

Pesquisa

A pesquisa tem como prioridade identificar em uma base de dados de genes, quais são os candidatos os quais podem ser genes *HouseKeeping*, para isso, no começo, a gente tem que gerar mais dados os quais são considerados como *HouseKeeping* pela literatura, a gente faz isso com as GAN's (*Generative adversarial network*), depois disso, mediante aprendizado de maquina a gente conseguiu dizer se o gen novo pode ser o não um bom candidato, isso já que se precisa reduzir o número de “candidatos”.

Protocolo do mapeamento sistemático

Foco da pesquisa: Identificar se um gene é o não candidato à ser *Housekeeping* baseado na utilização de redes GAN's.

Questão de pesquisa: Será que o gene é um gene candidato à ser *HouseKeeping*?

String de busca:

- **Passo 1 (termos-chave):** gene, candidato, *HouseKeeping*
- **Passo 2 (sinônimos):**
 - **gene:** DNA, genético, gênico
 - **candidato:** competidor, pretendente, postulante, aspirante, pretendedor, solicitante, proponente.
 - **HouseKeeping:** constitutivo, integrante, característico, peculiar, distintivo, típico.
- **Passo 3 (operador OR):**
 - gen OR DNA OR genético OR gênico
 - competidor OR pretendente OR postulante OR aspirante OR pretendedor OR solicitante OR proponente.
 - Housekeeping OR constitutivo OR integrante OR característico OR peculiar OR distintivo OR típico.
- **Passo 4 (operador AND):**

(gen OR DNA OR genético OR gênico OR gen*) AND
(competidor OR pretendente OR postulante OR aspirante OR pretendedor OR solicitante OR proponente) AND
(Housekeeping OR constitutivo OR integrante OR característico OR peculiar OR distintivo OR típico)

(gen* OR DNA) AND (candidate OR competitor OR suitor OR postulant OR aspirant OR proponent OR identify*) AND (housekeeping OR HKG)

Bases de dados: As seguintes bases de dados são escolhidas já que são as bases nas quais são publicados mais artigos que tentam com inteligencia artificial.

- IEEE
- Springer Link

- NIPS (Nueral Information Processing Systems)

Critérios de inclusão e exclusão:

- **Inclusão:**
 - Tem que estar escritos em português, espanhol ou inglês.
 - Tem que ter resultados comparáveis.
 - Tem que ter as palavras chaves.
- **Exclusão:**
 - Trabalhos publicados antes do ano 2014.
 - Trabalhos que não empreguem métodos de inteligencia artificial.

String: (gen* OR DNA) AND (candidate OR competitor OR suitor OR postulant OR aspirant OR proponent OR identify*) AND (housekeeping OR HKG) AND (in-silico)

Resultados da busca

Base de dados	Passo 1	Passo 2	Passo 3
IEEE	11	1	1
Springer Link	1.754	953	2
Busca Manual	-	-	1
Total	1.765	954	4

Passo 1: Pesquisa inicial. **Passo 2:** Critérios de inclusão. **Passo 3:** Critérios de exclusão.



Estudos por ano na literatura sobre estimação de genes *housekeeping*.

Extração dos dados

Título: *A Computational Approach Using Ratio Statistics for Identifying Housekeeping Genes from cDNA Microarray Data*

Fonte: IEEE

Resumo:

Os autores do artigo tentam predizer se um gene é ou não *housekeeping*, para isso eles empregam *Ratio Statistics Based Normalization Strategy*, jogam os dados que conhecem que não são genes *housekeeping* e fazem um modelo gaussiano para cada *feature* do gene para assim ajustar o limite da gaussiana com os genes que são *housekeeping*. O limite é ajustado com o produtório das probabilidades gaussianas de cada *feature*. Ao final os autores conseguem dizer se o gene é ou não *housekeeping*.

Título: *RNA-sequence data normalization through in silico prediction of reference genes: the bacterial response to DNA damage as case study*

Fonte: Springer

Resumo:

Título: *Identification and validation of quantitative real-time reverse transcription PCR reference genes for gene expression analysis in teak (Tectona grandis L.f.)*

Fonte: Springer

Resumo:

Título: *Elucidating tissue specific genes using the Benford distribution*

Fonte: Springer

Resumo: