

Anotação de genomas

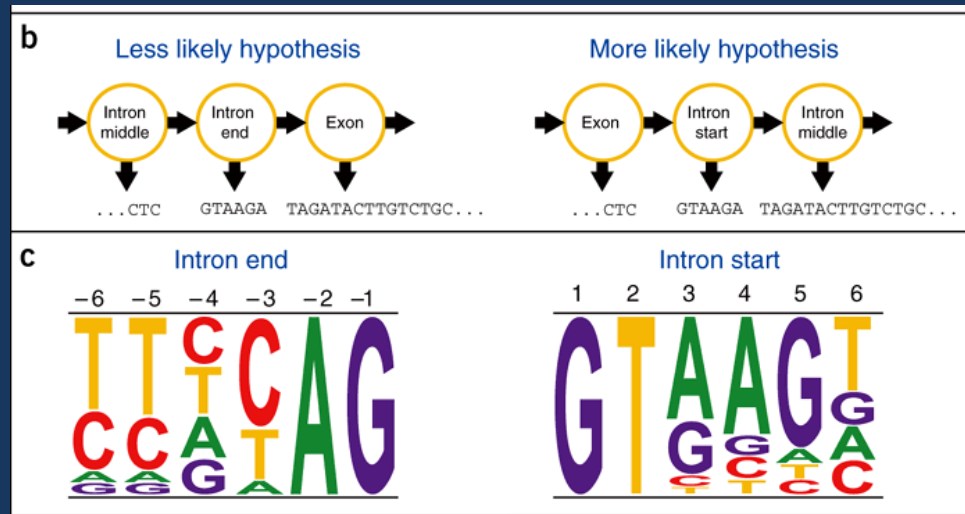
Predição *de novo* de genes

- Conforme vimos na aula anterior o simples seqüenciamento de um genoma de um organismo eucarioto não fornece de modo imediato a informação sobre as proteínas que são codificadas
- O seqüenciamento de moléculas de mRNA, seja via estratégia de seqüenciamento de ESTs ou moléculas inteiras, permite que haja esta dedução. No entanto normalmente não existem seqüências suficientes para a dedução de todos os genes de um organismo.
- Devido a isso desenvolveu-se algoritmos para a predição de genes a partir da seqüência genômica a partir do reconhecimento de padrões de sinalização de splicing e de características de seqüências codificadoras para proteínas

Predição *de novo* de genes

- Início e término de genes são realizados através da busca de elementos como TATA-box (início de transcrição) e sinais de poliadenilação (final de transcrição)

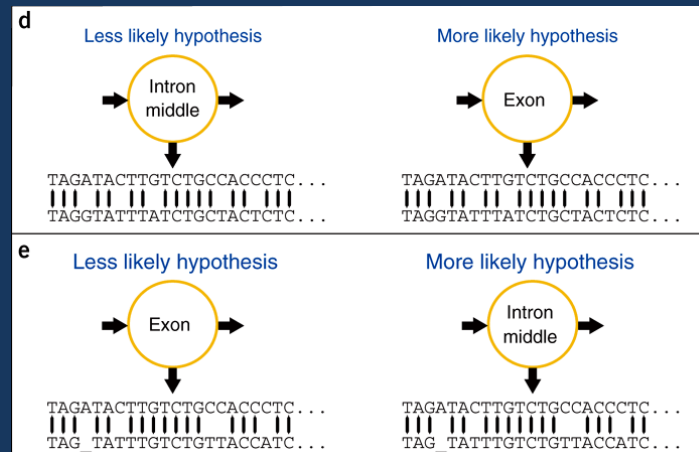
Predição *de novo* de genes



Reconhecimento de padrão de splicing- Busca por sítios canônicos de splicing determinados a partir de exemplos reais

Além disso busca-se seqüências que possuam um quadro de leitura aberta (ORF) compatível com a estrutura do gene

Predição *de novo* de genes



- Comparação com seqüências de genoma de outros organismos relacionados pode ajudar a selecionar regiões codificadoras de proteínas. Fatores evolucionários tendem a fazer com que regiões codificadoras tenham uma maior conservação entre organismos e terceira base do codon tende a ser a mais variável.

GeneScan



[For information about Genscan, click here](#)

This server provides access to the program Genscan for predicting the locations and exon-intron structures of genes in genomic sequences from

This server can accept sequences up to 1 million base pairs (1 Mbp) in length. If you have trouble with the web server or if you have a large num this page) or use the [GENSCAN email server](#). If your browser (*e.g.*, Lynx) does not support file upload or multipart forms, use the [older version](#).

Organism: Suboptimal exon cutoff (optional):

Sequence name (optional):

Print options:

Upload your DNA sequence file (one-letter code, upper or lower case, spaces/numbers ignored):

Or paste your DNA sequence here (one-letter code, upper or lower case, spaces/numbers ignored):

```
ATTTCAGTCATGGATGCGAAGGTAAGTTTCATATTGTTTTTTTTTGTAAATTACTTTATTACTTACTTACT
```



To have the results mailed to you, enter your email address here (optional):

[Back to the top](#)

GeneScan

Parameter matrix: HumanIso.smat

Predicted genes/exons:

Gn.Ex	Type	S	.Begin	...End	.Len	Fr	Ph	I/Ac	Do/T	CodRg	P....	Tscr..
1.03	PlyA	-	63	58	6							1.05
1.02	Term	-	664	428	237	1	0	42	48	178	0.037	4.38
1.01	Init	-	6863	6696	168	2	0	83	82	149	0.361	13.39
1.00	Prom	-	9494	9455	40							-4.75
2.00	Prom	+	10132	10171	40							-8.25
2.01	Init	+	12862	12921	60	0	0	82	61	64	0.423	4.40
2.02	Intr	+	14220	14360	141	2	0	4	95	131	0.716	5.03
2.03	Intr	+	14470	14590	121	0	1	57	61	35	0.693	-3.15
2.04	Intr	+	15320	15397	78	0	0	45	113	170	0.985	13.80
2.05	Intr	+	15975	16279	305	1	2	79	50	96	0.172	0.28

Click [here](#) to view a PDF image of the predicted gene(s)

Click [here](#) for a PostScript image of the predicted gene(s)

Predicted peptide sequence(s):

