

Apresentação v0.6

GENE ONTOLOGY

Outubro/2015

IN1047 Tópicos Avançados em Lógica e Sistemas Dedutivos 1

Docente: Fred Luiz (fred@cin.ufpe.br)

Discentes: Cloves Rocha (car2@cin.ufpe.br)



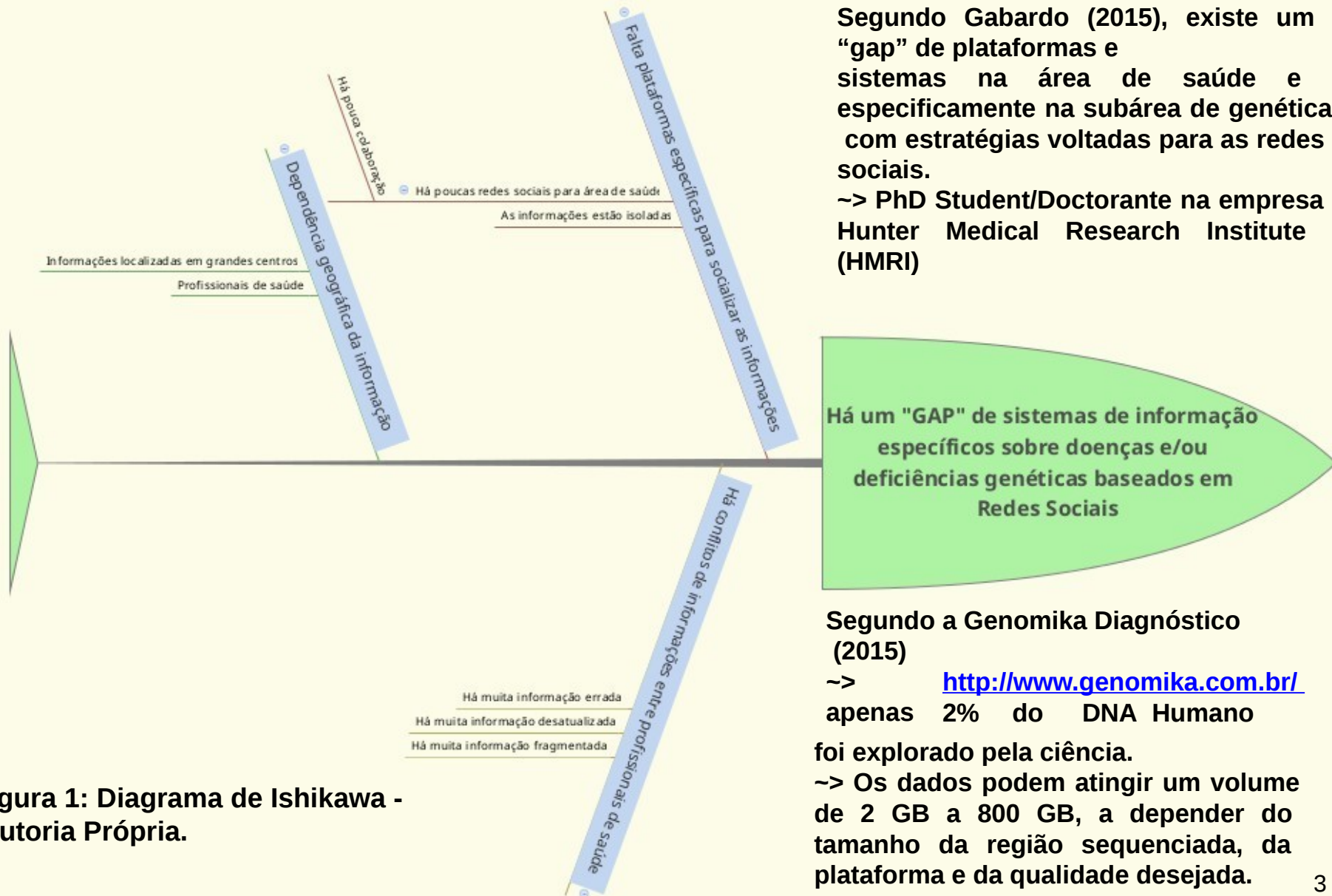
UNIVERSIDADE
FEDERAL
DE PERNAMBUCO

Cln.ufpe.br

Sumário:

