

ORGANIZAÇÃO MOLECULAR DOS GENOMAS

2.1 Dimensão dos genomas

Verdadeiro/Falso

1. Teoricamente, a quantidade de DNA em células diplóides do mesmo indivíduo deverá ser aproximadamente a mesma, dentro dos limites do erro experimental. V
2. Num mesmo indivíduo, a quantidade de DNA nas células germinais deverá ser aproximadamente metade da que existe nas células somáticas. V
3. O número de intrões por gene é um dos factores que pode contribuir para a variação de dimensão dos genomas. V
4. A dimensão dos intrões é geralmente inferior à dos exões. F
5. A dimensão do genoma de organismos filogeneticamente próximos e com o mesmo grau de complexidade é sempre muito aproximada. F
6. Genoma e transcrito são termos que se aplicam ao mesmo tipo de moléculas. F

Escolha múltipla

1. A dimensão dos genomas está estreitamente relacionada com: a) o número de genes; b) a complexidade do organismo; c) a proximidade filogenética; d) a quantidade de sequências de DNA repetitivo; e) a dimensão do DNA intergénico. d;
2. O genoma de organismos procarióticos: a) é haplóide; b) pode conter profagos; c) pode conter elementos móveis; d) apresenta uma maior compactação da informação genética relativamente ao genoma dos organismos eucarióticos; e) pode conter plasmídeos. a, b, c, d, e;
3. O paradoxo C é a: a) falta de relação observada entre a dimensão do genoma e o número de genes, em diferentes organismos; b) falta de relação observada entre o grau de complexidade de um organismo e a dimensão do seu genoma, relativamente a outros organismos; c) falta de relação observada entre o número de genes e a complexidade de um organismo; d) falta de relação observada no número total de genes em organismos filogeneticamente próximos; e) relação observada no número total de genes em organismos filogeneticamente próximos. b;
4. No genoma eucariótico: a) o DNA cromossómico pode estar na forma haplóide ou diplóide; b) as moléculas de DNA cromossómico podem ser circulares ou lineares; c) o DNA mitocondrial é de dimensões semelhantes às do DNA cromossómico; d) o DNA mitocondrial é circular; e) a dimensão pode variar entre 1×10^2 Mb – 1×10^5 Mb. a, d, e;

Questões básicas

1. Estabeleça a diferença entre os termos genoma e cromossoma, com base no modo como podem ser aplicados aos organismos eucarióticos e procarióticos?

R: O termo genoma engloba toda a informação genética de uma célula ou organismo, enquanto que o termo cromossoma se refere à estrutura genética, com capacidade de replicação, constituída por DNA que, por sua vez, contém na sua sequência nucleotídica o conjunto de genes dispostos de forma linear. As células eucarióticas, para além do DNA cromossómico (nuclear), também contêm DNA mitocondrial e podem ainda conter DNA cloroplastidial, que constituem a globalidade do genoma da célula. No caso das células procarióticas, que apresentam uma grande diversidade genética, alguns autores consideram que o genoma corresponde unicamente ao cromossoma bacteriano, enquanto outros incluem no conceito de genoma bacteriano as moléculas de DNA plasmídico, quando estas são indispensáveis à sobrevivência da bactéria em questão.

2. Aponte a principal razão molecular para o facto de organismos filogeneticamente próximos, nomeadamente em plantas (ex. arroz e milho), terem genomas com dimensões muito diferentes.

R. A principal razão molecular é o elevado conteúdo em DNA repetitivo na espécie com o genoma de maior dimensão (milho). Neste caso, 66% do seu genoma corresponde a DNA repetitivo, sobretudo retro elementos. Contudo, se durante muito tempo a grande diferença de dimensão se justificava com a presença de DNA repetitivo, sendo o número de genes semelhante entre as duas espécies, e apresentando estas uma considerável sintenia, a sequenciação completa do genoma de ambas as espécies revelou que o aumento na dimensão do genoma também se devia a um maior número de genes no milho (em 30 % dos genes do milho não foi encontrada homologia com os genes do arroz).

3. O que poderia avançar sobre a relação entre quantidade de DNA codificante e não codificante no conjunto de todos os organismos procariotas e eucariotas?

R. Globalmente, nos organismos procariotas, a quantidade de DNA codificante é superior à do DNA não codificante. Na realidade, nestes organismos há uma grande economia de espaço, encontrando-se a informação génica muito compactada e, praticamente, sem regiões intergénicas. À medida que se sobe na escala evolutiva, desde os eucariotas unicelulares, como a levedura, até aos seres multicelulares, a extensão do DNA codificante tende a ser progressivamente maior devido à presença de DNA repetitivo, genes com intrões e extensas regiões intergénicas.

