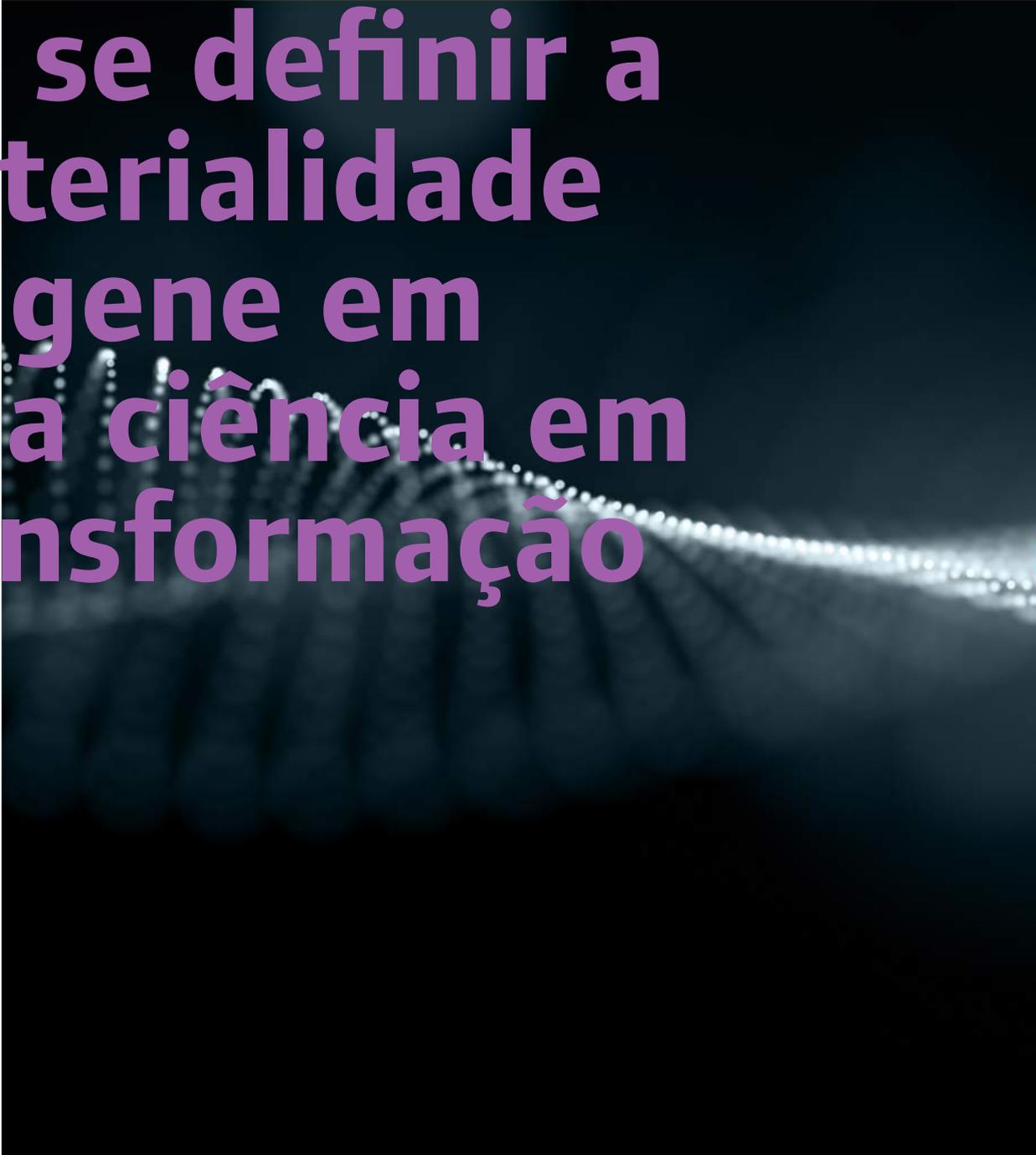


As dificuldades em se definir a materialidade do gene em uma ciência em transformação



Marcelo de Macedo Brígido

Universidade do Brasília, Departamento de Biologia Celular, Instituto de Biologia, Brasília, DF

Autor para correspondência - brigido@unb.br

Palavras-chave: conceito de gene, RNA mensageiro, regulação gênica, estrutura de genes, fluxo de informação, epigenética



O termo *gene* foi proposto para explicar os fatores propostos por Gregor Johann Mendel e a herança biológica. Aspectos como a consolidação do DNA como o material genético e seu modelo em dupla hélice mudaram a forma de ver o gene, causando intensos debates sobre seu significado. Neste artigo o autor discute marcos na Genética ao longo dos últimos cem anos que conduziram tal discussão, e como o projeto Genoma Humano ajudou a aquecer ainda mais o debate.

Motivação

A Genética faz parte do currículo do ensino fundamental e médio e dos cursos das áreas das ciências da vida no ensino superior, no entanto, a compreensão dos estudantes, a respeito de conceitos genéticos, principalmente no ensino médio, é limitada. A dificuldade em se conceituar o gene para além de metáforas, como elemento que contribui para o desenvolvimento do ser biológico, acaba por atrapalhar a comunicação entre aluno e professor. Observa-se que o aluno do ensino médio tem dificuldade em associar o conceito de gene com a molécula do DNA e, principalmente, em assimilar as regras de segregação genética dentro de um contexto molecular.

Neste artigo, o conceito de gene é trabalhado dentro de uma perspectiva histórica, focando nos avanços científicos que impuseram uma constante atualização conceitual e na sua delimitação molecular. A perspectiva histórica, para o entendimento de marcos científicos que impactaram o conhecimento do papel do gene no funcionamento do organismo biológico pode ajudar a compreensão dos conceitos de Genética tratados no ensino fundamental e médio.

O gene, um conceito em construção

Desde a antiguidade a herança biológica era entendida como uma mistura de características dos pais. Darwin, por exemplo, sugeriu a hipótese da pangênese, que previa que minúsculas partículas geradas nas diversas partes do corpo se reuniriam para compor o sêmen. Por outro lado, Mendel mostrou, em seus experimentos, que a herança biológica era particulada, consistindo de fatores unitários transmitidos pelos pais. Gregor Mendel era um monge austríaco que passou muito tempo pesquisando experimentalmente os padrões que governavam a formação e o desenvolvimento de híbridos, especialmente sobre como as características morfológicas

de ervilhas do gênero *Pisum* eram herdadas ao longo das gerações. As observações de Mendel resultaram na base da Genética Clássica que conhecemos.

Ao se falar em genes, a primeira ideia que vem à cabeça de quem já assistiu aulas de genética são os fatores hereditários propostos por Mendel, representados pelas letras **A**, quando o fator é dominante e **a**, quando recessivo. As plantas híbridas de ervilha por ele estudadas possuíam dois fatores, **AA** ou **aa**, quando homozigóticas ou **Aa**, quando heterozigóticas.