- Contexto: Qual é o Problema de Pesquisa?
- Motivação e Justificativa
- Grande área Ontologias Biomédicas
- O que é Gene Ontology (GO)?
- Objetivos
- Anotar Genes e Produtos de Genes
- Exemplos
- Conclusão
- Referências

Contexto: Qual é o Problema de Pesquisa?



Segundo Gabardo (2015), existe um "gap" de plataformas e sistemas na área de saúde e especificamente na subárea de genética com estratégias voltadas para as redes sociais.

~> PhD Student/Doctorante na empresa Hunter Medical Research Institute (HMRI)

Há um "GAP" de sistemas de informação específicos sobre doenças e/ou deficiências genéticas baseados em Redes Sociais

Segundo a Genomika Diagnóstico (2015)

~> <http://www.genomika.com.br/> apenas 2% do DNA Humano

foi explorado pela ciência.

~> Os dados podem atingir um volume de 2 GB a 800 GB, a depender do tamanho da região sequenciada, da plataforma e da qualidade desejada.

Figura 1: Diagrama de Ishikawa - Autoria Própria.

Motivação e Justificativa

~> As redes sociais estão ganhando destaque pela capacidade de disseminar, concentrar e distribuir informações. A utilização destes meios para busca de produtos e/ou conhecimento relacionado a empresas, marcas e produtos mostrou-se crescente.

~> De acordo com Cipriani (2011), a possibilidade de contato com pessoas desconhecidas é grande no ambiente de sites de redes sociais;

~> Identificar quais são os principais problemas enfrentados pelas pessoas que de alguma forma estão envolvidas com doenças e/ou deficiências genéticas na obtenção de informações específicas sobre esse tema.

~> A finalidade é elencar os requisitos para o desenvolvimento de um sistema de descoberta informações genéticas para que essas pessoas tenham esses problemas atenuados, ou seja, com as informações menos descentralizadas.



~> G6PD <~

Grande área Ontologias Biomédicas

OBO Foundry

Os formulários da biblioteca OBO Ontologia base na OBO Foundry, uma experiência de colaboração que envolve um grupo de desenvolvedores da ontologia que concordaram em avançar para a adoção de um conjunto cada vez maior de princípios que especifiquem as melhores práticas no desenvolvimento de ontologias.

Projetos:

1. A Ontologia de Investigações Biomédicas (OBI) é um acesso aberto
- 2. Gene Ontology Consortium**
3. Sequência Ontologia
4. Genéricos organismo modelo Databases
5. Normas e ontologias para a genómica funcional
6. FGED
7. Ontologia de Investigações Biomédicas
8. Ontologia planta
9. Phenoscape
- 10. OBO e Semantic Web**
11. OBO & coruja ida e volta Transformações

O que é Gene Ontology (GO)?

Ontologia Gene (GO) é um dos principais temas da bioinformática, que trata de unificar a representação do gene e produto do gene, ou seja, atributos em todas as espécies. **Consórcio Gene Ontology (Janeiro de 2008)**

Mais especificamente, o projeto tem como objetivos:

- 1) Manter e desenvolver o seu **vocabulário controlado de gene e produto do gene atributos;**
- 2) **Anotar genes e produtos de genes** e **assimilar e divulgar dados** de anotação;
- 3) Fornecer ferramentas para facilitar o acesso a todos os aspectos dos dados fornecidos pelo projeto, e de permitir a interpretação funcional dos dados experimentais utilizando o GO, por exemplo, **através de análise de enriquecimento.**

