicos** e **marcadores moleculares**.

1.1. Marcadores morfológicos - Os marcadores morfológicos baseiam-se em características fenotípicas como a cor da flor, forma do fruto, altura da planta ou outras características agronómicas importantes. Os marcadores morfológicos são de fácil utilização, sem necessidade de recorrer a tecnologia específica. No entanto, são amplamente influenciados pelos fatores ambientais e dependem da fase de desenvolvimento da planta, por vezes sendo necessário um longo período até que as características possam ser selecionadas.

1.2. Marcadores moleculares – Dizem respeito a variações no DNA ou RNA, não sendo influenciados pelos fatores ambientais, o que os torna muito confiáveis. Tendo em conta este tipo de classificação, os marcadores moleculares podem ser classificados em bioquímicos, baseados na técnica PCR (do inglês *PCR-based techniques*) e não baseados na técnica de PCR (do inglês *non-PCR-based techniques*).

1.2.1. Bioquímicos

Os marcadores bioquímicos, ou isozimas, são formas diferentes da mesma enzima, codificadas por genes diferentes, mas têm as mesmas funções. São variações alélicas de enzimas, ou seja, têm diferenças na sequência de aminoácidos e, por isso, as frequências genéticas e genotípicas podem ser estimadas com marcadores bioquímicos. Trata-se de marcadores codominantes, de fácil utilização e custo reduzido. No entanto, detetam um menor número de polimorfismos e são afetados pelos métodos de extração utilizados, o tipo de tecido do material vegetal e pelo estágio de crescimento da planta.

Exemplo: Algumas isozimas estão associadas a características específicas, como resistência a stresses abióticos (seca, salinidade, temperaturas extremas) ou bióticos (organismos patogénicos). A deteção dessas isozimas permite a seleção de plantas com maior resistência ou tolerância antes dos testes em campo, acelerando o processo de melhoramento.

1.2.2. non-PCR-based techniques – RFLP (Restriction Fragment Length Polymorphism)

Marcador baseado na variação no comprimento de fragmentos de DNA após a digestão com enzimas de restrição. Esta técnica baseia-se na análise de diferenças nas sequências de nucleótidos que afetam os locais de reconhecimento de enzimas de restrição, ou seja, polimorfismos que alteram uma sequência reconhecida por uma enzima de restrição.

O primeiro passo consiste na extração e purificação do DNA do organismo em estudo, sendo o mesmo posteriormente incubado com enzimas de restrição, utilizadas para cortar o DNA em *loci* específicos. Dessa incubação resulta um elevado número de fragmentos de DNA com diferentes tamanhos, sendo os mesmos separados num gel de

agarose, no qual serão visualizadas uma série de bandas. Cada banda representa um fragmento de diferente tamanho. Após a separação em gel de agarose, os fragmentos de DNA são transferidos do gel para uma membrana de nitrocelulose, um processo conhecido como *Southern blotting*. A membrana é então incubada com uma sonda de DNA marcada (radioativa ou fluorescente), sendo a sonda complementar a uma sequência específica de interesse no DNA. A sonda liga-se aos fragmentos de DNA na membrana que contêm a sequência alvo, sendo os mesmos visualizados.

Esta técnica apresenta como desvantagens o facto de necessitar de uma grande quantidade de DNA, ser muito dispendiosa e ser uma técnica relativamente complexa e demorada.

1.2.3. PCR-based techniques

A técnica PCR (do inglês *Polymerase chain reaction*) foi desenvolvida em 1983 por Cary Mullis e consiste na amplificação de fragmentos de DNA específicos. Desnaturação, *annealing* e extensão são as três etapas principais desta técnica. Existem vários tipos de marcadores moleculares que utilizam a técnica PCR para amplificação dos fragmentos, dos quais se destaca os **microsatélites** ou **SSRs** (do inglês *Simple Sequences Repeats*) e **SNP** (do inglês *Single Nucleotide Polymorphisms*).