>07:32:59|GENSCAN_predicted_peptide_1|134_aa
MENVEALDSRFVVLWNSSAAVRIHDPALTRFEPRTYINIVGCRLSGLSVKCSGARLRFFS
IIDEHSGSDADVRRARIGKARAAYSQLNNIWNKQLSVNQKDLNFQYKCQDSSTVWGVNL
ENYESHHPEDTSTVY

>07:32:59|GENSCAN_predicted_peptide_2|235_aa
MTIEKLVDGVESRLSMITTRCDDSKVFLSDNQHDNDAAFQGGVRLEKVDPKNSHLALS
HGYPSPAVIKTTSNPISFSSTSRRTLPRCGQPGSDISPHTRNCTRDQDEQPDERSEERE
DKAKRNEMTCSGEDSSAGYTCISELMFTLRWTRTQHLSLQTSSRYPLGHGVLITTCFLNG
VIFKNSFSIVRLNLLIDVFRTTACQSLIGIREYCAYCCLDIALIHKHCNACEFMN

- Prom- seqüência promotora
- PlyA- sinal de poliadenilação
- Init- exon inicial
- Term- exon terminal
- Intr- exon interno
- P.... -probabilidade do exon
- Scores:
 - >100- forte;
 - 100-50 modrado;
 - 50-0 fraco;
 - <0 improvavel

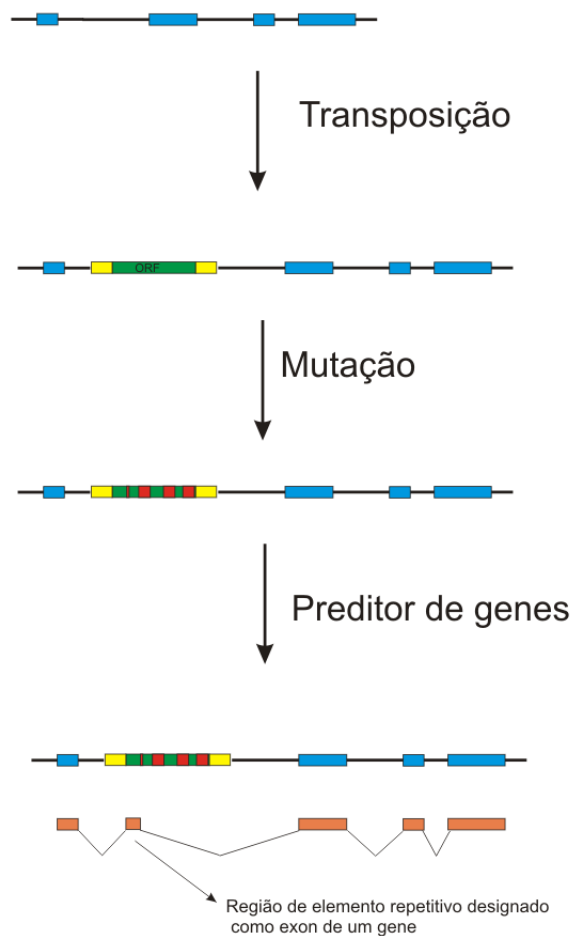
GeneScan

Probability	Predicted	Accuracy Class			
Range	Exons	Exactly Correct	Partially Correct	Overlapping	Wrong
0.00 - 0.50	248	29.8%	27.8%	4.0%	38.3%
0.50 - 0.75	362	54.1%	26.2%	2.2%	17.4%
0.75 - 0.90	337	74.8%	16.0%	1.2%	8.0%
0.90 - 0.95	263	87.8%	6.1%	0.4%	5.7%
0.95 - 0.99	551	92.4%	3.4%	0.2%	4.0%
0.99 - 1.00	917	97.7%	0.9%	0.0%	1.4%
Total	2,678	80.6%	9.7%	0.9%	8.8%

Elementos repetitivos

- Conforme visto anteriormente o genoma de organismos eucariotos possui uma grande fração de seu genoma composto de elementos repetitivos.
- Elementos repetitivos muitas vezes possuem repetições de padrões e por isso não comportam-se como seria esperado para uma sequência randômica
- Devido ao fato de grande parte destes elementos repetitivos serem derivados de elementos de transposição espera-se que alguns deles ainda apresentem regiões codificadoras que podem confundir os programas de predição de genes
- Por isso antes da predição de genes é necessário realizar uma mascaramento da sequência genômica que retire os elementos repetitivos da análise de predição de genes

Elementos repetitivos



Elementos de transposição podem se inserir nos introns de genes.

Como não são essenciais para os organismos estes elementos tendem a sofrer mutações que trunca sua mensagem gerando pequenas ORFs

Algumas destas ORFs podem ser confundidas pelo programa de predição de genes como sendo exons validos deste gene devido ao seu potencial de codificação.

Repeat masker

The screenshot shows the RepeatMasker Web Server interface. At the top, it says 'INSTITUTE FOR Systems Biology' with a logo and 'RepeatMasker Web Server'. Below this, a description states: 'RepeatMasker screens DNA sequences in FASTA format against a library of repetitive elements and returns a masked sequence. Reference: A.F.A. Smit, R. Hubley & P. Green, unpublished data. Current Version: open-3.2.6 (RMLib: 20080801)'. A link 'Check Current Queue Status' is provided. The 'Basic Options' section includes a file upload button 'Arquivo...', a text area for 'Sequence:', a 'Search Engine' dropdown set to 'wublast', 'Speed/Sensitivity' radio buttons for 'rush', 'quick', 'default' (selected), and 'slow', a 'DNA source' dropdown set to 'schistosoma mansoni', 'Return Format' radio buttons for 'html' (selected) and 'tar file', and 'Return Method' radio buttons for 'html' (selected) and 'email' with an adjacent email address field. At the bottom are buttons for 'Restaurar valores' and 'Submit Sequence'.