Objetivos

Vocabulário Controlado de Gene e Produto do Gene

```
format-version: 1.2
subsetdef: goantisl原因_grouping "Grouping classes that can be excluded"
subsetdef: gocheck_do_not_annotate "Term not to be used for direct annotation"
subsetdef: gocheck_do_not_manually_annotate "Term not to be used for direct manual annotation"
subsetdef: goslim_aspergillus "Aspergillus GO slim"
subsetdef: goslim_candida "Candida GO slim"
subsetdef: goslim_chembl "ChEMBL protein targets summary"
subsetdef: goslim_generic "Generic GO slim"
subsetdef: goslim_goa "GOA and proteome slim"
subsetdef: goslim_metagenomics "Metagenomics GO slim"
subsetdef: goslim_pir "PIR GO slim"
subsetdef: goslim_plant "Plant GO slim"
subsetdef: goslim_pombe "Fission yeast GO slim"
subsetdef: goslim_synapse "synapse GO slim"
subsetdef: goslim_virus "Viral GO slim"
subsetdef: goslim_yeast "Yeast GO slim"
subsetdef: gosubset_prok "Prokaryotic GO subset"
subsetdef: mf_needs_review "Catalytic activity terms in need of attention"
subsetdef: termgenie_unvetted "Terms created by TermGenie that do not follow a template and require additional vetting by editors"
subsetdef: virus_checked "Viral overhaul terms"
synonymtypedef: systematic_synonym "Systematic synonym" EXACT
ontology: go/subsets/goslim_generic
```

Fonte: ~> http://geneontology.org/ontology/subsets/goslim_generic.obo

Objetivos

Vocabulário Controlado de Gene e Produto do Gene

[Term]

id: GO:0001071

name: nucleic acid binding transcription factor activity

namespace: molecular_function

def: "Interacting selectively and non-covalently with a DNA or RNA sequence in order to modulate transcription. The transcription factor may or may not also interact selectively with a protein or macromolecular complex."

[GOC:txnOH]

comment: Note that this term is in the subset of terms that should not be used for direct gene product annotation. This term does not provide specificity with respect to the type of nucleic acid binding, e.g. RNA or DNA. Please use a child term that provides that specificity or, if no appropriate child term exists, please request a new term. Direct annotations to this term may be amended during annotation QC.

subset: gocheck_do_not_manually_annotate

subset: goslim_chembl

subset: goslim_generic

subset: goslim_metagenomics

subset: goslim_yeast

synonym: "transcription factor activity" BROAD []

is_a: GO:0003674 ! molecular_function

relationship: has_part GO:0003674 ! **molecular_function**

relationship: part_of GO:0008150 ! **biological_process**

created_by: kchris

creation_date: 2010-10-21T04:37:54Z

Fonte: ~> http://geneontology.org/ontology/subsets/goslim_generic.obo

Objetivos

Vocabulário Controlado de Gene e Produto do Gene

[Typedef]

id: ends_during

name: ends_during

namespace: external

xref: RO:0002093

[Typedef]

