



# Concurso Público Fiocruz 2023

## Pesquisador em Saúde Pública

### Prova Discursiva

#### PE41

## Biologia celular e molecular da interação vetor/patógeno/ hospedeiro

### Espelho de Resposta

**Pontuação de cada Questão Discursiva conforme Anexo II do Edital nº 3, de acordo com a Unidade detentora da vaga.**

Espera-se que o candidato, no desenvolvimento do tema, tenha feito considerações técnicas adequadas sobre os seguintes pontos:

#### Questão 01

1. O DNA é formado por nucleotídeos que são compostos por uma das 4 bases nitrogenadas adenina (A), guanina (G), citosina (C), ou timina (T), assim como pelo açúcar desoxirribose (uma pentose) e um ácido fosfórico. A e G são purinas (dois anéis aromáticos) e C e T são pirimidinas (um anel armático). Os nucleotídeos são unidos em cadeia por uma ligação forfodiéster entre o carbono 5' da desoxirribose de um nucleotídeo e o carbono 3' da desoxirribose do nucleotídeo anterior. O DNA é formado por duas cadeias de nucleotídeos que são unidas através de pontes de hidrogênio (interações fracas) na sua porção central. A citosina de uma cadeia se pareia com a guanina de outra cadeia por três pontes de hidrogênio, assim como a timina se pareia com a adenina por duas pontes de hidrogênio inter-cadeias. Esses dois filamentos formam uma hélice (a dupla-fita de DNA) com voltas para a direita e apresenta duas periodicidades. Uma delas reflete o empilhamento entre dois pares de bases consecutivos, que resulta em uma distância de 0,34 nm (ou 3,4 angstroms). A outra é consequência do empilhamento de 10 pares de bases, ou uma volta da hélice, que resulta em uma distância de 3,4 nm (ou 34 angstroms). O diâmetro da dupla hélice é de aproximadamente 2 nm. O empilhamento dos pares de bases e os ângulos da ligação química entre o açúcar e a base nitrogenada origina uma hélice com duas fendas, a maior e a menor. As cadeias de DNA possuem direção, sendo que no início existe um grupo fosfato unido ao carbono 5' da desoxirribose e no final uma hidroxila livre, por onde a cadeia cresce. Assim, o sentido da fita de DNA é 5' → 3'. As duas fitas de unem de modo antiparalelo para formar a hélice, sendo que o 5' de uma fita está pareado ao 3' da outra. Nas células, a dupla hélice do DNA é torcida, sendo que existem enzimas específicas para a atividade de adicionar ou retirar os giros ou torções no DNA.

2. A replicação do DNA se dá de forma semiconservativa, ou seja, cada uma das fitas mães contribuirá para a produção de duas fitas filhas e as duplas hélices de DNA resultantes da replicação são híbridas de fitas velhas e novas. Assim, a molécula de DNA materno é transmitida de forma fidedigna para cada uma das novas células filhas durante a divisão celular. A informação genética de cada uma das células filhas é idêntica à da célula mãe, salvo nos processos de mutação que podem ocorrer durante a replicação do DNA, como os erros de incorporação de bases na fita nascente, que são muito raros. A replicação é feita por um conjunto