INSTITUTE FOR
Systems Biology

RepeatMasker Web Server

[RepeatMasker](#) screens DNA sequences in FASTA format against a library of repetitive elements and returns a masked sequence.
Reference: A.F.A. Smit, R. Hubley & P. Green, unpublished data. Current Version: open-3.2.6 (RMLib: 20080801)
[Check Current Queue Status](#)

Basic Options

or

[Sequence:](#)

[Search Engine:](#) ☒ wublast ☐ cross_match

[Speed/Sensitivity:](#) ☐ rush ☐ quick ☒ default ☐ slow

[DNA source:](#)

[Return Format:](#) ☒ html ☐ tar file

[Return Method:](#) ☒ html ☐ email

- Ferramenta que compara seqüências contra um banco de dados pré-existente contendo elementos repetitivos previamente descritos de diversos organismos
- Busca também repetições simples que consistem na repetição contínua de motivos simples

Repeat masker

- Sumario mostra os elementos detectados bem como o numero de bases ocupadas por estes elementos

	number of elements*	length occupied	percentage of sequence
-----	-----	-----	-----
Retroelements	32	5668 bp	14.79 %
SINES:	14	2757 bp	7.20 %
Penelope	3	297 bp	0.78 %
LINEs:	8	1278 bp	3.34 %
CRE/SLACS	0	0 bp	0.00 %
L2/CR1/Rex	0	0 bp	0.00 %
R1/LOA/Jockey	0	0 bp	0.00 %
R2/R4/NeSL	0	0 bp	0.00 %
RTE/Bov-B	3	847 bp	2.21 %
L1/CIN4	0	0 bp	0.00 %
LTR elements:	10	1633 bp	4.26 %
BEL/Pao	0	0 bp	0.00 %
Ty1/Copia	0	0 bp	0.00 %
Gypsy/DIRS1	9	1399 bp	3.65 %
Retroviral	0	0 bp	0.00 %
DNA transposons	0	0 bp	0.00 %
hobo-Activator	0	0 bp	0.00 %
Tc1-IS630-Pogo	0	0 bp	0.00 %
En-Spm	0	0 bp	0.00 %
MuDR-IS905	0	0 bp	0.00 %
PiggyBac	0	0 bp	0.00 %
Tourist/Harbinger	0	0 bp	0.00 %
Other (Mirage, P-element, Transib)	0	0 bp	0.00 %
Rolling-circles	0	0 bp	0.00 %
Unclassified:	0	0 bp	0.00 %
Total interspersed repeats:		5668 bp	14.79 %
Small RNA:	14	2757 bp	7.20 %
Satellites:	0	0 bp	0.00 %
Simple repeats:	0	0 bp	0.00 %
Low complexity:	8	244 bp	0.64 %
=====			

Repeat masker

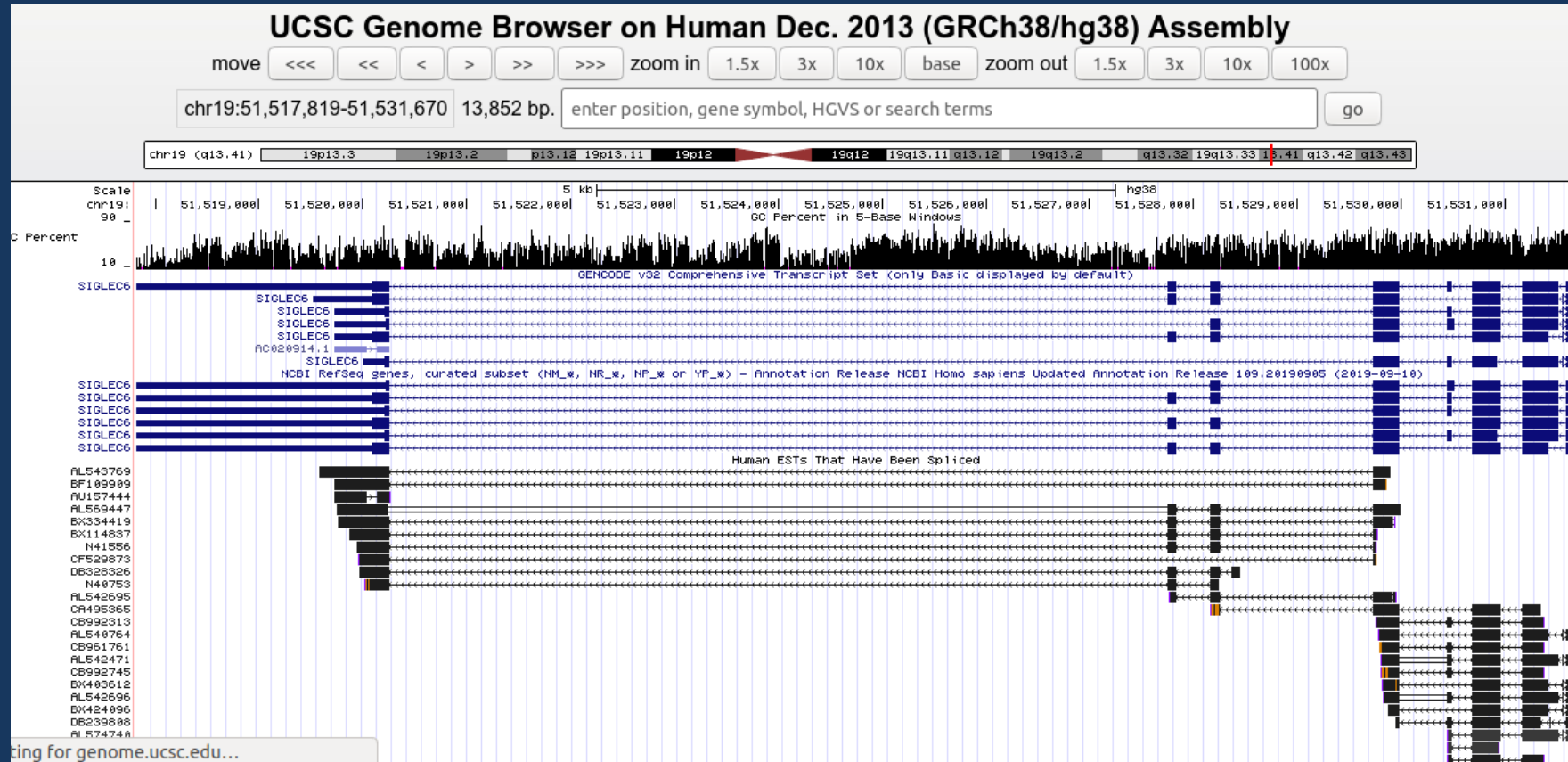
[illegible]

- Nova sequência é fornecida ao usuário contendo Ns no lugar das bases dos elementos repetitivos.
- Isso permite que programas de predição de genes ignorem estas bases, mas a informação de estrutura deste segmento de DNA não é perdida