id: happens_during

name: happens_during

namespace: external

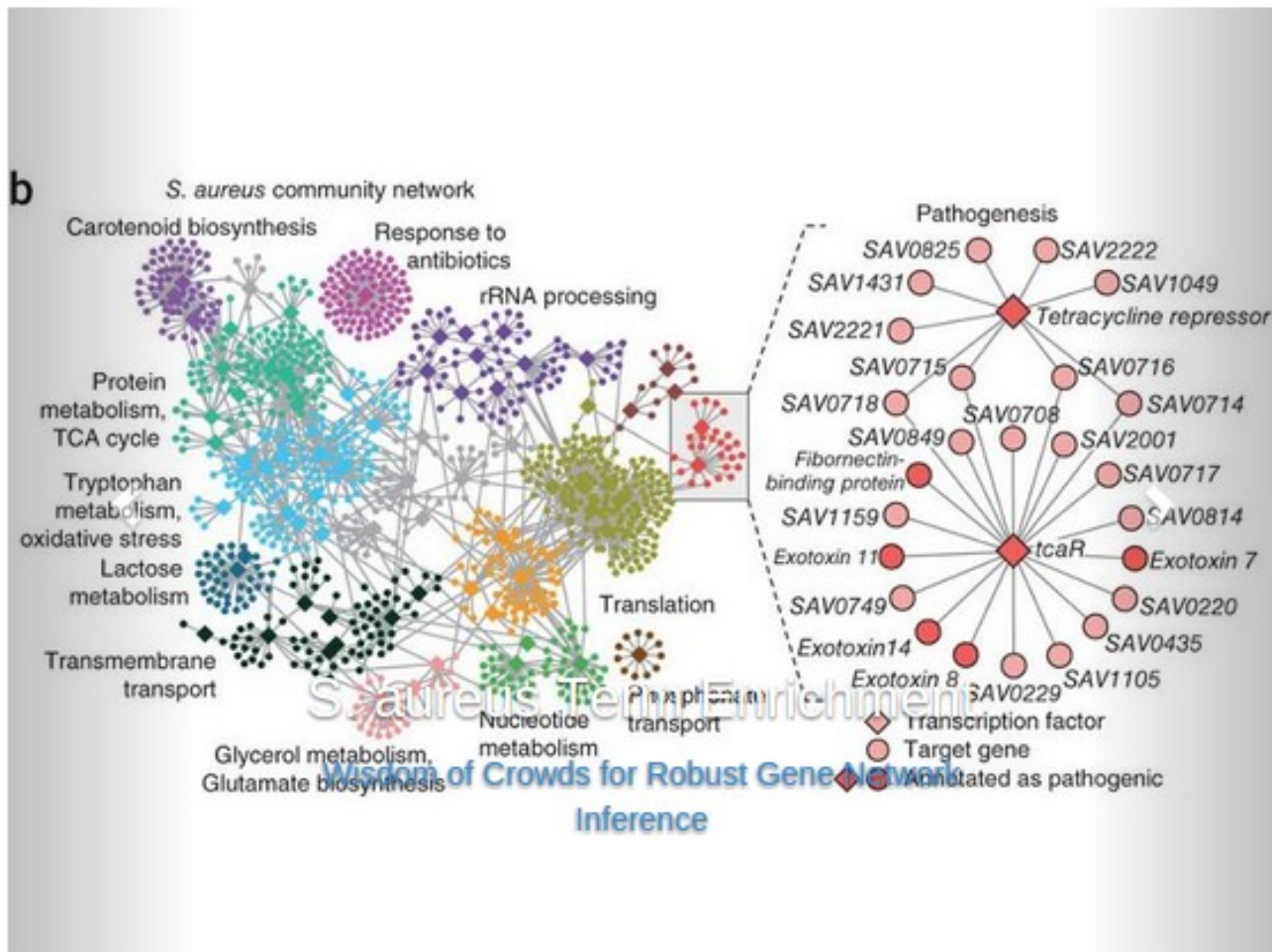
xref: RO:0002092

is_transitive: true

is_a: ends_during ! ends_during

Fonte: ~> http://geneontology.org/ontology/subsets/goslim_generic.obo

Anotar Genes e Produtos de Genes



Fonte: GeneOntology.org

Anotar Genes e Produtos de Genes

O código evidência vem do Evidence Código Ontologia, um vocabulário controlado de códigos que abrangem ambos os métodos de anotação manuais e automatizados.

Por exemplo, Declaração Autor rastreável (TAS) significa um curador tenha lido um artigo científico publicado e os metadados para essa anotação carrega uma citação para esse papel; inferida a partir de Sequência de Similaridade (ISS), um curador humano reviu a saída de uma semelhança de sequência pesquisar e verificou que ele é biologicamente significativa. Anotações de processos automatizados (por exemplo, as anotações remapeamento criados usando um outro vocabulário anotação) é dado o código inferida a partir de anotações eletrônico (IEA).