Apesar de muitos considerarem que os fatores hereditários de Mendel seriam determinados pelo que conhecemos hoje como genes, não foi ele quem cunhou o termo, mas o botânico dinamarquês Wilhelm Johannsen. Ele usou o termo *gene* para definir a unidade hereditária de Mendel, e cunhou o termo *genótipo*, que seria o conjunto de genes em um gameta. Johannsen também criou o termo *fenótipo* se referindo aos aspectos morfológicos de um indivíduo, facilmente observáveis e mensuráveis. Ele propôs que o genótipo, o componente genético, afetaria de alguma forma o fenótipo, garantindo sua manutenção ao longo das gerações. Para os primeiros mendelistas o gene era a unidade e herança.

Thomas Morgan e o colar de genes

Na virada do século dezanove ao vinte, o gene era uma unidade hipotética, em um conjunto de caracteres que eram herdados na progênie. Na esteira da descoberta dos cromossomas sexuais por Nettie Stevens em 1905, Thomas Morgan, um dos pioneiros no estudo genético de moscas, mostrou que os traços fenotípicos poderiam estar associados de alguma forma, como se ligados fisicamente. Depois de anos trabalhando com as frequências de permutação entre caracteres hereditários, o grupo de Morgan concluiu que os genes, localizavam-se de forma linear nos cromossomos, como um colar de contas. Os resultados consolidavam a jovem teoria cromossômica da herança, postulada anos



Thomas Hunt Morgan

antes por Walter Sutton e Theodor Boveri, e corroborada por Calvin e Bridges em 1916. Até então os cromossomos eram apenas corpúsculos que poderiam ser coloridos com corantes específicos, localizam-se no núcleo e eram ricos em ácidos nucleicos e proteínas. A teoria cromossômica colocou-os no centro da atenção. Eram esses cromossomos que, quando recombinados durante a formação dos gametas, misturavam genes dos pais, gerando variabilidade genética na prole.

Pensando o gene quimicamente

A nova geração de pesquisadores surgida na primeira metade do século XX apostava em novas abordagens experimentais para a biologia com forte influência na física e na química. Nesse ambiente, a ideia de uma molécula representar o gene era muito apropriada. Dois estudos dos anos 1940 marcaram essa onda de materialização química do gene. O primeiro deles foi conduzido por George Beadle e Edward Tatum, que geraram uma série de mutantes em *Neurospora crassa* com o uso de raios X. Eles conseguiram correlacionar as mutações obtidas, com a ausência de enzimas que causavam deficiência nutricional no fungo. Como as mutações eram herdáveis, assim como a ausência da enzima, deveria haver uma relação estreita entre o gene e a enzima, mais especificamente, propuseram que o papel do gene era determinar a síntese de uma enzima. A hipótese ficou conhecida como **um gene - uma enzima**. Apesar da fama, o trabalho da dupla apenas confirmou o trabalho pioneiro de Archibald Garrod, que chegou à mesma conclusão estudando doenças genéticas humanas. Juntos, eles mostraram que alterações genéticas podem estar relacionadas à ausência de enzimas específicas. Mais tarde, o estudo da estrutura química das proteínas demonstrou que as proteínas poderiam ser formadas por mais do que uma subunidade e a hipótese para o papel funcional do gene passou a ser um gene-um polipeptídeo.

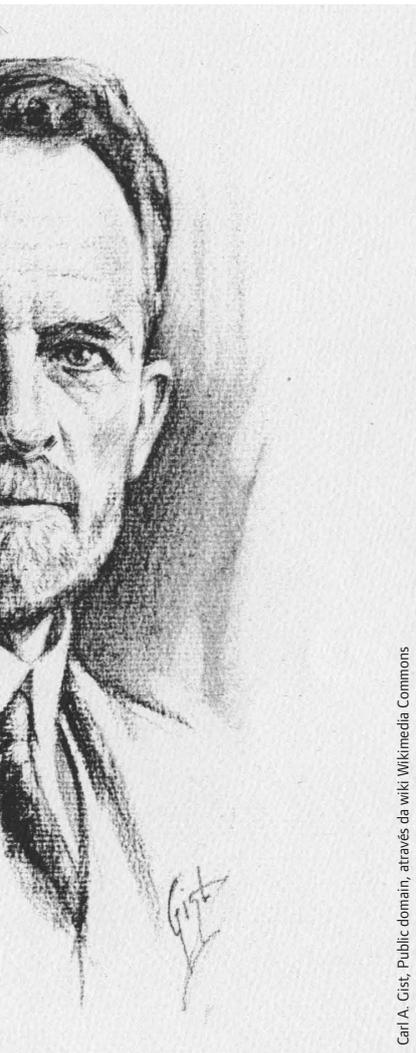
O segundo estudo que marcou a biologia nos anos 1940 mostrava a natureza química do

próprio gene, e foi inspirado nas pesquisas do inglês Frederick Griffith. As descobertas de Griffith sobre o princípio transformante, anos antes (1928), mostraram que era possível transformar uma bactéria aumentando sua virulência, isto é, a capacidade de produzir doença. Ele mostrou que um extrato de bactérias virulentas, mortas por calor, poderia tornar virulentas bactérias até então incapazes de provocar doença. Alguns anos mais tarde, em um experimento clássico, Avery, MacLeod e McCarty, avançaram no modelo experimental de Griffith e mostraram que era possível transferir características genéticas apenas com moléculas de ácido desoxirribonucleico. No entanto, a comunidade científica ficou reticente. Foi preciso mais alguns anos, e evidências experimentais de outros grupos de pesquisa, para que a nova ideia fosse finalmente assimilada: os genes eram feitos de DNA.

Hélices, intrigas e radiografias

Depois da segunda guerra mundial, os velhos radares que garantiram a defesa da Inglaterra não eram mais necessários e foram doados às universidades. O equipamento mostrou-se ideal para montar Difratômetros de Raios-X. Apesar do nome pomposo, o aparelho não passa de um tubo de raios catódicos (aqueles de televisão antiga e de radares também) onde uma amostra cristalina de material biológico é bombardeada com Raios-X, revelando sua natureza atômica mais íntima. Foi nessa época que começou a florescer uma nova área de investigação, a Biologia Molecular. Naquela época eram estudados os mais diversos materiais biológicos, como celulose, amido, proteínas e também ácidos nucleicos. O novo campo estudava como era a estrutura das moléculas biológicas e como elas funcionavam.