Anotação de genomas

- Uma vez descrito a seqüência do genoma de um organismo e realizada a anotação dos genes presentes obtemos uma extensa lista de potenciais transcritos. Como vimos anteriormente é necessário adicionar anotações de função para que esta lista possa ser informativa em relação aos diversos processos que estas proteínas estão mediando.
- Este tipo de anotação deve fornecer dados que permitam a comunidade científica (muitas vezes não especializada em genômica) realizar buscas de genes de interesse em seus trabalhos
- Devido ao alto numero de seqüências muitos dos processos devem ser automatizados em um primeiro momento, para permitir rapidamente com que o banco seja disponibilizado para comunidade científica realizar suas buscas

Genome Browser



Anotação de genomas

Human Gene SIGLEC6 (ENST00000346477.7) Description and Page Index

Description: Homo sapiens sialic acid binding Ig like lectin 6 (SIGLEC6), transcript variant 2, mRNA. (from RefSeq NM_198845)

RefSeq Summary (NM_198845): This gene encodes a member of the SIGLEC (sialic acid binding immunoglobulin-like lectin) family of proteins. The encoded transmembrane receptor binds sialyl-TN glycans and leptin. Placental expression of the encoded protein is upregulated in preeclampsia. [provided by RefSeq, Jul 2016].

Gencode Transcript: ENST00000346477.7

Gencode Gene: ENSG00000105492.16

Transcript (Including UTRs)

Position: hg38 chr19:51,519,525-51,531,717 **Size:** 12,193 **Total Exon Count:** 7 **Strand:** -

Coding Region

Position: hg38 chr19:51,520,082-51,531,648 **Size:** 11,567 **Coding Exon Count:** 7

Page Index	Sequence and Links	UniProtKB Comments	CTD	RNA-Seq Expression	Microarray Expression
RNA Structure	Protein Structure	Other Species	GO Annotations	mRNA Descriptions	Pathways
Other Names	Methods				

Data last updated: 2019-09-04

☐ Sequence and Links to Tools and Databases

Genomic Sequence (chr19:51,519,525-51,531,717)			mRNA (may differ from genome)		Protein (437 aa)
Gene Sorter	Genome Browser	Other Species FASTA	Gene interactions	Table Schema	BioGPS
CGAP	Ensembl	Entrez Gene	ExonPrimer	GeneCards	HGNC
HPRD	Lynx	MGI	neXtProt	OMIM	PubMed
Reactome	Stanford SOURCE	UniProtKB			

Anotação de genomas

Apesar deste tipo de anotação fornecer uma função inferida do produto gênico, ela apresenta alguns problemas:

- nem sempre é obvio para um não-especialista a função exata e contexto no qual atua a proteína codificada a partir do nome desta
- Pesquisadores buscando entender processos necessitam primeiro realizar uma busca bibliográfica para identificar os genes envolvidos em tais processos para então realizar uma busca no banco
- Muitas proteínas possuem nomes que são muito pouco informativos
- O nome dado as proteínas e domínios protéicos não é padronizado, sendo que muitas vezes a mesma proteína pode ter diversas denominações, o que pode gerar confusões

Anotação de genoma

- Deste modo, a agregação de anotações adicionais que indiquem claramente ao pesquisador a função e outros atributos do produto proteico é desejável
- Um dos problemas para anotação de genes é que um mesmo processo pode ser descrito através de diversos termos. (por exemplo, tradução e síntese de proteínas)
- Deste modo surgiram novos métodos de anotação para tentar suprir esta deficiência na anotação de genomas

Gene ontology (GO)

[WordNet](#) - [Cite This Source](#) - [Share This](#)
ontology

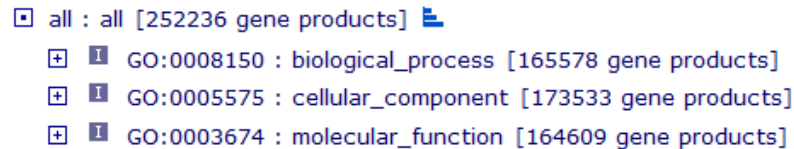
noun


1. (computer science) a rigorous and exhaustive organization of some knowledge domain that is usually hierarchical and contains all the relevant entities and their relations
2. the metaphysical study of the nature of being and existence







Gene ontology (GO)

- Serie de vocabulários controlados que apresenta uma relação hierárquica entre si.
- Através do uso deste termos é possível gerar diversas descrições do produto de um dado transcrito que podem ser buscadas através de palavras chaves ou navegando por uma arvore de termos
- Vocabulário descreve fenômenos moleculares e não objetos moleculares, não sendo substituto de outras anotações mais tradicionais

Gene ontology (GO)



all : all [252236 gene products] 

-   GO:0008150 : biological_process [165578 gene products]
-   GO:0005575 : cellular_component [173533 gene products]
-   GO:0003674 : molecular_function [164609 gene products]

- Dividido em três categorias principais:

Processo biológico- Indica associação a eventos gerados a partir de um ou mais conjuntos ordenados de funções moleculares (ex. transporte de lipídeos, geração de gametas, quimiotropismo)

Componente celular- Indicaria a possível localização do produto protéico (ex: núcleo, reticulo endoplasmático, proteassomo)

Função molecular- descreve atividades que ocorrem a nível molecular (ex: transporte de lipídeo acoplado a atividade ATPasica , atividade receptora de GABA)

GO Tree browser

- ▣ all : all [250922 gene products]
 - ⊕ ⓘ GO:0003674 : molecular_function [163897 gene products]
 - ⊕ ⓘ GO:0005488 : binding [45764 gene products]
 - ⊕ ⓘ GO:0043176 : amine binding [289 gene products]
 - ⊕ ⓘ GO:0042166 : acetylcholine binding [70 gene products]
 - ⊕ ⓘ **GO:0015464 : acetylcholine receptor activity [62 gene products]**
 - ⊕ ⓘ GO:0042165 : neurotransmitter binding [235 gene products]
 - ⊕ ⓘ GO:0042166 : acetylcholine binding [70 gene products]
 - ⊕ ⓘ **GO:0015464 : acetylcholine receptor activity [62 gene products]**
 - ⊕ ⓘ GO:0030594 : neurotransmitter receptor activity [218 gene products]
 - ⊕ ⓘ **GO:0015464 : acetylcholine receptor activity [62 gene products]**
 - ⊕ ⓘ GO:0060089 : molecular transducer activity [7611 gene products]
 - ⊕ ⓘ GO:0004871 : signal transducer activity [7611 gene products]
 - ⊕ ⓘ GO:0004872 : receptor activity [5255 gene products]
 - ⊕ ⓘ GO:0030594 : neurotransmitter receptor activity [218 gene products]
 - ⊕ ⓘ **GO:0015464 : acetylcholine receptor activity [62 gene products]**
 - ⊕ ⓘ GO:0004888 : transmembrane receptor activity [4439 gene products]
 - ⊕ ⓘ **GO:0015464 : acetylcholine receptor activity [62 gene products]**