2.2 Estrutura do genoma procariótico: cromossoma bacteriano e DNA plasmídico

Verdadeiro/Falso

1. O genoma bacteriano é diplóide. F
2. O DNA cromossómico bacteriano é constituído por cromatina. F
3. Ao cromossoma bacteriano estão associadas proteínas com funções no superenrolamento e condensação da molécula de DNA.V
4. O DNA cromossómico bacteriano pode ser linear. V
5. O DNA cromossómico de *E. coli* sofre dobras e voltas na sua estrutura gerando um conjunto de domínios que contribuem para a compactação do DNA.V
6. Nas bactérias, o DNA cromossómico circular está associado a certas proteínas, sofrendo dobras que formam múltiplos domínios, superenrolados e independentes, dando origem a uma estrutura chamada nucleóide. V

Escolha múltipla

1. Relativamente ao genoma bacteriano: a) tem a informação genética condensada devido à inexistência praticamente total de DNA repetitivo e intergénico; b) é maioritariamente uma molécula de DNA circular; c) é sempre composto por uma única molécula de DNA; d) é haplóide; e) é pelo menos 90% é codificante. a, b, d, e
2. O cromossoma bacteriano: a) está enrolado formando uma estrutura densa designada por nucléolo; b) está enrolado formando uma estrutura densa designada por nucleóide; c) está envolvido pela membrana nuclear; d) tem, associadas à molécula de DNA, proteínas semelhantes a histonas (*histone-like*), com funções na condensação da molécula de DNA; e) tem a informação genética condensada porque possui sequências intergénicas muito curtas. b, d, e

Questões básicas

1. Quais os componentes proteicos do cromossoma bacteriano?
R. O cromossoma bacteriano é essencialmente constituído por DNA ao qual estão associadas algumas proteínas semelhantes a histonas, como as proteínas HU, H-NS (H1), IHF, Fis, ainda não completamente caracterizadas no que se refere à especificidade da sua função, mas que asseguram a proteção, a compactação e estabelecimento das dobras e enrolamento do cromossoma.
2. O que são plasmídios? Onde foram identificados? Defina-os do ponto de vista estrutural e funcional.
R. Os plasmídios existentes na natureza / naturais são, na sua grande maioria, moléculas de DNA circulares, de cadeia dupla, que geralmente codificam proteínas que conferem uma vantagem para a célula onde residem. Existem em praticamente todo o tipo de bactérias. Por este motivo estão frequentemente associados a fenómenos de adaptação e evolução bacteriana (ex. a resistência aos antibióticos dissemina-se rapidamente numa população bacteriana porque os genes que lhes conferem lhes resistência localizam-se frequentemente em plasmídios). Os plasmídios possuem uma origem de replicação própria, independente da origem de replicação do cromossoma bacteriano, o que permite não só a sua manutenção extracromossómica como controlar o número de cópias que têm por célula. Embora muito raramente, existem plasmídios em cadeia simples (ex. detectados em *Bacillus subtilis* e *Staphylococcus aureus*) e lineares (ex. observado em *Borrelia burgdorferi*). Os plasmídios recombinantes já são os plasmídios engendrados/manipulados pelos investigadores, resultando assim da junção *in vitro* de diferentes sequências de DNA que em si mesmas conferem uma função ao plasmídio.

2.3 Estrutura do cromossoma eucariótico

Verdadeiro/Falso

1. A cromatina é o termo genérico utilizado para designar o complexo DNA/histonas/proteínas não-histónicas encontrado no núcleo das células eucarióticas. V
2. A cromatina dos cromossomas na metáfase está menos condensada do que na interfase. F
3. A condensação da cromatina não afecta a função das sequências codificantes. F
4. As histonas são proteínas compostas maioritariamente por aminoácidos de carácter ácido. F
5. O DNA de ligação (DNA *linker*) é definido como o segmento de DNA que liga dois nucleossomas. V
6. Observou-se que o DNA dos cromossomas em metáfase, superenrolado em *loops*, se encontra ligado a um esqueleto proteico (*chromosome scaffold* ou *nuclear-scaffold*), que apresenta uma estrutura semelhante no par de cromatídios-irmãos. V
7. Os locais MARs (*Matrix-Associated Region*) ou SARs (*Scaffold-Attachment Region*), são sequências na molécula de DNA do cromossoma eucariótico que se ligam à matriz nuclear (*nuclear-scaffold*). V
8. Diversas proteínas não-histónicas participam na dinâmica e organização dos *loops* de cromatina dos cromossomas em interfase e metáfase. V
9. A localização de genes em regiões eucromáticas é suficiente para garantir a sua expressão. F
10. O centrómero só tem relevância funcional no processo de separação dos cromossomas durante a meiose (e não durante a mitose). F

