

Projeto **Origem de Contigs** de IBI5035/QBQ2507 2020s2

1. Dados para o projeto
 - a. conjunto de sequências de DNA: 10 contigs genômicos de tamanhos entre 10 e 30kb
 - b. cada aluno terá seu próprio conjunto
 - c. disponível em <http://www.iq.usp.br/setubal/bmc/2020> (tabela com links para dados de cada aluno)
2. Para entregar
 - a. um **relatório individual** em PDF descrevendo sua análise das suas sequências
 - b. **o nome do PDF contendo o relatório deve ser seu nome e sobrenome concatenados**
 - i. Exemplo: joaoSetubal.pdf
 - c. esse arquivo deve ser enviado ao professor por email até **meia-noite** de **30/11/2020**.
Use como assunto do email as palavras “projeto BMC”
3. Itens que o relatório deve conter
 - a. A estrutura do relatório obrigatoriamente deve ser a seguinte:
 - i. **introdução**: mencione o contexto do relatório (“relatório apresentado para a disciplina Biologia Molecular Computacional 2020s2”) e apresente sua tabela de contigs (identificador de cada contig e tamanho em bp)
 - ii. **conclusão** sobre a origem dos contigs. Esta parte pode e deve ser **curta**.
 - iii. **justificativa** sobre a conclusão apresentada, evidenciada pelos resultados de suas análises. Minha expectativa é de que esta parte será **detalhada**, incluindo tabelas.
 - iv. **Métodos** utilizados. Aqui vc deve descrever quais métodos usou para chegar à sua conclusão. (Exemplo: usei blastn [cite o paper ou URL] contra o banco nt do NCBI; interproscan [cite o paper ou URL] das ORFs de cada contig na versão online; etc). Os **critérios utilizados** para análise dos resultados também devem ser descritos.
 - v. **Referências**
 - b. **Cada relatório não pode ter mais do que 20 páginas**
 - c. Na parte de **conclusão**, vc deve apresentar **uma única** hipótese sobre a origem das suas sequências: 1) **compostagem**; 2) **fezes de macacos bugios**; 3) **água de um lago**; 4) **esponjas do mar**; 5) **solo permafrost**; 6) indeterminado (não foi possível chegar a uma conclusão). É importante notar que todas as sequências de um conjunto atribuído a um aluno vieram de um **mesmo ambiente**. Então as hipóteses de 1 a 5 são mutuamente exclusivas.
 - d. Parte de **justificativa**: esta é a parte **mais importante** do relatório. Vc deve não apenas apresentar resultados das análises computacionais que fez sobre os contigs, mas também informações biológicas sobre a fonte inferida que embasem sua conclusão. Por exemplo: se um dos conjuntos de dados fosse microbioma humano a partir de fezes, e se vc conseguiu identificar que um dos contigs pertence à bactéria *Bifidobacterium longum*, uma rápida pesquisa indicaria que essa bactéria é comumente encontrada no trato gastro-intestinal humano, sendo portanto uma boa evidência de que a amostra

analisada veio do microbioma humano. Se sua conclusão for (6), vc deve também ser capaz de oferecer bons argumentos sobre essa conclusão. Se esses argumentos forem considerados fracos (a análise deveria ter permitido concluir a origem), o aluno não terá crédito nesta parte do relatório.

Prazo de entrega do relatório: **30/11/2020**. Use como assunto do email as palavras “projeto BMC”

Critérios de avaliação dos relatórios

1. A conclusão foi certa ou errada? [20%]
2. qualidade da argumentação sobre sua conclusão. Aqui vou levar em conta pesquisa sobre os organismos e ambientes relevantes. [30%]
3. metodologia empregada na análise das sequências. Aqui quero demonstração de esforço para fazer a análise. Um aluno que se limite a fazer blast das sequências terá nota menor do que outro que use uma bateria extensa de ferramentas de análise (supondo que essa bateria seja apropriada). Também levarei em conta esforço no sentido de aproveitar todos os dados disponibilizados (ou seja, todos os 10 contigs e os genes neles contidos) [40%]
4. qualidade da apresentação do relatório em termos de escrita, figuras, tabelas, referências [10%]
5. atrasos na entrega implicarão em desconto na nota independente da qualidade do relatório. Esse desconto será diretamente proporcional ao número de dias de atraso.
6. Relatórios com mais do que o limite de 20 páginas também serão penalizados