O mesmo termo pode derivar de diferentes termos “pais”

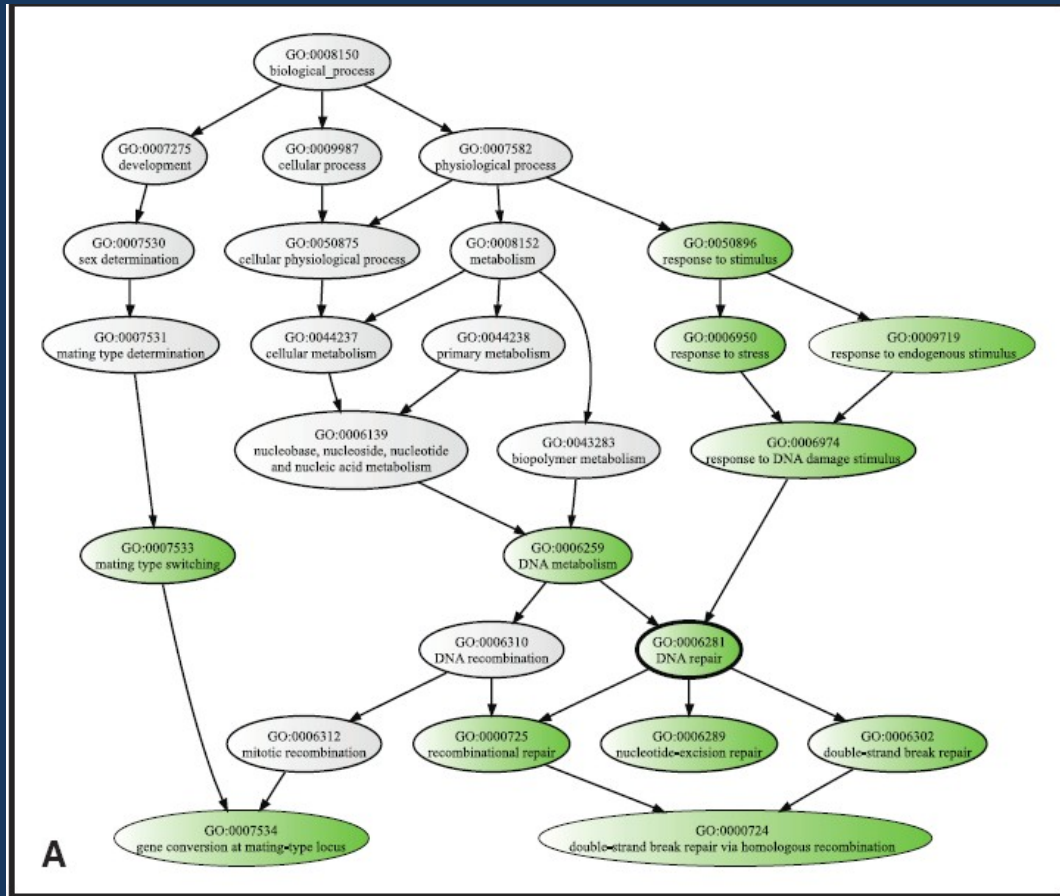
Detecção de GOs com representação aumentada

Comparações de dois conjuntos de transcritos diferentes (como transcritos expressos em tecidos diferentes) podem ser analisados em termos da diferença representatividade de termos de GO.


Este tipo de comparação busca definir termos de GO enriquecidos em determinado conjunto.

Cálculos estatísticos devem ser realizados para avaliar significância dos resultados.

Detecção de GOs com representação aumentada



KEGG



[Search KEGG](#)

[Get Entry](#)

KEGG Home
[Introduction](#)
[Overview](#)
[Release notes](#)
[Current statistics](#)

KEGG Identifiers

KGML

KEGG API

KEGG FTP

KegTools


[Feedback](#)
[GenomeNet](#)

KEGG: Kyoto Encyclopedia of Genes and Genomes

A grand challenge in the post-genomic era is a complete computer representation of the cell, the organism, and the biosphere, which will enable computational prediction of higher-level complexity of cellular processes and organism behaviors from genomic and molecular information. Towards this end we have been developing a bioinformatics resource named KEGG as part of the research projects of the Kanehisa Laboratories in the Bioinformatics Center of Kyoto University and the Human Genome Center of the University of Tokyo.

- Main entry point to the KEGG web service**
[KEGG2](#) [KEGG Table of Contents](#) [Update notes](#) [Help](#)
- Data-oriented entry points**
 - [KEGG Atlas](#) Global maps of cell/organism functions
 - [KEGG PATHWAY](#) Pathway maps and pathway modules
 - [KEGG BRTE](#) Functional hierarchies and ontologies
 - [KEGG GENES](#) Genomes, genes, proteins, and orthologs
 - [KEGG LIGAND](#) Chemical compounds, drugs, glycans, and reactions
- Organism-specific entry points**
[KEGG Organisms](#) Select (example) [hsa](#)
- Subject-specific entry points**
 - [KEGG DISEASE](#) Gene/molecule based disease information resource
 - [KEGG DRUG](#) Chemical structure based drug information resource
 - [KEGG GLYCAN](#) Glycome informatics resource
 - [KEGG COMPOUND](#) Knowledge base for biochemical compounds
 - [KEGG REACTION](#) Knowledge base for biochemical reactions
 - [KAAS](#) KEGG automatic annotation server

KEGG



Search KEGG Get Entry

KEGG Home
Introduction
Overview
Release notes
Current statistics

KEGG Identifiers

KGML

KEGG API

KEGG FTP

KegTools

Feedback
GenomeNet

KEGG: Kyoto Encyclopedia of Genes and Genomes

A grand challenge in the post-genomic era is a complete computer representation of the cell, the organism, and the biosphere, which will enable computational prediction of higher-level complexity of cellular processes and organism behaviors from genomic and molecular information. Towards this end we have been developing a bioinformatics resource named KEGG as part of the research projects of the Kanehisa Laboratories in the Bioinformatics Center of Kyoto University and the Human Genome Center of the University of Tokyo.