de proteínas e enzimas, sendo a enzima que produz as novas fitas denominada de DNA polimerase. Esta enzima adiciona novos nucleotídeos na fita filha em formação através da adição do nucleotídeo na ponta 3' livre da molécula em crescimento (polimerização). Cada nucleotídeo adicionado deve parear com o nucleotídeo da fita mãe molde, garantindo a formação do correto par de base na dupla-hélice. Nucleotídeos incorporados errados podem ser removidos pela própria DNA polimerase, que também possui uma atividade exonucleásica, garantindo a fidelidade da replicação. Assim, a taxa de erros de incorporação é muito baixa, cerca de 1 em 1 milhão de nucleotídeos incorporados. Para iniciar a replicação nas células, as origens de replicação são reconhecidas por proteínas específicas e as duas fitas são separadas pela ação de enzimas helicase, que rompe as pontes de hidrogênio, com a ajuda das topoisomerasas ou girases, que removem os giros ou torções no DNA. O DNA se abre no ponto da origem formando, assim, a forquilha de replicação, que se move nos dois sentidos da hélice, para formar, concomitantemente, as duas novas fitas de DNA. Existe uma helicase em cada uma das direções da forquilha de replicação. Para que a replicação ocorra, as duas fitas devem ser mantidas afastadas pela ação de proteínas que se ligam à DNA de fita simples, impedindo seu reanelamento. Como a DNA polimerase não consegue adicionar o primeiro nucleotídeo de DNA para a formação da fita-filha, uma enzima chamada primase entra em ação para produzir o *primer* ou iniciador para a polimerização do DNA. Esse *primer* é composto por ribonucleotídeos e o *primer* é uma molécula de RNA que será removida posteriormente. Após a síntese do *primer*, a DNA polimerase consegue adicionar o primeiro nucleotídeo de DNA na hidroxila 3' livre do último nucleotídeo do primer. Os nucleotídeos substratos para a síntese são sempre nucleotídeos trifosfatados. E a adição do nucleotídeo na fita filha libera um pirofosfato que será posteriormente hidrolisado e servirá para direcionar a reação no sentido da síntese da cadeia nova. A DNA polimerase polimeriza e produz a nova fita filha com uma velocidade muito grande, ou seja, a processividade da enzima chega a 1000 nucleotídeos adicionados do segundo. A polimerização das fitas novas acontece até a forquilha de replicação atingir a forquilha vizinha. As duas forquilhas se fundem, formando uma dupla fita mais longa. Como as fitas de DNA tem polaridades opostas ( $5' \rightarrow 3'$  e  $3' \rightarrow 5'$ ) e a polimerização é sempre no sentido  $5' \rightarrow 3'$ , põe-se um problema de polimerização concomitante das duas fitas novas no DNA em replicação. A natureza tratou de resolver esse problema pela polimerização de uma fita de modo processivo (fita *leader*) e da outra na forma de fragmentos (fragmentos de Okasaki). Essa última é a fita *lagging* ou retardada. Mesmo sendo sintetizada em fragmentos, a velocidade de síntese das duas fitas é muito semelhante. Ao final do processo de replicação, os *primers* são removidos (pela ação de uma outra DNA polimerase), nucleotídeos de DNA são adicionados no local do *primers* e, em seguida, os fragmentos são selados pela ação da enzima ligase, que finaliza o processo, produzindo uma única fita-dupla longa de DNA.

**3.** Os genes são sequências de nucleotídeos consecutivos na fita dupla de DNA. A diferença entre genes encontra-se na disposição e no número de seus 4 nucleotídeos constituintes ao longo da dupla-fita. Existem genes curtos (poucas centenas de bases) e genes longos (que podem atingir milhares de bases de extensão). Apesar do DNA ser composto por duas fitas complementares, a informação genética (posição dos genes) está contida em somente uma das fitas, que é denominada a fita codificadora, ou fita positiva. A outra fita é chamada de complementar, molde, negativa, ou não codificadora. Durante a transcrição do DNA em RNA, a fita não codificadora servirá de molde para a que a enzima RNA polimerase possa produzir um RNA a partir do DNA e a fita de RNA sintetizada será idêntica à fita codificadora, pois será produzida por complementariedade à fita molde, sendo sempre sintetizada na direção  $5' \rightarrow 3'$ . Os genes estão dispostos em ambas as fitas de DNA. Assim, no genoma, não existe uma única fita codificadora e outra molde. Ora uma fita pode conter a informação genética, ora outra fita a contém. Se estamos tratando de uma célula eucariota, os genes codificadores de proteínas podem ser compostos por regiões de DNA literalmente codificadoras (os exons) separadas por regiões não codificadoras (os introns). A estrutura de exons e introns fazem parte da maioria dos genes eucariotos, sendo que ao longo da evolução, os genes se tornaram mais longo e mais fragmentados, isto é, com mais exons e introns. Genes monoexônicos são raros em organismos complexos, como os humanos.