Escolha múltipla

1. Quais das seguintes características distinguem melhor os cromossomas eucarióticos dos procarióticos? a) os cromossomas eucarióticos são sempre lineares vs os cromossomas bacterianos são sempre circulares; b) muitos cromossomas eucarióticos vs um só cromossoma bacteriano; c) os cromossomas eucarióticos são constituídos por proteínas e DNA vs os cromossomas bacterianos são unicamente constituídos por DNA; d) as histonas são os principais componentes das proteínas cromossómicas eucarióticas vs as histonas não estão presentes nos cromossomas bacterianos; e) os cromossomas eucarióticos estão espacialmente confinados a um compartimento com membrana (o núcleo) vs os cromossomas bacterianos não estão rodeados por uma membrana, para além da membrana plasmática. d, e;
2. Um cromossoma eucariótico: a) também pode designar-se de cromatídio; b) é formado por uma única molécula de DNA; c) liga-se ao fuso acromático durante a prófase; d) contém sequências que nunca são transcritas; e) contém regiões que se mantêm condensadas durante a interfase. b, d, e;
3. A cromatina: a) composta por DNA e proteínas é característica dos organismos procarióticos; b) tem a ela associadas, proteínas não-histónicas, como DNA polimerases, topoisomerases, HATs; c) é constituída por subunidades, os nucleossomas, que têm o mesmo tipo de configuração em todos os eucariotas; d) é o mesmo que "código de histonas"; e) corresponde ao estado inalterado de enrolamento das moléculas de DNA cromossómico. b, c;
4. Os nucleossomas: a) correspondem a DNA repetitivo ao longo do cromossoma eucariótico; b) são estruturas condensadas de DNA localizadas no núcleo eucariótico; c) são unidades estruturais básicas da cromatina; d) uma das formas condensada têm um arranjo solenóide (fibra de 30 nm); e) correspondem ao complexo através do qual os cromossomas se ligam às fibras do fuso acromático. c, d;
5. O código de histonas: a) está associado à função da cromatina na transcrição dos genes; b) está associado à função da cromatina na tradução dos genes; c) refere-se ao diferente efeito que as modificações químicas das histonas têm na expressão génica; d) envolve a acetilação das lisinas; e) envolve essencialmente modificações químicas nos aminoácidos da porção C-terminal das histonas. a, c, d;
6. Relativamente às modificações químicas do DNA e das histonas, quais dos seguintes acontecimentos estão associados à regulação da expressão génica em eucariotas? a) as ilhas CpG localizadas em regiões promotoras, quando metiladas (^mCpG), conduzem frequentemente à activação transcricional; b) DNA metilado recruta complexos que contêm actividade de desacetilação de lisinas; c) numa histona, um mesmo aminoácido pode sofrer diferentes modificações químicas que têm efeitos opostos relativamente à activação da transcrição; d) as acetilases anulam a carga positiva das lisinas, favorecendo a condensação da cromatina; e) a acetilação de lisinas, geralmente na extremidade N-terminal das histonas, está associada à activação transcricional dos genes dessa região do cromossoma. b, c, e;

7. As histonas: a) são proteínas de baixa massa molecular; b) têm um motivo estrutural comum, designado de *histone fold*; c) formam o centro proteico do nucleossoma constituído exclusivamente por H2A, H2B; d) existe um elevado grau de homologia entre as histonas das diferentes espécies; e) ligam-se ao DNA carregado negativamente, neutralizando-o. a, b, d, e;
8. A histona H1: a) fixa o DNA ao centro proteico do nucleossoma formado pelas outras histonas; b) faz parte do centro proteico do nucleossoma; c) está protegida por DNA; d) possui uma extremidade N-terminal flexível que se projecta apresentando locais de modificação química importantes na função da cromatina; e) está associada ao DNA *linker*. a, e;
9. As sequências SARs: a) são regiões no DNA às quais se associam proteínas não-histónicas que organizam os *loops* de cromatina; b) são origens de replicação; c) são proteínas não-histónicas que permitem a formação do *chromosome scaffold*; d) são sequências de DNA às quais se ligam as histonas; e) significam o mesmo que *chromosome scaffold*. a
10. A heterocromatina: a) contém unicamente, genes inactivos; b) contém frequentemente, DNA repetitivo; c) é visível durante a metáfase e em núcleos interfásicos; d) pode ser activada pela metilação das lisinas das histonas; e) contém uma elevada proporção de RNA polimerases. b, c, d;
11. Os centrómeros: a) são necessários à segregação normal dos cromossomas durante a mitose e a meiose; b) são necessários à formação do cinetocoro nos eucariotas superiores; c) correspondem ao cromocentro; d) são essencialmente regiões eucromáticas; e) nos eucariotas superiores são regiões ricas em sequências repetitivas. a, b, e;
12. Os telómeros: a) asseguram a estabilidade das extremidades das moléculas de DNA cromossómico lineares; b) são necessários para a replicação de plasmídeos; c) contêm curtas sequências repetitivas que variam de organismo para organismo; d) contêm emparelhamentos de bases não-canónicas; e) são sintetizados por uma ribonucleoproteína. a, c, d, e;
13. A telomerase: a) controla o número de sequências repetidas dos telómeros; b) existe em baixas concentrações nas células estaminais; c) é um tipo de transcritase reversa; d) é uma ribonucleoproteína, cujo conteúdo proteico é responsável pela actividade catalítica; e) é uma ribozima. c, d;

