

CASA DE OSWALDO CRUZ – FIOCRUZ

PÓS-GRADUAÇÃO EM HISTÓRIA DAS CIÊNCIAS DA SAÚDE

GUSTAVO CIRAUDO FRAGA SOLHA

**OS GENES INTERROMPIDOS: INTRODUÇÃO HISTÓRICA AO IMPACTO DA
DESCOBERTA DOS ÍNTRONS (1977) NA CONTROVÉRSIA SOBRE A DEFINIÇÃO
DE GENE MOLECULAR CLÁSSICO (1960)**

Rio de Janeiro

2005

GUSTAVO CIRAUDO FRAGA SOLHA

**OS GENES INTERROMPIDOS: INTRODUÇÃO HISTÓRICA AO IMPACTO DA
DESCOBERTA DOS ÌNTRONS (1977) NA CONTROVÉRSIA SOBRE A DEFINIÇÃO
DE GENE MOLECULAR CLÁSSICO (1960)**

Dissertação de mestrado apresentada
ao Programa de Pós-Graduação em
História das Ciências da Saúde, da
Casa de Oswaldo Cruz/Fiocruz, como
requisito parcial para obtenção do
Grau de Mestre. Área de
Concentração: História das Ciências.

Orientador: Prof. Dr. Ricardo Waizbort

Rio de Janeiro

2005

GUSTAVO CIRAUDO FRAGA SOLHA

OS GENES INTERROMPIDOS: INTRODUÇÃO HISTÓRICA AO IMPACTO DA
DESCOBERTA DOS ÍNTRONS (1977) NA CONTROVÉRSIA SOBRE A DEFINIÇÃO DE
GENE MOLECULAR CLÁSSICO (1960)

Dissertação de mestrado apresentada
ao Programa de Pós-Graduação em
História das Ciências da Saúde, da
Casa de Oswaldo Cruz/Fiocruz, como
requisito parcial para obtenção do
Grau de Mestre. Área de
Concentração: História das Ciências.

Aprovado em maio de 2005.

BANCA EXAMINADORA

Prof. Dr. Ricardo Waizbort - Orientador
Casa de Oswaldo Cruz – FIOCRUZ

Profa. Dr^a. Nara Azevedo
Casa de Oswaldo Cruz – FIOCRUZ

Prof. Dr. Franklin Rumjanek
Instituto de Bioquímica Médica/UFRJ

Prof. Dr. Robert Wegner (suplente)
Casa de Oswaldo Cruz - FIOCRUZ

Rio de Janeiro
2005

Agradecimentos

Ao meu orientador, Ricardo Waizbort, pela orientação e socorro nos momentos mais críticos e por sua inacreditável paciência ao aceitar todos os prazos não cumpridos.

Aos professores da Casa de Oswaldo Cruz que sempre gentilmente me atenderam quando solicitados.

À Fundação Oswaldo Cruz, pelo apoio institucional e financeiro.

Ao Pedro “Mezenga”, em razão dos movimentos simulados.

Ao meu irmão, por ser meu irmão.

Ao atleta do estacionamento, Papai Smurf, She-Ha, Rafael Aquino, Léo Mané, Lango, Renilson, Rochinha, Tadeu e Itaberá, por serem seres mitológicos.

A Fernando Sabino, João Guimarães Rosa, Dostoiévsky, Kafka, Lewis Carrol, entre outros amigos mudos.

A meus pais

A inteligência existe precisamente para alcançar aquilo que se quer. Não podendo percorrer uma versta, andem-se apenas cem passos; sempre é melhor, chega-se mais perto do objetivo, se é que se caminha para um objetivo. E, se você quer a todo custo alcançar o objetivo com um só passo, isto, a meu ver, não é de modo algum inteligência. Isto se chama desamor ao trabalho. Não gostamos do esforço, não estamos acostumados a avançar passo a passo, queremos voar até o objetivo numa passada ou figurar entre os Régulos.

Fiódor Dostoiévsky, Notas de Inverno Sobre Impressões de Verão.

Sumário

RESUMO	ix
ABSTRACT	x
INTRODUÇÃO	1
Apresentando o problema: O gene, uma receita para sintetizar uma proteína	2
Os genes interrompidos	5
Revoluções kuhnianas na biologia?	8
Por uma história de problemas	14
CAPÍTULO I – O DOGMA CENTRAL DA BIOLOGIA MOLECULAR	17
O gene molecular clássico	18
As proteínas	22
Breves antecedentes ao conceito de gene molecular clássico	24
1940: A hipótese “um gene – uma enzima”	26
A base físico-química da hereditariedade e a dupla-hélice do DNA	28
1961: O código genético	34
O dogma central da biologia molecular	39
CAPÍTULO II – OS GENES INTERROMPIDOS E O DEBATE ACERCA DO CONCEITO DE GENE MOLECULAR CLÁSSICO	43
Apresentando o problema: Os genes interrompidos	44
Tecnologia no estudo do vivo: O nascimento da biologia molecular	46
Os genes interrompidos e o DNA recombinante	50
O processamento alternativo	61
Processamento alternativo e o número de genes	66
Outros eventos celulares	69
Evelyn Fox Keller e a dinâmica celular	75
Raphael Falk e os genes desorientados	82
Para onde se dirige a atenção de Keller	90
CONSIDERAÇÕES FINAIS	96

REFERÊNCIAS BIBLIOGRÁFICAS	102
-----------------------------------	-----

LISTA DE FIGURAS

Figura 1: A dupla-hélice	30
Figura 2: O “dogma central da biologia molecular”	39
Figura 3: Estrutura dos genes interrompidos I	52
Figura 4: Estrutura dos genes interrompidos II	59
Figura 5: Processamento alternativo do gene <i>sxl</i>	64
Figura 6: O “Operon lac”	70
Figura 7: O dogma expandido	94

Resumo

Durante o ano de 1977, equipes independentes de pesquisa como a do inglês Richard Roberts no *Cold Spring Harbor Laboratory*, do norte-americano Phillip Sharp no *Massachusetts Institute of Technology* (MIT), do francês Pierre Chambon no *Centre National de la Recherche Scientifique*, entre outras, descobrem que ao contrário do que era sabido em bactérias, o gene não era um trecho de DNA ininterruptos pronto para ser traduzido em proteínas. Em organismos eucariontes o RNA mensageiro precisa ser montado para dar origem à proteína final. Todavia, essa montagem do RNA mensageiro pode ocorrer de formas alternativas, gerando mais de uma proteína por gene. Esse fenômeno é chamado de processamento alternativo do RNA mensageiro. A descoberta deste, e de outros processos moleculares, são motivo necessário para que certos autores postulem a necessidade de uma mudança conceitual das pesquisas biológicas. Para outros autores, essas descobertas não são anomalias, mas mantêm o núcleo de um programa de pesquisa lakatosiano. O “dogma central da biologia molecular”, tratado ou como paradigma kuhniano ou como núcleo lakatosiano é debatido.

Abstract

During all the year of 1977, independent teams of research as of the English the Richard Roberts working at Cold Spring Harbor Laboratory, of the North American Phillip Sharp in the Massachusetts Institute of Technology (MIT), of the Frenchman Pierre Chambon in *National Centre of la Recherche Scientifique* among others, in contrast discover that of that was known in bacteria, the gene was not a piece of uninterrupted DNA soon to be translated proteins. In eukaryotic organisms the necessary messenger RNA to be mounted to give origin to the final protein. However, this assembly of the messenger RNA can occur of alternative forms, generating more than a protein for gene. This phenomenon is called alternative splicing of the messenger RNA. The discovery of this and other molecular processes is necessary reason so that certain authors claim the necessity of a conceptual change of the biological research. For other authors, these discoveries are not anomalies, but it keeps the inner core of a lakatosian program of research. The “central dogma of molecular biology”, treated or as kuhnian paradigm or as lakatosian inner core is debated.

INTRODUÇÃO

Apresentando o problema: O gene, uma receita para sintetizar uma proteína

“Quem é você?”

Perguntou a Lagarta.

Não era um começo de conversa muito animador. Alice respondeu, meio encabulada: “Eu... mal sei, *Sir*, nesse exato momento... pelo menos sei quem eu era quando me levantei esta manhã, mas acho que já passei por várias mudanças desde então.”

Carroll, Lewis, Aventuras de Alice no País das Maravilhas, p. 45.

O gene é uma das entidades teóricas mais básicas da biologia em geral e, particularmente, da biologia molecular (Mayr, 1998; Falk, 1986). Contudo, é notório que o conceito de gene, empregado por biólogos, goza igualmente de popularidade mesmo entre um público mais amplo (Eyck & Williment, 2003; Ridley, 2003; Epp, 1997). Frequentemente, tanto na imprensa especializada como na leiga, o gene é apontado como promessa para a erradicação de males que afligem a sociedade: da cura de doenças e distúrbios, tais como tumores, esquizofrenia e síndrome bipolar, à solução de conflitos sociais e inter-étnicos, passando pela possibilidade de aperfeiçoar a natureza de animais e plantas, resolver questões forenses e traçar genealogias. Até na vida cultural o gene tem exercido sua influência (Kemp, 2003). Contudo, qual a natureza do mecanismo que transmite através das gerações as características biológicas? Especificamente, o que são os genes?

Apesar da centralidade do termo “gene” para a biologia, sua definição permanece alvo de controvérsias e especulações (Griffiths & Neumann-Held, 1999). No livro *Nature via Nurture: Genes, Experience, and What Makes us Human* (2003), Matt Ridley, *Ph.D.* em Zoologia e professor visitante no *Cold Spring Harbor Laboratory*, lista nada menos que sete significados de “gene”. De acordo com Ridley, um gene pode ser definido como: 1) um arquivo hereditário mendeliano para um determinado traço físico ou comportamental; 2) um carreador de doença/saúde; 3) uma informação compartilhada entre diferentes espécies de seres vivos; 4) uma receita química para sintetizar proteínas; 5) um comutador molecular para ligar/desligar outros genes; 6) uma unidade de seleção natural; e 7) um dispositivo para extrair informação do ambiente. Cada uma dessas definições está relacionada a descobertas feitas ao longo do desenvolvimento da história da genética. Cada uma trilhou um caminho diferente e nenhuma delas, sozinha, poderia responder pela definição do que é o gene. Dessa forma, não constitui surpresa o fato de que os diversos significados possíveis atribuídos ao termo “gene” confundam o público leigo e mesmo alguns biólogos. Para sermos precisos, desde que foi cunhado pelo biólogo dinamarquês Wilhelm Johannsen, em 1909, o significado do termo “gene” jamais esteve ao abrigo de polêmicas dentro da história das ciências (Keller, 2000; Mayr, 1998; Falk, 1986; East, 1929).

No entanto, uma definição unívoca e inequívoca talvez não seja nem necessária e nem mesmo desejável. Diferentes conceitos de gene se prestam a diferentes áreas da biologia. No presente trabalho, nosso objetivo é investigar um conceito específico de

gene, denominado de “gene molecular clássico” (Griffiths, 2002; Griffiths & Neumann-Held, 1999; Falk, 1986). Segundo essa definição, o gene é “a stretch of DNA sequence that codes for a particular protein that has a particular function.” (Collins, 2001). Essa definição equivale ao quarto significado de gene listado por Ridley (2003), em que um gene aparece como uma espécie de receita química para produzir uma proteína. Embora oriunda de pesquisas realizadas ainda da década de 1960, essa definição de gene continua sendo a mais utilizada pelos autores nos livros modernos de biologia celular e molecular (Griffiths, 2002; Falk, 1986).

A idéia de que os genes codificam proteínas é derivada do chamado “dogma central da biologia molecular” (Mattick, 2003; Olby, 1975; Mayr, 1998). Essa sentença foi formulada, em 1957, pelo físico britânico Francis Crick, que, em trabalho conjunto com o biólogo norte-americano James Watson, revelou a estrutura físico-química da molécula de DNA, em 1953. No “dogma central” de Crick, o fluxo de informações genéticas segue *unidirecionalmente* do DNA para o RNA e daí para as proteínas (Silverman, 2004; Mattick, 2003; Olby, 1975; Mayr, 1998). Apesar das críticas de alguns biólogos sobre o emprego do termo “dogma” dentro das ciências biológicas, ele se mostrou – e ainda se mostra – fundamentalmente correto (Maynard-Smith, 2001). Para o filósofo e historiador da biologia Robert Olby, o “dogma central da biologia molecular” é o paradigma da biologia contemporânea, fornecendo a estrutura intelectual para que os praticantes da pesquisa genética resolvam os seus quebra-cabeças (Olby, 1975). Em 1968, Gunther Stent, biólogo molecular da Universidade da Califórnia em *Berkeley*, afirmou que para a maioria dos pesquisadores da época, o sentimento era de que os grandes problemas genéticos estavam resolvidos e o que restava era apenas o trabalho de “polir os detalhes” (Stent, 1968). No contexto da afirmação de Stent, a idéia de “polir os detalhes” pode ser tomada como análoga à atividade de pesquisa rotineira, uma característica que ocorreria durante a fase de ciência normal kuhniana (1994). Durante essa fase, os cientistas trabalham dentro dos limites prescritos pelo paradigma vigente.

Contudo, alguns autores contemporâneos acreditam que a primazia do gene como um conceito explicativo central das atividades biológicas está superado (Keller, 2000), e que tal conceito não reflete mais o estágio atual das pesquisas no campo da biologia molecular (Gelbart, 1998). A definição de gene ao qual esses autores se referem é, sem dúvida nenhuma, o “gene molecular clássico”, a quarta definição da lista de Ridley (2003). Essa contenda seria um reflexo de uma transformação conceitual no pensamento dos biólogos moleculares a respeito da passagem e do processamento da informação genética. Inevitavelmente, os principais atingidos são “o dogma central” e o próprio conceito de gene.

Para Richard Strohman (1997), biólogo molecular emérito da Universidade da Califórnia em *Berkeley*, o estudo dos genes, embora necessário, traz uma contribuição apenas parcial para o processamento da informação genética. A velha noção proclamada pelo “dogma central” que os genes codificam proteínas e estas, por sua vez, regulam as diversas reações químicas dos seres e compõem a arquitetura dos organismos não é suficiente para a explicação dos fenômenos biológicos por completo. Algo está faltando.

Os genes interrompidos

Durante o ano de 1977, pesquisadores anunciaram a descoberta de que a estrutura dos genes em organismos eucarióticos não é contínua, mas apresenta interrupções que, a princípio, não fariam nenhum sentido (Glover & Hogness, 1977; Berget *et al.*, 1977; Chow *et al.*, 1977; Sambrook, 1977; Aloni *et al.*, 1977; Breathnach *et al.*, 1977; Williamson, 1977; Brack & Tonegawa, 1977; Gilbert, 1978; Chambon, 1981). Como uma espécie de mosaico, os genes de organismos eucariontes são divididos em porções chamadas de íntrons, as regiões sem sentido, e pelas regiões da seqüência do DNA que seriam responsáveis pela especificação do produto gênico, os exons. Os genes que apresentam essa estrutura são os chamados “genes interrompidos”. O objetivo central da presente dissertação é investigar o impacto da descoberta dos genes interrompidos sobre o conceito de “gene molecular clássico”. Paralelamente, apresentaremos, de forma breve, algumas controvérsias sobre a definição de gene molecular nos dias de hoje.

Essas controvérsias poderiam, grosso modo, ser divididas em duas posições. De um lado estão aqueles que entendem que o conceito de gene, com a descoberta dos íntrons e de outros fenômenos moleculares, chegou ao limite de seu poder explicativo (Leite, 2003; Keller, 2000; Strohmman, 1997). No outro pólo estariam aqueles que compreendem que as dificuldades em definir (unívoca e inequivocamente) o gene são intrínsecas à natureza do próprio conhecimento científico (Downes, 2004; Falk, 1986; Griesemer, 2000; Griffiths & Neumann-Held, 1999). Nossa avaliação é de que essas controvérsias mostram que o conceito de gene e a teoria genética não estariam passando por uma revolução paradigmática kuhniana, mas tentando eliminar anomalias que ocorrem durante a experimentação biológica com os genes. O debate em torno do assunto pode ser entendido como uma passagem por uma fase progressiva de um programa de pesquisa lakatosiano.

Para Imre Lakatos (1979), as teorias são objetos de constante interpretação, revisão, reinterpretção. Nesta dissertação, utilizaremos o termo “teoria” com o sentido de “programa de pesquisa” na concepção de Imre Lakatos. Segundo o referido autor, uma teoria científica nunca é formada por uma única hipótese explicativa, mas, ao contrário, compõe-se de várias hipóteses que tentam dar conta de um determinado fenômeno da natureza. Essas hipóteses, articuladas entre si, formam a estrutura que constitui o “núcleo” de um programa de pesquisa.

O núcleo de um programa de pesquisa, de acordo com Lakatos, é sustentado por hipóteses auxiliares que formam um “cinto de proteção” para o programa. Enquanto o cinto de proteção garantir a corroborção do núcleo, este não é refutado, e seu conteúdo de informações aumenta. Se, de outra forma, anomalias no sentido kuhniano surgem e aumentam em número, de forma que não se consegue explicar através das hipóteses auxiliares o problema em questão, o núcleo deixa de ser válido e o programa é substituído por outro com maior conteúdo empírico. O novo programa de pesquisa terá que continuar a explicar aquilo que o anterior explicava e dar cabo das anomalias.

A estrutura das teorias constitutivas do núcleo de um programa de pesquisa pode se modificar ao longo do tempo. O próprio programa de pesquisas inclui regras metodológicas indicando os caminhos que podem ser trilhados e aqueles que devem ser evitados para que se possa “arquitetar conjecturas que tenham maior conteúdo empírico do que as suas predecessoras” (Popper *apud* Lakatos, 1979, p. 162). Em outras palavras, conjecturas que possam explicar melhor, que o programa utilizado anteriormente, um determinado fenômeno, ou que tragam uma contribuição maior para resolver um problema determinado. Mantido o núcleo do “dogma central”, o cinturão protetor do

programa genético admitiria o exponencial aumento de complexidade que se observa quando se manipula os genes experimentalmente.

A escolha do tema pode ser atribuída ao crescente reconhecimento da complexidade e diversidade do uso do termo “gene”. No atual estágio da pesquisa científica, a resposta para a pergunta aparentemente simples: “o que é o gene?” envolve um sem número de problemas, cada qual com um conjunto próprio de soluções necessariamente históricas. Desde sua descoberta por Mendel¹, em 1865, até os dias de hoje, esse conceito é alvo de constantes reformulações. A descoberta, em 1953, da estrutura físico-química do DNA, a molécula química em que se encontra o gene, implicou numa profunda mudança no modo como pensamos sobre muitas doenças, sobre o melhoramento animal e vegetal, e sobre a própria compreensão da história da espécie humana.

Engenharia genética, clonagem, projetos genoma (humano e de outros animais, plantas e microrganismos) e projetos proteomas, são apenas as partes mais visíveis de um *iceberg* que ganhou massa ao longo do século XX, principalmente a partir da década de 1970, quando passou a pressionar a sociedade a rever não só as velhas formas de conceber a saúde, mas também a própria reprodução humana e até mesmo a eugenia. Cumpre lembrar que o objetivo desta dissertação não é recontar essa longa história repleta de episódios já estudados por outros autores (Keller; 2000; Mayr; 1998; Jacob, 1998).

Alguns poderão se perguntar o porquê de – em meio a tantos outros eventos moleculares e tantas outras definições de gene – investigar a importância dos íntrons para um conceito específico de gene? Como já exposto acima, mesmo nos dias de hoje, a definição de “gene molecular clássico” ainda é a definição de gene mais comum. Sua permanência como tal é uma surpresa quando consideramos o expressivo número de descobertas feitas no campo da biologia molecular desde a década de 1960. Acreditamos que essa definição tenha se mantido sem muitos questionamentos em razão do fato de que até aqui o “dogma central” se mostrou fundamentalmente correto (Maynard-Smith, 2001).

O interesse em realizar uma historiografia sobre a descoberta dos íntrons se deve a hipótese de que a historiografia das ciências parece não ter sido capaz – como demonstram algumas evidências – de avaliar corretamente a sua importância para a história do conceito de gene. Nesse sentido, 1977 é um ano chave. Nesse ano, surgiu, pela primeira vez, a concepção de que os genes, sob o ponto de vista molecular, não são estruturas lineares sem interrupções. Essa descoberta surpreendeu os cientistas que se dedicavam a analisar a estrutura do gene e seus mecanismos de expressão (Aloni *et al.*, 1977; Berget *et al.*, 1977; Brack & Tonegawa, 1977; Breathnach *et al.*, 1977; Chow *et al.*, 1977, Glover & Hogness, 1977; Jeffrey & Flavell, 1977; Sambrook, 1977; Williamson, 1977; Gilbert, 1978; Chambon, 1981).

Revoluções kuhnianas na biologia?

¹ Como se sabe, Mendel descobriu como se comportava os elementos hereditários, mas não foi ele que os denominou de “genes” (Moss, 2001; Keller, 2000; Mayr, 1998; Falk, 1986; Mendel, 1865).

Com a morte, em 1996, do historiador e filósofo da ciência Thomas Kuhn, foram retomados os debates sobre a validade da idéia de revoluções científicas – desenvolvida, sobretudo, por Kuhn para lidar com as ciências físicas – às ciências biológicas (Strohman, 1997; Wilkins, 1996). Em *A Estrutura das Revoluções Científicas* (Kuhn, 1994), cuja primeira edição foi publicada no ano de 1962, Kuhn argumenta que a ciência compreende duas fases marcadamente diferentes: a ciência normal e a ciência revolucionária. Esta última, de acordo com o autor, envolve mudanças conceituais profundas no corpo do conhecimento científico vigente. São as chamadas “revoluções científicas”.

Durante o período de ciência normal, os cientistas trabalham dentro dos limites prescritos por uma estrutura conceitual – uma teoria-modelo ou uma técnica – que Kuhn denominou de paradigma. Todos os problemas e soluções encontradas devem estar contidos dentro do paradigma adotado. Kuhn classificou esse procedimento como o trabalho de resolução de quebra-cabeças (Alves-Mazzotti & Gewandszadner, 1998; Strohman, 1997; Wilkins, 1996; Kuhn, 1994; Olby, 1975). O termo “quebra-cabeças”, na concepção de Kuhn, tem origem na proposição de que durante a fase de desenvolvimento da ciência normal, a não resolução de um problema prescrito pelo paradigma é considerada uma falha do pesquisador e não um problema do paradigma empregado. “Em geral, o projeto cujo resultado não coincide com essa margem estreita de alternativas é considerado apenas uma pesquisa fracassada, fracasso que não se reflete sobre a natureza, mas sobre o cientista.” (Kuhn, 1994, p.58). A adesão ao referido paradigma abre caminho para a prática de uma pesquisa mais detalhada e eficiente. Esse contínuo trabalho dentro da estrutura conceitual conferida pelo paradigma tem o objetivo de fortalecê-lo. Contudo, esse mesmo trabalho tem caráter conservador, ou seja, não tem função de descobrir novidades substantivas. (Alves-Mazzotti & Gewandszadner, 1998).

No entanto, a prática continuada da busca de soluções para resolver os quebra-cabeças do paradigma, pode evidenciar as suas fraquezas. No momento em que surjam dados anômalos intratáveis para a estrutura conceitual trabalhada dentro do paradigma, este poderia ser sacrificado para dar lugar a um novo paradigma. Nesse momento, a ciência passaria por uma fase revolucionária. (Kuhn, 1994; Olby, 1975). Geralmente, essas mudanças conceituais não são realizadas por figuras já estabelecidas e acostumadas em resolver os quebra-cabeças propostos pelo paradigma, mas por cientistas mais jovens (Kuhn, 1994; Wilkins, 1996). A nova teoria será aceita pela comunidade se ela também for capaz de resolver uma parte significativa dos problemas já resolvidos pela teoria antiga. A posição de Kuhn é claramente instrumentalista: uma teoria é apenas uma ferramenta para produzir previsões precisas, não tendo qualquer relação com a verdade. Teorias não seriam verdadeiras nem falsas, mas eficientes ou não eficientes (Alves-Mazzotti & Gewandszadner, 1998).

Alguns historiadores e cientistas se perguntam se existiriam, nas ciências biológicas, revoluções científicas nos termos propostos por Kuhn (Strohman, 1997; Wilkins, 1996). A revolução darwiniana com a publicação, em 1859, de *Origem das Espécies*; a revolução na biologia molecular com a descoberta da estrutura físico-química da molécula de DNA em 1953, e; a revolução mendeliana de 1900 com a redescoberta de seus trabalhos, são os três exemplos imediatos (Wilkin, 1996). Contudo, segundo Wilkins, nenhum dos três parece encaixar-se perfeitamente no modelo de Kuhn. Neste trabalho, mencionamos apenas as justificativas para o modelo que nos interessa, a saber, a descoberta da estrutura da molécula de DNA.

Em 1953, a descoberta da estrutura da molécula de DNA seria um caso de revolução kuhniana devido ao seu efeito imediato e à sua receptividade dentro da comunidade científica. Rapidamente, e de forma universal, ocorreu uma aceitação geral de que o modelo de Watson e Crick poderia explicar o que era o gene² (Wilkins, 1996). No entanto, até então não havia idéias claras sobre a estrutura do gene e de seu modo de ação. E porque não havia paradigma prévio a ser destituído, não teria havido uma disputa entre cientistas proeminentes de algum velho paradigma lutando contra a instituição do novo. (Wilkins, 1996). Dessa forma, como não teria ocorrido nenhuma transição entre um velho e um novo paradigma, isso descaracterizaria a descoberta da estrutura do DNA como uma fase revolucionária da ciência.

Para Robert Olby (1975), o modelo de Watson e Crick de fato transformou conceitualmente o pensamento sobre a transferência da informação genética. Segundo Olby, a estrutura do DNA proposta por Watson e Crick em 1953 pode ser legitimamente considerada como um momento de crise, de revolução paradigmática, pois substitui as idéias prévias dos bioquímicos e dos virologistas de que as proteínas eram não somente as moléculas responsáveis pela catálise das reações químicas nas células, como também a própria substância da hereditariedade. A indicação das proteínas como a substância responsável pela transmissão das características hereditárias foi arduamente defendida como legítima representante do paradigma prévio³ (Olby, 1975).

Todavia, Wilkins (1996) considera que revoluções biológicas certamente também ocorreram nas ciências biológicas, mas devido às características diferentes entre as ciências físicas e as biológicas, as revoluções ocorridas nessas últimas não poderiam se ajustar perfeitamente nos termos propostos por Kuhn.

Segundo Richard Strohman (1997), como paradigma da biologia molecular, o “dogma DNA – proteína”, ou teoria do gene, promete o diagnóstico molecular e terapia para toda e qualquer manifestação biológica: desde o problema do nascimento prematuro até a cura de doenças genéticas que levam a morte nas idades mais diversas e das mais diversas formas. Em razão dessas promessas exageradas, os programas de investigação da biologia molecular captariam, segundo Strohman, os maiores fundos de pesquisas dentro das ciências da vida. Porém, como a biologia molecular está mais preocupada em produzir substâncias moleculares de importância econômica, ela se descarta de uma compreensão mais epistemológica da vida (Strohman, 1997). Nesse sentido, Strohman acredita que os biólogos moleculares precisam redescobrir a profunda complexidade da relação genótipo/fenótipo. O paradigma atual, o “dogma central”, é incapaz de equacionar essa situação porque algo está faltando.

Para Strohman (1997), todos os aspectos de praxe do paradigma kuhniano, desde sua ascensão à sua queda estão presentes no paradigma atual (DNA – proteína): 1) é criado um consenso no qual a ciência normal fornece explicações para os mecanismos genéticos moleculares da vida; 2) ocorre uma mudança na percepção do mundo das espécies e dos organismos que passam a ser entendidos como máquinas genéticas, ou seja, uma ênfase maior sobre mecanismos 3) há um rápido recrutamento de novos cientistas para a articulação do novo paradigma; 4) são desenvolvidas novas tecnologias e nessas tecnologias são treinadas as novas gerações de cientistas e, por último; 6) com o prosseguimento inexorável da ciência normal, surgem anomalias com uma frequência cada vez maior que expõem as fraquezas do paradigma vigente. Dessa forma, o

² Para Mayr, não houve essa aceitação imediata (Mayr, 1998).

³ Paul Berg (1980) relata situação parecida.

paradigma original cai e é substituído por um novo. Contudo, como o próprio Strohman (1997) afirma, o mais importante atributo das revoluções científicas está ausente, já que parece ainda não haver um novo paradigma pronto que substitua o velho, a saber, o “dogma DNA – proteínas”.