**4.** No núcleo das células eucariotas, cada dupla fita longa de DNA corresponde a um cromossomo. Se os organismos são diploides, existem pares de cromossomos homólogos, cada um herdado de um progenitor. Esses cromossomos não ficam dispostos aleatoriamente no núcleo. Eles possuem nichos específicos, os territórios cromossômicos, e os cromossomos homólogos tendem a ocupar posições

próximas. Essas posições podem variar ao longo do desenvolvimento ou condições a que a célula é submetida. Como os cromossomos são moléculas longas e torcidas, eles precisam ser empacotados para caberem dentro do núcleo. As principais proteínas de empacotamento do DNA são as histonas, que são proteínas pequenas (em torno de 100 resíduos de aminoácidos), extremamente conservadas entre organismos evolutivamente distantes e ricas em aminoácidos carregados positivamente (argininas e lisinas). Existem quatro tipos principais de histonas, que são as histonas H3, H4, H2A e H2B, mas também existem variantes, com funções específicas. Os quatro tipos de histonas se organizam em um octâmero, que é formado por dois tetrâmeros das histonas H3, H4, H2A e H2B. Esse octâmero forma o alicerce para que o DNA se enrole ao redor, formando o que chamamos de nucleossomo. O nucleossomo tem, em média, 146 pares de bases de DNA ao seu redor. Os nucleossomos são a base da estrutura da cromatina, que é o conjunto de DNA e proteínas no núcleo. A estrutura de DNA enrolado em histonas formando os nucleossomos, é chamada de contas de um colar, ou estrutura de fibra de 10nm. Essa é a cromatina aberta, ou descondensada e é encontrada em regiões transcricionalmente ativas da cromatina, a eucromatina. A fibra de 10nm pode se enrolar, graças à ajuda de uma histona adicional, chamada histona H1, ou *linker*, formando o solenoide (mola), ou fibra de 30nm, um grau mais compacto da cromatina, mas ainda visto em regiões eucromáticas do núcleo. Empacotamentos mais compactos da cromatina também podem ser vistos no núcleo, como a fibra de 300nm e a de 700nm. Esses graus maiores de empacotamento existem em regiões heterocromáticas (regiões inativas ou silenciadas da cromatina) e dependem da atuação de outras proteínas que formam o arcabouço dessas fibras. Após a replicação do DNA para a divisão celular, os cromossomos se enovelam e atingem seu maior grau de compactação (fibra de 1400nm) e a estrutura clássica de um cromossomo duplicado (em forma de X) pode ser vista por microscopia ótica no centro das células durante a metáfase (uma das fases da mitose).

## Questão 02

1. Os genes codificadores de proteínas são transcritos pela RNA polimerase 2 em eucariotos e são os mais bem conhecidos e estudados quanto ao controle de sua expressão. Para que um gene seja expresso é necessário que ele esteja numa região de cromatina ativa, como será explicado na questão 2.2, mas também é necessário a coordenação de fatores em cis e trans. Em cis, temos os elementos da região promotora ou controladora. Existe o promotor mínimo ou basal, onde se encontra a região de ligação da RNAPol2, muitas vezes representado pela sequência TATA. A RNAPol2 iniciará a transcrição do gene alguns nucleotídeos à frente do seu ponto de ligação ao DNA, sendo esse ponto conhecido como +1, ou TSS (sítio de início da transcrição, do inglês *transcription start site*). Lembrando que somente uma fita de DNA será utilizada como molde para a síntese do RNA, a fita molde. Logo atrás do promotor mínimo existe a região do promotor proximal, que varia de tamanho nos diferentes genes e que pode chegar a uns milhares de quilobases. Em média, considera-se essa parte do promotor como sendo uma região de 1000 a 2000 nucleotídeos que antecede o TSS. No promotor mínimo se encontram várias caixas (do inglês *boxes*), que consistem de sequências de DNA reconhecidas por fatores de transcrição, que são proteínas específicas de diferentes tipos e estruturas. Os fatores de transcrição, ou fatores transativadores, são os elementos em trans que disparam a expressão gênica. Se um gene tem sua expressão controlada por uma série de fatores de transcrição, estes precisam estar ligados em seus respectivos *boxes* na região promotora para recrutar a RNAPol2 para o promotor mínimo. Quanto maior o número e tipos de *boxes*, mais complexa é a regulação da expressão do gene. Estes fatores podem estar presentes no núcleo e serem levados para o promotor após a interação com um ligante, como no caso dos fatores de transcrição que respondem aos hormônios esteroides, ou serem proteínas citoplasmáticas, que são ativadas por cascatas de sinalização, como as das MAP cinases que ativam os fatores de transcrição após uma série de fosforilações, culminando com o seu transporte para o núcleo. Além do promotor proximal, um gene pode ter sua expressão controlada por regiões distais. Estas regiões podem estar a centenas de milhares de nucleotídeos à montante ou à jusante do promotor mínimo, ou podem estar contidas em regiões intrônicas do próprio gene, como demonstrado recentemente. Estes elementos em cis distais são conhecidos por acentuadores (do inglês *enhancers*) ou silenciadores (do inglês *silencers*). Nesses elementos em cis também se ligam fatores de transcrição. Quando isso ocorre, e o elemento é ativado, a cromatina faz uma alça que aproxima o elemento distal do promotor proximal e do promotor mínimo, para que uma interação física