Desde 1 de Abril de 2010, **mais de 98% de todas as GO anotações foram inferidas computacionalmente, não por curadores.**

Fonte: GeneOntology.org

Exemplo: Desenho na Ferramenta - Protege

The screenshot displays the Protege ontology editor interface. The top menu bar includes: Active Ontology, Entities, Classes, Object Properties, Data Properties, Annotation Properties, Individuals, OWLViz, DL Query, OntoGraf, Ontology Differences, and SPARQL Query.

The main window is divided into three panes:

- Class hierarchy (left):** Shows a tree structure starting with 'Thing'. Under 'Doença_Genética', there are several categories: 'Causas', 'Classificação', 'Distúrbios_Cromossômicos' (containing 'Síndrome_de_Cri-du-chat', 'Síndrome_de_Turner', 'Síndrome_de_Wolf-Hirschhorn', and 'Trissomia'), 'Distúrbios_Multifatoriais' (containing '08_Síndrome_de_Warkany', '13_Síndrome_de_Patau', '18_Síndrome_de_Edward', '21_Síndrome_de_Down', 'XXX_Síndrome_do_Triplo_X', and 'XXY_Síndrome_de_Klinefelter'), 'Alzheimer', 'cardiopatias_congênicas', 'Câncer', 'Diabetes_Mellitus', 'Hipertensão_Arterial', 'Mal_formações_congênicas', and 'Obesidade'. Other categories include 'Monogenéticas', 'Herança', 'Organizações_de_Apoio', 'População_Afetada', 'Referências', 'Sinais_e_Sintomas', 'Tratamento', 'Interesse', and 'Pessoa'.
- Annotations (top right):** Shows two annotations for 'Doença_Genética': 'CN_pl Doença_Genética' and 'CN_sg Doença_Genética'. Each annotation has control icons for adding, deleting, and refreshing.
- Description (bottom right):** Shows the 'Description: Doença_Genética' pane with various logical relationships: 'Equivalent To', 'SubClass Of', 'SubClass Of (Anonymous Ancestor)', 'Members', 'Target for Key', 'Disjoint With', and 'Disjoint Union Of'. Each relationship has a '+' icon to expand options.