De acordo com Strohman (1997), o paradigma científico atual omite em grande medida os sistemas dinâmicos:

There is a growing recognition that information for function may be not located solely in genomic databases. That is, it is becoming clear that sequence information in DNA, by itself, contains insufficient information for determining how gene products (proteins) interact to produce a mechanism of any kind. (Strohman, 1997, p.195).

Mesmo que ainda não haja um novo paradigma, todos os indicadores apontariam para a queda do paradigma atual, numa espécie de revolução incipiente. Nesse sentido, os danos sofridos pelo velho modelo, em algum momento, não poderão mais ser cientificamente suportados (Strohman, 1997).

De forma oposta a essa visão, o presente momento parece reunir todos os atributos de uma fase progressiva em um programa de pesquisa lakatosiano, como já indicado anteriormente. Para Lakatos (1979), a pesquisa científica seria mais bem explicada através de uma sucessão de teorias com algumas partes em comum: o cientista, a comunidade de cientistas, trabalha fazendo pequenas correções na teoria e substituindo-a por outra teoria ligeiramente modificada. Como também já mencionamos, essa sucessão de teorias é chamada de “programa de pesquisa científica”. O que não muda em um programa de pesquisa é o chamado “núcleo”. O núcleo é formado pelos princípios fundamentais de uma teoria. É ele que se mantém constante durante todo o programa de pesquisa, a medida em que as teorias são modificadas e substituídas por outras. O biólogo evolutivo Maynard-Smith (2001) afirmou que: “The dogma is perhaps the only statement in biology that is at the same time general, important, and – so far as we know – true.” (Maynard-Smith, 2001, p. 1496). Se tomarmos como referência a afirmação de Maynard-Smith, parece que podemos considerar o “dogma central” como o núcleo da biologia molecular.

Para tentar resolver as anomalias, as inadequações entre as previsões teóricas e as observações dos experimentos, modificam-se as hipóteses auxiliares (o cinto de proteção, que protege o núcleo contra refutações). No entanto, se houver mudanças no núcleo, estaremos, automaticamente, diante de um novo programa de pesquisa, com um outro conjunto de teorias, previsões e de testes (Alves-Mazzotti & Gewandszadner, 1998; Lakatos, 1979). Mas não parece ser esse o caso da biologia molecular.

A não alteração do núcleo permitiria uma melhor orientação da pesquisa científica, fornecendo sugestões sobre como mudar as hipóteses auxiliares, até que a observação esteja em concordância com o núcleo do programa. A recusa de um cientista em aceitar refutações ao núcleo central em seu programa de pesquisa será racional, segundo Lakatos (1979), enquanto o programa for capaz de modificar as hipóteses auxiliares de forma a gerar previsões de novos fatos. Assim, é racional recusar um programa não por causa de suas eventuais refutações ou por sua incapacidade de resolver anomalias, mas sim quando ele for incapaz de prever fatos novos. O que conta é o sucesso na previsão de fatos novos, conjecturas que possam explicar, melhor que o programa anterior, um determinado fenômeno ou dar uma contribuição maior para a solução de um determinado problema. Dentro desse quadro, estaria o gene molecular em crise?

Por uma história de problemas

O biólogo evolutivo e historiador da biologia Ernst Mayr (1998) sugere que a história das ciências revela diferentes conceitos de como escrever uma história. Embora todas as histórias sejam uma combinação de várias abordagens ou estratégias, Mayr indica que elas podem ser classificadas de cinco formas: 1) histórias lexicográficas, descritivas; 2) cronológicas, que abordam o que aconteceu em um determinado período histórico; 3) biográficas, que tratam os progressos científicos através das vidas dos principais cientistas; 4) culturais e sociológicas, que enfatizam que a ciência é uma forma de atividade humana; e 5) uma história de problemas (Mayr, 1998, pp.16-23).

Para Mayr (1998), uma história de problemas é particularmente apropriada para a biologia devido à longevidade de seus problemas. Esse é exatamente o caso do gene e da genética. Como em qualquer história, essa abordagem, obviamente, não descarta as outras formas de análises. No entanto, elas são tratadas como complementares. O tipo de abordagem histórica que estamos optando enfatiza a história das tentativas de solução dos problemas. No caso específico dessa pesquisa, investigaremos porque o ano de 1977 é crítico para a estabilidade do conceito de “gene molecular clássico”. Na história de problemas, a ênfase concentra-se no cientista atuante e no seu mundo conceitual. Aqui, peço licença para fazer minhas as palavras de Mayr. “Tendo em vista que eu sou um biólogo, sou mais bem qualificado para escrever uma história da biologia do que uma história biográfica ou sociológica.” (Mayr, 1998, p. 22).

Assim, no primeiro capítulo deste trabalho, apresentaremos uma história conceitual do estabelecimento do “dogma central da biologia molecular”, ou seja, a idéia de que a informação genética flui do DNA para a proteína. A descoberta da estrutura físico-química da molécula de DNA e a subsequente quebra do código genético, ocorridas durante as décadas de 1950 e 1960, talvez tenha trazido em seu âmago a idéia de que a genética, e, por conseguinte, o próprio gene molecular, precisava agora apenas de polimentos.

No segundo capítulo, apresentaremos o contexto científico da descoberta dos genes interrompidos durante o ano de 1977. Essa descoberta marcaria a passagem das pesquisas genéticas de organismos menos complexos, como as bactérias, para a organização mais complexa dos seres eucarióticos. Na medida em que bactérias não têm íntrons, os estudos genéticos nesses organismos indicavam os genes como uma estrutura linear, contínua. Os íntrons parecem ser elementos tardios em uma ontologia dos genes.

Ainda nesse capítulo, analisaremos uma das principais conseqüências da descoberta dos genes interrompidos, a saber, o processamento alternativo da molécula de RNA, responsável por mudanças significativas no dogma “um gene – uma proteína”, já que com o advento do processamento alternativo, um gene pode estar envolvido na síntese de muitas proteínas.

No final desse mesmo capítulo, apresentaremos dois pólos de um debate relativamente recente. Em *The Century of the Gene*, Evelyn Fox Keller (2000) supõe que o gene, como conceito explicativo, teve o seu reinado no século XX, mas que agora, dada à crescente complexidade do mundo molecular, “o próprio uso do termo gene se tornou um obstáculo à sua própria exposição” (Keller, 2002, p. 84). É certo que desde Mendel até os dias atuais as definições de gene têm se modificado. A leitura do texto de

Keller revela, no entanto, que ela não está pondo em xeque uma definição do gene em geral, mas a definição de “gene molecular clássico”, que ora nos propomos a analisar. A bibliografia que investigamos indica que discussão acerca da definição do gene nunca foi desprovida de problemas.

Do outro lado do debate, Raphael Falk (1986), geneticista emérito da *Hebrew University of Jerusalem* em Israel, acredita que o gene pode ser uma abstração, que pode ter ou não uma correspondência com a realidade. De qualquer forma, para Falk, o conceito de gene continua mantendo sua utilidade teórica para a explicação dos fenômenos biológicos.

Obviamente, a história da biologia molecular também é uma história de inúmeros interesses políticos, sociais e filosóficos. Na atualidade, a biologia molecular afeta as estruturas da sociedade e da indústria ao atrair pesados investimentos das principais multinacionais do setor de medicamentos e dos setores agrário e pecuário. O foco de nossa pesquisa, contudo, deverá estar centrado no desenvolvimento de um modelo molecular com grande impacto operacional dentro da biologia. O gene como uma entidade física passível de manipulação concreta e objetiva.

CAPÍTULO I
O DOGMA CENTRAL DA BIOLOGIA
MOLECULAR

O gene molecular clássico

*Optimists believe that ours
is the best of all possible
worlds. And pessimists are
those who fear that the
optimists are right.*

*Dobzhansky, Theodosius.
“Changing man.” Science,
vol. 155, 1967.*

Para a história do conceito de gene, o ano de 1977 é de capital importância. Foi ao longo desse ano que os genes interrompidos foram descobertos. Para os historiadores da biologia moderna, é fato que o conceito de gene foi, e ainda é, alvo de constantes reformulações (Keller, 2000; Chambon, 1981). Um levantamento da história desse conceito revela que desde que esse termo foi criado sempre houve conflitos em relação ao seu significado (Falk, 1986). Não é nossa intenção apresentar uma outra versão para a história do gene, mas sim examinar algumas implicações da descoberta dos genes interrompidos para a definição do chamado “gene molecular clássico” (CMG – *Classical Molecular Gene*) (Griffiths, 2002; Griffiths & Neumann-Held, 1999).

O termo “gene” foi cunhado, em 1909, pelo biólogo Dinamarquês Wilhelm Johannsen para denominar os “fatores” hereditários de Mendel (Moss, 2001; Keller, 2000; Falk, 1986; Chambon, 1981). A intenção de Johannsen era simplificar as idéias correntes envolvendo o termo “unidade de caráter”. Entendia-se por unidade de caráter dois conceitos distintos: 1) qualquer caráter visível de um organismo que se comporta como uma unidade de herança de Mendel; e 2) por consequência, aquilo que na célula produz o caráter visível. Com isso, Johannsen estabelece o que não havia ficado claro na obra de Mendel: a distinção entre o potencial para a característica (o genótipo) e o traço em si (o fenótipo) (Mayr, 1998; Falk, 1986). Johannsen sugere o termo “gene” para o determinante hipotético, na célula, responsável pelas características dos organismos. Etimologicamente, o termo “gene” origina-se de *génos*, radical do verbo grego *gígnesthai*, que significa “nascer”. Assim, *génos* pode ser entendido como “origem”, “o que gera”, “o que produz”. Nesse sentido, o gene é aquilo que gera, que dá origem, que produz alguma característica do organismo. Para Johannsen, não seria possível contabilizar quantos determinantes (genes) estariam envolvidos na produção de um único caráter, e tampouco o que seriam esses determinantes materialmente (cf. Falk, 1986). Dessa forma, Johannsen não procurou responder as perguntas de como agem tais determinantes para produzir ervilhas lisas ou rugosas? Qual a base física do gene?

Somente após quase meio século de indefinições quanto à natureza física e a função dos genes, essas começaram a ser descobertas. Em 1940, os bioquímicos norte-americanos George Beadle e Edward Tatum definem a função do gene através da hipótese “um gene – uma enzima”. Com essa hipótese, fica estabelecido que um gene é responsável pela produção de uma enzima específica, e essa, por sua vez, atuaria em etapas do metabolismo da célula, como, por exemplo, na degradação dos alimentos, na síntese de novos compostos para a célula, nos processos respiratórios, etc. (Jacob, 1983).

Em 1944, o bacteriologista canadense Oswald Avery e seus colaboradores identificam a molécula de DNA, o ácido desoxirribonucléico, como o material hereditário (Avery *et al.*, 1944). Contudo, somente em 1953 é que a estrutura físico-química do DNA foi revelada. Coube ao biólogo norte-americano James Watson e ao físico britânico Francis Crick descobrir a estrutura da molécula de DNA e especular como ela responderia aos requisitos necessários da replicação e armazenamento da informação genética (Watson & Crick, 1953a e 1953b). Em sua formulação, eles argüiram que uma das fitas serviria como um molde molecular para a replicação genética, e o armazenamento das informações genéticas estaria numa espécie de código inscrito nas seqüências de bases do DNA. Em 1957, antes mesmo da quebra do código genético, ocorrido durante a década de 60, Francis Crick já possuía dados suficientes para formular o controverso “dogma central da biologia molecular”, postulando que o fluxo de informações genéticas segue *unidirecionalmente* do DNA para o RNA e daí para as proteínas (Mayr, 1998; Olby, 1975).

Todos esses trabalhos consolidam a noção de “gene molecular clássico”, definição que se tornaria corrente no meio científico e que continua sendo empregada ainda nos dias de hoje (Griffiths, 2002). O gene emerge como um trecho, uma seqüência específica da molécula de DNA que possui a informação para a produção de uma proteína. Esta, por sua vez, está envolvida numa função particular dentro do ambiente celular. Esse gene, definido como uma região discreta de DNA (unidade de estrutura) e com funções específicas no desenvolvimento da célula por meio de um produto protéico (unidade de função), parecia perfeitamente adaptado para suportar as necessidades requeridas pelas pesquisas em biologia molecular (Falk, 1986).

Todavia, em 1968, ao resenhar uma história da biologia molecular por ocasião do lançamento de uma coleção de ensaios intitulado *Phage and the origins of molecular biology*, Gunther Stent (1968), professor emérito de biologia molecular da Universidade da Califórnia em *Berkeley*, afirma que como disciplina científica, essa mesma biologia molecular dava sinais de seu declínio. Segundo Stent, após 1963, essa disciplina entrava em sua fase acadêmica uma vez que para a maioria dos pesquisadores a expectativa era de que os principais problemas estavam resolvidos e o que restava era apenas o trabalho de “polir os detalhes” (Keller, 2000; Falk, 1986; Stent, 1968)⁴. Para Stent, muitos dos detalhes do código genético estavam resolvidos e a colinearidade entre o gene – a seqüência de nucleotídeos no DNA – e a seqüência de aminoácidos na proteína havia sido finalmente provada.

⁴ “(...) and what remained now was the need to iron out the details.” (Stent, 1968, p. 394).

Dez anos mais tarde, Walter Gilbert, professor de biologia molecular na Universidade de *Harvard*, afirma na primeira frase de um breve artigo na revista *Nature*, a mais importante publicação de divulgação científica do Reino Unido, que “our picture of the organisation of genes in higher organisms has recently undergone a revolution.” (Gilbert, 1978, p. 501). Para Walter Gilbert, a revolução atende pelo nome de “genes interrompidos”.

Durante o ano de 1977, diversas equipes de pesquisadores documentam que os genes apresentam inserções de trechos de DNA que não encontram correspondência no RNA mensageiro que especifica os aminoácidos na proteína (Glover & Hogness, 1977; Berget *et al.*, 1977; Chow *et al.*, 1977; Sambrook, 1977; Aloni *et al.*, 1977; Breathnach *et al.*, 1977; Williamson, 1977; Brack & Tonegawa, 1977; Jeffreys & Flavell, 1977; Gilbert, 1978; Chambon, 1981). Em suma, a seqüência de DNA cuja função é responder pela produção de uma proteína se encontra dividida (interrompida) por trechos de DNA, a princípio, sem nenhum sentido. Walter Gilbert (1978) sugeriu o termo *introns* (*intra-genic regions*) para as regiões de DNA sem sentido e *exons* (*expressed regions*) para as seqüências de DNA expressas em proteínas. É interessante notar, desde já, que os exons têm a possibilidade de serem combinados alternativamente, gerando a partir de uma mesma seqüência de DNA diferentes proteínas. Examinaremos, mais adiante, as conseqüências disso para o conceito de “gene molecular clássico”.

Nesta dissertação, analisaremos de que modo a descoberta dos genes interrompidos contribuiu para minar o conceito de “gene molecular clássico”, oriundo da década de 60. Obviamente, os genes interrompidos são apenas um entre os diversos elementos e processos moleculares que conflitam com o modelo proposto entre as décadas de 1940 a 1960. Contudo, a descoberta dos genes interrompidos, os íntrons, parece ser a primeira manifestação clara na direção de uma reformulação da definição de gene molecular, dado que o seu emprego, de forma irrestrita, não pode ser mais aceito. Nesse sentido, a descoberta dos íntrons é um sinal suficientemente claro de quão equivocado estava Gunther Stent (1968) ao afirmar a decadência da biologia molecular.

Para compreender o impacto da descoberta dos genes interrompidos sobre o conceito de “gene molecular clássico”, é imperativo examinar alguns eventos históricos que foram determinantes na constituição desse conceito de gene. Para tanto, escolhemos alguns

episódios importantes da construção desse conceito para ser apresentado no presente capítulo: 1) em 1940, o trabalho de George Beadle e Edward Tatum; 2) em 1953, a descoberta da estrutura físico-química da molécula de DNA como base da estrutura física do gene; 3) durante a década de 1960, a quebra do código genético e; 4) em 1957, o pronunciamento “dogma central da biologia molecular” de Francis Crick. Antes de discorrer sobre esses eventos, contudo, faremos um breve histórico sobre a importância das proteínas.

Além dos ácidos nucleicos (DNA e RNA), essa outra classe de moléculas orgânicas, as proteínas, é de importância fundamental para a compreensão dos sistemas vivos. Na medida em que estamos, freqüentemente, falando da relação entre genes e proteínas, não faria sentido deixar de mencionar essas moléculas. De uma certa forma, pode-se dizer que os genes são uma espécie de receita para sintetizar proteínas (Ridley, 2003).

As proteínas

As proteínas são a base material de tudo que é vivo. Entendidas como moléculas de importância essencial para a vida, as proteínas foram descobertas experimentalmente em finais do século XIX e começo do século XX, e marcaram o início da bioquímica moderna. Estudando as reações químicas que ocorrem nos corpos dos seres vivos, os bioquímicos logo se deslumbraram com as enzimas. A partir da análise cuidadosa de tecidos e células e o exame de seu conteúdo, as centenas de reações químicas que ocorrem nas células se mostraram acessíveis e passíveis de serem analisadas. Essas reações obedecem aos comandos específicos e precisos da ação enzimática, isto é, cada reação observada é

catalisada por enzimas específicas. Para os bioquímicos, a partir dessas descobertas, fica óbvio que estariam na natureza e na qualidade das proteínas os caminhos que levariam aos segredos da vida. (Jacob, 1983).

De modo análogo aos genes – as unidades de estudo da hereditariedade no começo do século passado –, as proteínas são as unidades de execução, promovendo as diversas reações químicas dos seres e compondo a arquitetura desses organismos. (Jacob, 1983). Desde a aparência e constituição da pele, órgãos, tecidos e membranas, até o funcionamento dos mesmos, regidos por enzimas e hormônios protéicos, as proteínas exercem um papel fundamental para os seres vivos. Obviamente, outras substâncias químicas participam também desses processos: a água, os lipídios (gorduras) e os sacarídeos (açúcares). A nossa constituição básica, no entanto, é feita de proteínas. Desse modo, é de todo justificável a importância que as proteínas exerceram e exercem.

No início do século XX, mediante a “teoria cromossômica da herança”, já havia suspeitas de que os cromossomos eram os portadores da base física da herança genética (Mayr, 1998). Basicamente, os cromossomos são constituídos de nucleoproteínas e ácido nucléico. O ácido nucléico era compreendido na época como “uma espécie molecular sem variedade nem fantasia, portanto sem aptidão para desempenhar qualquer papel na hereditariedade.” (Jacob, 1983, p. 248). Se assim eram vistos os ácidos nucléicos, as proteínas, por outro lado, eram cada vez mais estudadas e fascinavam os bioquímicos por suas múltiplas e importantes funções na química dos seres vivos. O DNA parecia servir como uma espécie de esqueleto onde se ancorariam as proteínas nos cromossomos. (Mayr, 1998).

A especificidade das funções biológicas deveria emanar das proteínas, o que equivaleria a dizer que o fluxo de informações genéticas deveria passar de proteínas para proteínas. Nesse sentido, provavelmente, seriam as nucleoproteínas as responsáveis por, de

alguma forma, especificar a ordem dos blocos de construção das proteínas, os aminoácidos. A hipótese “um gene – uma enzima” ajudava a confirmar ainda mais a importância das moléculas de proteínas.

A idéia da relação entre um gene e uma proteína é um dos pressupostos básicos da pesquisa genética que perdura até os dias de hoje. Embora caiba aos genes o armazenamento e a passagem da informação genética, são, principalmente, as proteínas que protegem, empacotam e regulam a atividade biológica.

Breves antecedentes ao conceito de gene molecular clássico

O conceito de gene, que em seus primórdios era considerado apenas a unidade de herança, sofreu, com o passar do tempo, contínuas transformações para acomodar novas estruturas e modos de ação (Fogle, 2000). A unidade de herança por excelência é o gene de Mendel, no qual sua suposta existência foi inferida conciliando modelos matemáticos e manifestações fenotípicas (Mayr, 1998; Dobzhansky, 1967). Mendel postulou a existência de genes que determinariam a cor das sementes, a forma das sementes, a forma da vagem, etc., e calculou, predizendo, com que frequência essas características apareceriam nas gerações seguintes – fornecendo a base para o desenvolvimento da genética mendeliana. Reconhece-se o gene mendeliano por meio de um determinado caráter visível, o fenótipo, ou mais especificamente, através das manifestações alternativas do fenótipo em questão, a variação: semente verde ou amarela; flores axiais ou distais, plantas baixas ou altas. Essa era a única forma de inferir e identificar a presença dos genes: pelos seus efeitos no organismo

(Falk, 1986). No entanto, Mendel não mencionou o termo “gene”, mas o termo alemão *Elemente* (Mayr, 1998, p. 798), traduzido para o inglês como *cell elements* (Stern, 1966, p. vi.). Tais termos correspondem ao que hoje denominamos de gene (Mayr, 1998). Contudo, para o monge austríaco não havia sentido em especular sobre alguma base física para esses *Elemente*.

No final da primeira década do século XX, sob a tutoria do embriologista norte-americano Thomas Hunt Morgan, da Universidade de *Columbia* em *New York*, formou-se o que ficou conhecido como o “Grupo das *Drosófilas*”, que desempenharia o papel de um dos principais articuladores da chamada “teoria cromossômica da herança”. A teoria cromossômica da herança surgiu após a devida percepção da semelhança entre o comportamento dos cromossomos durante a meiose⁵ e as previsões de Mendel, unindo a citologia e a genética. A referida teoria afirmava que os cromossomos eram os portadores das unidades hereditárias, dando, assim, um início de materialidade aos genes (Mayr, 1998; Reznik, 1995).

O aparecimento de mutantes, espontâneos e induzidos⁶, providenciou para o grupo de Morgan o material para analisar extensivamente a prole dos mais diversos cruzamentos. Isto tornou possível dissecar os aspectos mais importantes da transmissão das informações genéticas, permitindo

⁵ A meiose é o fenômeno de divisão das células germinativas responsável pela produção dos gametas, que possuem apenas a metade do material genético da célula original. Nós seres humanos temos 46 cromossomos em cada uma de nossas células organizados em 23 pares. Recebemos 23 cromossomos de nossa mãe e 23 de nosso pai. Durante a formação das células especializadas na reprodução, os gametas, os cromossomos dessas células sofrem um processo denominado de meiose, que tem por resultado a redução da quantidade de cromossomos à metade. Maduros os gametas contêm somente a metade da quantidade normal de cromossomos. A união dos gametas reconstitui o número normal de 46 cromossomos, 23 pares.

⁶ A descoberta de que a incidência de raios-X pode modificar o material genético valeu o Nobel de medicina ou fisiologia em 1946 à Hermann Joseph Muller, aluno de Morgan. Muller demonstrou que as mutações introduzem algum efeito nos genes que tem como consequência uma alteração dos padrões normais de desenvolvimento do organismo. No caso, alterações na forma do corpo e das asas e na cor dos olhos das moscas *drosófilas*.

estabelecer previsões sobre modos de hereditariedade e o teste imediato dessas previsões (Mayr, 1998). Mas a questão da natureza física do material hereditário não parecia relevante o suficiente para alterar o rumo dos trabalhos empreendidos pelo grupo de Morgan. Em sua conferência quando foi laureado com o Nobel de medicina ou fisiologia em 1933 pela demonstração do papel dos cromossomos na hereditariedade, Morgan afirmou que:

What is the nature of the elements of heredity that Mendel postulated as purely theoretical units? What are genes? Now that we locate them in the chromosomes are we justified in regarding them as material units; as chemical bodies of a higher order than molecules? Frankly, these are questions with which the working geneticist has not much concern himself, except now and then to speculate as to the nature of the postulated elements. There is not consensus of opinion amongst geneticists as to what genes are - whether they are real or purely fictitious - because at the level at which the genetic experiments lie, it does not make the slightest difference whether the gene is a hypothetical unit, or whether the gene is a material particle. (Morgan, 1934, p. 315).

1940: A Hipótese “Um Gene – Uma Enzima”

Em 1940, ainda não se sabia de que material eram feitos os genes. Mas foi precisamente nesse ano que os bioquímicos norte-americanos George Beadle e Edward Tatum respondem à pergunta: o que os genes fazem? As enzimas são proteínas que aumentam milhares de vezes a velocidade das reações químicas. Beadle e Tatum suspeitavam que os genes atuavam regulando seqüências de eventos químicos por intermédio da ação enzimática (Beadle, 1958).

Por uma série de transformações, os alimentos são tratados primeiro por enzimas específicas que os degradam, fragmentam, transformam em pequenas moléculas. Estas servem então de substrato a outras enzimas que os remodelam, adicionam átomos,

substituem radicais, dilatam, soldam; em suma, produzem os componentes característicos do organismo. Este torna-se um tipo de fábrica química onde pulula uma multidão de pequenas moléculas formadas, a partir dos alimentos, nas cadeias de degradação e transformadas em compostos específicos por cadeias de síntese. (Jacob, 1983, p. 244).

Nesse sentido, a formação de um determinado composto X passa por uma série de precursores mais simples, interconvertidos por ação enzimática (precursor W – precursor Z – precursor Y e finalmente X). Se uma dessas etapas é interrompida, não há formação do produto imediatamente posterior e, por conseguinte, do composto final.

Para investigar a interferência do controle gênico sobre as reações químicas, Beadle e Tatum utilizaram como organismo experimental o fungo do pão *Neurospora crassa* e nele incidiram Raios-X para provocar-lhe mutações. Nessa época, já se sabia que: A) os processos bioquímicos podem ser decompostos em uma série de etapas individuais; B) cada reação química é controlada por uma enzima; C) as mutações afetam o material genético. A partir desse experimento, Beadle e Tatum esperavam verificar se ocorreria uma alteração na habilidade da célula em prosseguir com suas reações químicas devido à incidência dos Raios-X (mutações). Como isso de fato aconteceu, puderam concluir que são os genes que controlam as reações químicas através da produção de enzimas específicas para cada tipo de reação química (Beadle, 1958). Como já era esperado, Beadle e Tatum notaram que as mutações interferiram na formação de enzimas intermediárias, necessárias para a interconversão dos precursores no composto final. Logo, se o gene é atingido, interrompe-se a cadeia de reações. A conexão entre genes e enzimas fornecida por Beadle e Tatum

ficou conhecida como a hipótese “um gene – uma enzima” (Keller, 2000; Mayr, 1998; Falk, 1986; Beadle, 1958).

Por esse trabalho Beadle e Tatum foram laureados com o Nobel de medicina ou fisiologia no ano de 1958. Contudo, como o próprio Beadle (1958) fez questão de reconhecer, a idéia seminal de que um gene é responsável pela produção de uma proteína específica já havia sido proposta, em 1908, pelo médico e bioquímico londrino Archibald Edward Garrod (Neel, 1994). Em suas pesquisas sobre o que ele denominou de erros inatos do metabolismo, Garrod reconheceu que a alcaptonúria – uma doença em que a urina fica escurecida pela não degradação de uma substância química, o alcapton – tinha um histórico familiar que se comportava como uma herança mendeliana recessiva. O resultado era um bloqueio em algum ponto do curso normal do metabolismo (Neel, 1994). Segundo Beadle (1958), Garrod foi o primeiro a pensar como genes, enzimas e reações químicas poderiam ser inter-relacionadas. Após Garrod, tal idéia foi paulatinamente desenvolvida por outros grupos de pesquisa.

What was later called the “one gene-one enzyme” concept was clearly in our minds at this time although as I remember, we did not so designate it. Ours was a scheme closely similar to that proposed by Garrod for alcaptonuria, except that he did not have genes that blocked an adjacent reaction in the sequence. But at the time we were oblivious of Garrod’s work, partly because geneticists were not in the habit of referring to it, and partly through failure of ourselves to explore the literature. Garrod’s book was available in many libraries. (Beadle, 1958, p. 592).

It is sometimes thought that the *Neurospora* work was responsible for the “one gene-one enzyme” hypothesis - the concept that genes in general have single primary functions, aside from serving an essential role in their own replication, and that in many cases this function is to direct specificities of enzymatically active proteins. The fact is that it was the other way around - the hypothesis was clearly responsible for the new approach. Although it may not have been stated explicitly, Ephrussi and I had some such concept in mind. (...) Regardless of when it was first written down on paper, or in what form, I myself am convinced that the one gene-one enzyme concept was the product of gradual evolution beginning with Garrod and contributed to by many including Moore, Goldschmidt, Troland, Haldane, Wright, Grüneberg and many others. (Beadle, 1958, p. 596-597).