PCR - Reação de Polimerização em Cadeia

Através da técnica PCR, um processo automatizado, é possível uma amplificação fácil, muito rápida e seletiva de fragmentos de DNA específicos.

A PCR é uma técnica muito utilizada em biologia molecular para amplificar sequências específicas de DNA, tornando possível a produção de milhões de cópias de uma região alvo do DNA num curto período. **Uma PCR necessita essencialmente de** (figura 43):

- uma amostra de **DNA ou cDNA** (DNA complementar resultante de uma reação de transcrição reversa utilizando mRNA como *template*) que apresenta a sequência que se pretende amplificar.

- dois oligonucleótidos iniciadores (do inglês **primers**) correspondentes a pequenas sequências de nucleótidos (~20 nts) que funcionam como ponto de partida para a síntese das novas cadeias de DNA. Para aquisição destes *primers* é, obviamente, fundamental definir o fragmento que se pretende amplificar de modo a que sejam desenhados os *primers* em regiões que flanqueiam esse fragmento. Na região a 5' do fragmento de interesse estará posicionado o *primer forward* e a 3' o *primer reverse*.

- quatro **nucleótidos – dNTPs**.

- uma enzima **DNA polimerase** responsável pela síntese das novas cadeias. Das DNA polimerases utilizadas em PCR destaca-se a *Taq* DNA polimerase. Esta enzima foi identificada na bactéria *Thermus aquaticus*, uma bactéria termófila descoberta nas fontes termais do Parque Nacional de Yellowstone (USA), e apresenta como característica o facto de resistir a temperaturas muito elevadas sem sofrer desnaturação.

- **tampão da enzima**, responsável pela manutenção de um ambiente ótimo para que a enzima apresente a sua máxima eficiência. Este tampão é específico de cada enzima sendo por isso fornecido pela empresa aquando da aquisição da enzima.

- **magnésio Mg²⁺** (na forma de MgCl₂) desempenha um papel crucial na PCR uma vez que é um cofator essencial para a DNA Polimerase, ou seja, é um cofator necessário para a atividade das DNA polimerases, ajudando a ativar a enzima e fazendo com que a mesma sintetize as novas cadeias de DNA.