Figura 01: Autoria Própria

Exemplo: Laboratório de Genética

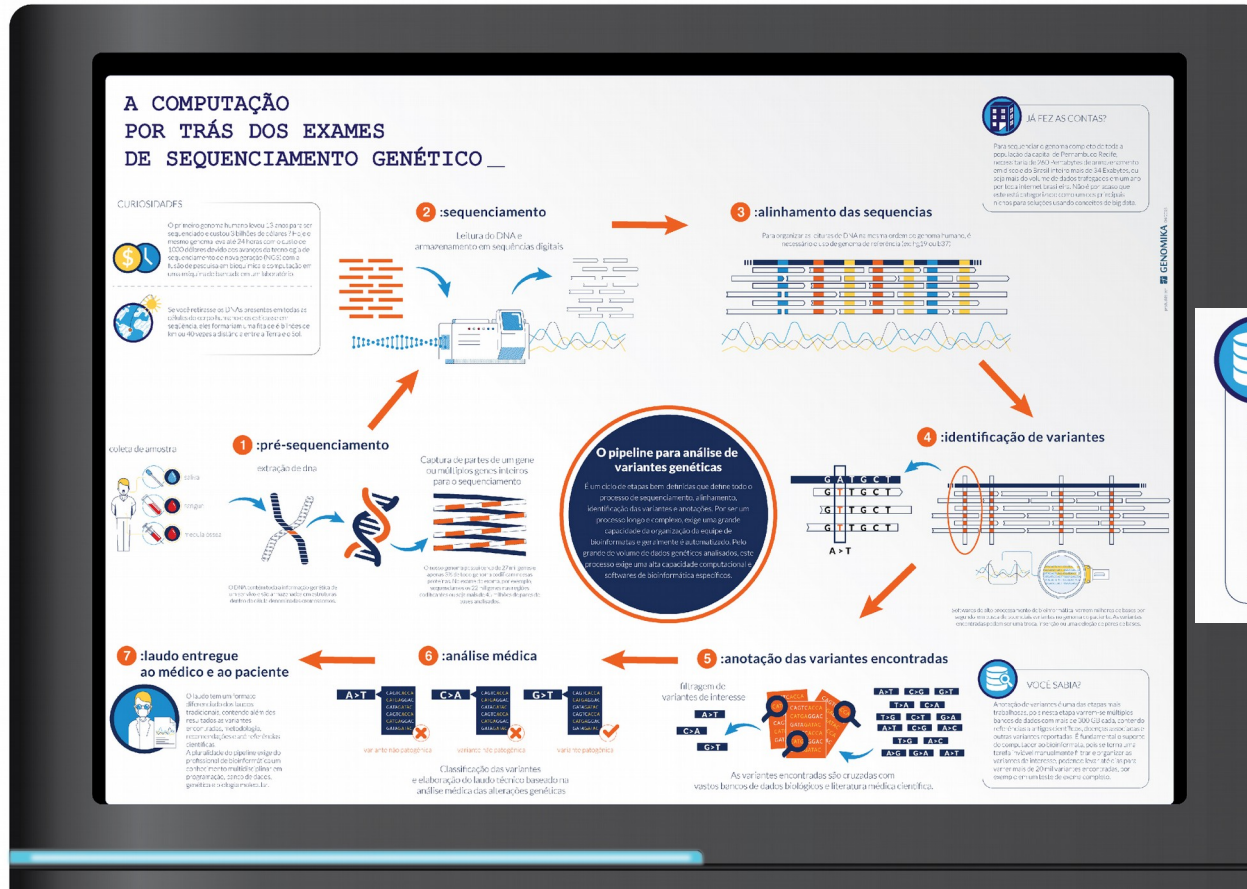


Exemplo: Laboratório de Genética

A computação por trás dos exames de sequenciamento genético : A Era do Big Data na Medicina Genômica

Publicado por em 28 de Abril de 2015 em: **Bioinformática Exames**

Exemplo: Laboratório de Genética



VOCÊ SABIA?

Anotação de variantes é uma das etapas mais trabalhosas, pois nesta etapa varrem-se múltiplos bancos de dados com mais de 300 GB cada, contendo referências a artigos científicos, doenças associadas e outras variantes reportadas. É fundamental o suporte do computador ao bioinformático, pois se torna uma tarefa inviável manualmente filtrar e organizar as variantes de interesse, podendo levar até dias para varrer mais de 20 mil variantes encontradas, por exemplo em um teste de exoma completo.

Conclusão

- Apesar dos esforços nos últimos anos, ainda há muitos GAP's na área de Gene Ontology.
- Consultas e Integração com Big Data é fundamental para o processo continuar evoluindo.
- Falta mais ferramentas para automatizar.

Referências

O Consórcio Gene Ontology (Janeiro de 2008). "O projeto Gene Ontology em 2008" Nucleic Acids Res 36 (problema de banco de dados):... D440-4 doi: 10.1093 / nar / gkm883. PMC 2.238.979. PMID 17984083.

~> <http://geneontology.org/page/download-ontology>

Jump up ^ Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, Goldberg LJ, Eilbeck K, Irlanda A, Mungall CJ, Leontis N, Rocca-Serra P, Ruttenberg A, Sansone SA, Scheuermann RH, Shah N , Whetzel PL, Lewis S (Novembro de 2007).

Wikipedia.org: Informações GO

~> <http://bit.ly/1JrK2BH> - Acesso em 27/10/2015, 8hs:49min

Laboratório Genomika

~> <https://www.genomika.com.br/blog> - Acesso em 27/10/2015, às 10hs:29min.