- Main entry point to the KEGG web service**
 - KEGG2** KEGG Table of Contents Update notes Help
- Data-oriented entry points**
 - KEGG Atlas** Global maps of cell/organism functions
 - KEGG PATHWAY** Pathway maps and pathway modules
 - KEGG BRTE** Functional hierarchies and ontologies
 - KEGG GENES** Genomes, genes, proteins, and orthologs
 - KEGG LIGAND** Chemical compounds, drugs, glycans, and reactions
- Organism-specific entry points**
 - KEGG Organisms** Select (example) hsa
- Subject-specific entry points**
 - KEGG DISEASE** Gene/molecule based disease information resource
 - KEGG DRUG** Chemical structure based drug information resource
 - KEGG GLYCAN** Glycome informatics resource
 - KEGG COMPOUND** Knowledge base for biochemical compounds
 - KEGG REACTION** Knowledge base for biochemical reactions
 - KAAS** KEGG automatic annotation server

1. Metabolism

1.1 Carbohydrate Metabolism

Glycolysis / Gluconeogenesis
Citrate cycle (TCA cycle)
Pentose phosphate pathway
Pentose and glucuronate interconversions
Fructose and mannose metabolism
Galactose metabolism
Ascorbate and aldarate metabolism
Starch and sucrose metabolism
Aminosugars metabolism
Nucleotide sugars metabolism
Pyruvate metabolism
Glyoxylate and dicarboxylate metabolism
Propanoate metabolism
Butanoate metabolism
C5-Branched dibasic acid metabolism
Inositol metabolism
Inositol phosphate metabolism

KEGG Orthology (KO)
KEGG pathway modules
Overview of biosynthetic pathways
Enzymes (+diseases)
Compounds with biological roles

1.2 Energy Metabolism

Oxidative phosphorylation
Photosynthesis
Photosynthesis - antenna proteins
Carbon fixation
Reductive carboxylate cycle (CO2 fixation)
Methane metabolism
Nitrogen metabolism
Sulfur metabolism

Photosynthesis proteins

1.3 Lipid Metabolism

Fatty acid biosynthesis
Fatty acid elongation in mitochondria
Fatty acid metabolism
Synthesis and degradation of ketone bodies
Biosynthesis of steroids
Bile acid biosynthesis
C21-Steroid hormone metabolism
Androgen and estrogen metabolism
Glycerolipid metabolism
Glycerophospholipid metabolism
Ether lipid metabolism
Sphingolipid metabolism
Arachidonic acid metabolism
Linoleic acid metabolism
alpha-Linolenic acid metabolism
Biosynthesis of unsaturated fatty acids

Lipids

- Possibilidade de pesquisar uma via metabólica determinada

KEGG Pathways

- Utilização do numero E.C. (Enzyme Commission), sistema de classificação estabelecido

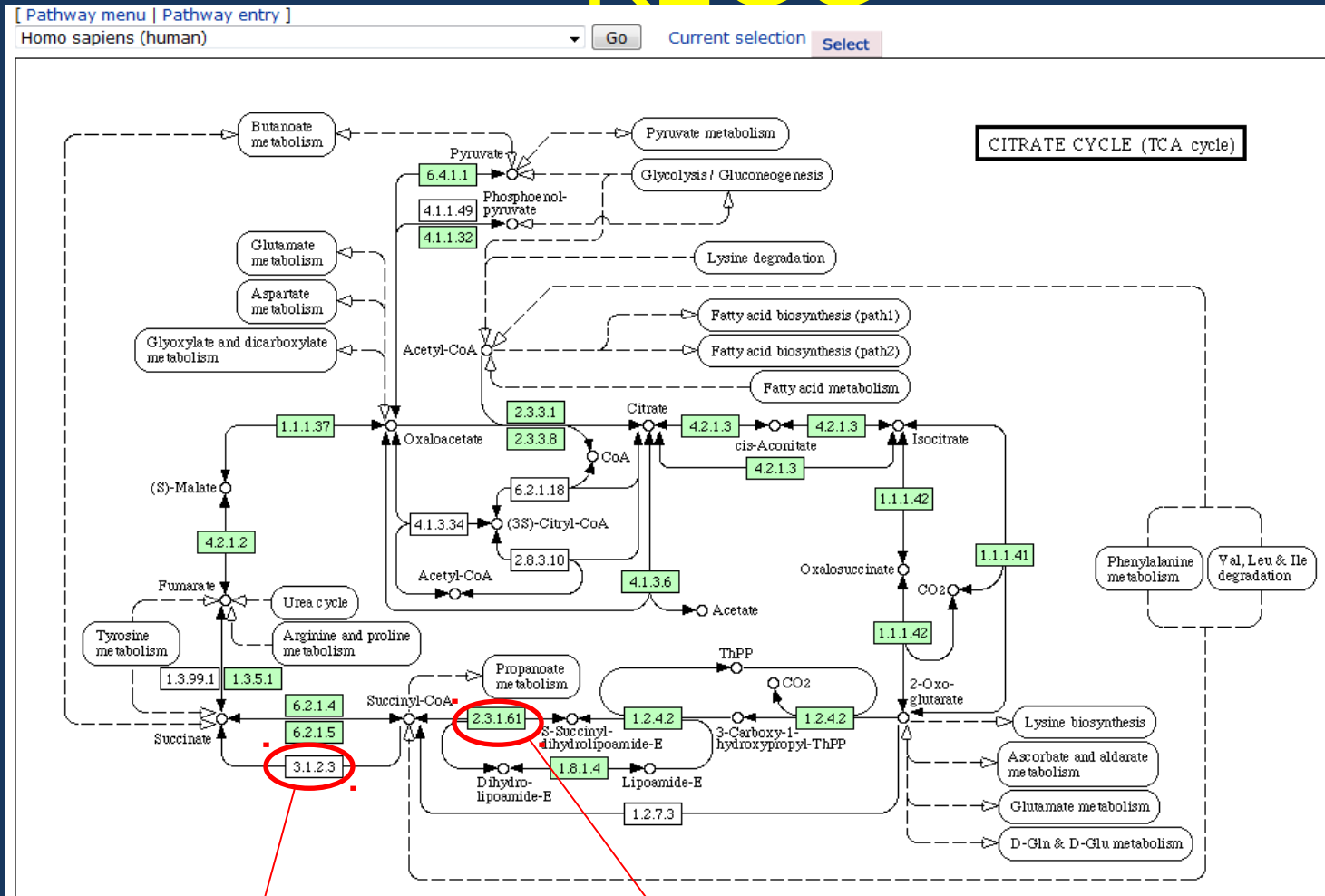
Top-level EC numbers ^[2]			
Group	Reaction catalyzed	Typical reaction	Enzyme example(s) with trivial name
EC 1 <i>Oxidoreductases</i>	To catalyze oxidation /reduction reactions; transfer of H and O atoms or electrons from one substance to another	$AH + B \rightarrow A + BH$ (reduced) $A + O \rightarrow AO$ (oxidized)	Dehydrogenase , oxidase
EC 2 <i>Transferases</i>	Transfer of a functional group from one substance to another. The group may be methyl-, acyl-, amino- or phosphate group	$AB + C \rightarrow A + BC$	Transaminase , kinase
EC 3 <i>Hydrolases</i>	Formation of two products from a substrate by hydrolysis	$AB + H_2O \rightarrow AOH + BH$	Lipase , amylase , peptidase
EC 4 <i>Lyases</i>	Non-hydrolytic addition or removal of groups from substrates. C-C, C-N, C-O or C-S bonds may be cleaved	$RCO_2COOH \rightarrow RCOH + CO_2$	Decarboxylase
EC 5 <i>Isomerases</i>	Intramolecule rearrangement, i.e. isomerization changes within a single molecule	$AB \rightarrow BA$	Isomerase , mutase
EC 6 <i>Ligases</i>	Join together two molecules by synthesis of new C-O, C-S, C-N or C-C bonds with simultaneous breakdown of ATP	$X + Y + ATP \rightarrow XY + ADP + P_i$	Synthetase