Questões básicas

1. Quais os componentes básicos dos cromossomas eucarióticos?
R. São os seguintes: DNA, histonas e proteínas não-histónicas, que globalmente formam a cromatina.
2. Quais as moléculas constituintes da cromatina?
R. Cromatina é o termo genérico utilizado para descrever o estado de enrolamento em que se encontram o DNA nuclear e as proteínas a ele associadas, durante a intérfase, entre mitoses, no ciclo celular eucariótico. Contém, para além das histonas, proteínas não-histónicas tais como as proteínas do *chromosome scaffold* e as proteínas HMP (*High-Mobility-Group*), correspondendo estas últimas a um pequeno grupo de proteínas em que algumas são factores de transcrição. A função das proteínas não-histónicas é controlar a expressão génica e os níveis de organização estrutural mais elevados. A cromatina, cuja unidade fundamental são os nucleossomas, apresenta uma estrutura com diferentes graus de compactação a que correspondem diferentes níveis de organização dos cromossomas eucarióticos.
3. O DNA das células eucarióticas está estreitamente associado a proteínas, formando um complexo designado por cromatina, que pode existir em formas mais ou menos condensadas. Descreva como é que a cromatina está organizada nos cromossomas, de forma a compactar o genoma no núcleo das células.
R. A cromatina adquire diferentes níveis de organização devido à acção de proteínas ainda não completamente definidas. O primeiro nível de organização corresponde à fibra de 10 nm que consiste essencialmente num filamento contíguo de nucleossomas. É visível *in vitro*, em condições de baixa força iónica, não se sabendo se, *in vivo*, apresenta esta forma. À medida que a força iónica aumenta, forma-se a fibra de 30 nm de diâmetro, com uma estrutura solenóide que equivale a um enrolamento do DNA de aproximadamente 40 vezes. Esta estrutura, visível tanto na cromatina interfásica como nos cromossomas mitóticos, corresponde ao segundo nível de organização com seis nucleossomas por cada volta. O terceiro nível é determinado pelo enrolamento da fibra de 30 nm, formando *loops* de 10 a 90 kb que se ligam ao *chromosome scaffold*, dando origem a uma fibra de 300 nm. O grau de empacotamento deste nível é de aproximadamente 1 000 vezes na eucromatina e de 10 000 vezes na heterocromatina, que mantém este grau de enrolamento quer na interfase quer no cromossoma metafásico. Estão ainda descritos mais níveis de enrolamento, 700 nm e 1400 nm, produzindo-se finalmente o cromossoma metafásico.

4. Em cada um dos pares de tipos de cromatina, qual é o elemento mais condensado?

- a) fibra de 100 Å ou fibra de 300 Å
- b) fibra de 300 Å ou *loop* de DNA associado ao *chromosome scaffold*
- c) eucromatina ou heterocromatina
- d) cromossoma em intérfase ou em metáfase

R. O elemento mais condensado é: a) fibra de 300 Å; b) *loop* de DNA associado ao *chromosome scaffold*; c) heterocromatina; d) cromossoma em metáfase.

5. Quais são os principais constituintes moleculares dos nucleossomas?

R. Os nucleossomas, cujos componentes e estrutura estão bem caracterizados, correspondem ao primeiro nível hierárquico no processo de enrolamento do DNA nuclear. Contêm aproximadamente 200 pb (145 pb + (10 a 100 pb)) de DNA, enrolados à volta de um octâmero proteico de histonas. O octâmero é constituído por duas moléculas de cada uma das histonas, H2A, H2B, H3 e H4. O conjunto, DNA (145 pb) + octâmero, representa a estrutura central do nucleossoma (*nucleosome core*). Para além disso, ainda existe uma histona H1 associada a um DNA *linker* de aproximadamente 55 pb, que mantém os nucleossomas adjacentes interligados. Os nucleossomas são os componentes invariáveis da eucromatina e da heterocromatina, no núcleo interfásico e nos cromossomas metafásicos.

6. Qual a consequência funcional da presença, em elevadas proporções, dos aminoácidos arginina e lisina nas proteínas histónicas?

R. A arginina e a lisina são aminoácidos positivamente carregados. Isto permite-lhes ligarem-se a fosfatos negativamente carregados no esqueleto da hélice de DNA.