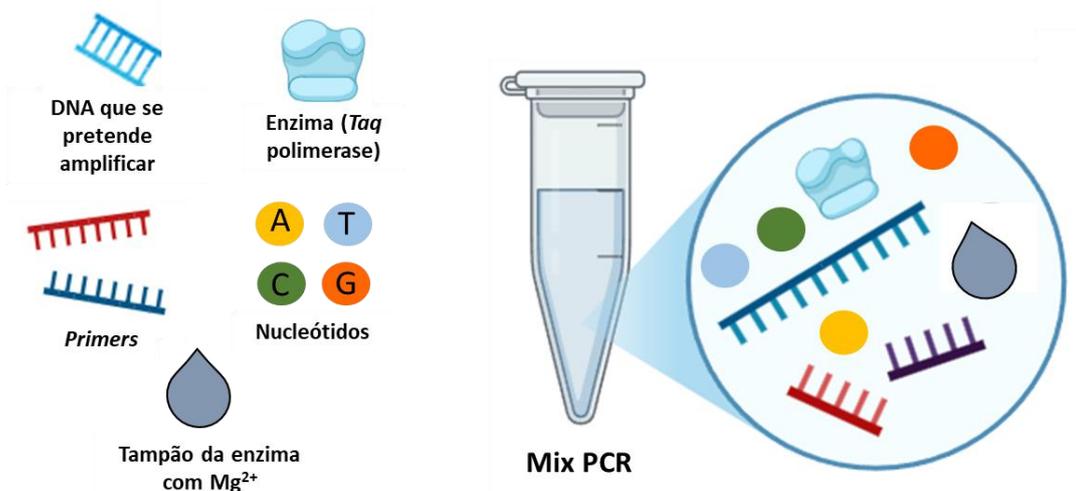
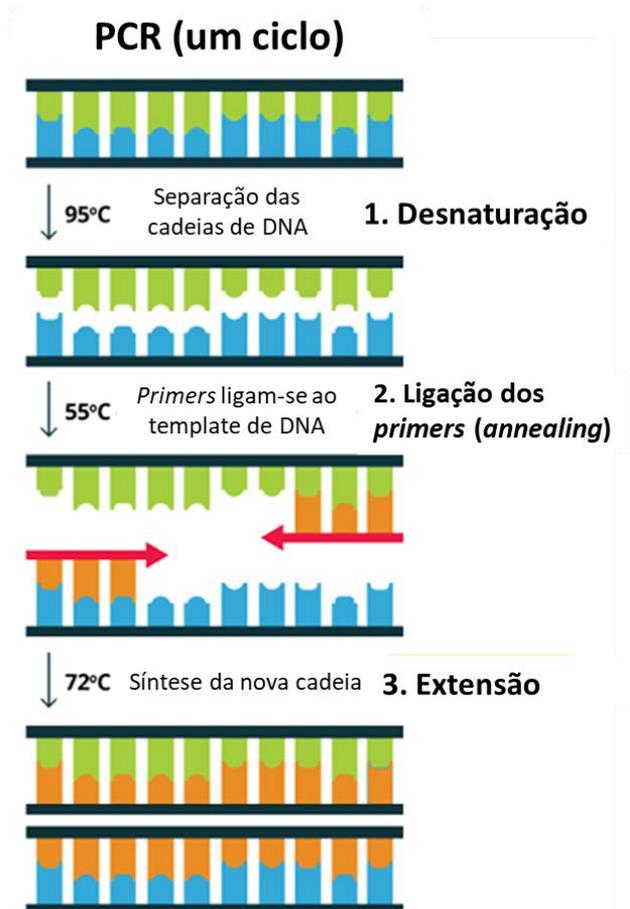


Figura 43 – Figura ilustrativa com os reagentes necessários para a realização de uma reação PCR. Imagem criada em BioRender.com.

Uma reação de PCR consiste numa série de ciclos que integram diferentes temperaturas. Cada ciclo envolve um passo a uma temperatura elevada promovendo a **desnaturação do DNA**, seguindo-se um passo que implica a descida da temperatura de modo a promover a **hibridação dos primers (annealing)** e um último passo que implica novamente a subida da temperatura para que ocorra a **síntese das novas cadeias de DNA (extensão)** (Fig. 44).

1. **Desnaturação (94-98°C):** Nesta etapa, a dupla hélice do DNA é separada em duas fitas simples através do aumento da temperatura, quebrando as ligações de hidrogénio entre as bases nitrogenadas.
2. **Annealing (50-65°C):** À semelhança do que acontece na replicação do DNA a nível celular, também aqui a enzima precisa de um ponto de partida para iniciar a síntese de DNA. Esse ponto de partida é dado pelos *primers* que se ligam a cada uma das cadeias simples de DNA. Para tal, é necessário diminuir a temperatura para intervalos entre 50-65 °C, consoante a composição da sequência dos *primers* a utilizar.



3. **Extensão (72°C):** O ciclo completa-se com a síntese de DNA pela enzima DNA polimerase. Utilizando os nucleótidos livres adicionados à solução, esta enzima faz a alongação dos *primers* usando como molde a cadeia a que cada *primer* está ligado, sempre no sentido de 5'→3' (tal como na replicação do DNA a cadeia sintetizada terá orientação antiparalela e será complementar à cadeia molde).

Durante cada ciclo adicional de PCR, as moléculas são duplicadas levando a uma amplificação exponencial do DNA. Cada n ciclos leva à produção de 2^n moléculas de DNA. Com esta técnica é possível obter uma grande quantidade de DNA alvo a partir de uma mistura complexa em que o mesmo pode estar presente numa cópia única.