KEGG Pathways

```
3. -. -. Hydrolases.
3. 1. -. Acting on ester bonds.
3. 1. 1.- Carboxylic ester hydrolases.
3. 1. 2.- Thiolester hydrolases.
3. 1. 3.- Phosphoric monoester hydrolases.
3. 1. 4.- Phosphoric diester hydrolases.
3. 1. 5.- Triphosphoric monoester hydrolases.
3. 1. 6.- Sulfuric ester hydrolases.
3. 1. 7.- Diphosphoric monoester hydrolases.
3. 1. 8.- Phosphoric triester hydrolases.
3. 1.11.- Exodeoxyribonucleases producing 5'-phosphomonoesters.
3. 1.13.- Exoribonucleases producing 5'-phosphomonoesters.
3. 1.14.- Exoribonucleases producing 3'-phosphomonoesters.
3. 1.15.- Exonucleases active with either ribo- or deoxyribonucleic acid
3. 1.16.- Exonucleases active with either ribo- or deoxyribonucleic acid
3. 1.21.- Endodeoxyribonucleases producing 5'-phosphomonoesters.
3. 1.22.- Endodeoxyribonucleases producing other than 5'-phosphomonoesters.
3. 1.25.- Site-specific endodeoxyribonucleases specific for altered bases.
3. 1.26.- Endoribonucleases producing 5'-phosphomonoesters.
3. 1.27.- Endoribonucleases producing other than 5'-phosphomonoesters.
3. 1.30.- Endoribonucleases active with either ribo- or deoxyribonucleic
3. 1.31.- Endoribonucleases active with either ribo- or deoxyribonucleic
3. 2. -. Glycosylases.
3. 2. 1.- Glycosidases, i.e. enzymes hydrolyzing O- and S-glycosyl
3. 2. 2.- Hydrolyzing N-glycosyl compounds.
3. 3. -. Acting on ether bonds.
3. 3. 1.- Thioether and trialkylsulfonium hydrolases.
3. 3. 2.- Ether hydrolases.
3. 4. -. Acting on peptide bonds (peptide hydrolases).
3. 4.11.- Aminopeptidases.
3. 4.13.- Dipeptidases.
3. 4.14.- Dipeptidyl-peptidases and tripeptidyl-peptidases.
3. 4.15.- Peptidyl-dipeptidases.
3. 4.16.- Serine-type carboxypeptidases.
3. 4.17.- Metallo-carboxypeptidases.
3. 4.18.- Cysteine-type carboxypeptidases.
3. 4.19.- Omega peptidases.
3. 4.21.- Serine endopeptidases.
3. 4.22.- Cysteine endopeptidases.
3. 4.23.- Aspartic endopeptidases.
3. 4.24.- Metalloendopeptidases.
3. 4.25.- Threonine endopeptidases.
3. 4.99.- Endopeptidases of unknown catalytic mechanism.
```

- Nomenclatura possui quatro níveis hierárquicos sendo que o primeiro nível divide entre as seis famílias principais