Diversos pesquisadores dedicaram-se a investigar um tipo específico de ácido nucleico, o ácido desoxirribonucleico, ou simplesmente DNA, seu nome mais popular. Dentre esses, James Watson e Francis Crick, da Universidade de Oxford, na Inglaterra, vieram a se



Carl A. Gist, Public domain, através da wiki Wikimedia Commons

tornar mundialmente famosos como aqueles que propuseram a estrutura em dupla hélice do DNA. Outros pesquisadores como Maurice Wilkins e Rosalind Franklin também contribuíram de forma fundamental nessa busca pela estrutura do DNA, apesar de até hoje serem tratados como coadjuvantes na descoberta. Todos publicaram as descobertas no mesmo número da conceituada revista inglesa *Nature*, mas somente Watson e Crick ficaram famosos. Talvez o fato de que Crick fosse físico, enquanto Watson era um jovem zoólogo, tenha fornecido a pimenta necessária para que eles fossem além da descrição estrutural da famosa dupla fita do DNA. Foram eles que sugeriram como a estrutura peculiar poderia ser responsável por replicar o próprio conteúdo genético, além de servir de molde para a síntese de proteínas, o assunto que naquela época era o que importava para a biologia.

A história indaga sobre a forma pela qual Watson e Crick conseguiram inspiração para propor a dupla fita do DNA: inicialmente obtendo acesso às radiografias de Rosalind; e também acesso às descobertas de Linus Pauling, um químico renomado que pesquisava sobre a estrutura do DNA do outro lado do oceano, nos Estados Unidos. O filho de Pauling, Peter Pauling, trabalhava no laboratório de Watson e Crick e teria mostrado cartas de seu pai com importantes achados sobre a estrutura do DNA. As cartas de Pauling continham informações técnicas que Watson e Crick souberam utilizar de forma magistral, mesmo que questionável, para propor o modelo do DNA. O legado de Watson e Crick não se limitou a um modelo físico da dupla fita, mas também incluiu a proposição de uma lógica mecânica para a replicação do DNA.

Alguns anos mais tarde, Watson e Crick esquematizaram o fluxo de informação genética na célula por meio de um esquema que ficou conhecido como o dogma central da biologia molecular. O gene é feito de DNA, uma molécula que pode ser facilmente duplicada e que se expressa na forma de proteínas que são codificadas na sequência de bases nucleotídicas.

Lactose e bactérias, enfim um modelo gênico regulatório

O *dogma central* proposto por Crick (Box 1) passou a ser referência à medida que o tripé: replicação, transcrição e tradução, passou a ser caracterizado bioquimicamente. O dogma central serviu como base de uma nova forma de enxergar o fluxo de informação dentro da célula. O DNA, sede do genótipo, é expresso dando origem a um fenótipo. Mas, poderia esse fenótipo ser modulado pelo ambiente? A resposta veio com um grupo de microbiologistas franceses, liderado por François Jacob e Jacques Monod, que começou a investigar a indução do fenótipo *lac+*, que possibilitava às bactérias responder à presença, e sobreviver, a partir do açúcar do leite, a lactose, como única fonte de carbono.

A enzima lactase (ou β -galactosidade) cliva a lactose e permite à bactéria *Escherichia coli* conseguir sobreviver em meio de cultura dispendo apenas de lactose como fonte de açúcar. A enzima, associada ao chamado fenótipo *lac+*, está codificada no genoma da bactéria, mas é produzida somente quando necessário (isto é, quando glicose não está disponível no meio). Jacob e Monod investigaram esse fenômeno e, depois de diversos anos de pesquisa, propuseram o modelo genético baseado no *operon*, uma unidade formada por genes que são funcionalmente relacionados. No caso do *operon lac*, as três enzimas produzidas para o metabolismo da lactose são controladas a partir de um único regulador. O regulador, uma proteína, liga-se fisicamente ao DNA, na região operadora, e impede a transcrição dos três genes envolvidos na via metabólica de degradação da lactose. O regulador pode controlar vários genes, que são transcritos como uma única unidade transcricional, isto é, um único RNAm. A região *operadora* controla a transcrição respondendo à presença de uma proteína reguladora. A associação física da proteína reguladora com a região operadora bloqueava a síntese do RNA e, consequentemente, da enzima lactase.

Um Dogma por Acaso

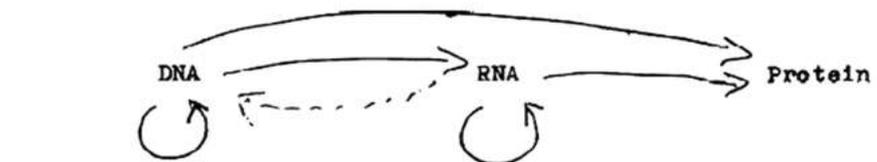
Entre os pensadores de sua geração, Francis Crick destacou-se pela genialidade. É dele a ideia do Dogma Central, um esquema que mostrava o fluxo de informação na célula. Como podemos observar no texto do próprio Crick de 1956, ele já previa que a informação genética fluía do DNA para as proteínas passando, ou não pelo RNA. Por outro lado, o fluxo contrário (de proteína para DNA) não acontecia. À época não se sabia a real participação do RNA como intermediário, por isso a conexão direta DNA-proteína. Mais tarde, com a demonstração da existência do RNA mensageiro (RNAm), o dogma foi atualizado. Fora isso, o esquema continua válido até hoje. Ele inclui até um fluxo de informação alternativo do RNA para o DNA previsto a época sem qualquer evidência experimental, mas a comprovação foi mostrada anos mais tarde. Hoje se diria que Crick lacrou com esse dogma. Outro aspecto curioso de Crick foi ter chamado seu esquema de dogma. Na ciência não existem dogmas, que são verdades que não podem ser invalidadas. Anos mais tarde ele mesmo comentou que usou o termo apenas por ênfase, por brincadeira, para chamar a atenção de “um fato que apesar de ainda não ter sido totalmente amparado experimentalmente, tinha tudo para estar certo”. Por isso, em respeito a Crick, continuamos a mencionar a expressão como o Dogma Central da Biologia Molecular.

Ideas on Protein Synthesis (Oct. 1956)

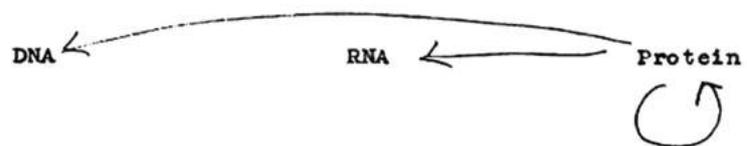
The Doctrine of the Triad.

The Central Dogma: "Once information has got into a protein it can't get out again". Information here means the sequence of the amino acid residues, or other sequences related to it.

That is, we may be able to have



but never



where the arrows show the transfer of information.

O modelo deu a Jacob e Monod o prêmio Nobel em 1966, e passou a ser considerado como o *paradigma da regulação gênica*. Eles haviam introduzido novos conceitos para elaborar um modelo *regulado* da expressão de genes: composto por genes estruturais, que codificam as enzimas envolvidas no meta-

bolismo da lactose e por genes regulatórios, que controlavam a expressão dos genes estruturais. O modelo do *operon Lac* continua sendo um modelo muito útil e didático para se entender a regulação gênica, e serve para explicar grande parte dos circuitos regulatórios bacterianos.