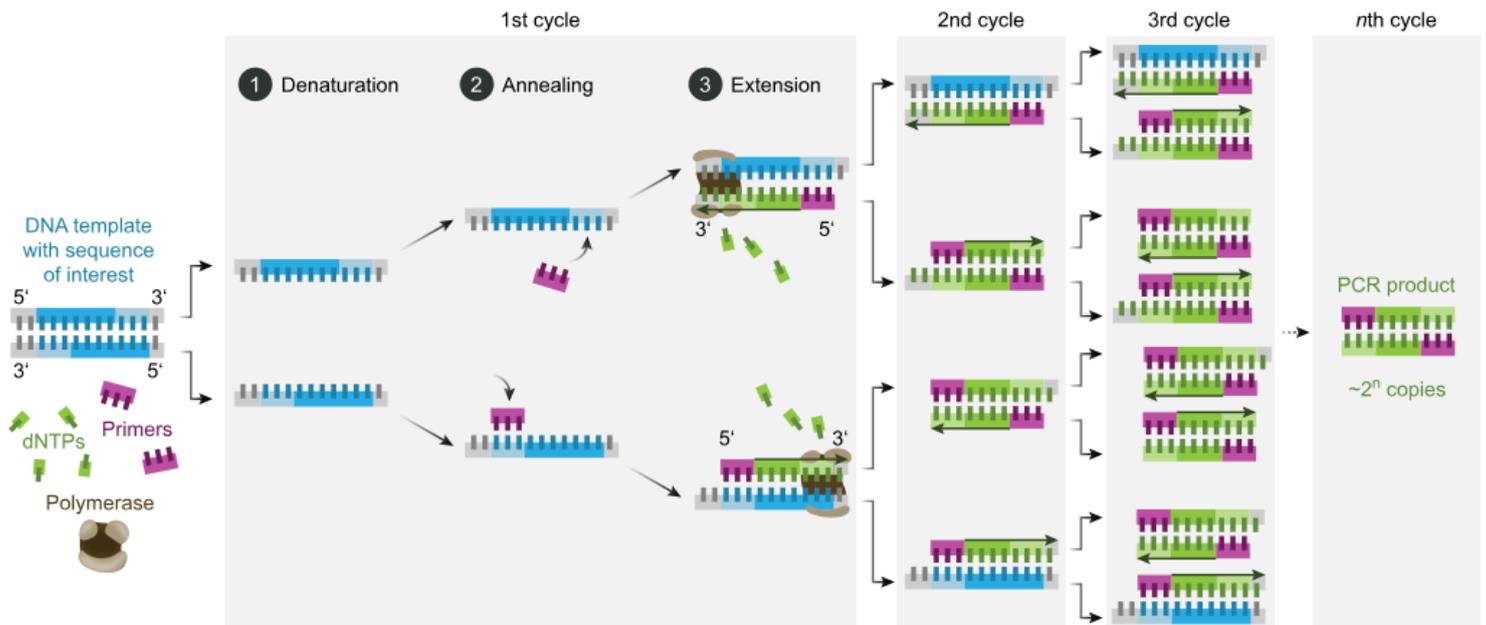


Figura 45. Esquema de amplificação de DNA por PCR.

Fonte: https://pt.wikipedia.org/wiki/Rea%C3%A7%C3%A3o_em_cadeia_da_polimerase

Para a realização de uma PCR foi desenvolvido um equipamento que permite realizar esta sequência de temperaturas várias vezes – ciclos de PCR. Estes equipamentos – termocicladores estão disponíveis em diversos formatos e são comercializados por várias marcas, tendo em comum o volume de reação, possibilitando apenas a utilização de tubos/strips ou placas de volume máximo de 200 μ L. A figura 46 ilustra o programa da reação PCR programado num termociclador.