7. O que significa dizer que uma sequência é conservada? Por que razão as proteínas histónicas são altamente conservadas nos diferentes organismos?

R. Uma sequência conservada apresenta pouca variação entre os organismos. As histonas são altamente conservadas porque a estrutura tridimensional do DNA é fundamentalmente a mesma em todos os organismos.

8. Por que razão os genes das histonas existem em mais do que uma cópia por cada histona (por exemplo, no genoma humano existem 60 genes, isto é, 10 a 15 cópias por cada tipo de histona)?

R. Para que as células tenham mais cópias dos genes das histonas, a serem simultaneamente transcritos e traduzidos, originando deste modo uma elevada produção de histonas.

9. Por que motivo não são encontradas mutações nos genes que codificam histonas, originando histonas não-funcionais?

R. Porque essas mutações são provavelmente letais. De facto, a sequência de aminoácidos das histonas é altamente conservada nos organismos eucarióticos, o que sugere que os organismos com histonas não funcionais não serão viáveis.

10. O que acontece aos cromossomas que perdem os telómeros?

R. Os cromossomas podem fundir-se entre si e recombinar com outros segmentos de cromatina. Tendem a quebrar e a diminuir de tamanho em ciclos celulares subsequentes, e eventualmente a serem degradados.

11. Qual o análogo dos telómeros, no cromossoma de *E. coli*?

R. Não existe qualquer análogo porque o cromossoma de *E. coli* é circular, e por isso desprovido de extremidades.

2.4 Estrutura e organização dos genes e das sequências genómicas

Verdadeiro/Falso

- 1. As proteínas são maioritariamente codificadas por sequências de DNA únicas. V
- 2. A maioria dos genes é de cópia única em bactérias e em eucariotas inferiores. V
- 3. A maior parte do genoma, nos eucariotas superiores, contém regiões codificantes de RNA e proteínas. F
- 4. As famílias multigénicas podem ser definidas como grupos de genes que codificam proteínas relacionadas ou idênticas. V
- 5. As famílias multigénicas têm que estar, necessariamente, agrupadas. F
- 6. O valor $Cot_{1/2}$ indica a concentração do DNA que reassociou (dsDNA) a metade do tempo total da reassociação. V
- 7. Quanto mais rápida a reassociação mais baixo é o valor de $Cot_{1/2}$. V

8. A colisão das cadeias complementares de segmentos de DNA cujas sequências existem em elevado número de cópias (ex. centenas de milhar de cópias) ocorre a uma maior frequência, sendo a velocidade de reassociação maior e como tal o valor de $Cot_{1/2}$ maior. F
9. Fragmentos de DNA longos e com a representação de uma ou duas cópias tendem a ter um elevado valor de $Cot_{1/2}$. V
10. A velocidade de reassociação depende do número de cada segmento e da frequência com o qual as cadeias complementares podem reagir e reassociar-se. V
11. Os microssatélites encontram-se em regiões heterocromáticas. V
12. No DNA eucariótico, a fracção de DNA que se reassocia mais rapidamente é a que corresponde aos microssatélites. V