O paradigma do *operon Lac* explicava como proteínas controlavam o funcionamento de genes, ligando-se a trechos específicos da molécula de DNA. Os genes são representados ao longo do DNA como se fossem caixinhas sobre uma linha sólida. Proteínas ou fatores proteicos controlam os genes ligando-se a regiões específicas do DNA. Resumindo, até aquela época os genes eram considerados trechos delimitados, dispostos sequencialmen-

te sobre uma molécula de DNA, que vem a ser considerado o componente genético do cromossomo. Por serem compatíveis com o modelo de colar de contas de Thomas Morgan, foram logo aceitas por grande parte dos investigadores. Aparentemente estava tudo resolvido: os genes haviam sido materializados na forma de moléculas de DNA; expressavam-se na forma de proteínas que, por sua vez, poderiam regular o próprio DNA.

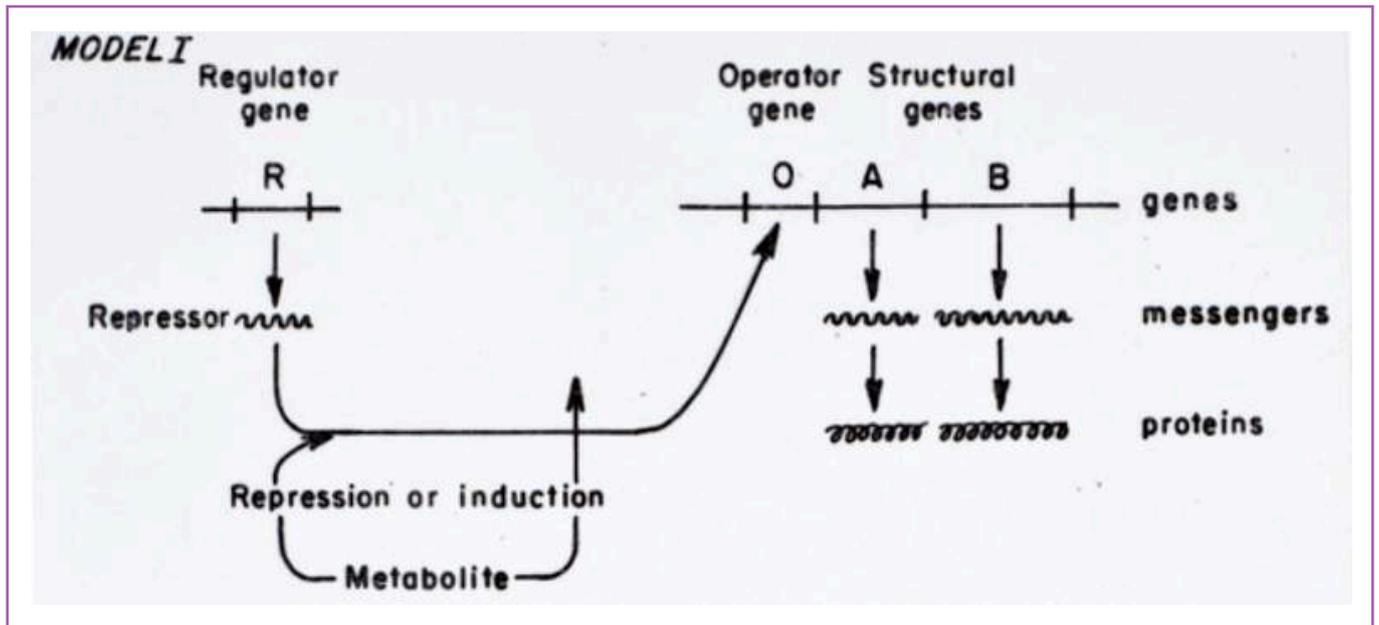


Figura 1. Esquema atribuído a François Jacob que descreve o operon Lac. Os genes estruturais de um operon são representados dispostos sequencialmente ao longo do cromossomo bacteriano (SG1, SG2), precedidos pela sequência operadora (op), local em que a molécula repressora pode se ligar. Se o repressor não estiver ligado ao operador, a transcrição ocorre normalmente, produzindo um RNA policistônico, ou seja, um único RNA mensageiro para todos os genes do operon. Na sequência, as proteínas relacionadas ao metabolismo da lactose podem ser traduzidas e a lactose utilizada como fonte de carbono. O repressor é codificado pelo gene regulador (RG), não contíguo ao operador. Adaptado de Watson, 2014.

Um dos legados de Jacob e Monod foi a noção de que havia um **RNA mensageiro** (RNAm), produzido pela ação de uma RNA polimerase e controlado por uma horda de proteínas reguladoras. Uma vez produzido, o destino era o transporte para o ribossomo onde servia de molde para síntese proteica. O RNA era o mensageiro da informação genética contida no DNA, uma mera cópia de um trecho específico de DNA.

Em eucariotos, o modelo do *operon* serve apenas como referência, como base da interação proteína-DNA, e como essa interação pode controlar os processos de transcrição. Por outro lado, modelo regulatório do *operon*, baseado na interação de fatores regulatórios e DNA, é muito simplista para explicar modelos regulatórios complexos como os observados em eucariotos.

O gene interrompido

Na década de 1970, a análise mais detalhada do RNAm rompeu com a ideia do gene como algo contínuo. Uma mudança fundamental em um conceito proposto inicialmente por Crick, baseado em uma suposta colinearidade entre as sequências nucleotídicas de DNA e RNAm. Ao observar em microscopia eletrônica híbridos entre moléculas de DNA, provenientes de genes eucariotos com o RNA dele transcrito, foi possível mostrar que o RNAm era formado por trechos do DNA que não eram contíguos no genoma. Portanto, a referida colinearidade não existia quando se considerava o RNA mensageiro maduro. Embora toda a sequência gênica fosse transcrita, originando um pre-RNA, o RNA maduro era formado pela junção de

pequenos trechos de sequência extraídos de uma região muito mais extensa na molécula do DNA.

O processo molecular que ficou conhecido como *splicing* mostrou-se um mecanismo geral, parte do processamento da molécula de RNA precursora produzida durante o processo de transcrição. O conceito de *gene interrompido*, como ficou conhecido, garantiu o prêmio Nobel para Philip Sharp em 1993 e mudou a forma de entender o gene eucariótico. Ele agora poderia ser dividido em trechos codificadores, conhecidos por **éxons**, e trechos intervenientes, os **íntrons**. Durante o processamento do RNA precursor, as sequências correspondentes aos íntrons eram removidas. O RNA maduro, o molde para a produção das proteínas, não era colinear com uma região única no DNA, mas uma colagem de pequenos trechos, agora conhecidos como éxons.

