

**Questão 3:** Na ausência de um genoma de referência, proponha um pipeline de análise para identificar genes diferencialmente expressos. Em sua resposta, descreva: a) as etapas necessárias para a montagem de novo do transcriptoma; b) as estratégias para quantificação da expressão dos transcritos; c) os principais cuidados estatísticos necessários para reduzir a ocorrência de falsos positivos na análise de expressão diferencial.

<b>Avaliador:</b>	<b>Valor</b>
<b>Conteúdo formal</b>	<b>0,0 a 6,0</b>
<b>a) Montagem de novo do transcriptoma:</b>	<b>0,0 a 2,5</b>
<b>a1. Controle de qualidade e pré-processamento das leituras</b> <b>O candidato deve mencionar:</b> - avaliação da qualidade dos reads; - remoção de adaptadores; Aprofundamento: - filtragem de bases/leitura de baixa qualidade; - eventual remoção de contaminantes.	0,0 a 0,5
<b>a2. Montagem <i>de novo</i> do transcriptoma</b> <b>O candidato deve mencionar:</b> - reconstrução de contigs/transcritos a partir das leituras; Aprofundamento: - possibilidade de isoformas redundância e fragmentação; - necessidade de parâmetros adequados de montagem.	0,0 a 1,0
<b>a3. Avaliação e refinamento da montagem</b> <b>O candidato deve mencionar:</b> - métricas de qualidade/completude (Busco) da montagem; - predição gênica; - anotação funcional para apoiar interpretação. Aprofundamento: - redução de redundância.	0,0 a 1,0
<b>b) Quantificação da expressão dos transcritos:</b>	<b>0,0 a 1,5</b>
<b>b1. Mapeamento ou alinhamento ou comparação das leituras contra o transcriptoma montado</b>	0,0 a 0,6
<b>b2. Estimativa de abundância de transcritos/genes</b> <b>O candidato deve mencionar:</b> - Obtenção de contagens de reads por transcrito. - Normalização para comparação entre amostras Aprofundamento: - TPM (transcripts per million), FPKM/RPKM (embora menos recomendados atualmente); - contagens brutas para métodos estatísticos.	0,0 a 0,6
<b>b3. Geração de matriz de contagens/abundâncias para comparação entre grupos</b> <b>O candidato deve mencionar:</b> - análise individual dos genes de interesse	0,0 a 0,3
<b>c) Cuidados estatísticos para redução de falsos positivos:</b>	<b>0,0 a 2,0</b>
<b>c1. Uso de réplicas biológicas e desenho experimental adequado (candidato deve mencionar ambos)</b>	0,0 a 0,7
<b>c2. Normalização apropriada entre amostras</b>	