Escolha múltipla

1. Numa única célula diplóide, eucariótica e em condições normais, os cromossomas homólogos: a) possuem os mesmos genes, e pela mesma ordem; b) apresentam a mesma sequência de nucleótidos no DNA; c) possuem os mesmos genes, mas frequentemente com uma ordem diferente; d) têm dimensão, forma e padrão de bandas idênticos, numa análise de cariótipo; e) possuem genes com uma organização estrutural idêntica. a, d, e;
2. Nas células diplóides de um organismo, o número de alelos diferentes num dado gene nuclear de cópia única é: a) só um; b) um ou dois; c) um, dois, três ou quatro; d) indeterminado; e) variável. b;
3. Numa população de organismos diplóides, o número de alelos de um dado gene nuclear de cópia única, pode ser: a) só um; b) um ou dois; c) um, dois, três ou quatro; d) muito elevado; e) todas as possibilidades anteriores. d;
4. Os genes que codificam os rRNAs 5S incluem-se em que tipo de DNA? a) DNA altamente repetitivo; b) DNA moderadamente repetitivo; c) família génica simples; d) família génica complexa agrupada; e) família génica complexa dispersa. b, c;
5. Nas células eucarióticas, o DNA designado de não-codificante pode incluir: a) intrões; b) pseudogenes; c) elementos móveis; d) DNA intergénico; e) SSRs. a, b, d, e;
6. A análise de reassociação $Cot_{1/2}$ de uma dada molécula de DNA: a) reflecte o número, extensão e tipo de sequências presentes na molécula de DNA; b) baseia-se no tempo de reassociação da molécula de DNA; c) apresenta um valor elevado em moléculas ricas em sequências SSR; d) indica que quanto mais baixo é o valor calculado, mais lenta é a velocidade de reassociação do DNA; e) indica que o DNA eucariótico de cópia única se encontra na fracção que reassocia mais lentamente. a, b, e;
7. O DNA satélite: a) pode ser separado do DNA cromossómico por centrifugação em gradiente de densidade (centrifugação isopícnica); b) apresenta uma menor condensação da cromatina; c) apresenta grupos de sequências repetitivas não-codificantes; d) corresponde ao DNA alfa; e) refere-se às sequências de DNA centroméricas. a, c, d, e;
8. Os minissatélites: a) têm um número variável de repetições que resulta essencialmente de acontecimentos de recombinação; b) são bons marcadores de DNA; c) localizam-se igualmente em regiões codificantes e não-codificantes; d) também são conhecidos por VNTRs (*Variable Number of Tandem Repeats*); e) podem ser detectados por PCR. a, b, d, e;
9. O genoma humano é predominantemente constituído por sequências: a) codificantes; b) exónicas; c) intrónicas; d) repetitivas; e) codificantes e exónicas. d;
10. Os marcadores genéticos ou marcadores de DNA: a) são *loci* identificáveis num cromossoma; b) são *loci* cuja hereditariedade não pode ser detectada; c) podem estar localizados em regiões codificantes e não-codificantes; d) correspondem a diferentes tipos de sequências de DNA sem função biológica atribuída; e) permitem distinguir indivíduos. a, c, d, e;
11. Um SNP (*single nucleotide polymorphism*): a) só é detectável por sequenciação do DNA; b) é um polimorfismo muito frequente nos genomas; c) origina sempre uma patologia; d) está na origem da *replication slippage*; e) é um polimorfismo que se refere a um único par de nucleótidos quando se comparam diferentes alelos. b, e;

Questões básicas

1. Em que medida é que a organização genética de *E. coli* difere da dos eucariotas?

R. A grande maioria dos genes em *E. coli* está organizada em operões, isto é, em unidades de DNA compostas por mais do que um gene, pertencentes à mesma unidade de transcrição. Assim, os operões estão sujeitos à mesma regulação transcricional, e os seus genes desempenham geralmente funções relacionadas. Nos eucariotas, os genes não estão organizados em operões, embora já tenha sido observada uma excepção em *Caenorhabditis elegans*. Em geral, os genes são de cópia única ou de baixo número de cópias, mas também há genes representados por um número variável de cópias (moderadamente repetitivos), podendo formar as chamadas famílias multigénicas (simples ou complexas, agrupadas ou dispersas).

2. Relativamente à organização genética, quais são os diferentes tipos de sequências de DNA que existem nos eucariotas? Exemplifique cada um dos tipos que referir.

R. Nos eucariotas existem diferentes tipos de sequências de DNA que, para além de poderem ser codificantes ou não, ainda podem estar representadas por um número variável de cópias e de uma forma agrupada ou dispersa. Sequências únicas ou com poucas cópias, são a maioria das sequências codificantes de proteínas, assim como uma série de sequências de função desconhecida. Sequências moderadamente repetitivas incluem algumas famílias génicas que codificam rRNAs e tRNAs (famílias génicas simples), mas a maioria consiste em elementos transponíveis e em vestígios de elementos transponíveis. Também se incluem nesta classe as famílias génicas complexas agrupadas (ex. genes das hemoglobinas: gene da hemoglobina α no cromossoma 16; gene da hemoglobina β no cromossoma 1) e complexas dispersas (genes da aldolase que se encontram em diferentes cromossomas). Por último, também existe DNA altamente repetitivo, representado pelas transposições e por DNA satélite. O DNA satélite consiste em curtas sequências repetidas agrupadas, presentes em centenas a milhares de cópias por genoma, localizadas na região do centrómero dos eucariotas superiores. Mais de 50% do genoma humano é constituído por sequências repetitivas: transposições (activos ou inactivos na maioria dos casos), pseudogenes, microssatélites, minissatélites, blocos de sequências repetitivas nos centrómeros (DNA alfa ou DNA satélite) e telómeros.

3. O que entende por polimorfismos de DNA?

R. Qualquer variação no DNA, uma vez descoberta, define um *locus* que pode estabelecer diferenças ainda que não afecte qualquer fenótipo. Polimorfismos de DNA são variantes de genes ou de sequências de DNA, geralmente não-codificantes, que possuem dois ou mais alelos. Um polimorfismo genético pode corresponder à existência de múltiplos alelos num determinado *locus*. Normalmente, um *locus* é considerado polimórfico se aparece em percentagem superior a 1% na população, o que o distingue de mutações pontuais. Os polimorfismos podem ser detectados a três níveis: a nível fenotípico quando a sequência afecta a função de um gene; a nível dos fragmentos de restrição quando afectam um local de restrição; e a nível da sequência de DNA através da análise directa do DNA.