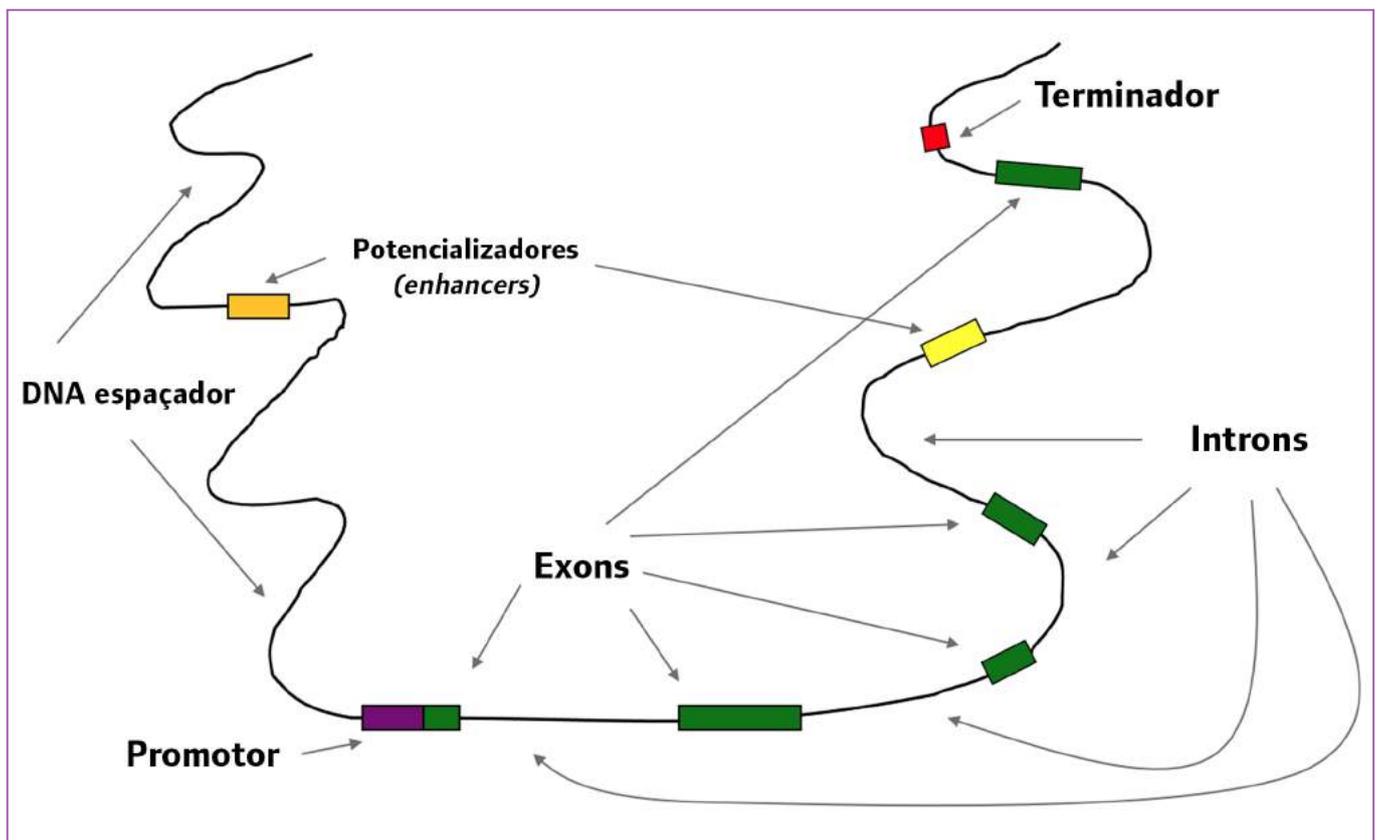


Figura 2. Estrutura geral de um gene eucariótico destacando os exons (em verde) e os íntrons (mostrados como linha em preto).

As petúnias pioraram as coisas

Em meados dos anos 1990 dois pesquisadores norte-americanos iniciaram um estudo sobre a coloração das pétalas em petúnias com um objetivo ousado, o de produzir petúnias com colorações mais intensas. Eles verificaram que, de alguma forma, a coloração dependia de ribonucleases, o que parecia ser um mecanismo novo de regulação gênica. A elucidação molecular desse mecanismo de regulação gênica foi realizada por Andrew Fire e Craig Mello, que usaram como modelo de estudo um pequeno nematoide, o *Caenorhabditis elegans*. O modelo garantiu-lhes o prêmio Nobel de 2004. O novo mecanismo dependia de pequenas moléculas de RNA

complementares a um determinado RNAm alvo. Ao formar moléculas híbridas com o RNAm, esses pequenos RNAs induziam uma degradação através de um complexo ribonucleásico conhecido hoje como RISC. O pequeno RNA (siRNA - a sigla em inglês para *small interfering RNA*) interferia com a meia vida do RNA mensageiro e esse mecanismo de silenciamento gênico ficou conhecido como interferência por RNA (RNAi).

O mecanismo de interferência por RNA é fruto de um mecanismo conservado em eucariotos e que permite o controle pós-transcricional da expressão gênica, ou seja, após sua síntese pela RNA polimerase. A expressão de um RNAm pode ser controlada por uma molécula de RNA regulador. Assim como as proteínas, os RNAs também poderiam atuar como reguladores da expressão de um gene.



O genoma humano – somos todos insetos

Em 2003, era anunciada a sequência de bases nucleotídicas de todo o DNA contido nas células humanas. Apesar do genoma não estar realmente completo naquele lançamento, a cerimônia marcava o cinquentenário da descrição da dupla hélice de Watson e Crick. A data simbólica também marcou a participação de James Watson no novo projeto, afinal fora ele o mentor intelectual da criação da Human Genome Organization (HUGO), no início da década de 1990. O objetivo da HUGO era desvendar completamente as 3 bilhões de bases do genoma humano. O projeto foi realmente colossal, transpôs diversos obstáculos técnicos que impediam o sequenciamento de DNA em escala genômica e, em apenas dez anos, permitiu chegar à quase completude do genoma humano. A iniciativa da HUGO criou as condições necessárias para a pesquisa genômica e, quando o sequenciamento do genoma humano foi anunciado, diversos outros genomas menores já haviam sido publicados. O genoma da bactéria *Haemophilus influenza* foi o primeiro genoma de um organismo celular publicado em 1995, com cerca de 1 milhão de pares de bases.