<p><b>O candidato deve mencionar:</b></p> <ul style="list-style-type: none"> <li>- <b>composição das bibliotecas;</b></li> </ul> <p>Aprofundamento:</p> <ul style="list-style-type: none"> <li>- profundidade de sequenciamento.</li> </ul>	<p>0,0 a 0,6</p>
<p><b>c3. Testes estatísticos apropriados para expressão diferencial</b></p> <p><b>O candidato deve mencionar:</b></p> <ul style="list-style-type: none"> <li>- <b>Controle de múltiplos testes: aplicação de correção de p-values para múltiplas comparações (FDR);</b></li> <li>- <b>Validação complementar: sugestão de validação experimental (RT-qPCR) para um subconjunto de genes diferencialmente expressos;</b></li> </ul> <p>Aprofundamento:</p> <ul style="list-style-type: none"> <li>- modelagem da variabilidade biológica;</li> <li>- Remoção de transcritos com baixa contagem total</li> </ul>	<p>0,0 a 0,7</p>
<p><b>Aprofundamento do conteúdo</b></p> <p><b>Para obter a pontuação completa, o candidato deve se aprofundar 5 vezes, conforme exemplificado nos gabaritos dos itens de conteúdo formal. Outros exemplos de aprofundamento:</b></p> <ul style="list-style-type: none"> <li>- dificuldade de separar isoformas em montagem de novo;</li> <li>- dificuldade de montagem com reads curtos (ambiguidade de mapeamento);</li> <li>- uso de uma plataforma ou mais de uma (Ex.: Illumina, Nanopore, PacBio); impacto de transcriptomas incompletos na quantificação;</li> <li>- diferença entre análise em nível de gene e de transcrito;</li> <li>- filtragem de baixa expressão antes do teste;</li> <li>- avaliação de batch effects;</li> <li>- inspeção exploratória dos dados, como PCA/clusterização;</li> <li>- validação funcional ou experimental posterior;</li> <li>- cautela na interpretação de “genes diferencialmente expressos” quando a reconstrução do transcriptoma é imperfeita.</li> </ul>	<p>0,0 a 2,0</p>
<p><b>Encadeamento lógico da resposta</b></p> <p><b>Critérios:</b></p> <ul style="list-style-type: none"> <li>- resposta apresentada em sequência coerente:</li> </ul> <p>Biblioteca → QC → montagem → avaliação/refino → quantificação → normalização → teste estatístico → controle de falsos positivos;</p> <ul style="list-style-type: none"> <li>- distinção clara entre etapas computacionais e estatísticas;</li> <li>- ausência de contradições importantes.</li> </ul>	<p>0,0 a 1,0</p>
<p><b>Correção linguística (ortografia, pontuação, concordância, regência, acentuação e construção frasal)</b></p> <p>Clareza na exposição, ausência de erros gramaticais ou ortográficos.</p>	<p>0,0 a 1,0</p>

**Questão 7:** Considere uma proteína-alvo de interesse biológico com estrutura experimental parcialmente resolvida, incluindo regiões sem coordenadas definidas e potencialmente relevantes para a ligação a ligantes. Considere também um conjunto de pequenas moléculas candidatas previamente selecionadas.

Proponha um fluxo de trabalho computacional para priorizar os ligantes com maior probabilidade de interação estável e específica com essa proteína, integrando modelagem estrutural, docking molecular e dinâmica molecular.

Em sua resposta, descreva:

- quatro etapas essenciais de preparação do sistema antes do docking;
- três critérios para avaliação crítica das poses obtidas no docking;
- como a dinâmica molecular pode ser utilizada para avaliar a estabilidade e a robustez de uma pose de ligação;
- as principais limitações do pipeline proposto.

<b>Avaliador:</b>	<b>Valor</b>
<b>Conteúdo formal</b>	<b>0,0 a 6,0</b>
<b>a) Quatro etapas essenciais de preparação antes do docking:</b>	<b>0,0 a 2,5</b>
<b>a1. Recuperação e inspeção da estrutura da proteína</b> <b>Deve incluir ambas:</b> - avaliação da estrutura experimental disponível; - identificação de lacunas, resíduos ausentes, cadeias incompletas ou baixa resolução.	0,0 a 0,5
<b>a2. Modelagem/refinamento das regiões ausentes ou mal resolvidas</b> <b>Deve incluir ambas:</b> - reconstrução de loops/regiões sem coordenadas; - justificativa de que essas regiões podem afetar o sítio de ligação.	0,0 a 0,5
<b>a3. Preparação da proteína para docking</b> <b>O candidato deve mencionar:</b> - adição de hidrogênios; - definição de estados de protonação; - tratamento de cofatores, metais, águas estruturais; Aprofundamento: - escolha da forma oligomérica relevante; - minimização/refinamento local.	0,0 a 0,5
<b>a4. Definição do sítio de ligação e preparação do espaço de busca</b> <b>Deve incluir ambas:</b> - uso de informação experimental, resíduos-chave, cavidades ou ligantes conhecidos; - definição do espaço experimental.	0,0 a 0,5
<b>a5. Preparação do conjunto de ligantes</b> <b>O candidato deve mencionar:</b> - geração de conformações; - protonação/tautomeria; Aprofundamento: - minimização.	0,0 a 0,5
<b>b) Três critérios para avaliação crítica das poses:</b>	<b>0,0 a 1,5</b>
<b>b1. Coerência fisicoquímica das interações com o sítio</b> <b>O candidato deve mencionar:</b> - avaliação das interações químicas; Aprofundamento: - ligações de hidrogênio plausíveis, interações hidrofóbicas, iônicas, coordenação com metais, complementaridade estérica.	0,0 a 0,5
<b>b2. Ocupação consistente do sítio e orientação plausível do ligante</b>	