seja feita entre essas duas regiões do DNA. Existem proteínas coativadoras ou mediadoras dessa interação e essas também são fatores em trans, que controlam a expressão do gene. Se um determinado elemento distal é um acentuador, a formação da alça da cromatina e aproximação dessa região ao promotor mínimo irá ativar fortemente a expressão do gene, mas se o elemento distal é um silenciador, o contrário acontecerá. Também já foi demonstrado que um gene pode ser controlado por mais de um acentuador e que existem “super acentuadores” que ativam a expressão de vários genes ao seu redor. Outra novidade na área é que o acentuador também pode ser transcrito pela RNAPol2 gerando um pequeno RNA, o eRNA, que será importante nesse processo de controle da expressão de genes distantes. Quanto mais eRNA produzido por um acentuador, mais forte será seu efeito na ativação da expressão de genes distais.

**2.** Genes próximos fisicamente no DNA muitas vezes tem níveis de expressão semelhantes, indicando que estes podem ter controles de expressão comuns. O motivo dessa semelhança de expressão pode estar associado à organização da cromatina em TADs, que são os domínios de associação topológica, do inglês, *topologically associating domains*. Os TADs foram descritos como unidades de organização tridimensional da cromatina, sendo regiões que podem estar mais frouxas (estruturas como as fitas de 10 e 30nm) e, portanto, mais ativas, ou estruturas de cromatina mais condensadas e silenciadas, como a heterocromatina. O TAD não possui um tamanho de DNA específico e, assim, pode albergar um ou mais genes nessa região, mas suas bordas são características e isoladas dos TADs vizinhos, onde os elementos de DNA no seu interior fazem contatos e interagem fisicamente com mais frequência do que com os elementos que estão fora do TAD. É possível, utilizando técnicas de captura de estrutura de cromatina, como o Hi-C, determinar as bordas de um TAD e a frequência dos contatos de DNA no seu interior. O TAD é um elemento dinâmico da cromatina e pode mudar de tamanho e estrutura dependendo de condições como estresse, doenças, desenvolvimento, dentre outras. Dentro do TAD se encontra um ou mais genes, um ou mais acentuadores e, em algumas vezes, silenciadores. Como dito no item 2.1, o acentuador é importante para ativar fortemente a expressão de um gene e, dentro do TAD, um acentuador pode ativar a expressão de um gene, ou de vários genes contidos nessa região delimitada, tornando a intensidade da expressão destes genes parecida (genes coexpressos). Além da organização da cromatina em TADs, o papel das histonas é crucial no controle da expressão gênica. As histonas fazem parte dos nucleossomos, como explicado no item 1.4, que são as unidades básicas da cromatina. Além da região de contato das histonas com o DNA, existem as regiões amino-terminais das histonas (caudas N-terminais), que não estão muitas vezes em contato direto com o DNA, pois alguns resíduos de aminoácidos dessas caudas podem estar quimicamente modificados. Estas são conhecidas como modificações pós-traducionais das histonas. Os principais resíduos de aminoácidos que sofrem modificações químicas são as lisinas e argininas, seguidos por serinas, dentre outros. As lisinas, por exemplo, podem estar acetiladas, metiladas, ubiquitinadas, ou sumoiladas e as serinas fosforiladas. Uma série de modificações pós-traducionais das histonas já foi descrita, demonstrando a complexidade da composição da cromatina. As histonas H3 e H4 são as que mais apresentam modificações químicas, sejam em número, ou em diversidade. Geralmente, as acetilações e fosforilações levam ao afrouxamento da cromatina e ativação da expressão dos genes na região daqueles nucleossomos modificados. Já a metilação às vezes é ativadora, às vezes é inibidora da expressão dos genes. Isso depende de qual resíduo de aminoácido será modificado. Como exemplo, citamos a marca H3K4Me como ativadora e H3K27Me como inibidora. A determinação dos diferentes papéis dessas modificações químicas das histonas é conhecida como o código das histonas.