A cerimônia de lançamento do genoma humano no laboratório norte-americano de Cold Spring Harbor foi uma grande festação pois era uma iniciativa global com participação de grupos de 50 países e envolvendo mais de 300 pesquisadores. Estranhamente, o anúncio do genoma humano chegou com pelo menos uma observação inusitada: o número de genes contidos no genoma estava na ordem dos 20 mil, o mesmo número observado no genoma de outro organismo que estava sendo concluído, o da mosca de frutas *Drosophila melanogaster*. O pequeno inseto teria basicamente o mesmo número de genes de um mamífero, cujo organismo era considerado muito mais complexo do ponto de vista de tecidos e estruturas biológicas. No entanto, em termos genéticos os seres huma-

nos não tinham qualquer diferença em relação a um inseto, por exemplo, uma barata.

O aparente paradoxo foi amplamente explorado pela mídia e levou a um questionamento sobre o conteúdo do genoma: por que tão poucos genes nos seres humanos e por que organismos mais complexos não apresentam um número proporcionalmente maior desses genes? Será que o número de genes não reflete a complexidade do organismo? Ou será que o modelo molecular do gene baseado na trilogia de Crick – DNA -> RNA -> proteína – seria muito simplista e precisaria ser revista? Considerando que apenas 1-2% do genoma codificam todas as proteínas, por que todo esse DNA, será tudo lixo evolutivo? Ou melhor, será que nós não sabemos para o que estamos olhando?

As respostas para diversas dessas perguntas começaram a surgir com um estudo mais profundo sobre a constituição e o funcionamento do genoma humano, o ENCODE (Enciclopédia de Elementos do DNA), exigindo uma revisão do conceito de gene e do seu papel na elaboração do organismo.

Enfim, uma enciclopédia de elementos gênicos

A descrição do genoma humano permitiu pela primeira vez a análise minuciosa de seu conteúdo. Finalmente seria possível observar os genes em sua aparência concreta, física, em pares de bases nucleotídicas. Mas, quais seriam os limites do gene? Como seriam regulados? A experiência conseguida com o sequenciamento e a montagem de genomas levou-nos a um outro patamar na pesquisa molecular. Em 2003, o projeto ENCODE foi lançado, uma iniciativa mundial liderada pelo Instituto Nacional de Saúde dos Estados Unidos, para estudar elementos funcionais do genoma. O objetivo era ir além dos genes codificadores de proteínas, identificados no genoma, mas que respondem por menos de 2% do genoma humano. O projeto mostrou que o genoma continha uma quan-

tidade gigantesca de elementos desconhecidos e que poderiam de alguma forma fazer parte do mecanismo molecular de controle da expressão do genoma, ou seja, como o genótipo daria origem ao fenótipo. O que o ENCODE descobriu mudou de novo a percepção sobre regulação genética e epigenética, mais ainda, mostrou a importância do RNA no processo de regulação gênica, sempre considerado secundário no metabolismo da informação genética.

O papel do RNA na síntese proteica é bem caracterizado: ele participa tanto como molécula intermediária no fluxo da informação entre o DNA e proteínas e também como fábrica de proteínas, enquanto componente fundamental dos ribossomos (RNAr). A importância do RNA como molécula regulatória só foi reconhecida depois dos anos 2000, e certamente o ENCODE teve papel substancial para esse entendimento. O ENCODE mostrou que praticamente todo o genoma é transcrito na forma de RNA. A transcrição não é restrita aos genes codificadores de proteína, e milhares de RNA não codificadores (ncRNA) são hoje conhecidos. Alguns deles já são bem descritos com funções regulatórias, ou mesmo estruturais.

O universo de moléculas de RNA regulatórias só agora começa a ser explorado e ainda é muito pouco compreendido. Talvez por isso tenha sido comparado com a *matéria escura* do genoma por um importante estudioso dos RNA não codificadores, o pesquisador australiano John Mattick. Este nome foi dado em alusão à mal compreendida teoria da física moderna sobre a matéria escura que permeia nosso universo.

Na última versão do genoma humano, divulgada em janeiro de 2018, foram contabilizados 20.376 genes codificadores de proteínas e 22.305 genes que transcrevem RNA não codificadores, isso tudo em pouco mais de 3 bilhões de pares de bases. Nessa última versão, mostra também que foram identificados 203.903 transcritos diferentes, ou seja, um universo de moléculas de RNA muito além do esperado, considerando uma relação dire-

ta entre o número de genes e o número de transcritos.

Para entender como são originados tantos RNAs diferentes, temos que lembrar que o RNA transcrito pela polimerase do RNA (o pré-RNA ou RNA precursor) é processado, gerando uma molécula madura de RNA contendo apenas éxons (Figura 2). Porém, o conteúdo de éxons em um dado transcrito é variável. Nem todo transcrito é processado de uma forma canônica (como é conhecida a forma contendo todos os éxons anotados naquela unidade transcripcional). Milhares de RNAs diferentes são fruto do uso alternativo de éxons em um processo conhecido por **splicing alternativo**. O padrão de uso desses éxons pode variar entre diferentes tipos celulares e pode ser regulado, introduzindo um componente a mais de complexidade na construção da célula. Esses RNAs associam-se a proteínas, a outros RNAs, e mesmo ao DNA, interferindo no funcionamento dos genes e ditando como os genes são expressos.

A célula como uma grande rede

A informação contida no DNA é transcrita na forma de RNA e parte dele serve de substrato para formar proteínas. As moléculas de RNA e de proteína interagem entre si e com o próprio DNA cromossômico, controlando a transcrição de novas moléculas que alimentam o ciclo (Figura 3). A informação inscrita no DNA serve para a construção do aparato molecular, um conjunto de moléculas que funcionaria como uma rede. A rede molecular, por sua vez, controlaria a síntese de novos transcritos que a alimentarão e controlarão a leitura da informação genética.

Sabemos hoje que o DNA sozinho não é capaz de fazer uma célula funcionar, a informação nele contida precisa ser interpretada. A rede desempenharia esse papel. Ela controla como a informação presente no DNA é lida, com impacto direto na conectividade desses componentes, e influenciando o aspecto final da própria rede.

Nesse modelo de funcionamento da célula, o DNA continua sendo a *casa* dos genes, mas o modo pelo qual esse DNA é lido e interpretado dependeria da rede. A rede é responsável pela identidade fenotípica da célula, modelando a forma como é lida a

informação contida no genoma, o genótipo. A rede funcionaria como um **epigenoma**, uma série de instruções herdáveis e que permitiriam manter a estabilidade de uma célula durante sucessivas divisões celulares.

