



Dicas para realizar projeto Origem de Contigs

João Carlos Setubal 2020 disciplina QBQ2507-IBI5035

Invstigue seus contigs com

- BLAST (NCBI)
 - blastn, blastx
- ORFfinder

ORFfinder

- Identificação de ORFs nos contigs
 - https://www.ncbi.nlm.nih.gov/orffinder/
- Para cada contig, analise tantas ORFs quanto necessário para chegar a uma conclusão robusta
 - use tamanho mínimo de 300 bp
 - Faça blastp nas ORFs promissoras
 - Rode Interproscan nas ORFs promissoras (para predizer função)

https://www.ebi.ac.uk/interpro/



InterPro: protein sequence analysis & classification

InterPro provides functional analysis of proteins by classifying them into families and predicting domains and important sites. We combine protein signatures from a number of member databases into a single searchable resource, capitalising on their individual strengths to produce a powerful integrated database and diagnostic tool. Read more about InterPro



Documentation

About InterPro: core concepts, update frequency, how to cite, team and consortium members.

<u>FAQs</u>: what are entry types and why are they important, interpreting results, downloading InterPro?

Web services documentation

Tools InterProScan



InterProScan is a sequence analysis application (nucleotide and protein

Protein focus

☑ Zika Virus and Microcephaly
Feb 2016

You have probably been as horrified and saddened as me to see the shocking abnormality that affects newborn babies whose ...

The InterPro Domain

Architecture (IDA) tool

allows you to search the

HTML | All protein focus

Tools IDA

Publications



The InterPro protein families
database: the classification
resource after 15 years
Our latest paper describing new
developments on the InterPro

website (*Nucleic Acids Research*, Jan 2015).

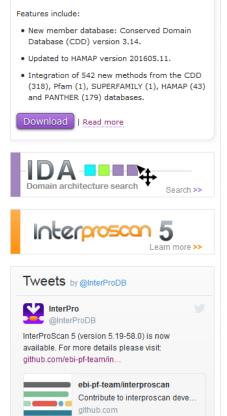
THTML | PDF (3,4Mb) | All publications

Bloa



A new InterPro member database
Jun 2016

CDD joins InterPro We are pleased



InterPro 58.0 9th June 2016

Artemis (opcional)

- browser de genomas (ou de contigs) que pode ser rodado localmente
- http://www.sanger.ac.uk/science/tools/artemis
- Tutorial disponibilizado na página da disciplina

Critérios de identificação de organismos com base em BLAST

- Se a comparação foi de nucleotídeos
 - Para <u>espécie</u>, no mínimo 97% de identidade e no mínimo 100 nt de alinhamento
 - Entre 90 e 96% pode-se fazer uma associação com o gênero ("parente")
 - Abaixo de 90% não é possível fazer afirmações categóricas; mas é possível especular
- Se a comparação foi de aminoácidos
 - Em geral não é possível inferir o organismo correspondente, mas é possível specular, especialmente levando em conta também os resultados de blastn

Espero ver informações sobre os micro-organismos identificados

- Pesquisar a literatura sobre organismos positivamente identificados (mas também os parentes) e apresentar os resultados dessa pesquisa no relatório (com referências)
- Há alguma coerência entre as diversas identificações dos diferentes contigs?

CDSs podem dar dicas adicionais

- Há uma coerência entre bons hits de blastp e bons hits de blastn?
- Há funções identificadas (dos produtos dos genes) que são coerentes com o ambiente da sua hipótese?

Outras ferramentas

- Existem diversas ferramentas de análises genômicas disponíveis na internet, além das aqui mencionadas
- Todos os alunos são encorajados a usar ferramentas não mencionadas e apresentar seus resultados, caso estes ajudem na geração da hipótese final

Apresente suas evidências!

- Exemplos
 - Mostre os alinhamentos <u>principais</u> que ajudaram vc a chegar em sua conclusão
 - Indique
 - % de identidade
 - E-value