Assim, sob o olhar retrospectivo, o trabalho de Beadle e Tatum pode ser considerado como o primeiro bloco de construção do “dogma central da biologia molecular” de Francis Crick, e da idéia de que o “gene molecular clássico” se manifesta sob a forma de produção de uma proteína. Como se pôde notar pelo relato de Beadle, esse resultado não foi exatamente surpreendente, mas antes, a confirmação de uma noção subjacente que se

desenrolava devido a já sabida importância atribuída às proteínas no que toca os fenômenos biológicos.

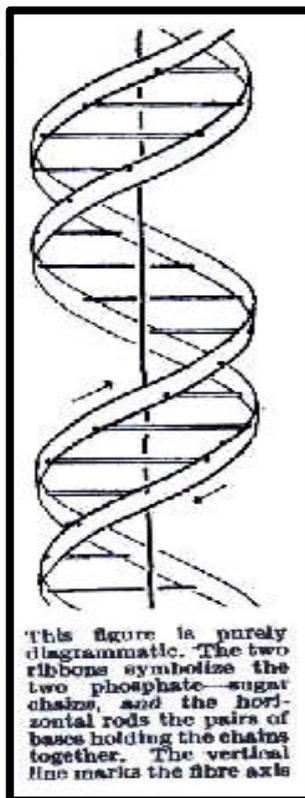
A base físico-química da hereditariedade e a dupla-hélice do DNA

Apesar de estabelecido que a função dos genes seria a responsabilidade pela produção de uma proteína, não se sabia qual era estrutura físico-química do gene, sua base material na célula. As proteínas, além de regular as atividades metabólicas das células e arquitetar a estrutura dos organismos, eram as moléculas sobre as quais recaíam as principais suspeitas de serem o suporte químico da hereditariedade (Olby, 1975). A hereditariedade é a relação de similaridade entre os pais e sua prole garantida pelos mecanismos reprodutivos (Griesemer, 2000). Para a teoria evolutiva de Darwin, o que é selecionado torna-se permanente através da hereditariedade. Somente as características dos organismos que podem ser transmitidas pela linhagem de células germinativas têm importância evolutiva. É responsabilidade do material hereditário armazenar e transmitir a experiência evolutiva das gerações passadas. Dessa forma, o material genético se encarrega, paralelamente, por duas causas: 1) as causas remotas (evolutivas), que transmitem aos seres a experiência acumulada pelo longo tempo evolutivo; e 2) as causas imediatas, referentes à informação que responde pelo desenvolvimento do organismo desde uma única célula ovo até o estágio adulto (Mayr, 1998). Para Jacob, “um organismo é apenas uma transição, uma etapa entre o que foi e o que será.” (Jacob, 1983, p.10).

A reprodução dos organismos e das moléculas que compõem os organismos depende, basicamente, dos ácidos nucleicos, a molécula responsável por transmitir através das gerações as instruções para a construção dos seres (Jacob, 1983). Contudo, somente a partir de 1944 com as pesquisas dos bacteriologistas canadenses Oswald Theodore Avery e

Collin MacLeod e do norte-americano Maclyn McCarty (1944) que o DNA começa a ser imaginado como a base química da hereditariedade. A tese segundo a qual a molécula de DNA era responsável pelo armazenamento e transmissão das informações hereditárias contrariava a noção vigente de que caberia às proteínas esse papel. Contudo, não foi sem resistência que a comunidade científica recebeu o anúncio de que as proteínas não tinham qualquer tipo de participação na hereditariedade (Friedman & Friedland, 2000; Mayr, 1998; Olby, 1975).

Mas, em 25 de abril de 1953, a revista científica britânica *Nature* publica, numa mesma edição, três artigos que se propõem a apresentar a estrutura físico-química do ácido desoxirribonucléico, o DNA: Watson & Crick, 1953a; Wilkins *et al.*, 1953; Franklin & Gosling, 1953. No entanto, foram os nomes de Watson e Crick que ficaram mais conhecidos na história das ciências como a dos descobridores da estrutura do DNA⁷ (Figura 1).



⁷ Adiantamos que não está no escopo dessa revisão os méritos pela descoberta da estrutura do DNA. Para mais informações, veja Friedman & Friedland, 2000; no próprio livro de bibliografia não se esgota nesses poucos nomes, pois na fisiologia ou medicina foram Watson, Crick e Franklin, e em câncer em abril de 1958.

essa revisão para quem se devem os créditos. Mais informações podem ser encontrados em Maddox (2003), sendo que essa estrutura foi proposta em 1962, os laureados ao Nobel de Física foram Watson, Crick e Franklin já havia falecido de

Figura 1: A estrutura físico-química do gene, o DNA. O desenho acima foi feito pela esposa de Crick, Odile Crick, e publicado no famoso artigo de 900 palavras da revista *Nature* de 25 de abril de 1953. Os significados biológicos dessa estrutura foram enfatizados por Watson e Crick nesse trabalho de abril e, principalmente, no trabalho imediatamente posterior em 30 de maio nesse mesmo periódico.

“We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.” (Watson & Crick, 1953, p. 737). Com esse parágrafo, que soa profundamente conjectural, James Watson e Francis Crick do *Cavendish Laboratory* em *Cambridge*, Inglaterra, iniciam o artigo de 25 de abril em que propõem um modelo para a estrutura físico-química da molécula de DNA. Cerca de um mês mais tarde, no artigo de 30 de maio dessa mesma revista, Watson e Crick (1953b) exploram as implicações biológicas da estrutura que propuseram, na medida em que: “The importance of deoxyribonucleic acid (DNA) within living cells is undisputed.” (Watson & Crick, 1953b, p. 964).

Já se sabia que a molécula de DNA era constituída por “blocos de construção” chamados de nucleotídeos, compostos, por sua vez, de um açúcar, um radical fosfato e de uma base nitrogenada. O que varia na composição de cada nucleotídeo do DNA são suas bases nitrogenadas. As bases nitrogenadas são de quatro tipos: Adenina (A), Citosina (C), Timina (T) e Guanina (G). A molécula de DNA é formada pelo emparelhamento de duas fitas complementares que se dispõem paralelamente (na verdade, anti-paralelamente), formando uma espécie de escada. Torcendo essa “escada”, chegamos à representação, já amplamente difundida, da dupla-hélice.

A partir dos dados do bioquímico austríaco Erwin Chargaff⁸, Watson e Crick perceberam que deveria ocorrer uma espécie de regra de pareamento entre os nucleotídeos que unem as fitas paralelas do DNA. De acordo com essa regra, A se parecia somente com T e C somente com G: “(...) if only specific pair of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.” (Watson & Crick, 1953a, p. 737). Dessa forma, cada fita apresenta uma seqüência de bases complementar à seqüência da outra fita. Se, por exemplo, uma fita contiver a seqüência (ATGCTTAG), na outra fita a seqüência complementar será (GACGAATC). Posteriormente, como consequência do pareamento específico das bases, Watson e Crick assinalaram que: “It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.” (Watson & Crick, 1953a, p. 737). Ou seja, o pareamento sugere um modelo de reprodução ou replicação que produz duas duplas-hélices a partir de uma única dupla-hélice.

Assim como haviam feito no início do primeiro artigo, Watson e Crick (1953b) voltam a adotar, no artigo de 30 de maio, um tom conjectural em algumas passagens. Nesse artigo mais recente, sugerem, por exemplo, especulações sobre quais seriam as implicações biológicas da estrutura da molécula de DNA: “Many lines of evidence indicate that it is the carrier of a part of *if not all* the genetic specificity of the chromosomes and thus of the gene itself.” (Watson & Crick, 1953b, p. 964. *Itálico nosso.*) Em outra

⁸ Chargaff estudou a composição química do DNA de vários organismos e percebeu que a quantidade de nucleotídeos adenina (A) era equivalente a de timinas (T) e de guaninas equivalente a de citosinas. Watson e Crick concluíram que onde tivesse uma adenina, haveria uma timina do outro lado da cadeia de DNA, e onde tivesse uma guanina haveria uma citosina.

passagem: “We have recently proposed a structure for the salt of deoxyribonucleic acid which, *if correct*, immediately suggests a mechanisms for its self-duplication.” (Watson & Crick, 1953b, p. 965. Itálico nosso.). Dois motivos parecem prováveis para essa avaliação de caráter ainda conjectural. O DNA, como estrutura portadora das informações hereditárias, ainda sofria algumas resistências decorrentes do fato de que até aquele momento não se admitia negar a participação das proteínas nas funções hereditárias. Na primeira passagem, ao incluir a ressalva “if not all”, os autores parecem indicar que naquele momento ainda consideravam as opiniões de que as proteínas poderiam, ao menos, dividir algum tipo de participação na especificação das informações hereditárias.

A nossa segunda suspeita recai no fato de que a elaboração da estrutura proposta foi baseada em evidências indiretas e cálculos, e, desse modo, ainda depender de futuras confirmações empíricas. A expressão “if correct”, assim como outras passagens do referido texto, parecem evidenciar isso: “Though the structure will not be completely proved until a more extensive comparison has been made (...)” (Watson & Crick, 1953b, p. 965). Todavia, em apenas um parágrafo, o artigo de maio de 1953 resume o significado biológico da estrutura do DNA.

It follows that in a long molecule many different permutations are possible, and it therefore seems likely that the precise sequence of the bases is the code which carries the genetical information. If the actual order of the bases were given, one could write down the exact order of the bases on the other one, because of the specific pairing. Thus one chain is, as it were, the complement of the other, and is this feature which suggests how the deoxyribonucleic acid molecule might duplicate itself. (Watson & Crick, 1953b, pp. 965-966).

Embora a replicação das características dos organismos seja uma

constante na natureza, é no surgimento de pequenas variações entre os organismos que depende o trabalho da seleção natural. Sem variação, não há seleção; sem mutação, não há variação. Watson e Crick também postularam que mutações pontuais poderiam ocorrer mediante alterações nas seqüências de bases. Devido à fidelidade do DNA durante o processo de replicação garantido pelo pareamento de bases, as alterações se perpetuariam após sofrerem mutação (Falk, 1986; Watson e Crick, 1953b). Como dito anteriormente, Watson e Crick deixam claro que o modelo proposto para a reprodução do DNA ainda deveria ser considerado como especulativo, e que restariam muitos detalhes a serem posteriormente esclarecidos (Watson & Crick, 1953b).

Um dos pontos deixados em aberto nessa discussão é aquele que dispõe sobre se haveria ou não a necessidade de algum tipo de enzima especial para a polimerização⁹ das novas cadeias de DNA, ou se as próprias hélices de DNA poderiam, de alguma forma, agir auto cataliticamente (Watson e Crick 1953b). Watson e Crick propuseram o modelo mecânico pelo qual ocorria a cópia do material genético. A descoberta do primeiro mecanismo químico pelo qual a molécula de DNA é construída na célula coube ao bioquímico norte-americano Arthur Kornberg que, em 1955, identificou na bactéria *Escherichia coli* uma das enzimas necessárias para que ocorresse a síntese do DNA (Kornberg, 1959).

Contudo, se o pareamento específico das bases mostrava como as instruções genéticas poderiam ser passadas de uma célula para outra, ainda

⁹ Polimerização é o processo no qual uma molécula polimérica é sintetizada. Uma molécula polimérica é formada por uma cadeia de monômeros, blocos de construção, quimicamente equivalentes. Assim, os blocos de construção do DNA são os nucleotídeos; enquanto os blocos de construção das proteínas são os aminoácidos.

não estava claro de que maneira, partindo da estrutura do DNA, poderia ser possível a codificação de uma grande variabilidade e quantidade de produtos diferentes e específicos. Nesse sentido, foi crucial o processo de quebra do código genético, a decodificação do sistema de passagem da informação genética para a molécula de RNA e desta para a síntese de proteínas.

1961: O código genético

Em 1943 no *Trinity College* em Dublin, Irlanda, o físico austríaco Erwin Schrödinger, um dos fundadores da mecânica quântica, ministra uma série de três palestras sob o título “*What is Life?*”. No ano seguinte, o material é reunido e publicado em um livro com o mesmo título. O volume abordava dois temas principais: a hereditariedade e a termodinâmica. No que toca a questão da hereditariedade, uma das idéias investigadas por Schrödinger em seu trabalho era como os organismos transmitem sua informação hereditária através das gerações, como é mantida a ordem a partir da ordem. Para Schrödinger, o gene armazenaria suas informações por meio de uma espécie de código, contido em sua estrutura. O livro de Schrödinger influenciou grandemente o meio científico da época ao estimular novas pesquisas no campo da hereditariedade, e, principalmente, “convocando” os físicos a procurar as bases físicas da vida, a saber, a informação genética (Murphy & O’Neill, 1997; Stent, 1968).

Desde a década de 40 que já estava confirmada a tese de que a síntese de proteínas ocorria fora do núcleo das células, em estruturas citoplasmáticas chamadas de ribossomos. Os ribossomos são constituídos de moléculas de RNA e de proteínas. Na medida em que o DNA dos organismos eucariontes está no núcleo das células, e que os ribossomos – onde

se dá a síntese protéica – se localizam fora do núcleo, a idéia de que a síntese protéica ocorreria diretamente a partir do DNA perdia consistência. O biólogo francês François Jacob postulou a existência de uma espécie de intermediário, ou mediador, entre o DNA e o local da síntese protéica. Esse intermediário seria o responsável em levar a informação genética do núcleo para o citoplasma, local onde ocorre a síntese de proteínas (Jacob, 1965). No início da década de 60, o biólogo molecular sul-africano Sydney Brenner foi um dos pesquisadores que demonstraram experimentalmente a existência de uma molécula de ácido nucléico que ficou conhecido pelo nome de RNA mensageiro, ou RNAm (Davies, 2001).

Sydney Brenner and I had decided to spend the month of June 1960 with M. Meselson, hunting for the messenger in the laboratory of Max Delbrück at the California Institute of Technology. The best candidate for the role of messenger seemed to us to be the RNA detected by Hershey and later by Volkin and Astrachan in bacteria infected with T2 phage. Thanks to the extraordinary intellectual and experimental agility of Sydney Brenner, we were able to show, within a few weeks, that the RNA formed by the phage associates with ribosomes synthesized wholly before infection, to produce on them phage proteins. The same ribosomes can thus make either phage or bacterial proteins, depending on the messenger with which they associate. Accordingly, it is the messenger which brings to the ribosomes a specific program for synthesis¹⁰. At this time, another member of our group at the Pasteur Institute, François Gros, had gone to spend several months at Harvard in the laboratory of J.D. Watson. With their collaborators, they rapidly succeeded in demonstrating the existence of a messenger fraction in the RNA of growing bacteria, and in establishing its principal properties. (Jacob, 1965, p. 153).

Em 1961, o zoólogo e bioquímico norte-americano Marshall Nirenberg do *National Institutes of Health* (NIH) inicia os trabalhos de decifração do código genético. Para que seu trabalho pudesse chegar a bom termo, esse RNA intermediário foi um elemento chave de suas pesquisas (Nirenberg, 1968). De acordo com Severo Ochoa:

There are indications that most, if not all, of the cytoplasmic RNA is synthesized in the nucleus and subsequently transported to the cytoplasm. In

¹⁰ Na medida em que os ribossomos são formados por uma associação de proteínas e RNA, o RNA presente nos ribossomos também era candidato a servir de molde para a especificação das estruturas protéicas.

transmitting genetic information, nuclear DNA is supposed to determine the nature of the nuclear RNA which, on entering the cytoplasm, determines in turn the nature of the proteins synthesized. (Ochoa, 1959, p.645).

Tomando-se a citação anterior, pode-se afirmar que o conceito de uma molécula de RNA como um molde para a síntese protéica já estava estabelecido. Em 1955, o biólogo espanhol Severo Ochoa (1959) da *New York University*, construiu os primeiros RNAm sintéticos. A síntese de filamentos de RNA se tornou possível quando Ochoa conseguiu isolar uma enzima bacteriana capaz de catalisar a síntese de ribonucleotídeos, a polinucleotídeo-fosforilase. (Ochoa, 1965).

Esse foi o primeiro passo para a decodificação do código genético. A idéia subjacente era simples¹¹. Primeiro, construir RNAs mensageiros sintéticos de seqüências específicas. Depois, fornecer esses RNAs à maquinaria de síntese protéica na célula (os ribossomos) e verificar quais aminoácidos eram formados a partir de uma determinada seqüência. Para isso, Marshall Nirenberg e Heinrich Mathaei arquitetaram um sistema celular artificial (*in vitro*) que sintetizasse proteínas. O sistema era, basicamente, constituído da maquinaria de síntese protéica da bactéria *Escherichia coli*, ou seja, um extrato de seus componentes celulares (Nirenberg, 1965).

O RNAm especifica as proteínas a serem formadas através de sua seqüência de nucleotídeos. Vale mencionar que no final da década de 50, pouco antes da decifração do código propriamente dito, Francis Crick e Sidney Brenner demonstraram que uma combinação de três nucleotídeos, chamadas de códon, correspondente a um aminoácido dos cerca de vinte aminoácidos existentes que compõem uma cadeia polipeptídica. O primeiro RNAm sintético construído por Nirenberg e Mathaei consistia numa seqüência composta apenas pelo nucleotídeo uracil. Esse RNA mensageiro foi chamado de poli-U. Dessa forma, uma longa molécula de RNA constituída apenas por uma base (UUUUUU) deveria codificar apenas um aminoácido. O resultado obtido pelo fornecimento do RNA mensageiro sintético poli-U à maquinaria de síntese protéica foi a formação de um polipeptídeo composto apenas pela repetição do aminoácido fenilalanina. O código genético começava a ser quebrado. Nesse código, o códon UUU seria responsável pela codificação de um único aminoácido, a fenilalanina.

Porém, ainda restava um longo trabalho: construir RNAs mensageiros com diferentes composições que permitissem descobrir o código inscrito nas seqüências do material

¹¹ A idéia do experimento era simples, mas não a execução.

genético. Uma expressiva parte do trabalho restante de codificação e ajuste do código genético coube ao biólogo molecular indiano da *University of Wisconsin*, Har Gobind Khorana (1968). Esse pesquisador planejou métodos de síntese de ácidos nucléicos bem definidos – derradeiro pré-requisito para a solução do código genético. Entre 1961 e 1966, os trabalhos de Nirenberg, Mathaei, Khorana, e suas respectivas equipes alcançaram seu objetivo: decifram o código genético, identificando os aminoácidos especificados por cada trinca de nucleotídeos no RNA mensageiro (os códon).

Pela interpretação do código genético e a revelação de o funcionamento da síntese protéica, Marshall W. Nirenberg e Har Gobind Khorana recebem o prêmio Nobel de 1968 de medicina ou fisiologia. No discurso que antecedeu a entrega do prêmio, o professor P. Reichard, membro do comitê Nobel, compara a síntese protéica com uma máquina de tradução. Segundo Reichard, essa máquina é alimentada pela seqüência escrita no alfabeto dos ácidos nucléicos, que, uma vez de posse dessa seqüência, alcança os meio de indicar os aminoácidos corretos para a formação da seqüência protéica. A seqüência de nucleotídeos no DNA seria colinear às seqüências de aminoácidos nas proteínas. De posse das seqüências do gene no DNA, poderíamos prever as seqüências de aminoácidos na proteína correspondente. Não é difícil imaginar porque esta idéia iria capturar a imaginação da comunidade científica. Com a descoberta dos íntrons, contudo, esse quadro logo começaria a mudar.

Com a elucidação do código genético, uma espécie de linguagem da vida, ampliavam-se as perspectivas de que a biologia molecular pudesse explicar mais detalhadamente os mecanismos da herança e avançar na compreensão das causas de muitas das doenças hereditárias. Além disso, a decifração do código genético parecia garantir para os biólogos moleculares a disponibilização de uma poderosa ferramenta cibernética. As informações genéticas podem, a partir desse importante marco das ciências, ser comparadas à uma espécie de linguagem molecular. Os nucleotídeos, blocos de construção do DNA, desempenhariam, dentro dessa nova perspectiva, o papel de letras desse alfabeto. Com isso, ganhava força, naquele momento, a convicção de que as células guardam as instruções para a fabricação das proteínas de nosso corpo e, por extensão, das nossas características. Em outro desdobramento da quebra do código genético, a segunda linguagem presente nas células, referente às proteínas, teria nos aminoácidos as suas letras (Davies, 2000; Reichard, 1968). São multidões de proteínas diversas, presentes aos milhares nas células, as responsáveis por executar as reações químicas necessárias para o desenvolvimento dos organismos e de suas respectivas arquiteturas (Jacob, 1998; Mayr,

1998; Jacob, 1983; Jacob, 1977). E, por último, fica estabelecido que são os ácidos nucleicos que dirigem a síntese dessas proteínas. A estrutura química do DNA determina a estrutura química das proteínas por intermédio do código genético, uma espécie de dicionário que traduz as informações de um alfabeto para o outro (Reichard, 1968). Com tudo isso, o esquema da síntese protéica parecia concluído. Contudo, ainda mesmo antes do código ser quebrado, tais idéias já se resumiam na influente expressão cunhada por Francis Crick em 1957, o “dogma central da biologia molecular” (Olby, 1975).

O dogma central da biologia molecular

Quando em 1940 George Wells Beadle e Edward Lawrie Tatum investigam a suposição de que os genes atuavam regulando os processos bioquímicos, não se suspeitava de que o DNA pudesse ser o dirigente da especificidade biológica. Com a descoberta da estrutura do DNA em 1953, a hipótese “um gene – uma enzima” adquire um novo sentido, pois passa a relacionar as seqüências de bases do gene às seqüências de aminoácidos em uma proteína (Keller, 2000; Falk, 1986). Com a descoberta da estrutura físico-química da molécula de DNA, o gene adquire materialidade. Assim, agora é possível estabelecer como a estrutura do gene (as seqüências de bases da molécula de DNA) responde por sua função (a especificidade pela produção de uma proteína). A partir de 1953, a estrutura e a função do gene adquire um mesmo significado. A existência de uma espécie de intermediário, ou mediador, entre o DNA e o local da síntese protéica, também foi satisfatoriamente confirmada. Em 1957, de posse desses novos dados, Francis Crick os reúne sob a forma do “dogma central da biologia molecular”, que resume numa simples figura como se dá o fluxo de informações genéticas (Figura 2).

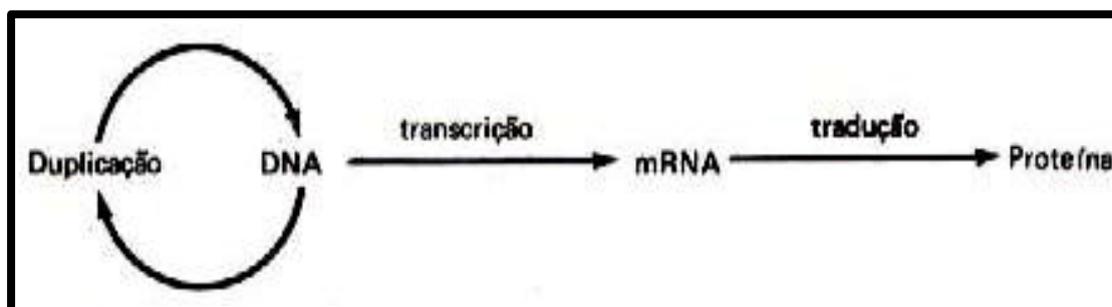


Figura 2: No esquema as setas indicam a direção da transferência da informação genética. A seta envolvendo o DNA significa que ele é o próprio molde de sua duplicação (auto-replicação do DNA). A seta entre o DNA e o RNAm (RNA mensageiro) indica que a molécula de RNA se origina a partir do molde da molécula de DNA. E a última seta indica que a seqüência protéica é derivada da molécula molde de RNA

O “dogma” afirma que, uma vez passada a informação para a proteína, ela não pode retornar. Ou seja, o fluxo de informação genética é unidirecional. Como indicado na figura, a transferência da informação segue de ácido nucléico para ácido nucléico, ou de ácido nucléico para proteína. No entanto, a transferência de informação de proteína para proteína ou da proteína para o ácido nucléico é impossível. Segundo esse modelo, as proteínas nunca servirão de molde para as seqüências de RNA, e o RNA por sua vez, jamais será o molde de uma seqüência de DNA. (Crick *apud* Olby, 1975).

Na presente dissertação, o “dogma central da biologia molecular” serve como instrumento para a abordagem de questões relevantes, para o teste de nossa hipótese. Pode, nesse sentido, ser utilizado como uma espécie de modelo básico de como os genes atuam a partir da estrutura do DNA, e de como essas relações foram paulatinamente modificadas e dotadas de maior complexidade. Para sermos justos, desde a sua elaboração, Francis Crick nunca negou que o “dogma” era um modelo aproximado de um quadro bem mais complexo (Olby, 1975).

A transmissão de características hereditárias ao longo de gerações, de forma praticamente imutável, confere ao gene uma estabilidade resultante de sua própria capacidade de autogeração. Além de ser capaz de fazer cópias de si mesmo, através da auto-replicação, processo sem o qual a informação seria perdida no decorrer da passagem de gerações, o gene também é responsável pela expressão das características que transmite. De acordo com Stent (1968), o “dogma” afirma que o DNA possui tanto funções autocatalíticas quanto heterocatalíticas. Autocatalítica ao servir de molde para a síntese de réplicas de cadeias de nucleotídeos, mediante a formação de ligações de complementaridade entre cadeias de DNA (replicação) e RNA (transcrição). Heterocatalítica quando as réplicas de cadeias de RNA são traduzidas em polipeptídios mediante o código genético (tradução). Desse modo, os processos de transferência de informação podem ocorrer em dois níveis. No primeiro, a informação genética presente no DNA ou é copiada para produzir mais DNA, ou é traduzida em proteínas. No segundo, as proteínas respondem pela estrutura e coordenação do metabolismo dos seres vivos.

O RNA, depois de transcrito, atravessa o núcleo. No citoplasma da célula ocorre o processo chamado de tradução. A tradução, como indicado anteriormente, é uma espécie de leitura química da informação genética. As bases nitrogenadas, como se fossem letras, são “lidas” nos ribossomos. A longa cadeia de aminoácidos, chamada de cadeia polipeptídica, constitui as proteínas, que são as responsáveis pelo funcionamento e estrutura dos seres vivos. Dessa forma, o RNA é uma espécie de mediador entre a informação contida no núcleo e a leitura dessas informações no citoplasma da célula.

Vale observar que no modelo de funcionamento gênico proposto por Crick existe, implicitamente, uma definição de gene. Essa definição se baseia numa visão molecular de gene em que este é materializado na forma de seqüências de DNA. Nesse sentido, o

gene é entendido como um pedaço de DNA, uma seqüência definida, discreta, de DNA responsável pela informação de uma proteína. Essa conceituação de gene molecular clássica está claramente vinculada à definição da estrutura do DNA. É interessante notar que, dentro desse conceito molecular, já está especificada uma “função” para o gene, ou seja, o gene não pode ser considerado como uma simples “região do DNA”, mas como uma região do DNA que contém informação.

Para o filósofo e historiador da ciência Robert Olby (1975) o “dogma central da biologia molecular” pode ser entendido como a matriz conceitual da biologia contemporânea, pois resume o modo de ação gênica. Segundo Olby, a identificação do DNA como material hereditário implicou numa notável reestruturação do pensamento biológico. Até o final da década de 1950, as proteínas ocuparam um papel central na compreensão da especificidade biológica. A noção de que as proteínas não poderiam servir como moldes para a determinação das seqüências protéicas e dos ácidos nucléicos representa, para Olby, pensar o DNA como uma transformação conceitual no sentido kuhniiano. Para Olby, o “dogma” de Crick poderia ser legitimamente considerado como o paradigma da biologia contemporânea (Olby, 1975).

Tais séries de descobertas contribuíram para a disseminação de uma espécie de sentimento que fazia crer aos pesquisadores que os conhecimentos sobre a genética estavam praticamente completos (Jacob, 1998), traduzindo-se numa precoce simplicidade da genética e da biologia molecular. Como já dito no começo deste capítulo, para Gunther Stent (1968), isto trouxe uma espécie de euforia antecipada de que o que restava era apenas polir os detalhes. Como afirmou Stent:

The first attempts to write the history of scientific discipline often presage its imminent senescence. And so the appearance a year ago of the collection of autobiographical essays entitled, *Phage and the origins of molecular biology* is probably symptomatic of the decline of molecular biology, only yesterday an avant-garde but today definitely a workaday field. (Stent, 1968, p. 160).