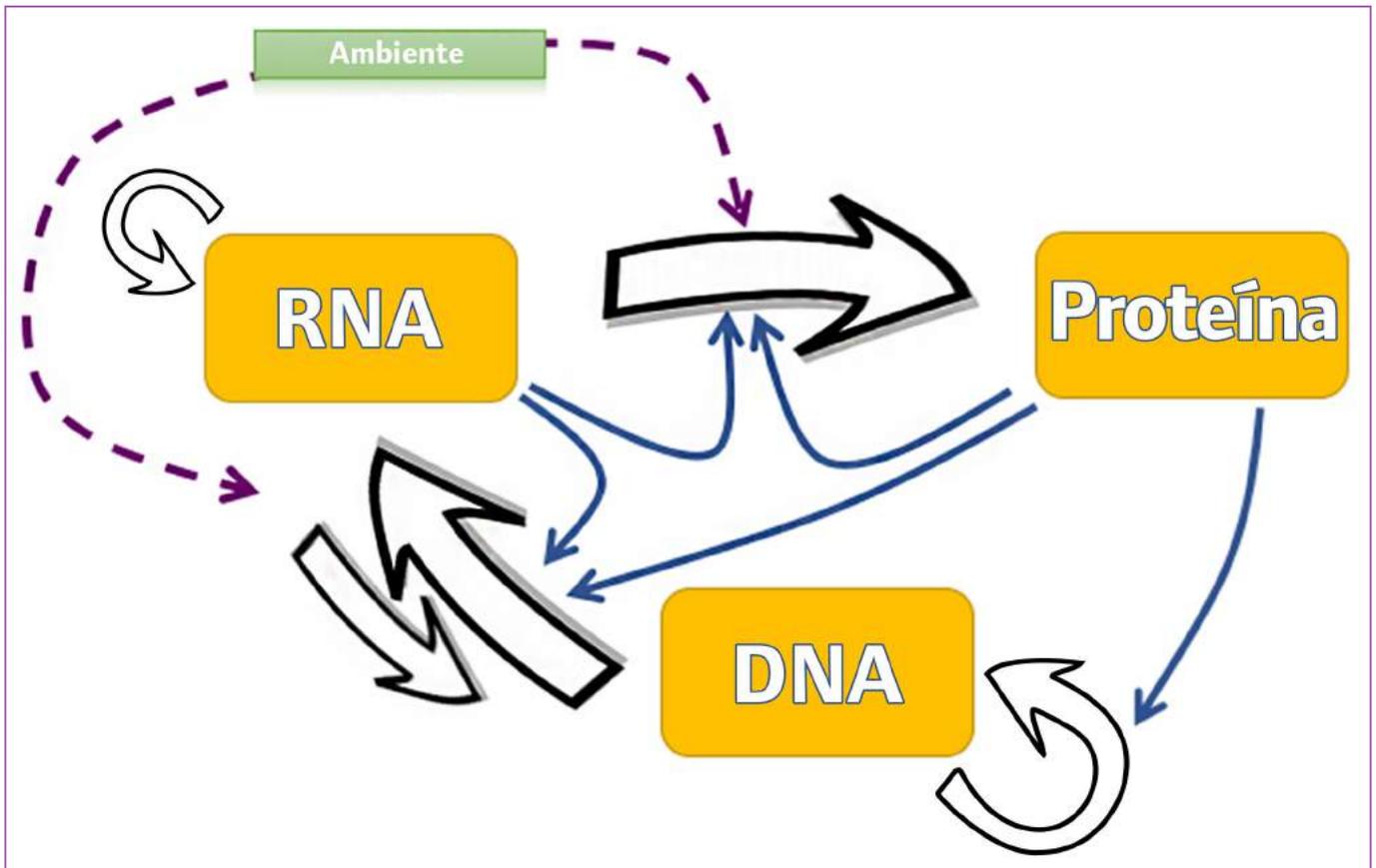


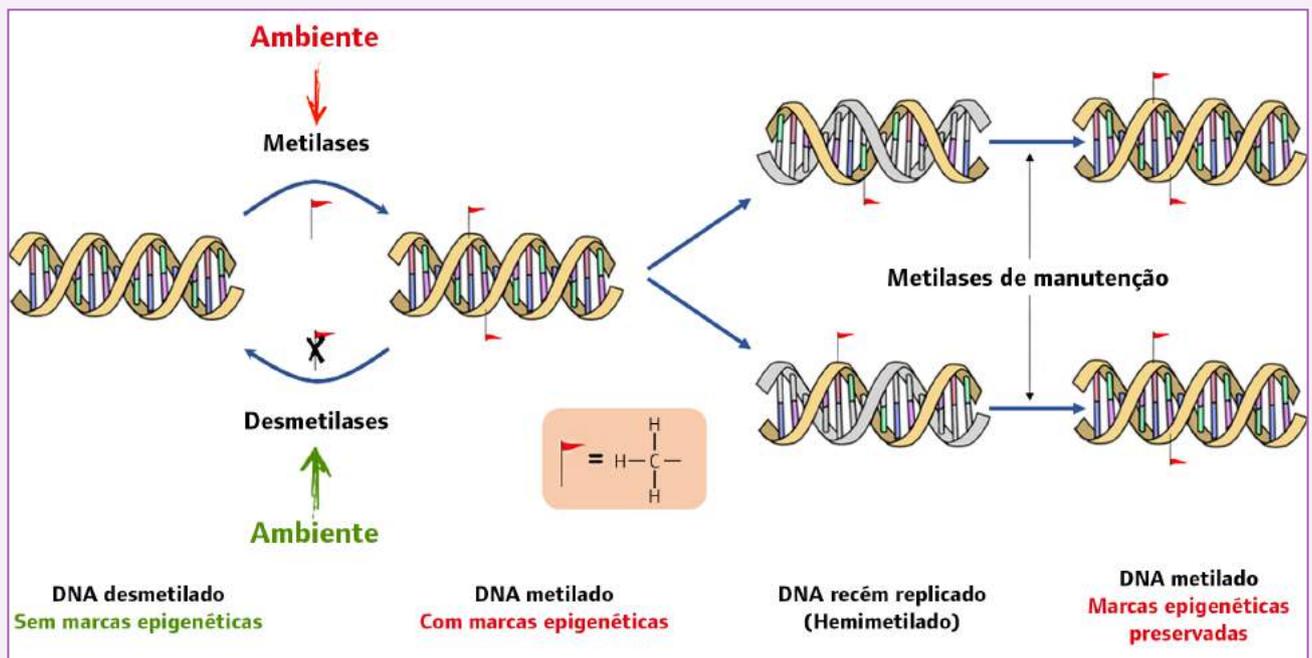
Figura 3. Uma versão estilizada e livre do Dogma Central proposto por Crick. Os elementos foram reposicionados para ressaltar o papel do RNA e das proteínas. As setas pretas mostram o fluxo de informação;

as azuis, a regulação do fluxo de informação controlado por proteínas e RNA. As setas lilases mostram o papel do ambiente influenciando os processos moleculares de regulação gênica.