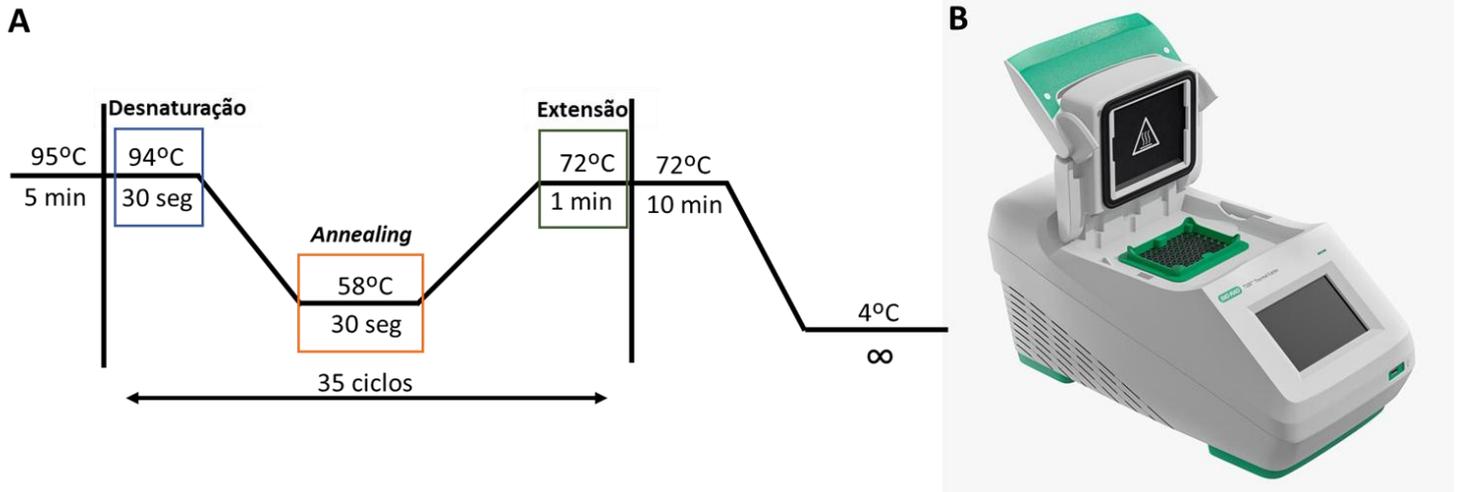


Figura 46 – A) Exemplo de programa de reação da PCR com identificação das três fases principais. B) Termociclador, equipamento utilizado em PCR.

A amplificação do DNA por PCR, assim como a qualidade e tamanho dos fragmentos é, na grande maioria das vezes, avaliada através da realização de uma eletroforese em gel de agarose. Esta técnica permite a separação do DNA amplificado num gel de agarose. A presença de uma banda no tamanho esperado confirma a amplificação do fragmento de DNA específico.

SNP (do inglês *Single Nucleotide Polymorphisms*) - Os Polimorfismos de um único nucleótido são uma das formas mais comuns de variação genética nos organismos. Representam alterações numa única base do DNA em posições específicas no genoma. Por exemplo, numa determinada sequência, a troca de uma citosina (C) por uma timina (T) caracteriza um SNP. Os SNPs são muito utilizados no melhoramento de plantas para identificar associações entre variações genéticas e características fenotípicas, como resistência a agentes patogénicos, produtividade agrícola e adaptação a condições ambientais.

Detecção e Análise de SNPs

A técnica de PCR é um dos métodos mais rápidos e mais utilizados para detetar SNPs, sendo utilizada para amplificar uma região específica do genoma que contém o SNP de interesse. Para tal são utilizados *primers* específicos para os diferentes alelos do SNP.