Talvez, por esse sentimento de que o trabalho restante da biologia molecular era somente o de investigar os detalhes do funcionamento genético, o “gene molecular clássico” resistiu por tanto tempo inalterado. Contudo, já por volta dos anos das décadas de 1960 e 1970 já começavam a se acumular informações empíricas que levantariam novos obstáculos à compreensão excessivamente simplificada dos mecanismos de expressão genética (Keller, 2000; Falk, 1986). Obviamente que essas novas descobertas não passaram despercebidas, e, hoje, talvez não sejam mais vistas como meros detalhes de um funcionamento que se mostra claramente muito mais complexo.

CAPÍTULO II

OS GENES INTERROMPIDOS E O DEBATE

ACERCA DO CONCEITO DE GENE

MOLECULAR CLÁSSICO

Apresentando o problema: Os genes interrompidos

*Como todos os homens da
biblioteca, viajei na minha
juventude; peregrinei em
busca de um livro, talvez o
catálogo de catálogos;
agora que meus olhos
quase não podem decifrar
o que escrevo, preparo-me
para morrer, a poucas
léguas do hexágono que
nasci.*

*Borges, Jorge Luis, A
biblioteca de Babel,
Ficções.*

No primeiro capítulo da presente dissertação abordamos alguns episódios do período que Gunther Stent (1968) denominou de fase dogmática da biologia molecular, cujo emblema é o “dogma central da biologia molecular” de Francis Crick. No dogma, Crick afirma que a passagem das informações genéticas segue do DNA para o RNA e daí para as proteínas. Representadas no dogma estão as duas funções primordiais da molécula de DNA, a base material da hereditariedade: 1) a auto-reprodução das informações genéticas e; 2) a indução da síntese de proteínas que regem as reações químicas e que constituem a arquitetura dos organismos (Olby, 1975). O conceito de “gene molecular clássico” que emerge dessa fase dogmática procura definir e conciliar uma unidade estrutural de DNA (uma seqüência específica de DNA) – com uma unidade funcional no organismo (uma proteína particular responsável por algum aspecto do funcionamento do organismo) (Griffiths, 2002).

Devido às dificuldades técnicas para analisar a organização mais complexa do genoma dos seres eucarióticos (animais, plantas e fungos) – notadamente as ordens taxonômicas mais altas desses organismos –, muito do conhecimento genético estabelecido

em relação aos processos de regulação e estrutura do gene foi obtido em sistemas bacterianos, sobretudo na bactéria *Escherichia coli*, e estendidos para os demais organismos (Osawa *et al.*, 1992; Berg, 1980). A partir da possibilidade de investigar mais diretamente a organização do material genético dos organismos eucariontes, algumas suposições anteriores tiveram que ser revistas.

Convém enfatizar que a *Escherichia coli* é um microorganismo importante não só como modelo genético, mas também como causa etiológica de algumas importantes doenças humanas e animais. Por exemplo, a colibacilose viária é uma das principais doenças da avicultura industrial moderna e causa anualmente grandes prejuízos econômicos. Nos seres humanos, é a causa de colites hemorrágicas (inflamação da mucosa intestinal) e gastroenterites (diarréias), entre outras afecções, sendo uma das maiores causas de mortalidade infantil por diarréia bacteriana, em países de terceiro mundo.

Uma das primeiras suposições que foram revistas se refere à descoberta, em 1977, de que os genes se dispõem linearmente sem interrupções em uma seqüência de bases do DNA. A estrutura do gene nos organismos eucarióticos é interrompida por trechos de DNA denominados de íntrons. Os íntrons são removidos do RNA mensageiro recém transcrito, e as seqüências retidas – os exons – dão origem a um só pedaço de RNA maduro a ser traduzido em proteína.

No presente capítulo desta dissertação, examinaremos o contexto da descoberta dos genes interrompidos, e a implicação dessa descoberta na síntese dos RNAs mensageiros. A inesperada maneira de como se procede a síntese dos RNAs mensageiros nos organismos eucariontes, tem conseqüências para a definição de “gene molecular clássico”.

A descoberta dos genes interrompidos apresenta relação direta com o aparecimento de novas técnicas de investigação no cenário da biologia molecular, notadamente o advento da tecnologia do DNA recombinante (Chambon, 1981). Os historiadores da

ciência sempre souberam destacar a importância do desenvolvimento de novas tecnologias (Mayr, 1998). Nesse sentido, é exemplar a relação entre a biologia e as técnicas físico-químicas, que começam a se tornar notórias por volta do final do século XIX, muito antes do advento da tecnologia do DNA recombinante e da descoberta dos genes interrompidos.

Tecnologias no estudo do vivo: O nascimento da biologia molecular

No final do século XIX, começavam a minguar certas teorias associadas ao pensamento biológico, como, por exemplo, o vitalismo¹², para dar lugar ao desenvolvimento de uma ciência mais experimental, como a bioquímica e a genética (Jacob, 1983). A união entre os métodos de análises físico-químicos e as questões biológicas foi o esboço do que viria a ser a chamada biologia molecular (Mayr, 1998). Assim, a biologia molecular ocupa uma região fronteira entre a biologia, a química e a física (Mayr, 1998). Francis Crick, um dos descobridores da estrutura da molécula de DNA, afirmou:

I myself was forced to call myself a molecular biologist because because when inquiring clergymen asked me what I did, I got tired of explaining that I was a mixture of crystallographer, biophysicist, biochemist and geneticist, an explanation which in any case they found too hard to grasp. (Crick *apud* Stent, 1968, p. 390).

Parafrazeando François Jacob (1983, p. 253), a bioquímica e a física, a genética e a fisiologia fundem-se numa só prática. Isto quer dizer que a biologia molecular não pode ser realizada por indivíduos isolados, cada um preocupado com seu problema e seu organismo. Ela exige um esforço conjugado dos homens e das técnicas. Em um mesmo instituto de

¹² O vitalismo foi uma doutrina formulada entre meados do século XVIII e meados do século XIX. Defendia que fenômenos referentes aos seres vivos seriam controlados por uma força exterior alheia ao domínio das leis físico-químicas. Desse modo, o vitalismo se opunha às idéias mecanicistas e cartesianas (Mayr, 1988; Jacob, 1983).

pesquisa, em um mesmo laboratório passam a cooperar especialistas separados por sua formação de origem, mas unidos por um mesmo tema de análise e um mesmo material.

Em 1912, o desenvolvimento da técnica de análise de moléculas por raios-X por dois físicos ingleses, William e Lawrence Bragg, pai e filho, é considerado como o marco inicial da biologia molecular propriamente dita (Friedman & Friedland, 2000; Mayr, 1998; Stent, 1968). O advento da técnica de difração por raios-X passou a permitir a análise das estruturas das moléculas que compõem os organismos, ao fornecer a posição exata de cada átomo. Contudo, devido à especificidade da técnica, esse período permaneceu reservado aos físicos estruturais. Pode-se dizer que a biologia molecular não nasceu exatamente do interesse de biólogos, mas, principalmente, de uma legião de físicos que acreditavam que podiam dar contribuições de suas áreas de pesquisa para a resolução de algumas questões biológicas. O acesso à intimidade das estruturas biológicas só se tornou possível porque os físicos passaram a se interessar pela complexidade das moléculas de importância biológica, na medida em que eles acreditavam que somente a configuração tridimensional destas moléculas poderia desvendar as funções fisiológicas da célula (Stent, 1968).

O interesse nas estruturas das moléculas que compõem os organismos deu origem ao que se convencionou chamar de “escola estruturalista” da biologia molecular. Esta escola dava ênfase às estruturas das proteínas – moléculas de predileção dos bioquímicos, devido a sua já citada importância nos mais diversos fenômenos biológicos. De acordo com Stent (1968), a convicção dessa escola era que a física poderia dar contribuições significativas para a biologia.

A escola estruturalista se valeu, principalmente, dos extratos celulares como fonte dos primeiros componentes passíveis de análise. A física e a química forneceram as ferramentas que permitiram purificar, determinar a composição e organização das macromoléculas, com as técnicas da cromatografia (que permite purificar e isolar

moléculas e substâncias químicas); da eletroforese (que permite averiguar o tamanho e as propriedades elétricas das moléculas); da centrifugação (separação por densidade dos extratos celulares); do uso de radioisótopos (elementos que emitem radiações distinguíveis e que possibilita o acompanhamento da sua trajetória na célula e/ou no organismo); de técnicas de microscopia (novos métodos de fixação de vários tipos de materiais biológicos e corantes com afinidade altamente específica em relação a certos componentes celulares ou moleculares); de microscópios cada mais eficientes, como os eletrônicos – que permitem observar organelas e certas macromoléculas. (Mayr, 1998; Reznik, 1995). Essas técnicas também foram utilizadas na escola funcionalista, que mencionaremos a seguir. Para Stent (1968), a influência da escola estruturalista na biologia não foi em nada revolucionária.

Todavia, a outra escola, denominada de “informacional” ou “funcionalista”, foi motivada por uma noção inovadora: de que a própria biologia forneceria notáveis contribuições à física, e, até mesmo, a descoberta de outras leis da física (Lwoff, 1965; Jacob, 1965). Essa escola tinha seu foco, sobretudo, na informação genética, a saber, as bases físicas do armazenamento da informação genética. Considera-se que tal escola teve início em 1938, a partir do físico alemão Max Delbrück, do *California Institute of Technology* (Caltech). Delbrück estudava a genética de bacteriófagos (fagos), vírus que infectam bactérias – ou de forma literal, “comedores de bactérias”. Para as décadas de 1920, 1930 e até mesmo 1940, o termo “comedores de bactérias” parece ser particularmente apropriado. De acordo com Lwoff (1965), não se sabia ao certo o quê eram os bacteriófagos. Pequenos micróbios, pequenas bactérias, proteínas, ou vírus, adicionando-se o fato de que havia uma certa confusão em relação ao que era um vírus. Foi Hershey, em 1952, que demonstrou serem os vírus partículas organizadoras de sua própria reprodução a partir de seu próprio material genético, mas usando a maquinaria celular da

célula hospedeira (Lwoff, 1965). Stent (1968) afirma que nesse período se chegou ao acordo de que fagos são realmente vírus que se multiplicam autonomamente dentro das células bacterianas hospedeiras, assunto que estava sob debate por dez anos ou mais.

Assim, Delbrück, Alfred Hershey (químico e bacteriologista norte-americano, do *Carnegie Institution of Washington*) e Salvatore Luria¹³ (microbiologista italiano naturalizado norte-americano, do MIT, *Massachusetts Institute of Technology*) foram os fundadores do que ficou conhecido como o “grupo americano dos fagos” (Stent, 1968). Esse grupo acreditava que os fagos eram os objetos ideais para o estudo da auto-replicação biológica e das bases físicas da hereditariedade, do gene. Dessa forma, pode-se dizer que muitos dos pesquisadores envolvidos no estabelecimento da noção de “gene molecular clássico” são oriundos do “grupo dos fagos” (Griffiths, 2002).

Ao mencionar os fagos, vale lembrar que muitas vezes o sucesso da pesquisa reside na habilidade – e sorte – do pesquisador em trabalhar com um organismo experimental adequado. Assim foi, por exemplo, com as ervilhas de Mendel (que não produziram os mesmos resultados com plantas que possuíam padrões de herança genética mais complexa); com o fungo *Neurospora* de Beadle e Tatum; e com as drosófilas de Thomas Hunt Morgan¹⁴.

Esse período é caracterizado por Stent (1968) como “fase romântica” da escola informacional da biologia molecular. Isto porque, embora tenha contribuído em desenvolver o sistema “fago hospedeiro” como um modelo de organismo experimental para a genética¹⁵, essa fase romântica de fato não resolveu o problema principal a que se propôs (a descoberta da estrutura responsável pela hereditariedade). A principal “falha”

¹³ Luria foi orientador de James Watson, um dos descobridores da estrutura físico química do DNA.

¹⁴ Entre outros autores, um relato bastante interessante sobre organismos experimentais pode ser encontrado no livro *O rato, a mosca e o homem* de François Jacob (1998).

¹⁵ Por suas descobertas em relação ao mecanismo de replicação e estrutura genética dos vírus, Delbrück, Hershey e Chase foram laureados com o Nobel em medicina ou fisiologia de 1969 (www.nobel.se).

desses pesquisadores residiu no fato de terem acreditado que as proteínas eram o material genético viral. Quando Hershey e Marta Chase (1952) concluem que é o genoma do vírus (fago) é o seu DNA, encerra-se essa fase para dar lugar à “fase dogmática” da biologia molecular¹⁶, em que estariam envolvidos dois físicos experientes em difração por Raios-X: Francis Crick e Maurice Wilkins (Stent, 1968).

Os Genes Interrompidos e o DNA recombinante

Durante as décadas de 1960 e 1970, devido à sua relativa simplicidade, facilidade de manipulação e alta capacidade reprodutiva, entre outros atributos, as bactérias estavam entre os principais organismos de pesquisa disponíveis para biólogos moleculares e geneticistas – notadamente, a bactéria *Escherichia coli* (Chambon, 1981). Muito do que hoje se sabe sobre como a informação genética é traduzida em proteínas foi aprendido com os estudos em bactérias:

Until a few years ago, much of what was known about the molecular details of gene structure, organization and function had been learned in studies with prokaryote microorganisms and the viruses that inhabit them, particularly, the bacterium *Escherichia coli* and the T and lambdoid bacteriophages. These organisms were the favorites of molecular biologists because they can be propagated readily and rapidly under controllable laboratory conditions. (Berg, 1980, pp. 386).

Antes de 1977, as moléculas de RNA que transportam a informação genética para os ribossomos (estruturas celulares onde as células sintetizam as proteínas) eram consideradas cópias fiéis do DNA, na qual cada molécula de RNA alinhava-se exatamente com a fita de DNA codificante. A seqüência de bases nitrogenadas do gene apresentava uma correspondência direta em relação à seqüência de aminoácidos a qual especificava. Esse

¹⁶ Que já apresentamos no primeiro capítulo.

princípio de colinearidade em bactérias foi assumido, tacitamente, para os demais organismos (Chambon, 1981).

De acordo com a biologia evolutiva e a bioquímica comparativa, certos processos biológicos são conservados durante o curso evolutivo e se tornam praticamente universais para todos os organismos. Isto permite inferir que estes processos sejam quase os mesmos tanto em organismos mais simples como nos mais complexos. Pode-se dizer que cada indivíduo traz consigo a experiência – evolutiva – de seus antepassados. Embora a diversidade de espécies do mundo vivo seja evidente, subjacentes a esta diversidade existem notáveis semelhanças entre as espécies. Tanto é assim, que é possível construir organismos tão heterogêneos fisicamente como ornitorrincos e flamingos, como moscas, polvos e macacos a partir de um conjunto não muito dessemelhante de proteínas, lipídeos, carboidratos (as classes de moléculas encontradas nos organismos vivos representam uma variedade muito restrita dos objetos químicos, teoricamente, possíveis); se valendo de um código genético, praticamente, universal; e uma maquinaria de tradução quase idêntica. O repertório químico e a função das estruturas biológicas são, basicamente, as mesmas, não importando tratar-se de células bacterianas, células de leveduras ou das diferentes células humanas. Nesse sentido, os processos fundamentais da vida, como a maior parte das reações centrais do metabolismo, permanecem fundamentalmente os mesmos das bactérias ao homem (Jacob, 1983).

Assumindo isso, o biólogo francês Jacques Monod – Nobel em 1965 pela pesquisa sobre o funcionamento genético em bactérias, conjuntamente com François Jacob e André Lwoff – resume esse pensamento com a declaração de que o que é verdade para a bactéria *Escherichia coli* também é verdadeiro para os elefantes. Admitir isso como verdade evidenciam as vantagens de estudar processos biológicos em organismos experimentais menos complexos.

Obviamente, como em todo modelo experimental, existem limites impostos pelas propriedades de cada organismo, suas características fisiológicas, estruturais e genéticas particulares. O surgimento de organismos experimentais que pertencem a grupos taxonômicos de organismos particulares tem proporcionado o estudo de processos biológicos específicos, como, por exemplo: a levedura *Saccharomyces cerevisiae* (um fungo unicelular utilizado como fermento biológico na fabricação de pães e na fermentação de bebidas alcoólicas, transformando quimicamente o açúcar em álcool); o verme *Caenorhabditis elegans*, utilizado para o estudo de processos universais da biologia do desenvolvimento em eucariontes; a erva *Arabidopsis thaliana*, utilizada como o modelo de estudo de plantas superiores.

Contudo, apesar da importância das pesquisas microbiológicas em contribuir para desvendar os mecanismos de controle e regulação da atividade da celular ainda nos dias de hoje (Shapiro, 1999; Strauss, 1995), durante o ano de 1977, uma série de trabalhos independentes demonstra que a organização dos genes em organismos eucariontes e em vírus é fundamentalmente diferente daquela encontrada em bactérias. As seqüências de DNA que especificam a produção de um determinado polipeptídio (cadeia de aminoácidos) não são totalmente colineares a esses polipeptídeos. De fato, os genes não são contínuos, mas uma espécie de mosaico em que uma matriz de leitura é intercalada por seqüências silenciosas, não codificantes (Sambroock, 1977; Gilbert, 1978) (Figura 3).

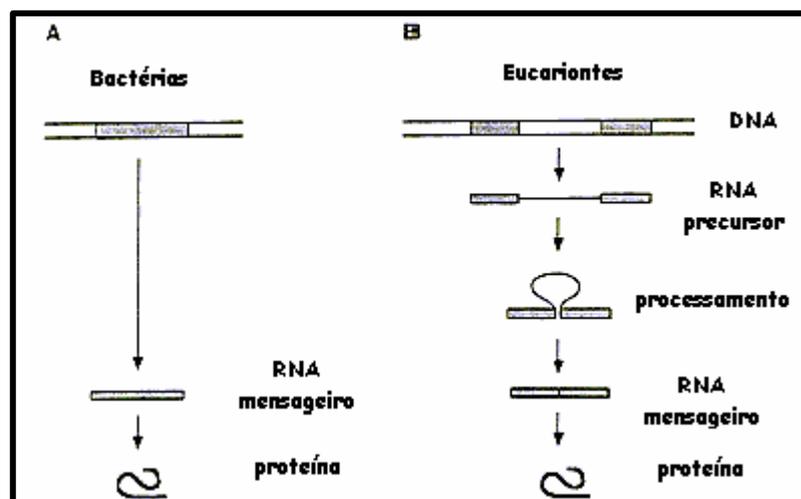


Figura 3: Estrutura do gene e fluxo da informação genética em bactérias (A) e em organismos superiores (B). Em bactérias a informação genética é armazenada em segmentos contínuos de DNA e o RNA mensageiro tem correspondência direta com a proteína final. Muitos genes de organismos eucariontes e virais apresentam-se interrompidos, dividido. A seqüência primária do RNA mensageiro precisa passar por um processo de corte e reunião antes de sua tradução em proteínas (adaptação a partir da figura encontrada no sítio eletrônico do *Nobel e-Museum*, www.nobel.se/medicine/laureates/1993/press.html).

A descoberta dos genes interrompidos está diretamente relacionada com o desenvolvimento das novas técnicas de investigação molecular do início da década de 1970, como a descoberta das enzimas de restrição e a tecnologia do DNA recombinante (Chambon, 1981). Essas são duas das ferramentas mais representativas de uma nova era da biologia molecular, a pesquisa genômica¹⁷. No final da década de 1960, as primeiras enzimas de restrição começam a ser descritas em *Escherichia coli* (Meselson & Yuan, 1968) e isoladas e caracterizadas na bactéria *Hemophilus influenzae* (Smith & Wilcox, 1970). O nome dessas enzimas advém do “fenômeno de restrição ao hospedeiro”. Esse fenômeno foi observado em algumas linhagens bacterianas que apresentavam imunidade ao ataque de vírus invasores. Averiguou-se que essas espécies bacterianas possuíam enzimas (de restrição) que destruíam o DNA invasor, ao reconhecendo e cortando seqüências específicas de DNA. Dessa forma, pode-se dizer que essas enzimas atuam como uma espécie de tesoura molecular. O próprio DNA bacteriano se protege do ataque dessas enzimas através de modificações químicas nos sítios de clivagem (de corte) da ação enzimática.

A partir do advento da tecnologia do DNA recombinante torna-se possível cortar genes de uma espécie e inseri-los (recombiná-los) em vetores. Os vetores podem ser

¹⁷ O termo *genomics* (genômica) foi cunhado em 1986 por Thomas Roderick para descrever a disciplina científica de mapeamento, seqüenciamento e análise de genomas. (Hieter & Boguski, 1997).

pequenas moléculas de DNA circular existentes em bactérias, chamados de plasmídios. Após a recombinação entre os vetores e os fragmentos de DNA de interesse, os vetores são introduzidos e propagados em organismos hospedeiros, como bactérias e leveduras. Dentro dos organismos hospedeiros, um fragmento específico da molécula de DNA pode ser clonado indefinidamente. O objetivo dessa tecnologia é obter em grande quantidade uma seqüência de DNA desejada, para posteriores estudos moleculares, com fins mais precisos e específicos.

As primeiras moléculas de DNA recombinante foram criadas em 1972 (Jackson *et al.*, 1972) e em 1973 (Cohen *et al.*, 1973). O trabalho de 1972 rendeu ao bioquímico Paul Berg, da *Stanford University School of Medicine*, Califórnia, o Nobel em química de 1980. Em 1974, Paul Berg em conjunto com mais dez pesquisadores redige uma carta onde solicita que o NIH (*National Institutes of Health*) regule o uso da tecnologia do DNA recombinante. E até que se tenha um melhor entendimento quanto à segurança dessa nova técnica, deveria ser interrompidos alguns experimentos envolvendo essa nova tecnologia (Berg *et al.*, 1974). Em fevereiro de 1975, decorrente da proposta de moratória nas pesquisas que envolvessem a manipulação genética a partir da tecnologia do DNA recombinante, realizou-se o que ficou conhecido como “Conferência de Asilomar”. Essa conferência reuniu cerca de 140 cientistas de diversas nacionalidades no Centro Convenções de Asilomar, em *Pacific Grove*, Califórnia. Nessa reunião científica discutiu-se a segurança da manipulação do DNA de diferentes espécies de seres vivos, sugerindo que se distinguisse e classificasse os experimentos com a tecnologia do DNA recombinante de acordo com o nível de risco imaginado. A conferência de Asilomar é um marco na história da preocupação ética aplicada à pesquisa. Foi a primeira vez que se discutiu aspectos de proteção aos próprios pesquisadores e demais profissionais envolvidos nas áreas em que se realiza o projeto de pesquisa, sem nenhum tipo de pressão pública –

provavelmente pelo fato de que nessa época as pesquisas genéticas não eram um assunto de domínio público, como o são nos dias de hoje. Dois pioneiros na pesquisa com DNA recombinante, Stanley Cohen da *Stanford* e Herbert Boyer da *U.C. San Francisco* pediram, e lhes foi concedida, a patente sobre a tecnologia do DNA recombinante. Isto garantiu para a *Stanford* e para a *U.C.S.F.* royalties de empresas de biotecnologia e companhias farmacêuticas. Estima-se que a patente Cohen-Boyer já tenha gerado cerca de 200 milhões de dólares para essas universidades (Zweiger, 2001).

Com a tecnologia do DNA recombinante é possível, por exemplo, estudar os mecanismos de expressão dos genes¹⁸ e a produção de animais e plantas transgênicos¹⁹. Além disso, a tecnologia do DNA recombinante teve um profundo impacto na indústria de fármacos. A inserção de genes de interesse em bactérias e outros microorganismos faz desses seres verdadeiras fábricas em miniatura, permitindo a clonagem e expressão genética de proteínas e de hormônios humanos de interesse médico, como, por exemplo: a insulina humana e fatores de coagulação sanguínea, substâncias de considerável valor econômico²⁰.

Trata-se, em suma, da possibilidade de cortar, manipular, analisar e reproduzir pedaços do material genético. Embora não possuam muitas das propriedades biológicas complexas de outros organismos, as bactérias e as leveduras tornam-se verdadeiros tubos de ensaio vivos, em que é possível reconstituir reações químicas e biológicas complexas e, com isso, saber mais sobre os processos de regulação e organização do gene das células eucarióticas – que já eram relativamente bem estabelecidos nas células procarióticas. De acordo com o bioquímico Paul Berg, um dos pioneiros na tecnologia do DNA recombinante:

The astounding successes in defining the genetic chemistry of prokaryotes during the 1950's and 60's were both exhilarating and challenging. Not surprisingly, I and others wondered whether the more complex genetic structures of eukaryote organisms, particularly those of mammalian and human cells, were organized and functioned in analogous ways. Specifically, did the requirements of cellular differentiation and intercellular communication, distinctive characteristics of multicellular organisms, require new modes of genome structure, organisation, function and regulation? Were there just variations of the prokaryote theme or wholly new principles waiting

¹⁸ Em 1973, pesquisadores da *Stanford* e *UCFS* fundiram um segmento de DNA contendo um gene de rã do gênero *Xenopus* e inseriram em *Escherichia coli*. O DNA da rã foi replicado, a medida em que a bactéria se replicava, e o gene que dirige a síntese de uma proteína específica da rã foi expresso. Esse foi o primeiro gene animal clonado, ou seja, isolado e propagado (Morrow *et al.*, 1973).

¹⁹ Em 1981 e 1982 foram produzidos os primeiros animais transgênicos. A introdução de um gene de coelho em camundongos (Constantini & Lacy, 1981) e a introdução de um gene que conferiu a coloração rósea aos olhos de drosófilas (Rubin & Spradling, 1982).

²⁰ Em 1976 foi fundada a primeira companhia de biotecnologia (engenharia genética), a *Genentech*. Ela tinha a experiência técnica de Herbert Boyer e o capital privado de Robert Swanson (Zweiger, 2001; Wilmot *et al.*, 2000).

to be discovered in explorations of the genetic chemistry of higher organisms? It seemed important to try to find out (Berg, 1980, p. 386).

Todavia, tão logo se tornou possível pesquisar e intervir mais diretamente nos genes averiguou-se que a estrutura dos genes em organismos eucariontes, e em vírus, apresentava características até então insuspeitadas.

No início da década de 1970, o então diretor do *Cold Spring Harbor Laboratory*, James Watson, concentra as pesquisas do instituto na genética de tumores e escolhe como sistema modelo certos vírus causadores de tumor (www.cshl.org/History/history.html). Os vírus como organismos experimentais são interessantes porque o seu material genético assemelha-se ao das células que infectam (hospedeiras) (Aloni *et al.*, 1977; Berget *et al.*, 1977; Chow *et al.*, 1977, Sambrook, 1977). Por tal característica, Watson esperava entender, por meio do estudo dos vírus, como os genes atuam na transformação de células humanas normais em células com tumores.

Em fevereiro de 1977, os bioquímicos norte-americanos David Glover e David Hogness da *Stanford University School of Medicine* na Califórnia foram os primeiros a documentar que os genes eucarióticos são interrompidos (Glover & Hogness, 1977). Essa descoberta se deu em genes codificadores de RNAs ribossomais, e parece não ter chamado muita atenção. Contudo, quando nos meses de agosto e setembro de 1977 as equipes de pesquisa do inglês Richard John Roberts do *Cold Spring Harbor Laboratory* (Chow *et al.*, 1977) e do americano Phillip Allen Sharp (Berget *et al.*, 1977) do *Center for Cancer Research* no *Massachusetts Institute of Technology* (MIT) descobrem que as seqüências de genes virais codificadores de proteínas estruturais também contêm interrupções (mais tarde denominadas por Walter Gilbert (1978) de íntrons), esse achado parece não ter passado despercebido. Isso indicava que, se os genes codificadores de proteínas dos vírus contêm

seqüências intercalares, o mesmo poderia acontecer para os organismos eucariontes hospedeiros.

