

eP2400**Análise de Off-Targets in Silico para pacientes com Mucopolissacaridose tipo I usando o sistema CRISPR/CAS9**

Paola Barcelos Carneiro; Martiela Vaz de Freitas; Ursula da Silveira Matte
 UFRGS - Universidade Federal do Rio Grande do Sul

Introdução: Mucopolissacaridose tipo I (MPS I) é uma doença autossômica recessiva relacionada com o depósito de glicosaminoglicanos devido à deficiência da enzima lisossômica alfa-L-iduronidase (EC. 3.2.1.76) codificada pelo gene IDUA. Dentre as variantes para a doença, p.Trp402* é a mais encontrada em pacientes com MPS I em diferentes populações. CRISPR (do inglês: Clustered Regularly Interspaced Short Palindromic Repeats) é um mecanismo imunológico de procarionotes que promove a clivagem de uma região mediada por uma sequência de RNA com 20 nucleotídeos complementar ao sítio-alvo adjacente a sequência PAM. A edição gênica usando o sistema CRISPR possibilita o desenvolvimento de uma nova alternativa terapêutica aos pacientes com MPS I. Contudo, sequências similares a do sítio-alvo podem ser alvos de clivagem sendo denominadas de off-targets. **Objetivo:** Neste contexto, o objetivo deste trabalho é avaliar potenciais off-targets in silico para a variante p.Trp402* utilizando o sistema CRISPR/Cas9. **Metodologia:** Para isso 5 softwares públicos (CHOCHOP, COSMID, CRISPOR, Cas OFFinder, CCTOP) foram utilizados para a predição. A avaliação de potenciais off-targets considerou sequências com até 6 pares de bases diferentes (mismatches) e até 2 inserções ou deleções (indels) similares a do sítio-alvo que as tornem potenciais sítios de clivagem. Conjuntamente, deu-se preferência para sequências off-targets sem mismatches e/ou indels nos 5 bp adjacentes a PAM, posto que essa região está relacionada com a atividade da Cas9 no sítio-alvo. **Resultados:** Assim, 63 sequências foram obtidas como potenciais regiões de clivagem no genoma humano além do sítio-alvo sendo 21 sequências com até 6 mismatches e nenhum indel, e as demais contendo ambos. **Conclusão:** A avaliação de off-targets preditos será realizada a partir da montagem de um painel para sequenciamento de segunda geração após cultivo celular in vitro utilizando fibroblastos humanos. **Palavras chave:** MPS I, CRISPR/Cas9, Sequenciamento de nova geração.

eP2672**Aplicação de um Machine Learning no ciclo estral**

Lauren Ferro da Silva; Débora Barrogi Constantino; Nicoli Bertuol Xavier; Maria Paz Loayza Hidalgo; Marco Idriart
 HCPA - Hospital de Clínicas de Porto Alegre