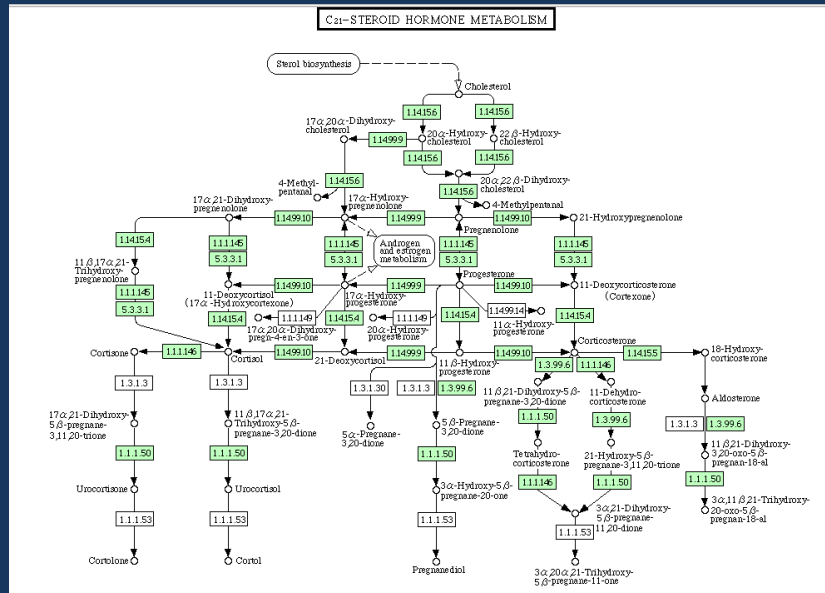
KEGG



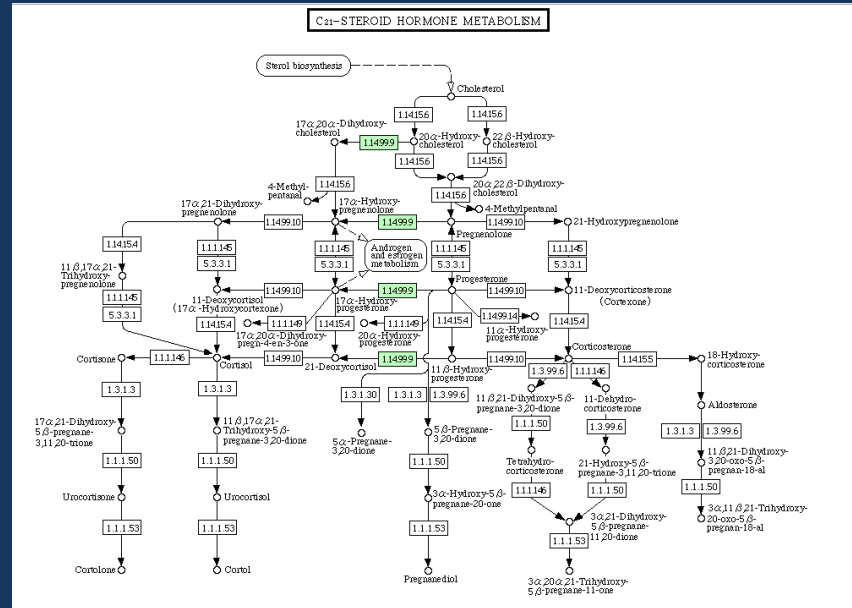
E.C. em branco indica ausência no organismo

E.C. em verde indica presença no organismo

Comparando vias entre espécies

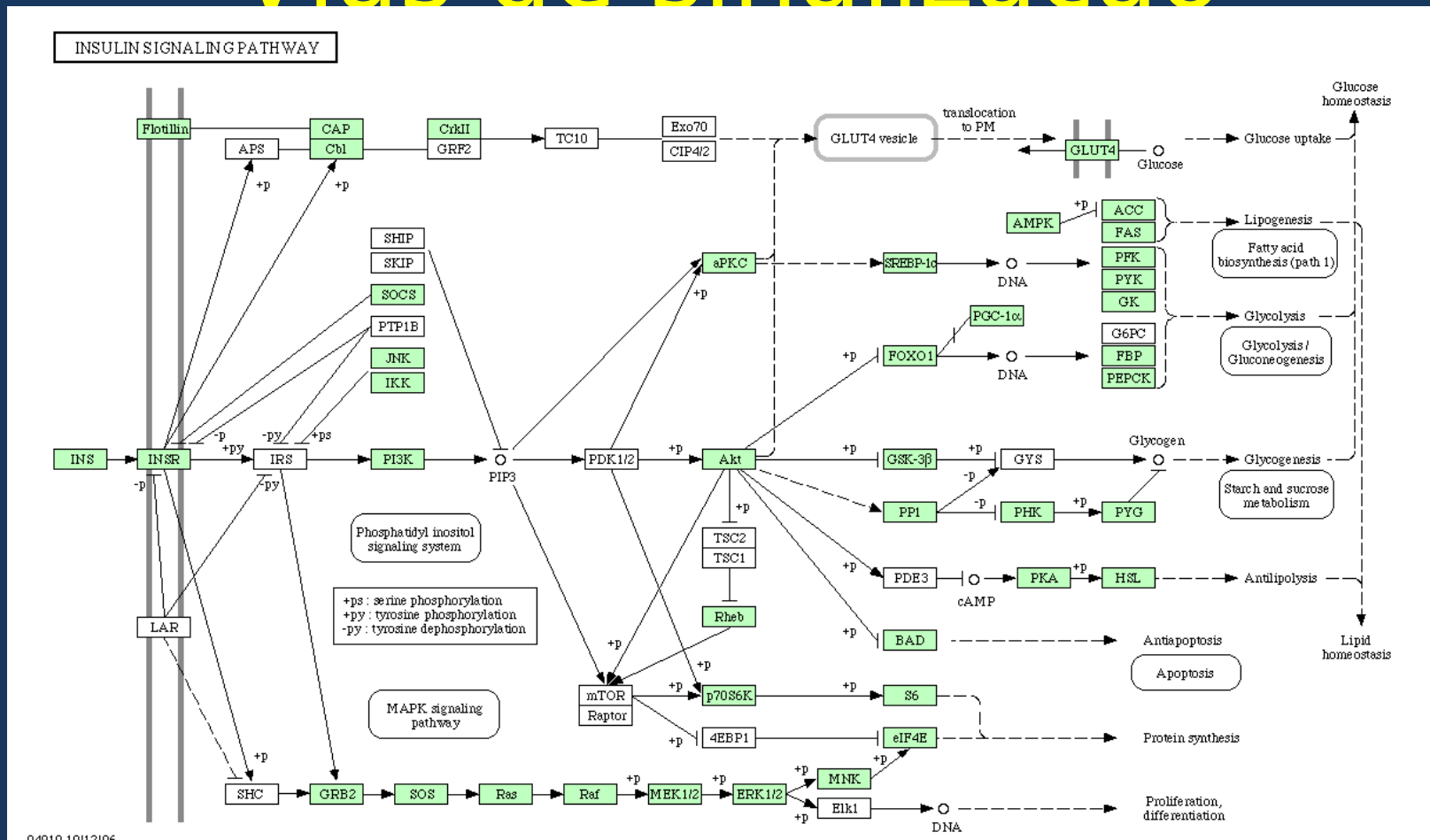


Homo sapiens



Aspergillus fumigatus (fungo)

Vias de sinalização



- Não utiliza E.C.s pois não se trata de atividade enzimática