Embora as instituições às quais pertenciam estivessem envolvidas em pesquisas sobre cânceres, Phillip Sharp e Richard Roberts não investigavam propriamente vírus tumorais, mas a expressão gênica em adenovírus 2 (Ad2). O Ad2 é um vírus de DNA que infectam as células humanas e que é um dos causadores do resfriado comum (Berget *et al.*, 1977). Em vista do maior conhecimento desse processo em células procarióticas, mas do quase desconhecimento em células eucarióticas, Robert e Sharp almejavam saber mais sobre os processos de transcrição em células eucarióticas. Devido às maiores complexidades dos organismos eucarióticos, escolheram estudar esse processo em um sistema hospedeiro-vírus, valendo-se das semelhanças existentes entre a estrutura do material genético desses organismos. Não há dúvida de que o desenvolvimento dessa pesquisa contribuiria também para o programa sobre cânceres (Chow *et al.*, 1977). A escolha do mesmo organismo modelo por essas duas equipes não parece ser exatamente uma coincidência pois, em 1972, Phillip Sharp e Richard Roberts estiveram pesquisando juntos esse mesmo organismo (o Ad2) no *Cold Spring Harbor* (www.cshl.org/History/history.html).

A expressão gênica dos adenovirus 2 (Ad2) está sobre controle temporal. Alguns genes são expressos continuamente, enquanto outros apenas em um período mais avançado da infecção (as proteínas tardias). As equipes de Phillip Sharp e Richard Roberts investigavam essa regulação gênica, principalmente a relacionada às proteínas tardias. Uma das tarefas a ser feita para realizar o estudo dos mecanismos de expressão gênica é o que chamamos de “mapa de restrição do gene”. De forma sucinta, essa técnica consiste em construir um mapa físico das regiões do gene cortadas por essas enzimas de restrição, para que se possa analisar o gene mais detidamente. Com esse mapa de restrição é possível encontrar o tamanho do fragmento de DNA contendo o gene e mesmo determinar a localização dos diferentes genes no genoma. Os RNAs mensageiros produzidos durante o período de infecção tardia foram isolados e comparados com diferentes fragmentos previamente mapeados no genoma viral. A comparação dos RNAs mensageiros com os fragmentos de DNA permitem identificar os locais do genoma responsáveis pela síntese correspondente.

Os resultados apresentados pela equipe de Roberts (Chow *et al.*, 1977) e de Sharp (Berget *et al.*, 1977) demonstraram que as seqüências que se apresentavam de uma forma contínua no RNA mensageiro são, na verdade, complementares a seqüências presentes em regiões espaçadas no genoma do Ad2. Quando a região do genoma do vírus a partir da qual o RNA mensageiro é transcrito, e o próprio RNA mensageiro correspondente foram hibridizados (alinhados, de acordo com as regras de complementaridade de bases sugerida por Watson e Crick), foram observadas regiões do DNA complementares (que se hibridizavam), intercaladas por seqüências no DNA que não se hibridizaram à molécula de RNAm. Ao permitir que o RNAm se pareie com a região do DNA do qual foi derivado, os segmentos de DNA intervenientes, as regiões não hibridizadas, foram observados ao microscópio eletrônico como alças. A formação dessas alças (*loops*) foi o indicativo de que essas seqüências não tinham correspondentes na molécula de RNAm (Figura 4). Resultados similares foram encontrados utilizando vírus animais (SV40, um vírus encontrado em símios) como sistema modelo também para a investigação dos mecanismos de regulação da expressão gênica nas células eucarióticas (Aloni *et al.*, 1977).

Se o mesmo fenômeno valeria para os organismos eucarióticos hospedeiros ou se essa era apenas uma propriedade das células quando infectadas, não era possível saber até que mais organismos fossem estudados (Chow *et al.*, 1977; Berget *et al.*, 1977). Meses antes, durante a realização do XLII *Cold Spring Harbor Symposium on Quantitative Biology* no início de Junho, Sharp, Roberts e outros pesquisadores do *Cold Spring Harbor* já haviam comunicado que os mensageiros maduros de RNA são constituídos de seqüências espaçadas no DNA.

Ainda durante o ano de 1977 foram anunciados diversos outros trabalhos que denunciavam que não se tratava de um fenômeno isolado, restrito ao sistema hospedeiro-vírus. Foi documentada a existência de genes interrompidos nos mais diversos grupos de

organismos, tais como: em insetos (Glover & Hogness, 1977), em aves (Breathnach *et al.*, 1977) e em mamíferos (Jeffreys & Flavell, 1977; Brack & Tonegawa, 1977). A existência dos genes interrompidos mostrava ser um fenômeno espalhado entre os organismos eucarióticos.

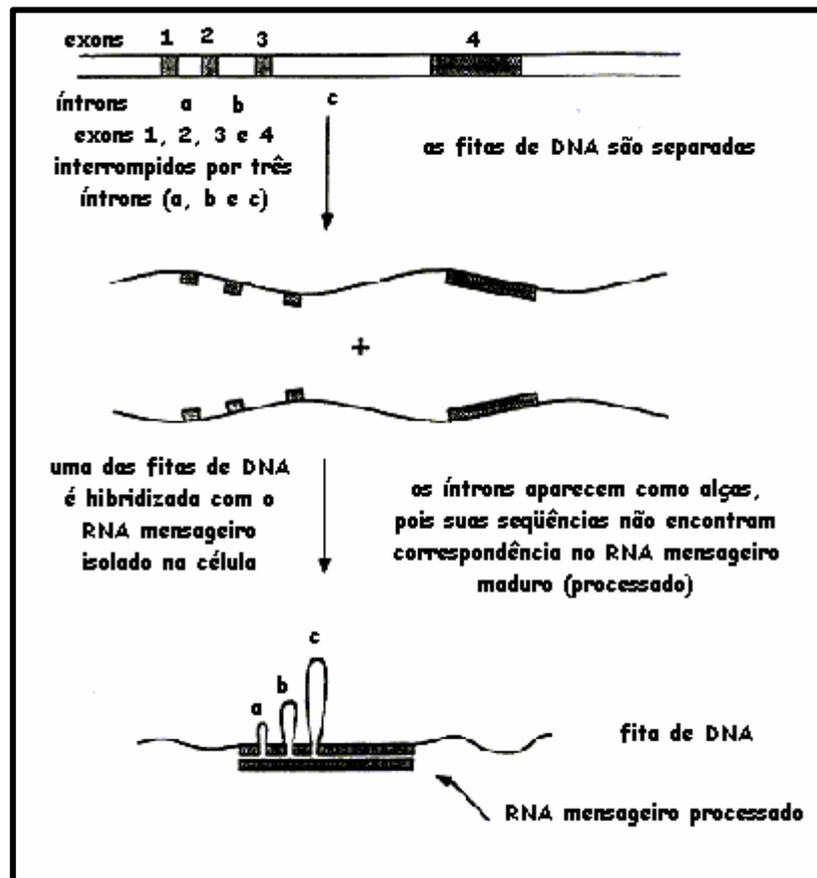


Figura 4: Representação de um modelo de genes que contêm regiões no DNA que não se encontram presentes no RNA mensageiro correspondente. Nesse esquema, o RNAm corresponde a quatro seqüências de DNA (exons 1, 2, 3 e 4) que se encontram espaçadas no genoma viral. Os segmentos adicionais de DNA (íntrons a, b e c) aparecem como alças quando hibridizados a fitas simples de DNA com a molécula de RNAm correspondente. (modificado a partir da figura encontrada em: www.nobel.se/medicine/laureates/1993/press.html).

Vale mencionar que dos trabalhos citados acima, apenas o do grupo liderado por Pierre Chambon (Breathnach *et al.*, 1977; Chambon, 1981) não focava diretamente nos mecanismos de regulação gênica. A equipe de Chambon investigava a diferenciação

celular em galinhas poedeiras. Mais especificamente, Chambon pretendia verificar se nos diferentes tipos celulares, a estrutura dos genes era a mesma. Em caso afirmativo, averiguar se a modificação da própria estrutura do gene era um dos mecanismos envolvidos na diferenciação celular. Para investigar isso, células imaturas recebiam estimulação hormonal e se diferenciavam em células altamente especializadas na síntese de proteínas presentes no ovo. Além da descoberta dos genes interrompidos, Chambon e sua equipe observaram que a estrutura do gene é a mesma, não importando o estado de diferenciação celular. Existe uma exceção para isso. Já em 1976 havia evidências de que as células do sistema imune, os linfócitos maduros, apresentam genes com uma organização diferente dos genes das células embrionárias. Nas células linfocitárias, ocorre uma espécie de “descarte” genético nos genes das imunoglobulinas. Dessa forma, esses genes sofrem uma espécie de rearranjo somático de seu material genético, durante sua passagem de células embrionárias para células maduras. Contudo, esses genes também possuem íntrons (Brack & Tonegawa, 1977). Mas isso é um detalhe a parte.

O processamento do RNA mensageiro está envolvido em algumas doenças. A β -talassemia é uma forma de anemia comum em países mediterrâneos. É uma doença hereditária causada por um defeito na síntese de uma das cadeias da hemoglobina, a proteína existente nas hemácias responsável pelo transporte de oxigênio. As hemácias normais são compostas por duas cadeias da proteína *alfa*-globina e duas da proteína *beta*-globina. Como na maioria dos genes eucarióticos, os genes para essas proteínas precisam ser processados para a retirada dos íntrons. A β -talassemia é causada por uma mutação que cria sítios alternativos para a maquinaria celular responsável pelo processamento. Esse processamento indevido gera uma proteína *beta*-globina defeituosa, que acarreta em hemácias também defeituosas, causando a anemia. Cerca de 15% das mutações que provocam doenças genéticas, provocam falhas no processamento do RNA mensageiro,

acarretando, por exemplo, certos tipos de cânceres e doenças neurológicas (Modrek & Lee, 2002; Graveley, 2001).

O processamento do RNA mensageiro também envolve a seleção de alguns segmentos de DNA disponíveis. Nesse sentido, os mecanismos da célula podem “decidir” como o RNA mensageiro poderá ser processado, combinando alternativamente determinados segmentos do RNA mensageiro (os exons). Esse processamento alternativo apresenta conseqüências bastante claras para a definição de “gene molecular clássico” que indicamos no primeiro capítulo. Examinaremos a seguir o processamento alternativo do RNA mensageiro, e também indicaremos outros processos moleculares que, além dos íntrons, têm gerado um intenso debate sobre o que vem a ser um gene.

O Processamento Alternativo

Segundo Mayr (1998), quando Nirenberg e Matthaei, em 1961, conseguiram desvendar o código genético, acreditava-se amplamente que o último grande problema da biologia molecular acabava de ser solucionado²¹. Acumularam-se desde então muitas descobertas totalmente inesperadas, cujo significado maior, até agora, se relaciona com os aspectos da fisiologia do gene, mas que, evidentemente, também tem importância evolutiva, como acabou por ficar demonstrado quando foram mais bem conhecidos os processos moleculares envolvidos na evolução.

A descoberta dos genes interrompidos, além de outros achados moleculares, indica a existência de mecanismos genéticos que fazem a célula mais dinâmica e complexa do que se suspeitava antes (Chambon, 1981). Como foi exposto, a região do DNA contendo o gene é transcrita numa longa molécula de RNA precursora, que é processada para dar lugar

²¹ O livro de Mayr (1998) foi originalmente publicado em 1982.

à molécula de RNA mensageiro maduro. Este mensageiro mais curto é traduzido em proteína.

Todos os organismos eucariontes contêm íntrons em pelo menos alguns de seus genes, embora esse número varie consideravelmente de organismo para organismo. Por exemplo, estima-se que apenas 250 dos cerca de 6.000 genes do genoma da levedura *Saccharomyces cerevisiae* contenha íntrons (cerca de 4%), enquanto a maioria dos aproximadamente 35.000 genes do genoma humano é composta por genes interrompidos.

Walter Gilbert (1978) conjecturou que uma consequência do modelo íntrônico é que o dogma “um gene, uma cadeia polipeptídica” desaparece. O gene, como uma região de DNA, corresponderia agora a uma unidade de transcrição. Todavia, esta unidade de transcrição, esta região, pode corresponder a não apenas um, mas também a diferentes cadeias polipeptídicas. E estas cadeias polipeptídicas podem tanto possuir funções celulares semelhantes como não possuir (Gilbert, 1978). Embora previsto por Walter Gilbert já em 1978, só em 1987 a existência do processamento alternativo foi verificada experimentalmente (Croft *et al.*, 2000).

O processamento alternativo, ou emenda alternativa, é apenas um entre diversos eventos que ocorrem entre a transcrição do DNA e o produto gênico final. Apenas em termos de modificações após a tradução, já são conhecidos cerca de duzentos tipos diferentes de processos (Brett *et al.*, 2001). Todavia, o processamento alternativo é considerado como um dos mais notáveis componentes da complexidade funcional do genoma humano (Modrek & Lee, 2002; Graveley, 2001; Brett *et al.*, 2001).

Estimativas recentes apontam que cerca de 40 a 60 % dos genes humanos são processados alternativamente. E este número tende a crescer, na medida em que se promova uma análise mais acurada do genoma (Modrek & Lee, 2002). De uma forma geral, os eventos de emenda alternativa são bastante específicos. Na maior parte das vezes,

ocorrem em tecidos particulares e em momentos precisos do desenvolvimento do organismo e/ou somente sob certas condições fisiológicas. Esta é uma das dificuldades de se estimar a proporção destes eventos no organismo (Graveley, 2001).

A ovalbumina, por exemplo, é a principal proteína presente no ovo de galinhas. Foi estudando esta proteína que a equipe de pesquisa liderada por Pierre Chambon fez a sua descoberta particular dos genes interrompidos (Breathnach *et al.*, 1977). O tamanho total do gene da ovalbumina (isto é, estamos considerando a região do DNA contendo toda a sequência que codifica o RNA mensageiro, incluindo os oito íntrons que não tem contraparte no mensageiro final) é de cerca de 7.700 pares de bases nitrogenadas. O RNA mensageiro maduro contém cerca de 1.872 nucleotídeos. Destes 1872 nucleotídeos, 64 pertencem a uma região que antecede a região codante²², que é denominada de sequência líder não traduzida. Uma outra região que também não é traduzida se encontra no final do mensageiro. Esta região tem cerca de 650 nucleotídeos. No final das contas, de toda a região do DNA transcrita, apenas 1.158 nucleotídeos são responsáveis pelos 386 aminoácidos da proteína. Em suma, a parte do mensageiro que é realmente traduzido em proteína é quase sete vezes menor que a sequência do gene no DNA (Chambon, 1981).

Provavelmente, o mais famoso e citado exemplo de processamento alternativo é aquele que tem implicações na determinação das características sexuais na mosca-de-frutas *Drosophila melanogaster* (Graveley, 2001; Baker, 1989). O produto do processamento diferencial do gene *sxl* dispara vias alternativas de genes em cascatas (Figura 5). Um efeito de cascata é um processo que ocorre quando os genes estão ligados de tal modo que a ativação ou desativação de um acarreta na ativação ou desativação do outro, que por sua vez, tem efeito no gene seguinte. Os diferentes genes disparados pelo produto diferencial do gene *sxl* resultam no estabelecimento das características sexuais femininas ou masculinas (Baker, 1989). Desse modo, a determinação sexual em drosófilas é controlada pela regulação do processamento do mensageiro de RNA²³.

²² Região da molécula do mensageiro que especifica os aminoácidos na proteína.

²³ O determinante principal do sexo em drosófilas é a taxa X:A, ou seja, a relação entre o número de cromossomos X e o conjunto de autossomos (A) (cromossomos não sexuais). Se esta taxa for igual a 1 originam-se a fêmeas, se for igual a 0,5 machos. Por exemplo, moscas com um cromossomo X e dois conjuntos de autossomos são machos. Não se sabe exatamente como essa taxa X:A ativa o gene *Sxl*, mas é provável que os fatores de ativação estejam presentes na célula ovo, produzido por genes maternos e por genes do próprio zigoto. (Baker, 1989).

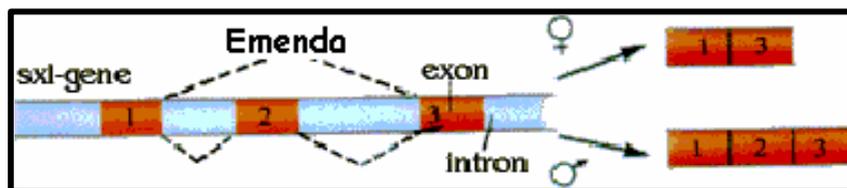


Figura 5: A molécula de RNA precursora pode dar origem a RNAs mensageiros com estruturas alternativas. A mosca-de-frutas *Drosophila melanogaster* é um exemplo em que o processamento alternativo tem conseqüências dramáticas para o seu desenvolvimento. Note que a seleção dos exons dá origem a proteínas com seqüências ligeiramente diferentes. Todavia, esta pequena diferença tem profundos comprometimentos para o organismo. Nesse caso, o processamento alternativo define a funcionalidade da proteína *Sxl*. Esta proteína regula uma cascata de eventos que determina o sexo desses organismos. Note que a proteína *Sxl* não contém o exon 2. (Baker, 1989). (figura obtida e modificada a partir do endereço eletrônico: www.nobel.se).

No exemplo anterior, observamos a geração de apenas duas emendas alternativas a partir de um mesmo mensageiro, ou um mesmo gene. Mas quantas formas alternativas podem ser geradas a partir de uma única região de transcrição, a partir de um único gene? Estão sendo descobertos diversos genes que são responsáveis não por apenas duas ou três emendas alternativas, mas por centenas e talvez até milhares de diferentes RNAs mensageiros. Menciono a seguir apenas dois exemplos nos quais o processamento alternativo de mensageiro de RNA é utilizado na produção de uma incrível diversidade molecular.

No ouvido interno de galinhas, um gene denominado *slo* pode gerar mais de 500 diferentes RNA mensageiros mediante o processamento alternativo (Graveley, 2001). A proteína gerada por esse gene desempenha um papel fundamental na habilidade das células ciliares do ouvido de captar variações na frequência de ressonância da célula e, assim, afetar a sensibilidade sonora do animal. Dessa forma, cada proteína variante (em que foi alterada a sua seqüência de aminoácidos devido ao processamento diferencial) responde a uma frequência sonora específica, que permite a interpretação de sons diferenciados captados coletivamente (Graveley, 2001, Keller, 2000).

Contudo, até o momento, talvez o exemplo mais notável de geração de diversidade através do processamento alternativo está relacionado com o gene *Dscam*. O *Drosophila Dscam* é o equivalente em Drosófilas (homólogo) ao gene humano *Dscam* (*Down syndrome cell adhesion molecule*) (Schmucker *et al.*, 2000)²⁴. Um dos eventos mais complexos durante o desenvolvimento é a migração e a conexão entre os neurônios. Este processo necessita de um sistema preciso para que o axônio – o prolongamento ou o “pescoço” do corpo de uma célula neuronal, por onde é conduzido o impulso nervoso – em desenvolvimento se dirija corretamente ao seu alvo. As moléculas de adesão celular são uma espécie de proteínas guias. Elas são necessárias para que os neurônios encontrem seus alvos durante a formação do cérebro (Ridley, 2003; Wang, 2002; Schmucker *et al.*, 2000; Serafini, 1999; Wu & Maniatis, 1999).

O processamento alternativo no RNA mensageiro do gene *Drosophila Dscam* pode, potencialmente, gerar 38.016 proteínas alternativas (Schmucker *et al.*, 2000). É possível que esta diversidade molecular contribua para a especificidade da conectividade dos neurônios. A estrutura do gene *Dscam* parece fornecer uma base molecular para o reconhecimento celular e, assim como a de outras proteínas relacionadas com o reconhecimento e formação das conexões nervosas, esta estrutura é parecida com a das imunoglobulinas, as proteínas altamente variáveis do sistema imune (Wang, 2002; Schmucker *et al.*, 2000; Serafini, 1999; Wu & Maniatis, 1999).

O processamento alternativo parece ter uma notável importância em organismos mais complexos. Os dados indicam que o processamento alternativo parece aumentar em razão da maior complexidade dos organismos. Nos organismos multicelulares mais complexos, a informação precisa ser processada diferentemente em distintas situações ou em situações em que é requerido um alto nível de diversidade, como em sistemas de reconhecimento e sinalização celular (Brett *et al.*, 2002; Modrek & Lee, 2002). O número de genes não é, necessariamente uma medida confiável da complexidade dos organismos, mas sim o modo como eles são combinados. O processamento alternativo permite que vários tipos de tecido possam trabalhar com um número limitado de genes, pois a

²⁴ O *Dscam* causa alguns dos sintomas da síndrome de Down por um mecanismo ainda desconhecido (Ridley, 2003).

informação armazenada nos genes de organismos complexos pode ser editada de várias formas.

Processamento alternativo e o número de genes

O anúncio, em fevereiro de 2001, de uma primeira estimativa do número de genes humanos em cerca de 30.000 genes (Lander *et al.*, 2001; Venter, *et al.*, 2001) foi um dos dados do Projeto Genoma Humano que causou uma grande repercussão (Downes, 2004; Ridley, 2003). Talvez, parte dessa surpresa tenha sido devido ao fato de que esse número não parecia muito mais alto, se comparado a outros organismos considerados menos complexos, como, por exemplo, os cerca de 26.000 genes da erva *Arabidopsis thaliana*, ou os aproximadamente 13.000 genes da mosca-de-fruta *Drosophila melanogaster*.

Todavia, se mantida a definição de que cada gene codifica apenas uma única proteína, a estimativa de que os genes humanos possuem cerca de 30.000 genes torna-se surpreendente, na medida em que o conjunto das proteínas do organismo humano estaria, provavelmente, subestimado, mesmo se considerando que o número de proteínas que um organismo pode produzir é um guia, relativamente, tosco da complexidade geral do organismo (Downes, 2004). Segundo Ridley (2003), John Sulston, um dos líderes do Projeto Genoma Humano, afirmou que somente 33 genes, cada um deles em duas variedades (ativo e inativo), seriam suficientes para tornar único cada ser humano no mundo, pois há mais de 10 bilhões de maneiras de lançar uma moeda 33 vezes. Desse modo, 30.000 genes não parece um número tão pequeno, afinal.(cf. Ridley, 2003, p. 10).

Entretanto, o que realmente parece ser mais surpreendente, ao nosso ponto de vista, é que cerca de vinte anos antes deste anúncio, Pierre Chambon (1981), um dos envolvidos na descoberta dos genes interrompidos durante o ano de 1977 (Breathnach *et al.*, 1977),

argumentou que diversos indícios sugeriam que o número de proteínas em mamíferos, e por conseqüência, o número de genes, dificilmente passaria de 150.000. Segundo Chambon, não seria improvável que o número real fosse tão reduzido quanto 30.000 genes. Não estamos querendo defender qualquer tipo de posição visionária de Pierre Chambon, mas apenas argumentar que já em 1981, uma análise mais atenta da estrutura do material genético indicava a idéia de que o número de genes não é, necessariamente, uma boa medida da complexidade dos organismos.

Chambon estava estimando o número de genes para proteínas e não genes para outros tipos de RNA, como por exemplo, RNAs ribossomais. Nessa época, em 1981, já se sabia da discrepância entre a quantidade total de DNA nas células eucarióticas e o a quantidade de RNAs mensageiros e, assim, de proteínas. Uma das explicações seria a redução do mensageiro através do processamento da molécula de RNA, além de outras características do material genético, como, por exemplo, um grande número de seqüências repetitivas, espaços entre um gene e outro, regiões reguladoras não transcritas. (Chambon, 1981).

O gene molecular utilizado para estimativas do número de genes é apenas um parâmetro para a análise genética (Fogle, 2000). Cada método de estimativa escolhe certos parâmetros, como, por exemplo, marcas estruturais do DNA que indicam o possível local de um gene (espécies de etiquetas moleculares indicando a possível existência de um gene) ou o número de RNAs expressos²⁵. Cada método produz estimativas ligeiramente diferentes e, obviamente, cada método apresenta vantagens e desvantagens um em relação ao outro. Contudo, numa primeira aproximação na contagem do número de genes é estimado, principalmente, o número de genes envolvidos na síntese de proteínas. Estes

²⁵ Existe uma discrepância entre as estimativas baseadas na análise da seqüência no DNA e na análise de RNAs expressos que não é devido somente ao processamento alternativo. Veremos isso, quando mencionarmos alguns outros processos celulares.

dados iniciais ignoram, por exemplo, os transcritos possíveis por processamento alternativo. Conta-se, sobretudo, a região do DNA como uma possível unidade de transcrição (Mattick, 2003; Fogle, 2000). Neste sentido, o número de genes depende, em grande parte, dos parâmetros escolhidos. A medida em que novas e mais análises vão sendo realizadas, este número tende a variar.

As duas análises, de Chambon e a do próprio Projeto Genoma Humano, parecem privilegiar o gene como uma unidade de transcrição de uma proteína. Ao nosso ver, isso nos traz de volta ao conceito de “gene molecular clássico”. É um dos motivos para a disparidade entre o número de genes e a cifra de proteínas necessárias para o desenvolvimento e manutenção do organismo parece residir justamente a que está se referindo o termo “gene” (Downes, 2004). Múltiplos transcritos derivados de um único *locus* não são considerados como múltiplos genes, mas variações de uma mesma unidade de transcrição (Miklos & Rubin, 1996).

Dois temas inevitavelmente nos confundem quando propomos a descoberta dos íntrons como um evento crucial para a reformulação do conceito de “gene molecular clássico”. O primeiro, obviamente, é a exatamente a que se refere o termo “gene”. E o segundo, tem a ver com o significado evolutivo dos íntrons e do processamento alternativo. Seria enriquecedor para a dissertação uma breve discussão sobre esse ponto. Infelizmente, deixaremos isso para uma outra oportunidade. Contudo, ao menos mencionaremos que o processamento alternativo teria algum tipo de relação com a geração de uma incrível diversidade molecular (Brett *et al.* 2002). Compreender como os genes são montados pode ser crucial para o entendimento de princípios biológicos verdadeiramente fundamentais para os organismos e para a evolução das espécies.

Outros eventos celulares

Em 1961, portanto na época em que o código genético começava a ser decifrado, os biólogos franceses François Jacob e Jacques Monod do Instituto *Pasteur*, propõem o sistema “Operon lac” para explicar a regulação genética em bactérias. Com este sistema estabelecem a distinção entre genes estruturais (genes encarregados da produção de proteínas necessárias ao organismo) e genes reguladores (genes que regulam a taxa de expressão dos genes estruturais).

Na medida em que os genes reguladores também sintetizam proteínas, à primeira vista, esta distinção não parece modificar em nada as idéias prévias sobre o gene. Contudo, para alguns autores essa foi uma primeira mancha na face do dogma central (Keller, 2000), o momento em que a idéia de gene molecular já começa a ter que localizar outros referentes para explicar o seu funcionamento (Reznik, 1995).

François Jacob e Jacques Monod propõem através do modelo “Operon lac” um mecanismo responsável por controlar a síntese da enzima β -galactosidase, que degrada β -galactosídios²⁶. De forma breve, o modelo é composto, basicamente, por três elementos: o gene regulador, o repressor e o operador. O gene regulador sintetiza um repressor específico. Este repressor bloqueia a região do DNA denominada de operador. Esta enzima repressora ligada na região do operador impede o reconhecimento da RNA polimerase, e, desse modo, bloqueia a produção do mensageiro e, por conseqüência da enzima (Keller; 2000; Falk, 1986; Jacob, 1965).

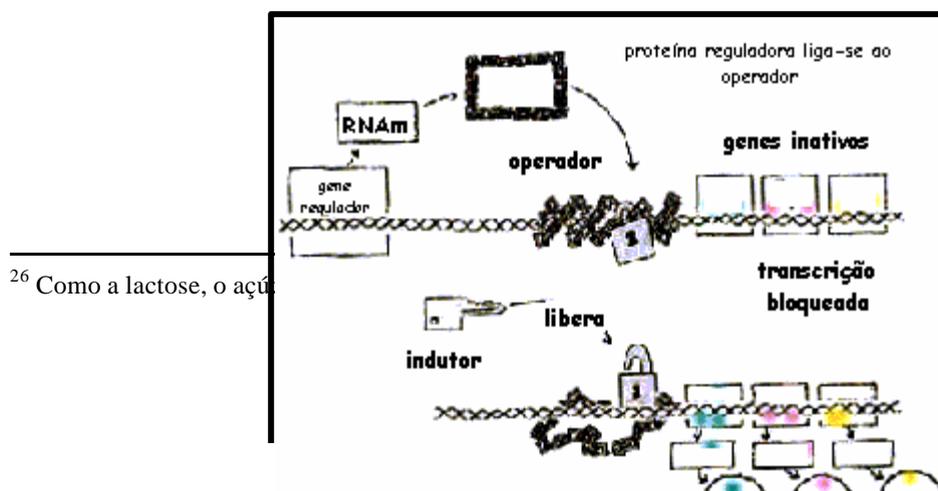


Figura 6: O modelo “Operon lac” de Jacob e Monod, de 1961, estabeleceu a distinção entre genes estruturais (genes encarregados pela produção de enzimas, que são proteínas, necessárias à construção do organismo) e genes reguladores (genes que regulam a taxa metabólica destes genes). Uma proteína reguladora liga-se ao operador (uma região do DNA), evitando que a polimerase do RNA transcreva os genes estruturais. Quando o indutor está presente, liberta o repressor do operador pela ligação direta com a proteína, e assim mudando sua conformação. Logo que fica livre do repressor, a polimerase do RNA pode continuar a transcrição dos genes estruturais do operon em RNA mensageiro. (figura retirada de Keller, 2002).