<p>Aprofundamento:</p> <ul style="list-style-type: none"> <li>- ausência de choques estéricos;</li> <li>- encaixe em região biologicamente relevante;</li> <li>- orientação compatível com resíduos-chave.</li> </ul>	<p>0,0 a 0,5</p>
<p><b>b3. Consistência entre escore e interação estrutural</b></p> <p><b>O candidato deve explicar que não se pode aceitar ranking apenas pelo escore</b></p> <p>Aprofundamento:</p> <ul style="list-style-type: none"> <li>- comparar poses, consenso entre métodos, reprodutibilidade da pose</li> <li>- comparação com ligantes de referência;</li> <li>- diversidade/convergência de poses;</li> <li>- estabilidade preliminar após minimização;</li> <li>- compatibilidade com dados experimentais existentes.</li> </ul>	<p>0,0 a 0,5</p>
<p><b>c) Uso de dinâmica molecular para avaliar estabilidade/robustez:</b></p> <p><b>O candidato deve mencionar:</b></p> <ul style="list-style-type: none"> <li>- observação da permanência do ligante no sítio;</li> <li>- análise de estabilidade da pose ao longo do tempo;</li> <li>- monitoramento de RMSD, contatos, ligações de hidrogênio;</li> <li>- eventual estimativa energética comparativa entre candidatos;</li> </ul> <p>Aprofundamento:</p> <ul style="list-style-type: none"> <li>- inserção do complexo proteína-ligante em sistema solvatado, com íons;</li> <li>- minimização, equilíbrio e produção da trajetória.</li> </ul>	<p>0,0 a 1,0</p>
<p><b>d) Limitações do pipeline:</b></p> <p><b>O candidato deve mencionar:</b></p> <ul style="list-style-type: none"> <li>- dependência da qualidade da estrutura inicial;</li> <li>- amostragem incompleta de conformações da proteína e do ligante;</li> <li>- ausência de validação experimental;</li> </ul> <p>Aprofundamento:</p> <ul style="list-style-type: none"> <li>- incerteza da modelagem das regiões ausentes;</li> <li>- custo computacional da dinâmica molecular;</li> <li>- dificuldade de inferir afinidade real apenas por métodos computacionais.</li> </ul>	<p>0,0 a 1,0</p>
<p><b>Aprofundamento do conteúdo</b></p> <p><b>Para obter a pontuação completa, o candidato deve se aprofundar 5 vezes, conforme exemplificado nos gabaritos dos itens de conteúdo formal. Outros exemplos de aprofundamento:</b></p> <ul style="list-style-type: none"> <li>- discussão da flexibilidade do receptor;</li> <li>- justificativa da importância das regiões ausentes para o sítio de ligação;</li> <li>- uso de mais de um modelo estrutural/conformação (ensemble);</li> <li>- necessidade de controles positivos/ligantes conhecidos;</li> <li>- réplica de simulações ou comparação entre múltiplas poses;</li> <li>- cautela em interpretar estabilidade em DM como prova de afinidade.</li> </ul>	<p>0,0 a 2,0</p>
<p><b>Encadeamento lógico da resposta</b></p> <p><b>Critérios:</b></p> <ul style="list-style-type: none"> <li>- resposta organizada em fluxo:</li> </ul> <p>inspeção estrutural → modelagem/refinamento → preparação de proteína e ligantes → docking → análise crítica de poses → dinâmica molecular → priorização final → limitações;</p> <ul style="list-style-type: none"> <li>- separação clara entre preparação, avaliação e limitações;</li> <li>- resposta compatível com o que foi perguntado em cada item.</li> </ul>	<p>0,0 a 1,0</p>
<p><b>Correção linguística (ortografia, pontuação, concordância, regência, acentuação e construção frasal)</b></p> <p>Clareza na exposição, ausência de erros gramaticais ou ortográficos.</p>	<p>0,0 a 1,0</p>

**Questão 10:** Proponha uma estratégia didática para o ensino de biologia computacional no nível superior, considerando a diversidade de formação prévia dos estudantes.