**3.** Quando o DNA é transcrito em RNA pela RNAPol2, uma série de processamentos ocorrem na molécula de RNA para sua maturação e exportação do núcleo para o citoplasma. Quando a RNAPol2 inicia a transcrição, sua cauda carboxi-terminal é intensamente fosforilada. É nessa cauda que se ligam os fatores de processamento do RNA, à medida que a transcrição ocorre. A primeira modificação química no RNA nascente é a adição do capacete, ou CAP na extremidade 5' do RNA. Esta adição serve como proteção da ponta do RNA e para que este possa ser traduzido no citoplasma pela forma convencional de tradução ribossomal (reconhecimento e entrada da subunidade menor do ribossomo pelo CAP). O CAP é um nucleotídeo de guanina modificado quimicamente por metilação (o 7-metil-guanosina), sendo ligado na ponta do RNA nascente por uma ligação não convencional 5' → 5'. Após a adição do CAP, o alongamento da transcrição está ocorrendo e o RNA nascente será produzido contendo por exons e introns. Para a geração do RNA maduro, os introns são removidos por um processo denominado *splicing*, que leva à produção de um transcrito maduro, só contendo exons, para

ser futuramente transportado para o citoplasma. O *splicing* é realizado pelo spliceossomo, uma estrutura composta de pequenos RNAs (snRNAs) e proteínas, as ribonucleoproteínas (RNPs). O spliceossomo também fica ligado à cauda C-terminal fosforilada da RNAPol2. O processo do *splicing* consiste no ataque nucleofílico feito por um nucleotídeo de adenina presente no ponto de ramificação dentro do intron ao sítio doador do *splicing* presente na junção exon à montante-intron. Com isso, ocorre a primeira reação de transesterificação e forma-se uma estrutura em anel dentro do intron. O exon à montante fica livre. A hidroxila livre da extremidade 3' do exon à montante ataca agora o sítio acceptor do *splicing* presente no final do intron e início do exon à jusante. Faz-se a segunda reação de transesterificação e os dois exons são unidos. O intron é eliminado na forma de laço (com a estrutura em anel interno). O *splicing* do RNA é um processo que ocorre concomitante à síntese do RNA pela RNAPol2 (processo co-transcricional). Ao chegar próximo ao final do gene, a RNAPol transcreve um sinal de parada transcricional. Esse sinal é chamado de sinal de poliadenilação, cuja sequência é geralmente AAUAAA. Essa sequência no RNA recruta um complexo de proteínas de clivagem e poliadenilação do RNA, que estão ligadas às cauda C-terminal da RNAPol2. O RNA é clivado poucos nucleotídeos à frente desse sinal e a poliA polimerase inicia a adição da cauda de poli adeninas na extremidade final do RNA. O tamanho da cauda adicionada é variável, mas tem, em média 200nt. A cauda poliA também serve para proteger a ponta final do mRNA, assim como permitir sua tradução ribossomal no citoplasma. Ao final da adição da cauda de poliA, o mRNA está pronto para ser transportado para o citoplasma.