Microssatélites ou SSRs (do inglês *Simple Sequence Repeat*) - Trata-se de pequenas sequências de DNA repetidas, ou seja, sequências de nucleótidos que se repetem em *tandem*. A unidade repetida pode ser composta por mononucleótidos (1 nucleótido em repetição), dinucleótidos (2 nucleótidos em repetição), trinucleótidos (3 nucleótidos em repetição), tetranucleótidos (4 nucleótidos em repetição), pentanucleótidos (5 nucleótidos em repetição) ou hexanucleótidos (6 nucleótidos em repetição).

SSRs são marcadores com um nível muito elevado de polimorfismo, ou seja, há uma grande variação no número de repetições entre diferentes indivíduos dentro de uma população, o que os torna excelentes para estudos de diversidade genética. Encontram-se amplamente distribuídos pelo genoma, tanto em regiões codificantes quanto em não codificantes, contudo, são mais comuns em regiões não codificantes. SSRs são marcadores co-dominantes, o que significa que é possível distinguir entre indivíduos heterozigotos e homozigotos.

Deteção e Análise de SSRs

A primeira etapa para detetar microssatélites consiste na extração do DNA genómico das células do organismo em estudo. Em seguida, as regiões que contêm os SSRs são amplificadas por **PCR**, utilizando *primers* específicos para as regiões flangeadoras da sequência repetida. A visualização dos fragmentos é feita em gel de agarose e a identificação da região polimórfica é realizada por sequenciação SANGER.

Na agricultura, SSRs são usados para identificar e selecionar plantas ou animais com características desejáveis, acelerando os programas de melhoramento genético. Estes marcadores têm como vantagens o facto de terem uma sensibilidade e reprodutibilidade muito elevadas e apresentarem heritabilidade codominante. Como desvantagens destaca-se o elevado custo e complexidade de desenvolvimento da técnica.

Tipos de Marcadores Moleculares com Base na Heritabilidade:

- **Marcadores dominantes:** não permitem distinguir entre indivíduos heterozigóticos e homozigóticos, ou seja, apenas indicam a presença ou ausência de um alelo dominante. São utilizados em estudos em que não é necessário ocorrer uma distinção precisa entre genótipos, ou em programas de seleção que priorizam a presença de características dominantes.

- **Marcadores codominantes:** permitem discriminar ambos os alelos expressos, permitindo distinguir indivíduos homozigóticos de heterozigóticos.

Bibliografia

Agarwal, M., Shrivastava, N., & Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports*, 27, 617-631.

Al-Samarai, F. R., & Al-Kazaz, A. A. (2015). Molecular markers: An introduction and applications. *European Journal of Molecular Biotechnology*, 9(3), 118-130.

Arnaldo, V. (2011). *Engenharia genética: princípios e aplicações* (2ª ed.). Lisboa: Lidel.

Babbitt, G. A. (2008). How accurate is the phenotype? – An analysis of developmental noise in a cotton aphid clone. *BMC Developmental Biology*, 8, 1-9.

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., Zipursky, S. L., & Darnell, J. (2005). *Biologia celular e molecular* (5ª ed.). Porto Alegre: Artmed.

Morley, S. A., & Nielsen, B. L. (2017). Plant mitochondrial DNA. *Molecules*, 15(17), 10-2741.

Muto, A., & Ushida, C. (2002). Transcription and translation. *Molecular Biology and Pathogenicity of Mycoplasmas*, 323-345.

Nelson, D. L., & Cox, M. M. (2014). *Princípios de bioquímica de Lehninger* (6ª ed.). Porto Alegre: Artmed.

Oliveira, P. (1996). *Manual de genética*.

Ribeiro, M. (2009). *Genética molecular*. Universidade Federal de Santa Catarina.

Souza, P., Silva, H., Leite, B., Maia, D., Garcia, L., & Montes, A. (2015). *Genética geral para universitários* (1ª ed.). Universidade Federal Rural de Pernambuco.

Susman, M. (2001). Genes: Definition and structure. *Encyclopedia of Life Sciences*, 1-7.