4. O que é uma família multigénica? Dê exemplos baseados no genoma humano.

R. A designação, família multigénica, refere-se a uma forma de organização dos genes em eucariotas, em que cada família é definida como um grupo de genes que codificam proteínas idênticas ou, na grande maioria dos casos, relacionadas, apresentando pequenas diferenças funcionais. Trata-se geralmente de informação génica moderadamente repetitiva. Os membros da família multigénica derivam de um ancestral comum, por duplicação génica, seguida de acumulação de mutações (alterações na sequência nucleotídica) nas cópias do gene. É o caso do gene que codifica o rRNA 5S, com 2000 cópias em tandem, isto é, agrupadas no cromossoma 1. Por outro lado, na família de genes dos rRNAs 28S, 5,8S e 18S, os genes encontram-se em tandem (50 a 70 cópias), e simultaneamente, localizam-se de forma dispersa porque um número de cópias idênticas, está presente em vários cromossomas (13, 14, 15, 21 e 22); cada cópia contém um só gene que será transcrito e depois processado, gerando os diferentes rRNAs. Se considerarmos a variabilidade da sequência nucleotídica dos genes que compõem as famílias multigénicas, estas podem ser simples, como as dos dois exemplos anteriores, mas também podem ser compostas por vários genes, com sequências semelhantes, embora suficientemente diferentes para que os produtos génicos tenham propriedades bioquímicas diferentes, respondendo adequadamente a alterações fisiológicas durante o desenvolvimento. Estas famílias podem ser agrupadas (por exemplo, os genes da α e β globinas, localizados nos cromossomas 16 e 11, respectivamente), ou dispersas (genes da aldolase, envolvidos na produção de energia e localizados em diferentes cromossomas, 3, 9, 10, 16 e 17). Uma família multigénica pode conter pseudogenes.

5. Que vantagem metabólica, as sequências moderadamente repetidas dos genes de rRNA e das histonas representam para a célula?

R. Tipicamente, os produtos génicos destas sequências são necessários em grandes quantidades para assegurar um metabolismo celular óptimo. Assim, a célula pode atingir elevada produção de rRNA e histonas por via cópias múltiplas de genes em células metabolicamente activas.

6. O que entende por hiperchromicidade? De que forma este conceito está relacionado com a desnaturação do DNA?

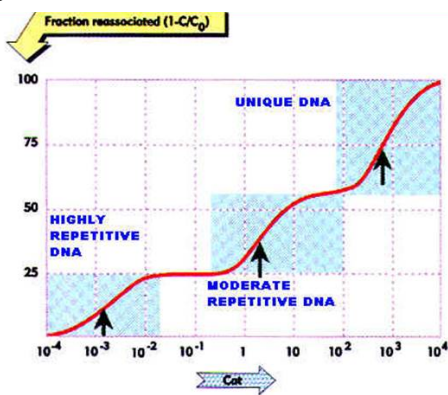
R. A hiperchromicidade é o aumento da capacidade de absorção de luz (densidade óptica), de um material. A hiperchromicidade do DNA ocorre quando a dupla hélice do DNA é desnaturada por quebra das ligações de hidrogénio entre os pares de bases complementares. A desnaturação, que se realiza sob a acção do calor, da adição de um

agente desnaturante ou do aumento do nível de pH, expõe as bases de purina e pirimidina no DNA que absorvem fortemente a luz ultravioleta (UV). O DNA de cadeia dupla absorve menos fortemente a radiação UV devido ao empilhamento e interação entre as bases que se encontram no interior da molécula. Um gráfico relacionando a densidade óptica com a temperatura é conhecido como uma "curva de fusão". Esta relação é extremamente útil nos estudos de composição em bases da molécula do DNA assim como nas tecnologias que envolvem hibridação de ácidos nucleicos. O oposto, uma diminuição da absorção, é chamada de hipocromicidade.