Em sua resposta, descreva como essa estratégia pode incorporar:

- a) a heterogeneidade do perfil dos alunos;
- b) o uso de dados reais;
- c) atividades de aprendizagem prática;
- d) a formação em boas práticas computacionais e reprodutibilidade.

**Avaliador:**

**Valor**

**Conteúdo formal**

**0,0 a 6,0**

**a) Heterogeneidade do perfil dos alunos:**

Espera-se que o candidato reconheça que turmas de biologia computacional frequentemente incluem estudantes com formações distintas, por exemplo:

- biologia;
- computação;
- estatística;
- química;
- áreas da saúde.

0,0 a 1,5

Devem ser mencionadas pelo menos 3 estratégias, como:

- diagnóstico inicial de conhecimentos;
- nivelamento ou trilhas de aprendizagem;
- progressão do simples ao complexo;
- integração entre conceitos biológicos e computacionais;
- trabalho colaborativo entre perfis distintos.

**b) Uso de dados reais:**

O candidato deve mencionar pelo menos 2:

- utilização de bases públicas ou conjuntos reais de dados biológicos;
- formulação de perguntas biológicas autênticas;
- exposição dos alunos a dados imperfeitos e problemas reais;
- aproximação entre ensino, pesquisa e análise crítica.

0,0 a 1,5

**c) Atividades de aprendizagem prática:**

O candidato deve mencionar:

- inclusão de aulas práticas computacionais;
- resolução de problemas e estudos de caso;

Aprofundamento:

- uso de scripts, notebooks, pipelines ou ferramentas apropriadas;
- eventual desenvolvimento de mini-projetos.

0,0 a 1,5

**d) Boas práticas computacionais e reprodutibilidade**

O candidato deve mencionar:

- documentação de procedimentos;
- controle de versão;

Aprofundamento:

- organização de arquivos e dados;
- registro de parâmetros;
- compartilhamento de código e resultados;
- noções de ambiente computacional reprodutível;
- importância ética e científica da reprodutibilidade.

0,0 a 1,5

**Aprofundamento do conteúdo**

<p><b>Para obter a pontuação completa, o candidato deve se aprofundar 4 vezes, conforme exemplificado nos gabaritos dos itens de conteúdo formal. Outros exemplos de aprofundamento:</b></p> <ul style="list-style-type: none"> <li>- proponham uma estrutura pedagógica concreta, por exemplo em módulos;</li> <li>- articulem objetivos de aprendizagem, avaliação e prática;</li> <li>- distingam competências mínimas e avançadas.</li> <li>- avaliação formativa e somativa;</li> <li>- aprendizagem baseada em problemas ou projetos;</li> <li>- uso de Jupyter, R Markdown, Git ou ferramentas equivalentes como exemplos;</li> <li>- discussão do equilíbrio entre ferramentas gráficas e programação;</li> <li>- atenção à acessibilidade e acolhimento de estudantes com pouca experiência computacional;</li> <li>- tratamento e uso de IA.</li> </ul>	<p>0,0 a 2,0</p>
<p><b>Encadeamento lógico da resposta</b></p> <p><b>Critérios:</b></p> <ul style="list-style-type: none"> <li>- apresentação de uma proposta pedagógica coerente;</li> <li>- articulação entre perfil dos alunos, atividades, dados reais e reprodutibilidade;</li> <li>- ausência de resposta fragmentada ou apenas enumerativa</li> </ul>	<p>0,0 a 1,0</p>
<p><b>Correção linguística (ortografia, pontuação, concordância, regência, acentuação e construção frasal)</b></p> <p>Clareza na exposição, ausência de erros gramaticais ou ortográficos.</p>	<p>0,0 a 1,0</p>
<p><b>Nota final da questão</b></p>	