A nosso ver, uma das contribuições mais importantes do modelo “Operon lac” de François Jacob e Jacques Monod, não foi a distinção entre genes reguladores e genes estruturais: sua notável relevância foi interpretar as seqüências de DNA como possuindo diferentes funções. Nem todas as regiões do DNA que tem importância para a célula codificam proteínas. A existência de um inibidor específico dentro da célula, o repressor, só faz sentido, se houver um local no DNA em que este repressor possa reconhecer e bloquear a síntese da proteína (Jacob, 1965). Existe uma espécie de sintaxe molecular, um conjunto de regras químicas que regulariam a organização molecular. O próprio aparecimento do produto a ser sintetizado ou degradado, envia mensagens para que a célula produza as enzimas necessárias para o seu metabolismo.

Regiões não transcritas do DNA como o operador, promotores²⁷ e *enhancers*²⁸, sinais de término²⁹ possuem informações cruciais para o reconhecimento e ligação de proteínas necessárias para ativação e/ou inativação do gene. Na medida em que a presença dos ?-galactosídios liberam a ligação do repressor ao operador, Jacob e Monod apresentam um modelo de como a célula encontra um mecanismo capaz de reconhecer quando há necessidade de expressar os genes devido às diferentes condições ambientais e execução das funções vitais celulares.

²⁷ Seqüências de DNA, geralmente anteriores ao gene a ser transcrito e que age como o local de reconhecimento e ligação da enzima responsável pela transcrição, a RNA-polimerase.

²⁸ Poderiam ser traduzidos como acentuadores. São elementos, seqüências de DNA, que potencializam a ação dos promotores, as taxas de transcrição. Contudo, os *enhancers* não são específicos aos promotores que controlam e podem se localizar em regiões variadas em relação ao promotor.

²⁹ Assim como possui os promotores como sítio de reconhecimento para que a RNA polimerase inicie a transcrição, esta enzima também deve reconhecer sinais de término no DNA que indicam o fim da transcrição. Não confundir regiões terminadoras, com códons finalizadores. Os códons finalizadores indicam o fim da tradução, término da proteína, e não da transcrição. Desnecessário dizer que cada um desses processos tem uma história específica (Mayr, 1998).

Certos autores cogitam que um gene abrange as seqüências reguladoras necessárias à sua expressão (Moss, 2001; Moss, No Prelo; Griffiths & Neumann-Held, 1999), enquanto outros acreditam que a inclusão de seqüências reguladoras introduz uma complexidade considerável ao termo “gene”. Isto porque, existem muitos tipos diferentes de elementos reguladores, e estes, atuam mediante complexas combinações (Epp, 1977).

Além da informação biológica aparente, existe uma espécie de camada de informação biológica oculta no DNA (Mattick, 2003). Vimos no primeiro capítulo dessa da dissertação que o “dogma central da biologia molecular” de Francis Crick sustenta que a informação genética, normalmente, flui do DNA para o RNA, e desse para as proteínas. Com isso, e a subsequente descoberta do código genético, foi assumido, de uma forma geral, que os genes eram sinônimos de proteínas, isto é, os genes seriam repositórios de seqüências codificantes de proteínas, apenas excetuando-se os tipos de RNAs necessários para a infraestrutura e funcionamento da maquinaria celular. Os genes sintetizariam as proteínas, e essas cumpririam as funções estruturais, catalíticas e reguladoras necessárias para o funcionamento da célula.

Estimativas recentes apontam que as seqüências codificantes de proteínas constituem apenas cerca de 2% do genoma (Mattick, 2003). No entanto, o genoma abriga, além dos genes codificadores de proteínas, uma infinidade de RNAs que não codificam nenhuma proteína: os chamados ncRNA – *non-protein-coding RNA* – nome genérico para diversos tipos de RNA, como os já conhecidos RNAs ribossomais, RNAs transportadores, RNAs spliceomais³⁰ e numerosos outros tipos de RNAs. A discrepância nas estimativas do número de genes em mamíferos baseadas em análises da seqüência do DNA (30.000-40.000) e análises de RNAs transcritos na célula (65.000-70.000) ocorre por causa dos diferentes tipos de ncRNAs (Mattick, 2003;. Storz, 2002).

E mais uma vez os íntrons, estão envolvidos nessa história. Íntrons derivados tanto de genes codificadores de proteínas, quanto dos próprios ncRNAs parecem exercem algum tipo de regulação na taxa de expressão desses transcritos. Mas, a figura que mais parece surpreender os cientistas são os pseudogenes, cópias defeituosas ou encurtadas de genes

³⁰ Não mencionamos, mas o spliceossoma é o nome que recebe a maquinaria celular responsável pelo processamento dos transcritos de RNA. O spliceossoma é formado por moléculas de RNA (que reconhecem algumas seqüências nos limites entre íntrons e exons) e proteínas.

que se encontram ativos nas espécies de organismos. Existem muitos deles espalhados pelo genoma e em número quase igual ao dos genes ativos. Os pseudogenes foram considerados como uma espécie de resquício molecular, genes que perderam a funcionalidade devido a algum tipo de mutação e que ficaram abandonados no genoma. Entretanto, a eliminação de pseudogenes do genoma tem provocado a desativação de sua contraparte funcional. Assim, ao que parece, os pseudogenes podem controlar a expressão do gene do qual são cópias. Muitas vezes, esses genes se encontram até mesmo em cromossomos diferentes (Mattick, 2003; Storz, 2002).

Tem-se observado que esse montante de transcritos de RNA não traduzido – que não codifica proteína alguma – apresenta diferentes funções estruturais, catalíticas e, principalmente reguladora. (Mattick, 2003; Riddihough, 2002; Storz, 2002; Eddy, 2001). Ao que parece, tais RNAs deveriam ser legitimamente considerados como genes (Mattick, 2003; Eddy, 2001; Fogle, 2000).

Quando temos em mente alguns dos eventos moleculares que citamos anteriormente, a definição de “gene molecular clássico” – como uma fita específica do DNA que codifica uma única cadeia polipeptídica (Griffiths & Neumann-Held, 1999) – não nos parece muito satisfatória. Mas por que ela deveria ser? A definição de conceitos parece uma tarefa a rondar mais as cabeças de filósofos e historiadores, mas que não parece realmente preocupar os biólogos nas suas rotinas de laboratório (Ridley, 2003; Keller, 2000).

Não vemos melhor forma de justificar o nosso interesse do que nos valer das palavras de um geneticista que se auto-intitula “card-carrying”. De acordo com William Gelbart (1998), do Departamento de Biologia Celular e Molecular da Universidade de *Harvard*. Segundo Gelbart, “(...) we may well have come to the point where the use of the term ‘gene’ is of limited value and might in fact to be a hindrance to our understanding of the genome.” (Gelbart, 1998, p. 660). Gelbart considera que essa sua afirmação pode soar como heresia, contudo “the realities of genome organization are much more complex than can be accommodated in the classical gene concept.” (Gelbart, 1998, p. 660). O geneticista

afirma incisivamente que essa não é uma discussão puramente abstrata, mas que o conceito de gene como comumente utilizado pode confundir o entendimento preciso dos dados obtidos pela pesquisa genômica.

Nossa pesquisa indica que esses problemas já começaram bem antes, a saber, durante a passagem das pesquisas em organismos procariontes para organismos eucarióticos que possuem uma organização do material genético mais complexa. Quando se pesa a importância da descoberta dos genes interrompidos e de alguns outros processos moleculares que mencionamos, surgem visões conflitantes sobre a que se refere o termo “gene”. Apresentaremos, sucintamente, o que acreditamos serem os dois pólos dessa discussão. De acordo com Evelyn Fox Keller (2000), professora de história e filosofia da ciência no *Massachusetts Institute of Technology* (MIT) e Ph.D. em física, os genes tiveram uma vida gloriosa no século XX, mas, o gene, como um conceito explicativo central para a biologia é uma característica do século passado que dificilmente se sustentará no século XXI. Para o geneticista Raphael Falk (1986), do Departamento de Genética, da *Hebrew University of Jerusalem*, Israel, a flexibilidade desse conceito é um componente essencial para que ele acomode todo o recente desenvolvimento da pesquisa experimental.

Para acompanhar esse debate, além de outros artigos, estaremos armados de uma espécie de minuta que antecedeu um *Workshop* realizado no ano de 2003, na Universidade de *Pittsburgh*. O encontro foi inteiramente voltado para a discussão dos problemas relativos ao conceito de gene. A minuta pode ser encontrada no seguinte endereço eletrônico: <http://www.pitt.edu/~kstotz/genes/minutes.html>. Após a realização desse *Workshop* em 2003, parece ter se organizado formalmente, um grupo de pesquisa sobre o tema. Está prevista para o ano de 2005 a realização de outro *Workshop*.

Nossa escolha recaiu sobre esses dois autores por razões bastante claras. O livro de Keller, *The Century of the Gene*, teve repercussão tanto negativa quanto positiva no meio acadêmico, porém é, sobretudo, um trabalho de popularização científica. O artigo de Raphael Falk³¹, “What is a Gene?”, é, de acordo com nosso levantamento bibliográfico e leituras prévias, o mais antigo a abordar o conceito de gene com relação aos problemas que queremos tratar. De acordo com a bibliografia apresentada na minuta anteriormente citada, o artigo de Falk também é o mais antigo dentre os estudados por esses pesquisadores, que consideram que apesar da data de publicação, tal trabalho ainda pode ser lido como o “estado da arte”, ou seja, se mantém à par do nível de desenvolvimento atual das pesquisas biológicas. Embora Keller enfatize o Projeto Genoma Humano como centro de sua discussão, a autora parece discutir as mesmas questões colocadas por Falk. Apesar disso, esses dois autores chegam a conclusões bastante distintas.

Evelyn Fox Keller e a dinâmica celular

Em *The Century of the Gene* (2000), Evelyn Fox Keller defende a tese de que o conceito de gene molecular pode ter se tornado um obstáculo para a compreensão dos biólogos. Para Keller, o responsável por desencadear tal crise teria sido o Projeto Genoma Humano.

O objetivo deste livro é celebrar os efeitos surpreendentes que os sucessos desse projeto tiveram sobre o pensamento biológico. Ao contrário de todas as expectativas, em vez de apoiar noções familiares de determinismo genético que adquiriram tão grande poder na imaginação popular, esses resultados criaram desafios críticos a essas noções clássicas. Hoje, a proeminência dos genes, tanto na mídia em geral quanto na imprensa científica, sugere que nessa nova ciência da genômica, a genética do século XX atingiu sua apoteose. Porém, o próprio sucesso que tanto agitou nossa imaginação também minou de forma radical seu conceito fundamental, o conceito de gene. (Keller, 2002, p. 17).

³¹ Raphael Falk é um dos organizadores de um livro, recentemente publicado, sobre o conceito de gene. *The Concept of the Gene in Development and Evolution*. Cambridge: Cambridge University Press. 2001.

Retraçando toda a sua história, desde Mendel, Keller argumenta que desde cedo se atribuiu aos genes diversas funções, como: explicar a constância da herança (a passagem das características ao longo das gerações), a variabilidade necessária ao processo evolutivo, além da condução dos processos do desenvolvimento e do metabolismo. Segundo Keller era necessário que os genes apresentassem uma estrutura que respondesse por todas essas funções. Os pesquisadores imaginavam que para responder o que um gene faz, precisavam descobrir a estrutura do gene e investigar como tal estrutura explicaria sua função (Keller, 2000). Segundo Keller, após a redescoberta dos trabalhos de Mendel em 1900, a grande questão era: afinal, o que é um gene? Que tipo de entidade é capaz de regular o desenvolvimento dos organismos e ao mesmo tempo assegurar a estabilidade da passagem dessas informações de uma geração para outra? Contudo, como expomos e como Keller também não negligenciou, isso de forma alguma impediu o desenvolvimento da genética. Podemos citar o desenvolvimento da “teoria cromossômica da herança” por parte do grupo de Morgan (a indução de mutações e previsão de seus efeitos nos organismos) e os estudos da genética de populações. Essas e outras suposições levaram adiante os estudos genéticos sem que nenhuma elucidação da base física da herança fosse necessária. Como demonstramos, o interesse nas bases físico-químicas da vida parece ter nascido mais do interesse dos físicos e dos químicos, do que propriamente dos biólogos. Uma contundente afirmação desse aparente desinteresse é a declaração de Morgan na página vinte e seis da presente dissertação, que não iremos replicar aqui.

Como já expomos, coube a estrutura da molécula de DNA responder por essas funções. A partir da década de 1960, começa-se a entender o que um é gene do ponto de vista de seu funcionamento molecular, de sua capacidade de informar a produção de proteínas. A natureza molecular dos genes trouxe novas questões a respeito do funcionamento gênico. À medida que mais se estudavam os processos de replicação,

transcrição e tradução, ficava evidente tratar-se de um processo bastante complexo, envolvido em delicados mecanismos de regulação. Nesse sentido, a passagem e a manutenção das informações manifestam-se como altamente dependentes de enzimas, fatores protéicos, fontes energéticas e moléculas diversas de RNA, implicando na participação de outros agentes, não apenas genes (Keller, 2000; Lewin, 2000).

A partir disso, a tese de Keller defende que o gene não pode ser o único responsável pela origem das atividades biológicas. Não deve ser esquecida a importância da célula, como um todo, para determinação de sua própria atividade. Não seria possível concentrar toda a responsabilidade, por exemplo, de nossas características físicas sobre trechos de DNA. Para a autora, um dos principais sinais dessa dificuldade reside no fato de que já não é tão fácil definir e localizar um gene, na medida em que as seqüências do DNA sofrem diversos processamentos, que acarretam na divergência do produto gênico final.

Evelyn Fox Keller conjectura que existem lacunas entre a informação genética contida nas seqüências de DNA e o seu significado biológico. Para suprir essas lacunas – evidenciadas pela divulgação do Projeto Genoma Humano – há a necessidade da instituição de um novo paradigma que conduza as pesquisas genéticas, na medida em que o dogma DNA-RNA-PROTEÍNAS não é mais suficiente para sustentar uma explicação satisfatória. Para a autora, esse projeto transformou em dúvidas o nosso seguro conhecimento biológico. Nesse sentido, os genes, a molécula de DNA, devem ser entendidos apenas como parte da explicação e não a própria explicação para os fenômenos biológicos. Nas palavras de Keller:

(...) o gene não pode mais estar acima e afastado dos processos que especificam a organização celular e intercelular. Esse gene é, ele próprio, parte e parcela de processos definidos e criados pela ação de um sistema complexo, dinâmico e auto-regulatório no qual,

e para qual, o DNA herdado fornece a matéria prima crucial e absolutamente indispensável, mas não mais que isso (Keller, 2002, p. 83-84)³².

Haveria uma revolução na biologia em curso: a necessidade de mudar o paradigma atual e promover transformações no modelo que guia as pesquisas científicas. O “dogma central da biologia molecular”, representante legítimo do paradigma das pesquisas genéticas, se encontra abalado por uma avalanche e um acúmulo de novos dados genéticos moleculares que se mostram desarticulados, e desmistificam antigas idéias a respeito do funcionamento dos genes. Ao invés de genes, a atenção deve agora ser deslocada para a interação entre os vários componentes da célula.

Isto porque se observa que a suposta estabilidade da molécula de DNA é muito mais complexa do que se supunha. A estratégia de Keller parece sempre ser uma comparação entre o que se sabe hoje e as expectativas anteriores: a evidente distância crescente entre as suposições iniciais da década de 1960 e os dados efetivos que novas ferramentas moleculares permitem obter. Numa metáfora exagerada, diríamos que essa descoberta molecular traria à tona a verificação de que elefantes não são *Escherichia coli*.

Hoje, a estabilidade da molécula de DNA é percebida não como algo intrínseco, mas como parte de um processo dinâmico celular muito mais complexo. O pareamento específico das bases, por si só, não garante que a replicação ocorra de modo eficiente, sem erros. A fidelidade da replicação depende da ação das polimerases e de outras enzimas (sistema de reparo do DNA) que conferem e consertam os erros de pareamento dos nucleotídeos. Para a estabilidade da molécula de DNA durante a replicação e transcrição, é necessário um sofisticado sistema enzimático, uma maquinaria celular muito específica, de reparo e revisão, que diminui drasticamente a possibilidade de erros tanto na replicação, quanto na transcrição. E mais uma vez, é desnecessário dizer que para cada aspecto comentado há uma história que responde as questões básicas de Mayr (Quem? Quando? Onde? O quê? Como? e Por quê?). Apenas lembramos que no primeiro capítulo da dissertação mencionamos que durante a formulação da estrutura físico-química do DNA, Watson e Crick postularam que poderia haver algum tipo de enzima envolvida na replicação da molécula de DNA (conferir página 33 da presente dissertação).

Para Keller, a estabilidade da estrutura do gene passa a ser entendida não como um ponto de partida, mas como um produto final de um processo dinâmico de enzimas organizadas em redes metabólicas que asseguram a estabilidade e a fidelidade de

³² Devido a afirmação, a nosso ver, contraditória de Keller, reproduzimos a versão original em inglês para evitar qualquer especulação sobre erros de tradução na versão em português. “(...) the gene can no longer be set above and apart from the process that specify cellular and intercellular organization. That gene is itself part and parcel of a complex self-regulating dynamical system in which, and for which, the inherited DNA provides the crucial and absolutely indispensable raw material, but no more than that. (Keller, 2000, p. 71). Simplesmente, em termos lógicos, alguma coisa não pode ser ao mesmo tempo matéria prima crucial e indispensável e ser diminuída retoricamente pela expressão “but no more than that. O que significa isso?

replicação. “Atualmente, já abandonamos a esperança de encontrar na estrutura molecular de genes particulados uma explicação adequada para a estabilidade da organização biológica através de gerações.” (Keller, 2002, p. 53). Contudo, “para ser fiel à história” (Keller, 2002, p. 40), entre o final dos anos 50 e início da década de 60, foi demonstrado em bactérias e fagos (vírus) que lesões no DNA provocadas por radiações poderiam ser revertidas, o que significa que os ácidos nucleicos possuem mecanismos que garantem essa estabilidade, que estão aptos a reparar lesões ocorridas.

Além da questão da estabilidade da molécula de DNA, Keller questiona a função gênica. Vale lembrar que em 1940, Beadle e Tatum, através dos experimentos com os mutantes do fungo *Neurospora* estabelecem com a hipótese “um gene – uma enzima” que o gene responde pelo controle das reações metabólicas da célula. Com a descoberta da estrutura do DNA em 1953, a formulação do “dogma central da biologia molecular” de Crick em 1957 e a quebra do código genético, nos anos de 1961 a 1966, se estabelece a correspondência entre a seqüência de bases do DNA e a seqüência de aminoácidos de uma dada proteína.

Até que enfim, tivemos uma resposta para a questão de como atuam os genes. O que um gene faz? Ele faz (ou codifica) uma enzima. Mais ainda, um defeito no gene conduz a um defeito (ou talvez à ausência) da enzima, que, por sua vez, pode ser com freqüência diretamente correlacionado à anormalidade (ou ausência) de uma determinada característica. (Keller, 2002, pp. 66-67).

Para Keller, essa relação era espantosa pela simplicidade e elegância, equivalente a uma equação matemática. Todavia, essa simplicidade começa a ser posta em dúvida na medida em que, segundo Keller, rugas se transformam em fendas. Os segredos da vida que residiam na simplicidade da molécula de DNA, já estavam presentes nos anos 1960 e 1970, antes mesmo do código ser decifrado, paulatinamente e sem a devida percepção se transformando em quadros muito complexos e confusos. Como mencionado, para Keller, a primeira ruga ao “dogma central da biologia molecular” veio em 1959, por intermédio do trabalho de François Jacob e Jacques Monod. Estruturalmente, não ocorre nenhuma modificação no “dogma central”, mas funcionalmente se passou a estabelecer a distinção entre os “genes estruturais” (genes que se encarregaram da produção de enzimas

necessárias à construção do organismo) e “genes reguladores” (genes que regulam a taxa metabólica dos genes estruturais).

Segundo Keller, a hipótese “um gene – uma enzima” não implicava em nenhum tipo de regulação gênica e, dessa forma, a distinção entre “genes reguladores” e “genes estruturais” amplia os limites e o poder interpretativo das funções gênicas, embora, como a própria Keller admite, não institua a idéia de regulação como algo inteiramente novo.

Com a descoberta dos genes interrompidos, “uma pedra ainda maior foi atirada nas engrenagens do conceito de gene” (Keller, 2002, p. 72). Nesse sentido, Keller retoma sua tese. Para que o evento de montagem dos RNAs mensageiros ocorra de forma precisa, é necessário um complexo dispositivo de processamento constituído de proteínas e outros RNAs³³, o spliceossoma, como já mencionamos.

As leituras variáveis de um mesmo transcrito (processamento alternativo) são necessariamente frutos de uma cuidadosa regulação que depende do tipo de célula em que ocorre e em que estado ambiental se encontra. Se diferentes proteínas podem ser geradas, quais dos diferentes transcritos correspondem ao que deveríamos chamar de gene, questiona Keller? A autora entende que se o termo “gene” pretende se referir ao trecho original de DNA, isso equivale a abandonar a noção de que um gene faz uma proteína, pois, nesse sentido, um gene gera transcritos que podem gerar muitas proteínas. Isso significaria que o gene, além de perder o seu poder de especificidade, também perderia o seu poder de agência, pois a prerrogativa de escolher que proteínas devem ser feitas e sob quais circunstâncias não mais residiria no gene, e sim na complexa dinâmica reguladora da própria célula como um todo. O novo paradigma, que como afirma Strohman (1997) parece ainda não existir, terá a incumbência de desemaranhar a estrutura desses trajetos metabólicos.

³³ Um complexo de proteínas e SnRNAs, *small nuclear RNAs*, uma classe de ncRNAs (Mattick, 2003; Eddy, 2001).

Além do processamento alternativo, outros mecanismos de edição do RNA confundiram as perspectivas de uma relação simples entre genes e proteínas. Todos esses eventos são dependentes de um aparato celular específico, algo mais do que apenas a simples existência de fitas de DNA (Keller, 2000). Não seria mais possível conciliar a estrutura do gene, na medida em que não existiria no DNA qualquer sinal que o identifique como um gene particular, com uma função, pois ela passa a ser entendida não em termos de seqüência no DNA, mas em termos dinâmicos, através de um processo que envolve o contexto genético e o contexto citoplasmático. De acordo com Keller,

(...) a fonte da estabilidade genética não se encontraria na estrutura de uma entidade estática, mas que a própria estabilidade é o produto de um processo dinâmico. Aqui, aprendemos que a função do gene também precisa ser entendida em termos dinâmicos. Porque a atividade biológica é inerente à atividade de proteínas, e não à de genes, a quebra da hipótese um gene – uma proteína enfraquece criticamente a possibilidade de se atribuir função à unidade estrutural que foi tradicionalmente tomada como sendo o gene. Mesmo reconcebido como uma unidade funcional (por exemplo, a seqüência emendada e editada do RNAm), o gene não pode mais estar acima e afastado dos processos que especificam a organização celular e intercelular. (Keller, 2002, p. 83)

Mas tal situação necessariamente não é verdade: com a bioinformática, pode-se prever a existência de um gene sem nunca ter se manipulado uma pipeta ou entrado num laboratório de pesquisa experimental. Não podemos é exatamente prever o produto gênico. Pois a biologia não é uma ciência de previsão exata. Todavia, para Keller, o gene se tornou muita coisa³⁴; não mais uma única entidade, uma palavra de grande plasticidade definida somente pelo contexto experimental específico no qual é utilizada. Nesse ponto, Keller parece concordar com a abordagem de Raphael Falk, que veremos agora. Mas veremos também como as abordagens são distintas.

³⁴ Essa idéia parece remontar um pouco a definição de gene de Johannsen. Aquilo que gera, que dá origem. Nesse sentido, o que gera pode estar distribuído pela célula e não especificamente em seqüências do DNA. O “que gera” estaria disperso e não concentrado em estruturais materiais. Johannsen advertira aos outros pesquisadores que uma concepção do gene como uma estrutura material e morfológicamente caracterizada seria muito perigosa para o avanço seguro da genética (Mayr, 1998).

Raphael Falk e os genes desorientados

Raphael Falk é geneticista e professor emérito da Universidade Hebraica de Jerusalém, em Israel, cuja linha de pesquisa atual abrange história e filosofia da biologia, especialmente no que tange o significado dos conceitos em genética. Falk argumenta que há pelo menos dezenove anos os genes estão “desorientados” (1986). Essa é uma tradução nossa para o termo “the bewildering gene” utilizado por Falk (o gene desconcertante, desorientado, confuso, perplexo). Assim como Ridley em 2003, mas dezessete anos antes, Falk listou sete situações históricas envolvendo controvérsias acerca da definição de gene. Segundo a cronologia de Falk, antes dos “bewildering genes” veio “the DNA gene”: O gene de Watson e Crick, ou o “gene molecular clássico”. A desorientação começa em 1968, portanto, no mesmo ano em que Gunther Stent afirmava que o trabalho restante era o de “polir os detalhes”: com a descoberta de Britten e Kohne (1968) de que seqüências repetitivas do DNA e, portanto, seqüências não codificantes de proteínas, poderiam ter estar envolvidas na regulação dos genes. Entre uma série de outros achados, a descoberta, em 1977, dos íntrons culminaria o processo de desorientação dos genes durante esse período.

Os “genes desorientados” seguem-se aos “genes de DNA”. Essa não é uma constatação sem importância, mas o cerne da questão de Falk (1986). Durante todo o trabalho de Falk seguem-se inúmeras dicotomias propostas pelo próprio autor como, por exemplo: construtos hipotéticos (conceitos hipotéticos)/ variáveis intervenientes (conceitos abstratos)³⁵, reducionismo/ holismo, contextualizado/descontextualizado, gene molecular/gene evolutivo, entre outros (<http://www.pitt.edu/~kstotz/genes/minutes.html>). No entanto, para responder a pergunta “what is a gene?”, acreditamos que a questão crucial é a distinção entre genótipo e fenótipo. Essa distinção poderia ser interpretada como uma espécie de analogia entre o gene ontológico em sua expressão fenotípica e o gene epistemológico em sua suposta localização genotípica.

³⁵ Falk retira dois conceitos oriundos da psicologia que para nós, e também para os pesquisadores da minuta de 2003 anteriormente referida, não ficam muito claros. É a distinção entre construtos hipotéticos (conceitos hipotéticos) e variáveis intervenientes (conceitos abstratos). Se algum leitor estiver disposto a ajudar, segundo Falk (1986), esses conceitos seriam úteis para transmitir as diferenças de significado e os propósitos (explícitos ou implícitos) desses “genes” (Falk, 1986). Entende-se por variáveis intervenientes, conceitos abstratos ou conceitos absolutos, como uma manipulação dos valores de variáveis empíricas que não envolvem hipóteses de existência de entidades observáveis ou a ocorrência de processos não observados (Falk, 1986). Nesse sentido, hipotético é o que não é concreto, que opera unicamente com idéias, com associação de idéias, e não diretamente com a realidade sensível. Desvia-se do conceito as determinações empíricas ou acidentais, e assim, o conceito de gene pode ser entendido como uma idéia abstrata quando concebida em separado dos objetos empíricos que apresentam, ou podem acidentalmente se apresentar, como, por exemplo, um cromossomo, uma estrutura de DNA. Esta realidade só existiria no âmbito da subjetividade humana, sem existência material. Por si só o conceito de gene não estaria submetido às condições limitativas do sujeito que os opera ou manipula. Construtos hipotéticos ou conceitos hipotéticos não seriam definidos explicitamente por suas implicações ou relações empíricas (Falk, 1986), no entanto, precedem as suposições que orientam a investigação e podem antecipar características prováveis do objeto investigado. Eles são válidos na medida em que tanto podem confirmar essas características, quanto indicar novos caminhos de investigação. Provisoriamente ou não, se esses conceitos são aceitos como explicações dos fatos, isso deve ser ulteriormente testado de forma empírica. Apenas vale lembrar, que mesmo dentro da psicologia há controvérsias em relação à distinção desses dois conceitos (MacCorquodale & Meehl, 1948).

