

## Disciplina QBQ2507 / IBI5035 Biologia Molecular Computacional 2024s2

Prof. João Carlos Setubal

### Tarefa Prática (TP)

Esta tarefa é individual. Cada aluno poderá escolher entre as opções resumidas na tabela a seguir e explicadas depois.

opção	nome	versão local	versão web
1	análise de famílias gênicas ( <b>cap2</b> )	disponível	bv-brc.org
2	comparação entre genomas bacterianos ( <b>cap5</b> )	disponível	bv-brc.org
3	análise de MAGs ( <b>cap6</b> )	disponível	KBase, bv-brc.org, MGnify
4	comparação entre genomas bacterianos ( <b>artigo</b> )	não disponível	obrigatoriamente pela plataforma IMG/M

### Detalhamento

Na página da disciplina estão disponibilizados os seguintes textos:

- **Cap2:** Protein-Coding Gene Families in Prokaryote Genome Comparisons
- **Cap5:** Step-by-step Bacterial Genome Comparison
- **Cap6:** How to obtain and compare Metagenome-Assembled Genomes
- **Artigo:** Seshadri et al., Comparative Genomics Using the Integrated Microbial Genomes and Microbiomes (IMG/M) System: A *Deinococcus* Use Case

Para realizar esta tarefa prática cada aluno terá que escolher uma das 7 opções acima. Além disso, cada aluno terá que escolher seu próprio conjunto de genomas (*com pelo menos 7 genomas*), que deve ser **único e diferente** de todos os demais conjuntos escolhidos pelos outros alunos, para as opções de 1, 2 e 4. Para a opção 3 o aluno deverá escolher um conjunto de dados metagenômicos, **que também deve ser único**.

Para assegurar unicidade, cada aluno terá que informar à monitora da disciplina, Daniela Bizinelli, qual sua opção de tarefa e qual o conjunto de genomas ou metagenoma que analisará. Isto deve ser feito até dia **9/9**, meio dia, hora de São Paulo, por meio de e-mail à monitora ([daniela.bizinelli@usp.br](mailto:daniela.bizinelli@usp.br)). Caso a escolha de um aluno coincida com uma escolha já feita por outro aluno, o monitor informará o aluno, que terá que fazer nova escolha.

**NB1:** para a opção 1 é importante escolher organismos (bactérias ou arqueias) bastante diversos, como por exemplo organismos de filos diferentes. Para as opções 2 e 4, recomenda-se escolher diferentes cepas de uma mesma espécie ou diferentes espécies de um mesmo gênero.

**NB2:** Alunos que não informarem sua opção e seu conjunto de dados até o prazo de 9/9 sofrerão penalização na nota da TP.

### O que significa “versão local”

Os Capítulos 2, 5 e 6 apresentam comandos em linux e R para realizar as comparações descritas. Um aluno que escolha esta opção precisa saber linux e R e ter acesso a um computador com esses recursos, além da capacidade de instalar programas nesse computador.

### O que significa “versão web”

Neste caso, os alunos terão que seguir as instruções dos capítulos ou artigo indicados, realizando as análises apresentadas nos capítulos nas plataformas web indicadas. Caso haja uma análise do capítulo que não possa ser executada na plataforma escolhida, procure fazer de alguma outra forma (veja o exemplo abaixo de pangenoma). A falta de uma análise de sua opção poderá acarretar penalização, a menos que esta falta esteja bem justificada.

Plataformas web recomendadas:

- <https://www.bv-brc.org/>
  - complementado com <https://panexplorer.southgreen.fr/cgi-bin/home.cgi>
- <https://www.kbase.us/>
- <https://img.jgi.doe.gov/cgi-bin/m/main.cgi>
- <https://www.ebi.ac.uk/metagenomics>

## Relatórios

Cada aluno deve entregar um relatório conforme as instruções a seguir.

1. Relatórios são individuais
2. Seu relatório deve ser um único arquivo PDF
3. O nome desse arquivo deve ser o seu nome, sem mais nenhuma outra informação. Por exemplo: joasetubal.pdf
4. O arquivo deve ser enviado no prazo (24/10/2024, meio dia, hora de SP) para [setubal@iq.usp.br](mailto:setubal@iq.usp.br)
5. Use como **assunto** a seguinte cadeia de caracteres (use apenas minúsculas e inclua os asteriscos): **\*tarefapratica\***
6. Você receberá uma resposta automática, que será seu **recibo de entrega**. *Sem esse recibo, não considerarei que você entregou o relatório*

7. A primeira página do relatório deve informar seu nome, número USP, data, nome da disciplina, e qual das 7 possíveis opções foi escolhida (por exemplo: opção 1-web). **A primeira informação técnica que precisa constar do relatório é o conjunto de genomas escolhidos (opções 1, 2 e 4) ou o conjunto de dados metagenômicos escolhido (opção 3).** O resto do relatório é formato livre (mas veja os critérios abaixo).

#### **Critérios para correção dos relatórios**

- aderência ao formato acima especificado
- se as análises feitas correspondem ao esperado para a opção escolhida
- a existência de uma seção de discussão/interpretação dos resultados obtidos
- estruturação e legibilidade