Introdução: O ciclo estral é o ciclo reprodutivo das ratas, o qual consiste em mudanças fisiológicas e hormonais e pode ser dividido em 4 estágios: proestro, estro, diestro e metaestro. A identificação de cada fase é feita por avaliadores treinados com base no tipo e na proporção de células presentes em lavados vaginais vistas em microscopia ótica, sendo assim um processo que pode ter erros. Modelos de reconhecimento de imagem por machine learning possibilitam automatizar e padronizar o processo de identificação, diminuir a influência do observador e a necessidade de treinamento específico, podendo ser utilizado em estudos experimentais. **Objetivo:** Desenvolver um algoritmo de machine learning para classificar fases do ciclo estral com acurácia de 80%. **Metodologia:** As fotos utilizadas no projeto são oriundas do banco de dados do projeto "Avaliação do Efeito de Diferentes Padrões de Iluminação no Desenvolvimento Puberal de ratas Wistar", aprovado pelo Comitê de Ética (#HCPA 16-0378). O desenvolvimento do algoritmo ocorreu em duas fases. Na primeira, foram selecionadas 80 fotos de lavado vaginal com resolução de 800x800 pixels. Na segunda fase, foram utilizadas 500 novas imagens, com resolução de 400x400 pixels. Todas as fotos utilizadas foram previamente convertidas para escala de cinzas. Em ambas as fases, 80% das amostras foi utilizada para o conjunto de treino e 20% para o conjunto de teste. O tipo de técnica de machine learning utilizada foi o de redes neurais convolucionais, implementado pelo pacote Keras na linguagem Python 2.7. **Resultados:** O primeiro treinamento resultou em uma acurácia de 91%, porém o teste final teve uma acurácia de 60%. Na segunda fase foi obtida uma acurácia de 69% no teste, muito similar a acurácia do treinamento (73,6%), o que indica um menor sobreajuste e maior confiança em relação ao método. Além disso, foi analisado quais fases são classificadas incorretamente pelo algoritmo. Dentro destas, 53% dos erros pertencem a fase de diestro, 21% de proestro, 13% de estro e 13% de metaestro. **Conclusão:** Os resultados sugerem que o algoritmo foi capaz de identificar as fases do ciclo estral com uma acurácia de 69%. Como forma de aperfeiçoar este modelo, é fundamental a construção de um banco de dados maior e com imagens coletadas de forma padronizada. A aplicação dos resultados deste estudo pode ser construtiva para desenvolver futuros algoritmos capazes de identificar e diagnosticar processos biológicos em outros tipos de células.

eP2840**Perfil de expressão gênica associado ao prognóstico em Sarcoma de Ewing**

Matheus Gibeke Siqueira Dalmolin; Maurício Gomes de Queiroz; Ricardo Melo Ferreira; Caroline Brunetto de Farias; André Tesainer Brunetto; Mariane da Cunha Jaeger; Rafael Roesler; Marialva Sinigaglia; Rita Maria Cunha de Almeida
 Outras Instituições

Introdução: O Sarcoma de Ewing é um tumor altamente agressivo, afeta ossos e tecidos moles, sendo a segunda neoplasia óssea mais frequente em crianças e adultos jovens. Caracteriza-se pela presença de uma translocação envolvendo o gene EWS e outro da família ETS, geralmente FLI1. Apesar do tratamento para doença localizada apresentar uma eficácia comprovada, a sobrevida a longo prazo de pacientes com Ewing metastático ou que apresentam recidiva é ainda muito baixo. **Objetivo:** Selecionar genes marcadores de prognóstico em Sarcoma de Ewing. **Métodos:** Analisamos dados de expressão de Sarcoma de Ewing depositados no Gene Expression Omnibus (GEO) que continham metadados sobre o desfecho do paciente. Dados de três coortes de pacientes depositados no GEO: COG (GSE63155), EuroEwing (GSE63156) e Finlandeses (GSE17679) foram analisadas separadamente. Cada coorte foi classificada em dois grupos: Vivo (sobrevida acima de cinco anos) e Morto (sobrevida abaixo de 5 anos). A comparação entre os dois desfechos foi realizada utilizando o software The Transcriptogramer V.1. O enriquecimento funcional foi realizado no David Tools 6.8. **Resultados e discussão:** Foi observada diferença significativa entre o perfil de expressão de entre os grupos analisados apenas na coorte do COG. O grupo de genes diferencialmente expressos (GGDE) nesta coorte foi utilizado para realizar o enriquecimento funcional. As vias do Kegg, Reactome e os termos do Gene Ontology foram analisados e os dez genes mais significativos de cada via foram recuperados, totalizando 139 genes. Foi avaliada a curva de Kaplan Meier dos 139 genes na coorte do COG. Os genes significativos nesta coorte foram avaliados nos grupos do EuroEwing e dos Finlandeses. Ao final, 23 genes foram selecionados, destes, 17 genes tiveram alta expressão relacionada ao melhor prognóstico, e seis tiveram baixa expressão