8. O que entende por valor Cot? Que vantagens tem a sua determinação?

R. O valor Cot é um dado quantitativo que representa a percentagem de dsDNA reassociado relativamente a uma concentração inicial de DNA (C_0), durante um determinado período de tempo (t) e a partir de um tempo zero. A medida da reassociação é feita por DO_{260} em intervalos de tempo determinados. As reacções de reassociação podem ser graficamente representadas pelas curvas de Cot. A velocidade de reassociação depende do comprimento e da complementaridade dos fragmentos que se encontram na reacção de reassociação, mas também do número de cópias desses fragmentos. Por esta razão, esta medida permite-nos identificar 3 tipos essenciais de fragmentos (ou sequências): fragmentos de cópia única cuja velocidade de reassociação é muito lenta pois é mais difícil encontrar a sequência única complementar; sequências com sequências muito semelhantes, que redundam num elevado número de cópias e como tal a colisão entre cópias semelhantes é elevada e a reacção de reassociação muito rápida; entre estas duas classes, uma classe em que o número de cópias repetidas é moderado e que se reassociam a uma velocidade entre a primeira classe e a terceira. Por esta razão a determinação dos valores Cot de qualquer fonte de DNA, independentemente da sua origem e concentração, dão informação sobre o tipo de sequências que compõem essa molécula sendo o valor de Cot proporcional à complexidade das mesmas sequências. A complexidade de um genoma pode ser determinada comparando os seus valores de Cot com os valores de Cot de um genoma conhecido. As técnicas de reassociação podem ser usadas para descobrir relações filogenéticas entre os organismos o que é muito importante para compreender a evolução. Estes estudos, à época, foram úteis para entender o paradoxo C.

9. Observe o gráfico da cinética de reassociação de um DNA genómico, isolado de células eucarióticas. O DNA foi quebrado em fragmentos de cerca de 5 000 pb, desnaturado pelo calor em cadeias simples e colocado em condições de renaturação/reassociação de acordo com a complementaridade de bases. O que conclui relativamente à composição e proporção dos diferentes tipos de sequências nucleotídicas deste genoma?



R. Observando o gráfico da cinética de reassociação desta molécula de DNA, é evidente um perfil de curva Cot que reflecte a complexidade química do DNA analisado, i.e., as três composições principais/típicas de sequências de DNA existentes num genoma eucariótico. Cada componente representa uma diferente proporção de DNA genómico total que se reassocia. Observa-se assim um componente em que o valor de Cot é baixo e que corresponderá a DNA que se reassocia muito rapidamente, contendo como tal DNA altamente repetitivo. Um segundo componente em que o DNA já será só moderadamente repetitivo, demorando um pouco mais a reassociar e o valor de Cot é um pouco mais elevado. E um terceiro componente de cópias únicas de DNA, sendo necessário mais tempo para que estas se encontrem e se reassociem e como tal com um valor Cot mais elevado.

10. Dê exemplos, baseados no genoma humano, que permitam distinguir entre DNA satélite, minissatélite e microsatélite.

R. O DNA satélite corresponde a uma fracção do genoma constituída por múltiplas sequências, curtas, agrupadas, altamente repetitivas, que apresentam propriedades particulares quando se analisa DNA por centrifugação em gradientes de densidade. Geralmente, está associado a regiões mais ou menos inertes dos cromossomas, regiões heterocromáticas, como sejam os centrómeros e, neste caso particular, é por vezes designado de DNA alféide. Os microsatélites e minissatélites também são sequências de DNA que se repetem em tandem, mas em geral não originam bandas de DNA satélite em gradientes de densidade. Os microsatélites correspondem a di- tri- ou tetranucleótidos (frequentemente com menos de 10 pb), que se repetem 10-20 vezes em tandem; também se

designam de SSRs (*simple sequence repeats*) ou STRs (*simple tandem repeats*). Um exemplo de microsatélite é o DNA repetitivo do receptor β das células T. Os minissatélites correspondem a sequências de DNA repetitivo, com uma extensão até cerca de 20 kb; também se designam de VNTR. Um exemplo de minissatélite é o DNA repetitivo dos telómeros.

2.5 Elementos móveis

Escolha múltipla

1. Que particularidade(s) não se aplica(m) aos elementos transponíveis? a) podem deixar uma cópia no local original, após a transposição; b) estão presentes num genoma, em geral em cópia simples; c) encontram-se em todos os membros de todas as espécies; d) podem não melhorar o *fitness* dos organismos hospedeiros; e) podem reverter um gene para o alelo selvagem, após transposição. b, c;

Questões básicas

1. O que é um elemento transponível?

R. Um elemento transponível é qualquer segmento de DNA que se move no genoma, independentemente do mecanismo, da sua origem e da sua função. Estes elementos não têm de ser sequências com uma determinada função no organismo. A classificação dos elementos transponíveis pode ser feita com base no modo como estes se movem no genoma. Assim, existem os retrotransposões que se movem através de um intermediário de RNA, e os transposões de DNA, que movem directamente o seu DNA, sem necessidade de um intermediário de RNA. Para além destes, ainda são elementos transponíveis as sequências de inserção existentes em bactérias. Estas sequências de DNA móvel também podem ser transferidas de uma célula para outra.

2. O que são sequências de inserção?

R. As sequências de inserção são elementos móveis de DNA, muito simples, ou seja, são transposões simples, que estão ladeados de curtas sequências invertidas, em que a maioria dos elementos contém o gene que codifica a transposase, enzima responsável pela transposição conservativa ou replicativa destes elementos.