A rede controladora que envolve a interação de diferentes componentes celulares define como o DNA é lido, permitindo que tenhamos, por exemplo, células tão diferentes, como células musculares e neurônios. Em ambos os tipos celulares o DNA é o mesmo, o que varia é como a informação nele contida é interpretada. Trechos inteiros do DNA ficam adormecidos enquanto outros são ativamente expressos (Box 2). Mesmo

nas regiões ativas, os transcritos produzidos podem conter informações diferentes, produzindo proteínas e RNAs que são únicos para aquele tipo celular. Nessas condições é difícil definir as fronteiras físicas de um gene. Se um único trecho de DNA pode produzir transcritos e proteínas tão diferentes, será adequado dizer que se trata de um gene? Ou seria melhor considerá-lo genes diferentes?

Epigenoma é um termo relacionado à *Epigenética*, um conceito da década de 1940 que propunha um modelo para a modulação do fenótipo pelo ambiente. Conrad Waddington utilizou o termo Epigenética para se referir à expressão dos genes sob a influência do ambiente criando o fenótipo. Atualmente, são conhecidos mecanismos que modificam a molécula de DNA, pela adição de sinais químicos, atuando para ligar e desligar genes, o que propicia a aquisição de fenótipos celulares característicos durante o desenvolvimento do organismo. O conjunto dessas modificações dentro de um genoma seria o epigenoma. A modificação epigenética mais comum é a metilação da base nitrogenada citosina. Um grupo metila (-CH₃) pode ser adicionado a citosinas e se tal modificação do DNA ocorre em regiões promotoras, pode ocasionar o desligamento de um gene específico. Uma enzima chamada metilase é responsável pela modificação e atua em resposta a eventos ambientais (por exemplo, hormônios ou morfógenos). Quando o DNA é replicado, o padrão de metilação do DNA é mantido por metilases de manutenção, fazendo com que o epigenoma seja transmitido através das divisões celulares. Eventualmente desmetilases podem remover a marca (o grupo metil), alterando assim o padrão local de metilação, o que pode ter impacto no funcionamento gênico.



O gene, um conceito em mutação

A comunidade científica promoveu mudanças no significado do termo gene ao longo do século XX (p. ex. “um gene, uma enzima”, “um gene, um polipeptídeo”). Apesar disso, no senso comum, como na interpretação normalmente dada pela mídia, o termo é entendido sob forma metafórica apenas. O *gene da inteligência*, o *gene do alcoolismo*, são expressões comumente utilizadas na tentativa de associar um componente genético a de-

terminadas características biológicas. Essas associações, comuns num contexto social e político, são especulativas e podem, inclusive, gerar distorções com efeitos relativamente graves. Os genes estão associados a fenótipos, mas não são os únicos determinantes dos fenótipos. Para a chamada genética clássica, o gene é um caractere hereditário que contribui para a construção de um fenótipo juntamente com elementos epigenéticos e ambientais. O grau de determinação genética no fenótipo é muito variável. Já no âmbito da genética molecular, o gene toma a dimensão do DNA, mesmo sem uma definição clara de o que, no DNA, corresponderia aos genes conforme descritos pela genética clássica.

O conceito de DNA também é largamente utilizado no senso comum. É comum até mesmo as pessoas se referirem ao DNA de alguma coisa não biológica, como o *DNA de uma empresa*, no sentido de o que seria a essência ou missão da instituição. A popularização se deve a uma crescente exposição do DNA na mídia, fruto, em grande parte, do projeto genoma humano, bem como dos avanços da biologia molecular na biotecnologia e na medicina. Para evitar distorções igualmente indesejáveis, é importante que a popularização efetuada pela mídia seja cientificamente fundamentada.

Mas, como o DNA e o gene relacionam-se? A observação de que o DNA contém material genético sugere que os genes deveriam estar ali de alguma forma. Pensadores importantes como Francis Crick ajudaram a compreender o gene com uma argumentação simples: um gene, uma proteína, uma função. No entanto, descobertas recentes da biologia molecular não se encaixam mais nesse modelo.

A noção de que um gene é uma sequência de DNA que codifica uma proteína é normalmente encontrada nos textos básicos de biologia e representa uma visão inicial que está mudando muito rapidamente. *Um gene, uma proteína* é muito pouco para um genoma como o humano, onde apenas 2% do DNA codifica proteínas. *Um gene, um transcrito de RNA* parece mais adequado, já que quase todo o genoma é transcrito. Mas, com toda complexidade do *splicing* alternativo, como delimitar esses genes? Além disso, não podemos esquecer as regiões regulatórias que atuam na rede de proteínas e RNA que caracteriza a célula, e controla como o DNA será lido. Essas regiões também se comportam como caracteres herdáveis. Um geneticista de populações, que analisa os genes apenas pela sua transmissibilidade entre gerações, chamará qualquer uma das possibilidades acima como genes.

Importante ressaltar que a definição molecular do gene é uma definição funcional. Serve para identificar e delimitar sequências com funções biológicas em potencial, mas ela não

conversa com aspectos mais clássicos da genética, como a relação genótipo-fenótipo. É uma definição prática e facilita sobremaneira um entendimento mais genérico do gene, mesmo estando aquém do seu significado clássico. Afinal, mesmo depois de 100 anos de ter sido proposto, o gene continua sendo uma abstração.

Conclusão

A ideia de fatores discretos responsáveis pela transmissão da informação genética entre gerações existe há pouco mais de um século. Durante esse tempo a noção sobre a materialidade e o funcionamento desses caracteres foi se transformando, e essas novas ideias refletem o desenvolvimento, tanto científico quanto cultural, de nossa sociedade. Atualmente é comum se associar um gene a uma proteína, mas esse conceito se mostrou muito limitado, e uma definição mais generalista envolvendo DNA e herança se mostra muito mais adequada.

Para saber mais

- ANDRADE, L.A.B; SILVA, E.P. Mendel e Seus Abismos. *Genética na Escola*, v. 11, n. 2, p. 234-243, 2016.
- CRICK, F.H.C. On Protein Synthesis. *Symposia of the Society for Experimental Biology* v. XII, p. 139-163, 1956.
- GRIFFITHS, P.E., STOTZ K. Genes in the post-genomic era. *Theor Med Bioeth*. v. 27 p. 499-521, 2006.
- FALK, R. What is a gene? - Revisited. *Studies in the History and Philosophy of Science* v. 41 p. 396-406, 2010.
- JOAQUIM, L.M.; EL-HANI, C.N.. A genética em transformação: crise e revisão do conceito de gene. *Scientiae Studia*, v. 8, p. 93-128, 2010.
- JOHANNSEN, W. The genotype conception of heredity. *The American Naturalist* 45:129-159, 1911. Reimpresso em 2014, doi: 10.1093/ije/dyu063.
- MOSS, L. What Genes Can't Do. Cambridge-MA: MIT Press. 2003.
- PEARSON, H.. What is a Gene? *Nature* v. 44, p. 398-401, 2006.