A primeira situação histórica colocada por Falk, a distinção entre genótipo e fenótipo, equivale ao primeiro significado de gene de Ridley (2003) mencionado no início da introdução desse trabalho: o gene como um arquivo hereditário mendeliano para um determinado traço físico ou comportamental. Como mencionamos no primeiro capítulo da dissertação, o termo “gene” criado por Johanssen pretendia distinguir o que é visível no organismo (o caráter, o fenótipo) e a coisa na célula que produz o caráter (o gene, o genótipo). Desde então, como sabemos, o trabalho da genética tem sido tentar unir o que falta entre o gene e as características hereditárias, entre uma entidade que existiria no interior das células e sua expressão em termos de traços físicos como a cor da pele, a dureza das garras, o número dedos, etc. Até os dias de hoje essa questão ainda não foi completamente resolvida. Para alguns autores, como Keller, há uma brecha entre o gene e o traço. Ao nosso ver, entre o genótipo (o conjunto de todos os genes) e o fenótipo como um todo, há um verdadeiro abismo, apesar de todos os progressos das ciências que estudam o desenvolvimento, a embriologia.

O que se segue então, é um verdadeiro conflito, como Keller argumentou, entre a estrutura específica do gene e sua determinada função. A tentativa de conciliação entre a estrutura do gene e sua função tem levado a desdobramentos sucessivos na história da genética. Falk chamou essa situação de processo das “bonecas russas”, numa analogia com o tradicional brinquedo em que uma boneca se encaixa dentro da outra. Abrindo uma, encontra-se outra que, por sua vez, possui uma outra dentro e, assim sucessivamente. Tentativas de descobrir a função do gene tem levado a uma análise mais profunda de sua estrutura, que por sua vez leva a uma análise mais atenta de sua função e, assim sucessivamente.

Não devemos esquecer que Mendel reduziu ao mínimo sua especulação sobre a natureza do material genético (Mayr, 1998; Canguilhem, 1977; Dobzhansky, 1967; Fisher, 1966; Wright, 1966). Do mesmo modo, Johanssen não se preocupou em especular qual seria a base material do gene. Pelo contrário, como já mencionamos seu receio era que recaísse sobre o seu conceito de gene algum tipo de pensamento pré-formationista³⁶ (Mayr, 1998). O gene inicia sua vida como o que Falk denominou de uma “variável interveniente”, um “conceito abstrato”. Segundo o autor, não importa sua estrutura física, e sim que responda aos dados formulados pela teoria mendeliana de hereditariedade (função).

As pesquisas caminharam para elucidar que estrutura da célula poderia responder pela transmissão das características hereditárias. A resposta para essa pergunta, veio com o desenvolvimento da “teoria cromossômica da hereditariedade”, segundo a qual, os genes estariam arranjados linearmente nos cromossomos (estrutura) (Mayr, 1998). Note que, de certa forma, parece haver uma contradição. Afirmamos anteriormente que, para Morgan, o desenvolvimento da “teoria cromossômica da herança” não precisava especular em quê consiste materialmente o gene. No entanto, a história da genética, em

³⁶ Em meados do século XVII, os pré-formationistas postulavam que todos os organismos preexistem nas linhagens germinais de seus pais. Ou seja, dentro de cada indivíduo, homem ou baleia, já estaria pré-formada sua prole e assim por diante, exatamente do mesmo modo que bonecas russas. A mais conhecida expressão do pensamento pré-formationista é representada pela figura do homúnculo, um ser em miniatura que já existiria pré-formado em uma das células germinais, a saber, no espermatozóide (Pinto-Corrêa, 1999; Roe, 1983). Posteriormente, surgiram os epigeneticistas como escola rival aos pré-formationistas. Para os defensores da epigênese, o desenvolvimento do embrião é fruto de uma elaboração progressiva. A estrutura de um ser vivo se organiza pouco a pouco por uma série de dobras através de uma seqüência, no tempo e no espaço, sujeitas a operações mecânicas devido à interação com as forças físicas do ambiente (Jacob, 1983).

sua insondável complexidade, apresenta vários caminhos que se entrelaçam. Apenas estamos traçando um desenvolvimento simultâneo e paralelo às pesquisas de Morgan. Esse desenvolvimento levou à necessidade de perguntar, novamente, qual a função dos genes – como estruturas alinhadas nos cromossomos. Em 1940, veio a resposta com Beadle e Tatum. O gene responde pela especificação de uma proteína no metabolismo da célula (função).

Mas que estrutura seria responsável por armazenar a informação hereditária das proteínas de nossos organismos e as de incontáveis outras espécies de seres vivos? Em 1953, é descoberta a estrutura do DNA. Esta levou a um aprofundamento das pesquisas sobre a organização do material genético nos níveis físicos e químicos. Verificou-se então, que uma simples seqüência do DNA não responde pela codificação, pura e simplesmente, de proteínas. A função do gene em regular a atividade biológica é dependente de um sistema bem mais complexo e dinâmico que as seqüências de DNA não poderiam por si só exercer (Keller, 2000). Logo, novamente essa estrutura não é responsável por todas suas funções. Segundo Keller, a chave do processo – a função do gene – deveria ser procurada em outro lugar, não apenas no DNA.

Para Keller (2000), os problemas para se conceituar o gene vêm da dificuldade em conciliar sob uma mesma entidade a responsabilidade pela estrutura e função dos genes. Dada a função do gene, não podemos antever imediatamente sua estrutura. Ao mesmo tempo, a seqüência do DNA não é suficiente para responder por qual função ela é responsável. Enquanto Keller vê nisso um empecilho para o gene como um conceito explicativo, Falk tem uma visão oposta. Para Falk, o gene estaria apenas voltando às suas origens: o gene mendeliano. O gene nada mais é que um conceito instrumental/operacional para os biólogos. Aí reside sua importância, útil para a organização dos dados fornecidos pela pesquisas genéticas. Ontologicamente o gene talvez não possa ser conhecido nunca, como tanto outros objetos, mas epistemologicamente ele tem um valor inestimável. O que liga os genótipos aos fenótipos é a teoria genética e não uma garantida materialidade dos genes.

Aí parece residir o poder do gene de Watson e Crick, o “gene molecular clássico”. Pela primeira vez pareceu possível conciliar a estrutura do gene com uma função inequívoca. Deste então, uma espécie de zelo toma conta desse conceito. Gostaríamos de deixar claro que essa é uma opinião nossa e não de Falk. Contudo, segundo a minuta anteriormente citada, outro trabalho de Falk, publicado em 2000, descreve a tensão fundamental entre duas diferentes perspectivas de gene: a biológica e a molecular. O gene como uma entidade que corresponde a um certo traço físico ou comportamental, codificado em uma ou várias seqüências de DNA (uma função que esteja entalhada nas seqüências de DNA (biológica) e o gene que automaticamente e universalmente é identificado como uma seqüência particular de DNA, molecular.

(<http://www.pitt.edu/~kstotz/genes/minutes.html>). Falk vê de forma otimista e positiva as tensões que existem entre os diferentes conceitos de gene, porque embora elas sejam aparentemente incomensuráveis, a comunicação entre diferentes áreas das ciências biológicas pode ser possível. Permanece, assim, a possibilidade de se estabelecer essa ponte entre o epistemológico e o ontológico. A existência de diferentes conceitos não é um problema, sendo mesmo desejável (Griffiths & Neumann-Held; Falk, 1986).

Para Falk (1986), o conceito de gene evoluiu, principalmente, em dois níveis teóricos. Um que é baseado numa entidade reducionista experimental, sendo o gene mendeliano o exemplo clássico. E no outro, um modelo realista, em que os genes se baseiam numa

entidade física, como na estrutura físico-química da molécula de DNA. Ao examinar as causas e conseqüências da evolução do conceito de gene, Falk (1986) conjectura que com base no atual desenvolvimento da genética molecular o gene nada mais é do que um dispositivo intelectual bastante valioso na organização de dados. Nesse sentido, a resposta para “o que é um gene?”, parece que teria sido encontrada em 1912. “How far may we carry this conceptual notation? My answer is: just as far as the notation interprets the facts of breeding and is helpful . . .” (East *apud* Falk, 1986, p. 145).

Contudo, esses diferentes modos de interpretar os mesmos dados, não parecem ser uma novidade na história do conceito de gene. Segundo Falk, “Every experimental observation which was presented as evidence for the existence of genes, whether as hypothetical constructs or particulate material entities, was interpreted differently, and sometimes even more elegantly.” (Falk, 1986, p. 155).

Dessa forma, para os historiadores da biologia moderna, o conceito de gene nunca foi nem unívoco, nem estável. Contudo, para Keller (2000), as dificuldades do passado sempre puderam ser contidas, “o que é diferente hoje é que o progresso na biologia molecular tornou possível a quebra desse silêncio histórico” (Keller, 2002, p. 82). Mas como a biologia molecular quebrou esse silêncio histórico? Que silêncio histórico é esse? Qual distinção fundamental que faz com que as dificuldades do passado possam ser contidas e as de hoje não? Isso Keller não explica.

Para Keller, o conceito de gene seria um conceito em risco, à beira de um colapso. A confiança da realidade física do gene sempre foi acompanhada pela suposição de que sua composição material (estrutura) e sua função eram todas propriedades de um único e mesmo fenômeno – as seqüências de bases no DNA responsáveis pela codificação de uma proteína. Hoje, contudo, esta auto-identidade se encontraria desfeita. Quereria Keller dizer com isso que as seqüências de base no DNA não codificam mais proteínas em hipótese alguma? A função do gene não poderia mais ser compreendida pela sua estrutura. Sua existência é transitória, contingente, dependendo crucialmente da dinâmica funcional do organismo inteiro. (Keller, 2000).

Antes de chegar a seus limites, o léxico do gene teve primeiro que ser construído, e a história da genética nos primeiros três quartos deste século oferece um testemunho eloqüente da versatilidade e da força deste conceito central. A evidência acumulada no último quarto do século, porém, dá um testemunho diferente: mostra-nos que, mesmo em sua juventude, a “palavrinha” de Johannsen, tão inocentemente concebida no início do século, teve de suportar uma carga digna de Hércules. Uma única entidade foi tomada como garantia da estabilidade intergeracional, como fator responsável pelas características individuais, e, ao mesmo tempo, como agente diretor do desenvolvimento do organismo. Na verdade, pode-se dizer que a carga não parecia tão pesada – ao menos enquanto o gene era visto como uma entidade quase mítica. Mas, pela metade do século, o gene já tinha passado a ser reconhecido como uma molécula no sentido físico real – de fato, era apenas um pedacinho de DNA – e aqui, neste ponto do tempo, a história da genética faz a sua mudança mais surpreendente. Tanto o triunfo como a excitação da nova ciência da biologia molecular derivou da evidência notável que ela fornecia sugerindo que, por incrível que parecesse, o gene, agora compreendido como uma molécula de DNA auto-replicadora, era uma estrutura à altura de sua tarefa. Mas, com o amadurecimento da biologia molecular, a impraticabilidade (talvez a impossibilidade) dessa carga tem se tornado cada vez mais fácil de perceber. (Keller, 2002, p. 162).

Para contrabalançar a afirmação de Keller (que o Projeto Genoma Humano minou nossas expectativas sobre a concepção de gene ao evidenciar um funcionamento muito mais dinâmico da célula), observemos as passagens de François Jacob, originalmente escritas em 1970, antes da revolução na biologia molecular, sobretudo as descobertas do DNA recombinante e dos genes interrompidos.

Entre os componentes dos seres vivos, o material genético desempenha um papel privilegiado. Ele ocupa o pico da pirâmide e decide a respeito das qualidades dos organismos. Os outros componentes são encarregados da execução. Mas sem o citoplasma que o envolve, o núcleo nada pode fazer. É a célula inteira que constitui a unidade elementar do vivo, que detém suas propriedades, que assimila, cresce e se reproduz. O gene representa o elemento último da análise genética, mas não tem autonomia alguma. Sua expressão depende quase sempre dos outros genes que o circundam. É o material genético inteiro, é a combinação específica dos genes realizada em um organismo que determina seu desenvolvimento, sua forma e suas propriedades. (Jacob, 1983, p. 229).

Só há integração na medida em que a rigidez das instruções prescritas no programa genético é corrigida pela flexibilidade das informações recolhidas sobre a situação local, sobre o estado do sistema, sobre a natureza do meio. (...) Isto exige que aos órgãos de execução estejam atrelados, para dirigi-los, órgãos de percepção capazes de “sondar” o mundo exterior, de descobrir a presença de certos compostos que agem como sinais, de medir sua concentração. (...) É pela presença destas estruturas que se estabelece uma rede de comunicação entre célula e meio, entre genes e citoplasma, entre pares de componentes sem afinidade química. (Jacob, 1983b, pp. 283-284).

Vale mencionar que Keller escreveu, em 1983, *A Feeling for the Organism: The Life and Work of Barbara McClintock*. (Oxford, W.H. Freeman and Company). Também em 1983, Barbara McClintock foi laureada com o Nobel de medicina ou fisiologia pela descoberta, em 1944, dos “jumping genes”, algo como “genes saltitantes” ou “genes saltadores” (McClintock, 1950). Essa descoberta demonstrou que certas regiões do material genético poderiam se locomover dentro do genoma, saltar de um ponto do genoma para o outro. Note que em 1944, a estrutura físico-química da molécula de DNA ainda não havia sido descoberta. Devido a mecanismos desse tipo, cerca de 50% do nosso genoma é composto por material genético “invasor” de vírus e bactérias. Dessa forma, o genoma não seria exatamente estável. Hoje diversos tipos de elementos móveis no genoma são conhecidos e chamados, de um modo geral, de transposons, *transposable elements* (Prak & Kazazian, 2000). Embora o trabalho de McClintock tenha sido bastante conhecido, em grande parte foi ignorada a sua substancial contribuição para o corpo do conhecimento genético durante quase 50 anos. Para Raphael Falk (1986), essa recusa em assimilar o trabalho de McClintock durante as décadas de 1940 e 1950 foi um ato de defesa da comunidade genética. Naquele momento, a base material das entidades genéticas estava sendo finalmente estabelecida, como seqüências específicas e discretas de DNA. Os “genes saltitantes” promovem a

idéia de instabilidade dos genes tanto em sua localização quanto em sua função. Essa idéia destrutiva somente agora, na era anárquica da formulação instrumental das entidades genéticas e da heterodoxia dos genes, pode ser entusiasticamente admitida. Perguntamos se será que foi preciso os dados do Projeto Genoma para que Keller (2000) percebesse a importância das idéias de sua biografada?

Para Onde se Dirige a Atenção de Keller?

Se para Jacob (1983) o “material genético desempenha um papel privilegiado”, para Keller (2000) o DNA é apenas mais um, entre tantos outros componentes da célula, responsável pelas funções do organismo. O DNA é apenas “a matéria prima crucial e absolutamente indispensável, mas não mais que isso.” (Keller, 2002, p. 84). O objetivo de Keller parece ser desmerecer a importância do DNA, para que outros componentes celulares e epigenéticos sejam enaltecidos. No entanto, Keller não deixa claro como pode ser empiricamente testado esse processo dinâmico, ou o que exatamente é esse processo dinâmico.

Embora Keller em nenhum momento mencione, suas atenções parecem se dirigir às reflexões teóricas relacionadas à “Teoria de Desenvolvimento de Sistemas” (DST – *Developmental Systems Theory*). A DST sustenta que o desenvolvimento dos organismos não é o desdobramento de um “programa genético”, a execução de um programa pré-existente localizado nos genes ou em qualquer outro lugar (Lickliter & Honeycutt, 2003). Embora a noção de que os fenótipos surjam a partir da interação entre os genes e o ambiente – lembremos de uma velha fórmula empregada no ensino de biologia: genótipo + meio = fenótipo –, para os teóricos da DST isto significa que é impossível, *a priori*, e frequentemente suficientemente impossível *a posteriori*, especificar uma primazia causal tanto em relação aos genes quanto em relação ao ambiente, no que diz respeito ao resultado de um fenótipo específico, e, nesse sentido, a hierarquia usual entre causas externas e internas não se sustenta.

Para os teóricos da DST, não somente os organismos devem ser incluídos numa análise realista, mas também as feições ambientais extra-organísmicas. Segundo eles, deve ser privilegiada a noção de interação. Os sistemas analisados pelos teóricos da DST compreendem um conjunto não especificado que inclui todas as influências e entidades, em todos os níveis de análises, incluindo a molecular, a celular, a organísmica, a ecológica e a biogeográfica (Lickliter & Honeycutt, 2003 e 2003b).

De acordo com os teóricos da DST, os genes são apenas mais um dos muitos recursos utilizados durante o desenvolvimento. (Lickliter & Honeycutt, 2003b). A principal proposta dos teóricos da DST é integrar o pensamento evolutivo com as ciências psicológicas, a chamada psicologia evolutiva, se opondo às perspectivas centradas no gene para compreender o comportamento humano. Se valeria de dados advindos da genética, embriologia e biologia do desenvolvimento, entre outras áreas. Dessa forma, além da rede de interações entre genes e as complexas e multideterminadas interações moleculares dentro e fora das células, no caso do comportamento humano deve ainda ser avaliadas a natureza e a seqüência do ambiente físico, biológico e social através do qual passa o indivíduo. (Lickliter & Honeycutt, 2003).

Não é nosso objetivo entrar no debate natureza *versus* cultura, no qual estão embebidos os teóricos da DST. Nossas dúvidas, entretanto, estão em sintonia com a de outros tantos que questionam a proposta dos teóricos da DST (Bjorklund, 2003; Crawford, 2003; Krebs, 2003). Será possível oferecer uma estrutura coerente que consiga integrar fatores genéticos, ambientais e experiências individuais numa teoria do desenvolvimento do organismo e do comportamento humano? Que modelo diferente os teóricos da DST de fato apresentam?

Dessa forma, as críticas de Keller ao Projeto Genoma Humano como crucial para a desarticulação do conceito de gene parecem estar mal dirigidas. Suspeitamos que as críticas da autora ao Projeto Genoma Humano reflitam mais uma postura ideológica (política/econômica) do que propriamente científica. Para isso, possuímos apenas indícios, como as já citadas críticas ao determinismo genético: “Ao contrário de todas as expectativas, em vez de apoiar noções familiares de *determinismo genético* que adquiriram tão grande poder na *imaginação popular*, esses resultados criaram desafios críticos a essas noções clássicas.” (Keller, 2002, p. 17. Itálicos nossos.).

Essas duas posturas, como vimos antes, são as principais causas defendidas pelos teóricos da DST. Provavelmente não por acaso, um dos defensores da DST é o biólogo Richard Lewontin, crítico ferrenho do Projeto Genoma Humano e do determinismo genético em livros como: *A Tripla Hélice: Gene, Organismo e Ambiente* (2002) e *Biologia como Ideologia: A Doutrina do DNA* (2000). Segundo Lewontin (2000),

Essas três idéias – de que nos distinguimos nas habilidades fundamentais por causa das diferenças inatas, de que as diferenças inatas são biologicamente herdadas e de que a natureza humana garante a formação de uma sociedade hierárquica – quando reunidas, formam o que podemos chamar de *ideologia do determinismo biológico*. (Lewontin, 2000, p. 29)

Afirmamos novamente que não estamos defendendo nem o Projeto Genoma Humano nem a existência de algum tipo de determinismo biológico, mas apenas identificando para onde podem estar voltadas as atenções de Keller.

O que está faltando – e seria absolutamente necessário para uma compreensão da linguagem na pesquisa biológica – é uma investigação muito mais profunda do contexto material, econômico e social no qual a linguagem funciona.” (Keller, 2002, p. 156-157).

Minha esperança é que tais conceitos novos e novas maneiras de pensar logo conseguirão afrouxar o laço ainda mais apertado que, mais recentemente, os genes vieram a ter na imaginação popular. Pois se, como Gelbart sugere, o termo *gene* pode, de fato, ter-se tornado um obstáculo à compreensão dos biólogos, ele talvez tenha se tornado um obstáculo ainda maior para a compreensão dos leigos, criando equívocos, tanto quanto informando. Como conseqüência, ele dá forma às esperanças e ansiedades populares de maneiras que são muitas vezes desfocadas, e de fato contraprodutivas para a discussão efetiva de políticas públicas, mesmo onde as questões são reais e urgentes. No melhor de meu otimismo, imagino mesmo a possibilidade de que novos conceitos possam abrir terreno inovador no qual cientistas e leigos possam pensar e agir juntos para desenvolver uma postura que seja política e cientificamente realista. (Keller, 2002, p. 165).

O “dogma central da biologia molecular” incorpora a abstração apropriada para todas as questões biológicas, ou da biologia molecular, como a descoberta dos íntrons? Obviamente, a resposta é não. O “dogma central” não explica porque *Escherichia coli* sintetiza b-galactosidase na presença de lactose, mas não na presença de glucose; não explica como funciona o sistema imune; como o cérebro trabalha ou como as plantas se regeneram quando podadas. A importância do “dogma central” para a biologia molecular é como a de uma primeira aproximação a um processo extremamente complexo pode ser representado através de uma simples abstração. Um diagrama rigoroso de todas as reações, enzimas, cofatores envolvidos na expressão da informação genética não somente seria tão complexo quanto obscureceriam em parte, a elegância dessas idéias.

Um exemplo de um novo modelo desse tipo, muito mais complexo, e menos elegante, poderia ser exibido pelo esquema de Paul Silverman (2004), biólogo molecular emérito da Universidade da Califórnia, no artigo cujo título é “Rethinking Genetic Determinism: With Only 30,000 genes, What Makes Humans Human?” (Figura 7).

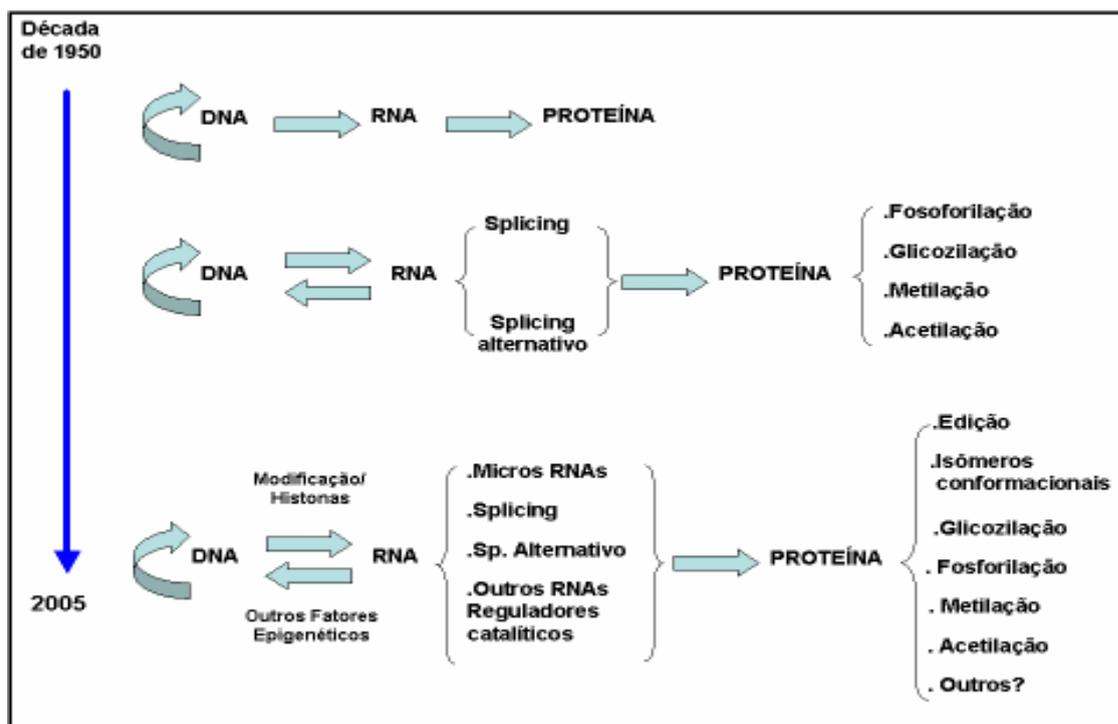


Figura 7: O dogma expandido. O “dogma central da biologia molecular” como formulado em 1957 propunha a unidirecionalidade da expressão gênica. Uma série de novos dados moleculares sofisticaram esse modelo. Transcrição reversa, modificações da proteína após esta ser traduzida, além de diversos fatores que afetam a expressão gênica em todos os estágios do processo que vai do gene no DNA até a proteína final.

Novamente podemos encontrar a afirmação de que a recente descoberta da quantidade de genes de nossa espécie, não suplantando a marca de 30.000, seria um indicativo da insuficiência de averiguar a complexidade, sobretudo humana, a partir do número de genes. Mas em que lugar estaria escrito que os biólogos moleculares e geneticistas sempre acreditaram dogmaticamente que o número de genes de uma espécie fosse uma medida de sua complexidade? Que há outros elementos envolvidos na transmissão das características hereditárias e do desenvolvimento dos traços fenotípicos, já é discutido desde muito tempo, como essa dissertação espera ter demonstrado. Obviamente, 33 genes são poucos para explicar o que nos faz humanos. Mas a cifra de algumas poucas dezenas de milhares não desmerece o lugar do conceito de gene na estrutura explicativa que a biologia vem procurando desenvolver, para tentar entender o fenômeno de estar vivo e dos diversos motivos que levam essa vida a se encerrar.

CONSIDERAÇÕES FINAIS

“Quando eu uso uma palavra’, disse Humpty Dumpty num tom bastante desdenhoso, ‘ela significa exatamente o que eu escolho que ela signifique: nem mais nem menos.’”

Carroll, Lewis. Através do Espelho.

Em agosto de 1926, o botânico da Universidade de *Harvard* E. M. East apresentou no *International Congress of Plant Sciences*, em Nova Iorque, o trabalho intitulado “The Concept of the Gene.” East (1929) acreditava que os genes eram apenas uma notação útil para os experimentos de cruzamentos de animais e plantas, sem realidade biológica. O gene era meramente um termo descritivo.

The argument was an elaboration of the proposition that the germ-cell unit of heredity, the gene, was an abstract, formless, characterless concept used for convenience in describing the results of breeding experiments. It was the ghost of an entity which might later be clothed with flesh, but its usefulness at the time was due to its adaptability to mathematical treatment. (East, 1929, p. 889).

Seu principal adversário era o geneticista W.E. Castle. Para Castle:

Before this theory could be accepted unreservedly, it has seemed desirable to know whether all observed inheritance phenomena can be expressed satisfactorily in terms of genes, which are supposed to be to heredity what atoms are to chemistry, the ultimate, indivisible units, which constitute gametes much as atoms in combination constitute compounds. (Castle *apud* Falk, 1986).

Por essa perspectiva histórica que remete à década de 1920, o conceito de gene nunca foi desprovido de controvérsias. O debate atual parece envolver duas situações:

1) A acumulação de uma série de dados a respeito do funcionamento e organização molecular dos genes, que não mais se encaixam nos parâmetros anteriormente aceitos para a definição de gene do início dos anos 1960. De acordo com Gelbart (1998), é essa definição que, fundamentalmente, ainda hoje é utilizada e que deve ser modificada para acompanhar a visão atual do genoma. O acúmulo não crítico de novos dados moleculares parece expandir os limites do conceito de gene para além da noção de gene implicitamente aceita, ou seja, a noção que indica o gene como segmentos de DNA específicos nos cromossomos responsáveis pela codificação de uma proteína. Novos dados moleculares, proporcionados por novas técnicas de investigação molecular, afetam o entendimento do que é um gene nos termos do “gene molecular clássico”.

2) O conceito de gene é apropriado com diferentes significados, em diferentes áreas da biologia, o que aumenta a confusão em torno de seu significado. Para alguns autores, isto não representa um problema e sim, algo mesmo desejável. (Griffiths & Neumann-Held, 1999; Falk, 1986). No entanto, se faz preciso que esses conceitos sejam claramente distinguidos e usados dentro de suas respectivas áreas (Griffiths e Neumann-Held, 1999), sendo isto necessário para que seja permitida a avaliação crítica dos resultados e observações experimentais e somente depois, finalmente, tentar-se uma síntese entre os conceitos diferentes de gene (Falk, 1986).

Aproximadamente a metade da vida do complexo de idéias que atende pelo nome de gene foi transcorrida sem que se identificasse a base material de sua existência, quer dizer, a estrutura físico-química do DNA. A descoberta dessa estrutura concedeu ao gene sua materialidade e a idéia de que a genética parecia ter resolvido quase todos os seus problemas. O conceito de “gene molecular clássico”, que se instaurou durante esse período, se tornou um dos conceitos mais influentes da história da genética. A descoberta dos íntrons, em 1977, é um dos elementos que contribuem para que não se possa aceitar

mais esse conceito de uma forma irrestrita. A chave de alguns princípios biológicos verdadeiramente fundamentais pode estar na resolução de como os genes são montados pelo processamento do RNA e de como as diversas espécies de RNAs se comportam na célula.

Como alguns autores postulam (Keller, 2000; Strohman, 1997), essas novas descobertas moleculares estariam demolindo o paradigma da biologia molecular contemporânea, o seu “dogma central” e, inevitavelmente, o próprio conceito de gene. Para esses autores, a imensa quantidade de trabalho em correlacionar a seqüência de DNA primária com função genética é essencial para trazer a relevância biológica completa ao fluxo de dados produzido pelos projetos de seqüenciamento de organismos. Somente um novo paradigma poderia dar cabo de resolver essa situação. Contudo, um dos principais atributos para uma revolução kuhniana, parece ainda não existir, a saber, um novo paradigma que substitua o antigo (Strohman, 1997). Mas não seria o próprio projeto proteoma a resolução para esse paradigma?

Ao nosso ver, os íntrons não podem ser visto como um dos elementos provocadores de uma mudança paradigmática. A descoberta dos íntrons seria mais um dos diversos elementos importantes surgidos recentemente que servem como uma espécie de aviso de que os genomas eucarióticos são bem mais complexos. Baleias e rinocerontes não são *Escherichia coli*. A descoberta dos íntrons, uma das conseqüências da revolução do DNA recombinante, parece representar um marco dessa passagem de organismos menos complexos para organismos notadamente mais complexos.

No centro dessa discussão está o conceito de gene. Jacob (1983) afirma que os conceitos têm a função de servir como operadores para o estudo do mundo vivo e, dessa forma, a importância de um conceito se mede pelo seu valor operatório, pelo papel que desempenha dirigindo a observação e a experiência.

Ao mesmo tempo em que se discute a materialidade dos genes, ou até mesmo a sua existência, parece no mínimo paradoxal que a inserção de genes em organismos de laboratórios nos forneçam provas suficientes de que eles de fato existem. Podemos não conhecer em detalhes sua estrutura e seu funcionamento. Mas a inserção de um gene e a conseqüente produção de uma proteína associada a ele em um construto biológico, parece confirmar o gene ontologicamente.

O conceito de gene parece ter evoluído, principalmente, em dois níveis teóricos, como uma entidade reducionista instrumental e como uma entidade material. Esses dois conceitos podem até ser competitivos, mas não são necessariamente incomensuráveis. A evolução do conceito de gene pode refletir um conflito entre diferentes interpretações dos resultados experimentais. Esse confronto levaria ao modelo de desdobramentos (bonecas russas) ditado por Raphael Falk (1986)

Hoje, o gene não é mais a unidade material ou unidade instrumental de herança, mas uma unidade, um segmento que corresponde a uma unidade de função, definido de acordo com as necessidades experimentais. Não é contínuo, pois existem os íntrons; não tem uma localização fixa, pois existem transposons; não tem uma função definida, pois existem os pseudogenes; e nem mesmo seqüências precisas, devido ao processamento alternativo. O gene se assemelha a situação dos elétrons, como um termo “trans-teórico”, ou seja, um conceito que apresenta o mesmo referente em diferentes teorias (Falk, 1986).

Enquanto a comunicação entre diferentes teorias for possível, a mudança do conceito de gene pode permitir uma síntese crítica dos diferentes conceitos. Os desenvolvimentos na genética durante esses últimos anos têm sido bastante perturbadores para o conceito de gene. Muitos assumem uma conotação estranha e até mesmo contraditória para a estrutura conceitual da teoria genética. No entanto, ela ainda não parece superada. O conceito de gene não é uma espécie de entidade sagrada para a teoria genética, mas pode ser visto

como parte do núcleo de um programa de pesquisa lakatosiano, ou até mesmo como parte de uma matriz conceitual do paradigma genético em termos kuhnianos.

Os confrontos entre as diferentes aproximações têm sido fundamentais na flexibilização necessária à acomodação dos desenvolvimentos das pesquisas experimentais. Contudo, também trazem dificuldade para evitar certos excessos veiculados, sobretudo, na imprensa não especializada, pelas teorias genéticas. Tais confrontos refletem também os encontros e desencontros entre nossos desejos científicos de conhecer mais sobre os processos hereditários e os dados empíricos oriundos das ciências genéticas.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALONI, Yosef; DHAR, Ravi; LAUB, Orgad; HOROWITZ, Mia; KHOURY, George
 “Novel mechanism for RNA maturation: The leader sequences of Simian Virus 40 mRNA are not transcribed adjacent to the coding sequences.” *Proceedings of the Natural Academy of Sciences USA*, vol. 74, pp. 3686-3600, 1977.
- ALVES-MAZZOTTI, Alda Judith & GEWANDSZNAJER, Fernando *O Método nas Ciências Naturais e Sociais: Pesquisa Quantitativa e Qualitativa*. São Paulo: Editora Pioneira. 1998.
- AVERY, O.T.; MACLEOD, C.M.; MCCARTY, M. “Studies on the chemical nature of the substance inducing transformation of *pneumococcal* types.” *J. Exp. Med*, vol. 98, pp. 137-158, 1944.
- BAKER, Bruce S. “Sex in flies: The splice of life.” *Nature*, vol. 340, pp.521-524, 1989.
- BEADLE, George W. “Genes and chemical reactions in *Neurospora*.” 1958. In: *Nobel Lectures, Physiology or Medicine 1942-1962*. Amsterdam: Elsevier Publishing Company, 1964. Disponível em: <http://nobelprize.org/medicine/laureates/1958/beadle-lecture.html>
- BERG, Paul. “Dissections and reconstructions of genes and chromosomes.” 1980. In: *Nobel Lectures, Chemistry 1971-1980*. Singapore: Editor-in-Charge Tore Frängsmyr, Editor Sture Forsén, World Scientific Publishing Co., 1993. Disponível em: <http://nobelprize.org/chemistry/laureates/1980/berg-lecture.html>
- _____; BALTIMORE, D.; BOYER, H.W.; COHEN, S.N.; DAVIS, R.W.; HOGNESS, D.S.; NATHANS, D.; ROBLIN, R.; WATSON, J.D.; WEISSMAN, S.; ZINDER, N.D. “Potentials biohazards of recombinant DNA molecules.” [Letter] *Science*, vol. 185, p. 303, 1974.
- BERGET, Susan M.; MOORE, Claire; SHARP, Phillip A. “Spliced segments at the 5’ terminus of Adenovirus 2 Late mRNA.” *Proceedings of the Natural Academy of Sciences USA*, vol. 74, pp. 3171-3175, 1977.
- BJORKLUND, David F. “Evolutionary psychology from a developmental systems perspective: Comment on Lickliter and Honeycutt (2003).” *Psychological Bulletin*, vol. 129, pp. 836-841, 2003.

- BRACK, Christine & TONEGAWA, Susumu “Variable and constant parts of the immunoglobulin light chain gene of a mouse myeloma cell are 1250 nontranslated bases apart.” *Proceedings of the Natural Academy of Sciences USA*, vol. 74, pp. 5662-5656, 1977.
- BREATHNACH, R.; MANDEL, J.L.; CHAMBON, P. “Ovalbumin gene is split in chicken DNA.” *Nature*, vol. 270, p. 314-319, 1977.
- BRETT, David; POSPISIL, Heike; VALCÁREL, Juan; BORK, Peer. “Alternative splicing and genome complexity.” *Nature Genetics*, vol. 30, pp. 29-30, 2002.
- BRITTEN, R.J. & KOHNE, D.E. “Repeated sequences in DNA.” *Science*, vol. 161, pp. 529-540, 1968.
- CANGUILHEM, G. *Ideologia e Racionalidade nas Ciências da Vida*. São Paulo: Edições 70, 1977.
- CHAMBON, Pierre. “Split genes.” *Scientific American*, vol. 244, pp. 60-71, 1981.
- CHOW, Louise T.; GELINAS, RICHARD E.; BROKER, Thomas R.; ROBERTS, Richard J. “An amazing sequence arrangement at the 5’ ends of Adenovirus 2 messenger RNA.” *Cell*, vol. 12, pp. 1-8, 1977.
- COHEN, S.N.; CHANG, A.C.; BOYER, H.W.; HELLING, R.B. “Construction of biologically functional bacterial plasmids *in vitro*.” *Proceedings of the Natural Academy of Sciences USA*, vol. 70, pp. 3240-32944, 1973.
- COLLINS, F. “Glossary of genetic terms: gene, [Website].” National Human Genome Research Institute. Available: <http://www.nhgri.nih.gov/DIR/VIP/Glossary>. 2001.
- COSTANTINI, F. & LACY, E. “Introduction of a rabbit *beta*-globin gene into the mouse germ line.” *Nature*, vol. 294, pp. 92-94, 1981.
- CRAWFORD, Charles B. “A prolegomenon for a viable evolutionary psychology – the myth and the reality: Comment on Lickliter and Honeycutt (2003).” *Psychological Bulletin*, vol. 129, pp. 854-857, 2003.
- CROFT, Larry; SCHANDORFF, Soeren; CLARK, Francis; BURRAGE, Kevin; ARCTANDER, Peter; MATTICK, John S. “ISIS, the intron information system, reveals the high frequency of alternative splicing in the human genome.” *Nature Genetics*, vol. 24, pp. 340-341, 2000.

- DAVIES, Kevin. *Decifrando o Genoma: A corrida para Desvendar o DNA Humano*. São Paulo: Companhia das Letras, 2001.
- DOBZHANSKY, Theodosius. “Looking back at Mendel’s discovery.” *Science*, vol. 156, pp. 1588-1589, 1967.
- DOWNES, Stephen M. “Alternative splicing, the gene concept, and evolution.” *History and Philosophy of the Life Sciences*, vol. 26, pp. 91-104, 2004.
- EAST, E. M. “The concept of the gene.” In: *Proceedings of the International Congress of Plant Sciences, Ithaca New York, August 16-23, 1926*, vol. 1 Menasha, WI: George Banta Publishing Co. pp. 889–895. 1929.
- EDDY, Sean R. “Non-coding RNA genes and the modern RNA world.” *Nature Reviews/Genetics*, vol. 2, pp.10-20, 2001.
- EPP, Christopher D. “Definition of a gene.” *Nature*, vol. 389, p. 537. 1997.
- EYCK, Toby A. & WILLIMENT, Melissa. “The national media and things genetic: Coverage in the New York Times (1971-2001) and the Washington Post (1977-2001).” *Science Communication*, vol. 25, pp. 129-152, 2003
- FALK, Raphael. “What is a gene?” *Studies in History and Philosophy of Science*, vol. 17, pp. 133-173, 1986.
- FISHER, R.A. “Has Mendel’s work been rediscovered?” In: Stern, C. & Sherwood, E.R. (orgs.). *The Origin of Genetics: A Mendel Source Book*. S. Francisco: W.H. Freeman & Company, 1966.
- FOGLE, Thomas. “The dissolution of protein coding genes in molecular biology.” In: BEURTON, Peter J. & FALK, Raphael & RHEINBERGER, Hans-Jörg (orgs.). *The Concept of the Gene in Development and Evolution*. Cambridge: Cambridge University Press, 2000.
- FRANKLIN, Rosalind. E. & GOSLING, R.G. “Molecular configuration in sodium thymonucleate.” *Nature*, vol. 171, pp. 740-741, 1953.
- FRIEDMAN, Meyer & FRIEDLAND, Gerald D. “Maurice Wilkins e o DNA.” In: FRIEDMAN, Meyer & FRIEDLAND, Gerald D. *As Dez Maiores Descobertas da Medicina*. São Paulo: Companhia das Letras, 2000.

- GELBART, William M. "Databases in genomic research." *Science*, vol. 282, pp. 659-661, 1998.
- GILBERT, Walter. "Why genes in pieces?" *Nature*, vol. 271, p. 501, 1978.
- GLOVER, David M. & HOGNESS, David S. "A novel arrangement of the 18s and 28s sequences in a repeating unit of *Drosophila melanogaster* rDNA." *Cell*, vol. 10, pp. 167-176, 1977.
- GRAVELEY, Brenton R. "Alternative splicing: increasing diversity in the proteomic world." *Trends in Genetics*, vol. 17, pp. 100-107, 2001.
- GRIESEMER, James. "Reproduction and the reduction of genetics." In: BEURTON, Peter J. & FALK, Raphael & RHEINBERGER, Hans Jörg (orgs.). *The Concept of the Gene in Development and Evolution*. Cambridge: Cambridge University Press, 2000.
- GRIFFITHS, Paul E. "Lost: One gene concept. Reward to finder." *Biology and Philosophy*, vol. 17, pp. 271-283, 2002.
- _____ & NEUMANN-HELD, Eva M. "The many faces of the gene." *BioScience*, vol. 49, pp. 656-662, 1999.
- HERSHEY, A.D. & CHASE, M. "Independent functions of viral protein and nucleic acid in growth of bacteriophage." *J. Gen. Physiol.*, vol. 36, pp. 39-56, 1952.
- HIETER, Philip & BOGUSKI, Mark. "Functional genomics: It's all how you read it." *Science*, vol. 278, pp. 601-602, 1997.
- JACKSON, D.A.; Symons R.H.; BERG, P. "Biochemical method for inserting new genetic information into DNA of Simian Virus 40: Circular SV 40 DNA molecules containing *Lambda* phage genes and the galactose Operon of *Escherichia coli*." *Proceedings of the Natural Academy of Sciences USA*, vol. 10, pp. 2904-2909, 1972.
- JACOB, François. *O Rato, a Mosca e o Homem*. São Paulo: Companhia das Letras, 1998.
- _____ *A Lógica da Vida: Uma História da Hereditariedade*. Rio de Janeiro: Edições Graal, 1983.
- _____ "Evolution and tinkering." *Science*, vol. 196, pp. 1161-1166, 1977.

- _____. “Genetics of the bacterial cell.” 1965. In: *Nobel Lectures, Physiology or Medicine 1963-1970*. Amsterdam: Elsevier Publishing Company, 1972. Disponível em: <http://nobelprize.org/medicine/laureates/1965/jacob-lecture.html>
- JEFFREYS, A.J. & FLAVELL, R.A. “The rabbit β -globin gene contains a large insert in the coding sequence.” *Cell*, vol. 12, p. 1097-1108, 1977.
- KELLER, Evelyn Fox. *O Século do Gene*. Belo Horizonte: Editora Crisálida, 2002
- _____. *The Century of the Gene*. Cambridge, Massachusetts, and London: Harvard University Press, 2000.
- _____. *A Feeling for the Organism: The Life and Work of Barbara McClintock*. Oxford: W.H. Freeman and Company, 1983.
- KEMP, Martin. “The Mona Lisa of modern science.” *Science*, vol. 421, pp. 416-420, 2003.
- KHORANA, Har Gobind. “Nucleic acid synthesis in the study of the genetic code.” 1968. In: *Nobel Lectures, Physiology or Medicine 1963-1970*. Amsterdam: Elsevier Publishing Company, 1972. Disponível em: <http://nobelprize.org/medicine/laureates/1968/khorana-lecture.html>
- KORNBERG, Arthur. “The biologic synthesis of deoxyribonucleic acid.” 1959. In: *Nobel Lectures, Physiology or Medicine 1942-1962*. Amsterdam: Elsevier Publishing Company, 1964. Disponível em: <http://nobelprize.org/medicine/laureates/1959/kornberg-lecture.html>
- KREBS, Dennis L. “Fictions and facts about evolutionary approaches to human behavior: Comment on Lickliter and Honeycutt (2003).” *Psychological Bulletin*, vol. 129, pp. 842-847, 2003.
- KUHN, Thomas S. *A Estrutura das Revoluções Científicas*. São Paulo: Editora Perspectiva, 1994.
- LAKATOS, Imre. “O falseamento e a metodologia dos programas de pesquisa científica.” In: _____ & MUSGRAVE, Alan. *A Crítica e o Desenvolvimento do Conhecimento*. 4º volume das atas do colóquio internacional sobre filosofia das ciências, realizado em Londres em 1965. São Paulo: Cultrix/EDUSP, 1979.
- LANDER *et al.* “Initial sequencing and analysis of the human genome.” *Nature*, vol. 409, pp. 860-921, 2001.

- LEITE, Marcelo. “Hegemonia e crise da noção de ‘gene’ nos 50 anos do DNA’.” Comunicação no 49º Congresso Nacional de Genética, Águas de Lindóia, 16 de Setembro de 2003.
- LEWIN, Benjamin. *Genes VII*. New York: Oxford University Press Inc., 2000.
- LEWONTIN, Richard C. *A Tripla Hélice: Gene, Organismo e Ambiente*. São Paulo: Companhia das Letras, 2002.
- _____. *Biologia como Ideologia*. Ribeirão Preto: FUNPEC-Editora, 2000.
- LICKLITER, Robert & HONEYCUTT, Hunter. “Developmental dynamics and contemporary evolutionary psychology: *Status quo* or irreconcilable views? Reply to Bjorklund (2003), Krebs (2003), Buss and Reeve (2003), Crawford (2003), and Tooby *et al.* (2003).” *Psychological Bulletin*, vol. 129, pp. 866-872, 2003b.
- _____. “Developmental Dynamics: Toward a biologically plausible evolutionary psychology.” *Psychological Bulletin*, vol. 129, pp. 819-835, 2003a.
- LWOFF, André Michel. “Interactions among virus, cell, and organism.” 1965. In: *Nobel Lectures, Physiology or Medicine 1963-1970*. Amsterdam: Elsevier Publishing Company, 1972. Disponível em: <http://nobelprize.org/medicine/laureates/1965/lwoff-lecture.html>
- MACCORQUODALE, Kenneth & MEEHL, Paul E. “On a distinction between hypothetical constructs and intervening variables.” *Psychological Review*, vol. 55, pp. 95-107, 1948.
- MADDOX, Brenda. “The double helix and the ‘wronged heroine’.” *Nature*, vol. 421, pp. 407-408, 2003.
- MATTICK, John S. “Challenging the dogma: The hidden layer of non-protein-coding RNAs in complex organisms.” *BioEssays*, vol.25, pp. 930-939, 2003.
- MAYNARD-SMITH, John. “Reconciling Marx and Darwin.” *Evolution*, vol. 55, 1496-1498, 2001.
- MAYR, Ernst. *O Desenvolvimento do Pensamento Biológico*. Brasília: Editora Universidade de Brasília, 1998.

- MCCLINTOCK, Barbara. “The origin and behavior of mutable *loci* in maize.” *Proceedings of the Natural Academy of Sciences USA*, vol. 36, pp. 344-355, 1950.
- MENDEL, G. “Experiments on plant hybrids.” 1865. In: Stern, C. & Sherwood, E.R. *The Origin of Genetics: A Mendel Source Book*. S. Francisco: W.H. Freeman & Company, 1966.
- MESELSON, M. & YUAN, R. “DNA restriction enzyme from *E. coli*.” *Nature*, vol. 217, pp. 1110-1114, 1968.
- MIKLOS, George L. Gabor & RUBIN, Gerald M. “The role of the genome project in determining gene function: Insights from model organisms.” *Cell*, vol. 86, pp. 521-529, 1996.
- MODREK, Barmak & LEE, Christopher. “A genomic view of alternative splicing.” *Nature Genetics*, vol. 30, pp. 13-19, 2002.
- MORGAN, Thomas Hunt. “The relation of genetics to physiology and medicine.” 1934. In: *Nobel Lectures, Physiology or Medicine 1922-1941*. Amsterdam: Elsevier Publishing Company, 1965. Disponível em: <http://nobelprize.org/medicine/laureates/1933/morgan-lecture.html>
- MORROW, J.F.; COHEN, S.N.; CHANG, A.C.; BOYER, H.W.; GOODMAN, H.M.; HELLING, R.B. “Replication and transcription of eukariotic DNA in *Escherichia coli*.” *Proceedings of the Natural Academy of Sciences USA*, vol. 71, 1743-1747, 1974.
- MOSS, Lenny. “Deconstructing the gene and reconstructing molecular developmental systems.” In: *Cycles of Contingency Developmental Systems and Evolution*. OYAMA, Susan; GRIFFITHS, Paul E.; GRAY, Russel D. (orgs.). Massachusetts: MIT Press, 2001.
- _____ “One, two (too?), many genes?” Revisão de: *The Concept of the Gene in Development and Evolution*. BEURTON, P.; FALK, R.; RHEINBERGER, H.J. (orgs.). Cambridge: Cambridge University Press, 2000. Para: *Quarterly Review of Biology*, (No Prelo). Disponível em: <http://www.nd.edu/~ndphilofaculty/lmo.htm>
- MURPHY, Michael P. & O’NEILL, Luke A.J. “O que é vida?” Uma introdução sobre os próximos 50 anos.” In: “*O que é vida?*” 50 anos depois: *Especulações sobre o*

- futuro da biologia.* _____ (orgs.). São Paulo: Fundação Editora da UNESP/Cambridge University Press, 1997.
- NEEL, James V. "Biomedical progenitor." *Science*, vol. 263, pp. 697-698, 1994.
- NIRENBERG, Marshall. "The genetic code." 1968. In: *Nobel Lectures, Physiology or Medicine 1963-1970*. Amsterdam: Elsevier Publishing Company, 1972. Disponível em: <http://nobelprize.org/medicine/laureates/1968/nirenberg-lecture.html>
- OCHOA, Severo. "Enzymatic synthesis of ribonucleic acid." 1959. In: *Nobel Lectures, Physiology or Medicine 1942-1962*. Amsterdam: Elsevier Publishing Company, 1964. Disponível em: <http://nobelprize.org/medicine/laureates/1959/choa-lecture.html>
- OLBY, Robert. "The protein version of the central dogma: A crisis for biologists." *Genetics*, vol. 79, pp. 3-14, 1975.
- OSAWA, Syozo; JUKES, Thomas H.; WATANABE, Kimitsuna; MUTO, Akira "Recent evidence for evolution of the genetic code." *Microbiological Reviews*, vol. 56, pp. 229-264, 1992.
- PINTO-CORREIA, Clara. *O Ovário de Eva: A Origem da Vida*. Rio de Janeiro: Editora Campus, 1999.
- PRAK, Eline.T.Luning & KAZAZIAN Jr., Haig.H. "Mobile elements and the human genome." *Nature Reviews/Genetics*, vol. 1, pp. 134-144, 2000.
- REICHARD, P. "Presentation speech: The nobel prize physiology or medicine 1968." 1968. In: *Lex Prix Nobel en 1968*. Stockholm: Editor Wilhelm Odelberg, [Nobel Foundation], 1969.
- REZNIK, Tânia. *O desenvolvimento do conceito de gene e a sua apropriação nos livros didáticos de biologia*. Tese, Mestrado em Educação/UFF/Niterói, 182 pp., 1995.
- RIDDIHOUGH, Guy. "The other RNA world." *Science*. vol. 296, p. 1259, 2002.
- RIDLEY, Matt. *Nature via Nurture: Genes, Experience, and What Makes us Human*. New York: HarperCollins Publishers Inc., 2003.
- ROBERTS, Richard J. "An amazing distortion in DNA induced by a methyltransferase." 1993. In: *Nobel Lectures, Physiology or Medicine 1991-1995*. Editor Nils Ringertz,

- World Scientific Publishing Co., 1997. Disponível em: <http://nobelprize.org/medicine/laureates/1993/roberts-lecture.html>
- ROE, S.A. “John Turbeville Needham and the generation of living organisms.” *Isis*, vol. 74, pp. 159-184, 1983.
- RUBIN, G.M. & SPRADLING, A.C. “Genetic transformation of *Drosophila* with transposable element vector.” *Science*, vol. 218, pp. 348-353, 1982.
- SAMBROOK, J. “Adenovirus amazes at Cold Spring Harbor.” *Nature*, vol. 268, p. 101-104, 1977.
- SCHMUCKER, Dietmar; CLEMENS, James C.; SHU, Huidy; WORBY, Carolyn A.; XIAO, Jian; MUDA, Marco; DIXON, Jack E.; ZIPURSKY S. Lawrence “*Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity.” *Cell*, vol. 101, p. 671-684, 2000.
- SERAFINI, Tito. “Finding a partner in a crowd: Neuronal diversity and synaptogenesis.” *Cell*, vol. 98, pp. 133-136, 1999.
- SHAPIRO, James A. “Views about evolution are evolving.” *ASM News*. volume 65, pp. 201-207, 1999.
- SHARP, Phillip A. “Split genes and RNA splicing.” 1993. In: *Nobel Lectures, Physiology or Medicine 1991-1995*. Editor Nils Ringertz, World Scientific Publishing Co., 1997. Disponível em: <http://nobelprize.org/medicine/laureates/1993/sharp-lecture.html>
- SILVERMAN, Paul H. “Rethinking genetic determinism: With only 30,000 genes, what is it that makes humans human?” *The Scientist*, vol. 18, pp.32-35, 2004. Disponível em: http://www.the-scientist.com/yr2004/may/research3_040524.html
- SMITH, H.O. & WILCOX, K.W. “A restriction enzyme from *Hemophilus influenzae*. Purification and general properties.” *Journal of Molecular Biology*, vol. 51, 379-391, 1970.
- STENT, Gunther S. “That was the molecular biology that was.” *Science*, vol. 160, pp. 390-395, 1968.

- STERN, Curt. "Foreword." In: STERN, Curt & SHERWOOD, Eva R. (orgs.). *The Origin of Genetics: A Mendel Source Book*. S. Francisco: W.H. Freeman & Company, pp. v-xii, 1966.
- STORZ, Gisela. "An expanding universe of noncoding RNAs." *Science*, vol. 296, pp. 1260-1263, 2002.
- STRAUSS, Bernard S. "Molecular pathologies." *Science*, vol. 270, pp. 1511-1512, 1995
- STROHMAN, Richard C. "The coming kuhnian revolution in biology." *Nature Biotechnology*, vol. 15, pp. 194-200, 1997.
- VENTER, Craig J.; *et al.* "The sequence of human genome." *Science*, vol. 291, pp. 1304-1351, 2001.
- WANG, Xiaozhong; SU, Hong; BRADLEY, Allan. "Molecular mechanisms governing Pcdh-? gene expression: Evidence for a multiple promoter and *cis*-alternative splicing model." *Genes and Development*, vol. 16, pp.1890-1905, 2002.
- WATSON, James Dewey. *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*. Touchstone Books, 2001.
- WATSON, James Dewey & CRICK, Francis Harry Compton "Genetical implications of the structure of deoxyribonucleic acid." *Nature*, vol. 171, pp. 964-967, 1953b.
- WATSON, James Dewey & CRICK, Francis Harry Compton "Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid." *Nature*, vol. 171, pp. 737-738, 1953a.
- WILLIAMSON, Bob. "DNA Insertions and gene structure." *Nature*, vol. 270, pp. 295-297, 1977.
- WILKINS, Adam S. "Are there 'kuhnian' revolutions in biology?" *BioEssays*, vol. 18, pp. 695-696, 1996.
- WILKINS, Maurice Hugh Frederick; STOKES, A.R.; WILSON, H.R. "Molecular structure of deoxypentose nucleic acids." *Nature*, vol. 171, pp. 738-740, 1953.
- WILMUT, Ian; CAMPBELL, Keith; TUDGE, Colin. *Dolly, A Segunda Criação*. Rio de Janeiro: Editora Objetiva, 2000.

- WRIGHT, S. "Mendel's ratio." In: Stern, C. & Sherwood, E.R. (orgs.). *The Origin of Genetics: A Mendel Source Book*. S. Francisco: W.H. Freeman & Company, 1966.
- WU, Qiang & MANIATIS, Tom. "A striking organization of a large family of human neural cadherin-like cell adhesion genes." *Cell*, vol. 97, pp. 779-790, 1999.
- ZWEIGER, Gary. "Human genome history." In: _____ *Transducing the Genome: Information, Anarchy, and Revolution in the BioMedical Sciences*. McGraw-Hill Companies, 2001.