



Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal

DECIFRAR AS PROPRIEDADES CHAVE PARA SUPERAR A RESISTÊNCIA A FÁRMACOS INDUZIDA POR BÁCTERIAS QUE POTENCIAM TAMBÉM A RESPOSTA IMUNE ANTICANCERÍGENA

BÁRBARA PINTO FREITAS, MARIA OLÍVIA PEREIRA E ANÁLIA LOURENÇO



4 JUNHO, 2025



TIMELINE OVERVIEW

01

DEFINIÇÃO DO PROBLEMA

02

REVISÃO DE LITERATURA

03

STATE OF ART

04

METODOLOGIA

05

APRESENTAÇÃO E
RESULTADOS PRELIMINARES

06

ANÁLISE PREDITIVA

07

RESULTADOS FINAIS E
CONSIDERAÇÕES

08

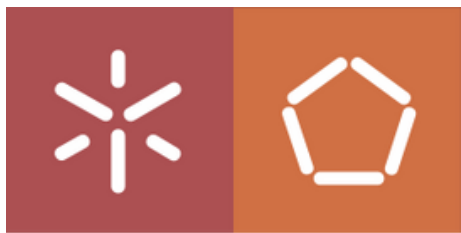
ARTIGO FINAL



BACKGROUND

- O uso excessivo e inadequado de antibióticos tem potenciado o desenvolvimento de resistência em microrganismos patogénicos.
 - Patógenos multirresistentes, nomeadamente os do grupo ESKAPE são uma das principais causas de infeções.
 - Paralelamente, o cancro continua a ser uma das principais causas de mortalidade, e a resistência a terapias convencionais constitui um dos maiores obstáculos ao sucesso clínico.
 - Torna-se, assim, imperativo identificar alternativas terapêuticas eficazes.
- A Organização Mundial da Saúde classifica a resistência antimicrobiana (AMR) como uma das 10 principais ameaças à saúde global, contabilizando em média 1,27 milhões de mortes por ano.





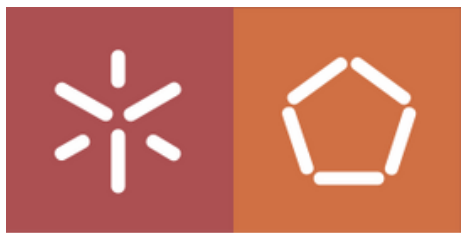
PÉPTIDOS ANTIMICROBIANOS- UMA ALTERNATIVA PROMISSORA

- Perante o evidente aumento da resistência aos antibióticos, os peptídeos antimicrobianos (AMPs) surgem como alternativas promissoras.

- Entre os AMPs, foi identificada uma classe com dupla função que será o alvo deste projeto:
 - Atividade antimicrobiana
 - Atividade anticancerígena

Esta dualidade torna-os candidatos ideais para enfrentar duas das maiores preocupações de saúde pública do século XXI:

- **RESISTÊNCIA
ANTIMICROBIANA**
- **RESISTÊNCIA
ANTICANCERÍGENA**



MECANISMOS DE AÇÃO

Os AMPs com atividade anticancerígena podem ser classificados, segundo Boohaker et al., em três categorias funcionais, com base nos seus mecanismos de ação:

Peptídeos formadores de poros → Magainin II

- Participam da defesa imune inata.

Peptídeos penetrantes de células → BR2

- Capazes de atravessar membranas celulares;
- Transportam compostos anticancerígenos para o interior das células.

Peptídeos direcionadores de tumor

- Reconhecem recetores de integrinas altamente expressos em células tumorais.



OBJETIVO

- O principal objetivo deste projeto é desenvolver ferramentas computacionais que possam identificar as propriedades que ajudam a superar quer a resistência antimicrobiana quer a anticancerígena.
- Inicialmente, o projeto centra-se na compilação de conjuntos de dados abrangentes de peptídeos antimicrobianos. Posteriormente, será desenvolvido um modelo para previsão.
- Em suma, pretende-se encontrar um modelo capaz de prever com precisão a atividade dupla dos candidatos presentes no conjunto de dados.



METODOLOGIA

1. RECOLHA DE DADOS

Extração de AMPs com atividade dupla da base de dados pública DRAMP e integração com peptídeos identificados através de revisão da literatura científica.

Remoção de duplicados.

2. FILTRAGEM E ANÁLISE DE SIMILARIDADE

Realização de um alinhamento global entre os peptídeos recolhidos e um conjunto de peptídeos apenas antimicrobianos, para identificar potenciais candidatos com atividade não anotada.

3. EXTRAÇÃO DE DESCRITORES DE CARGA

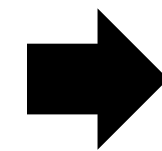
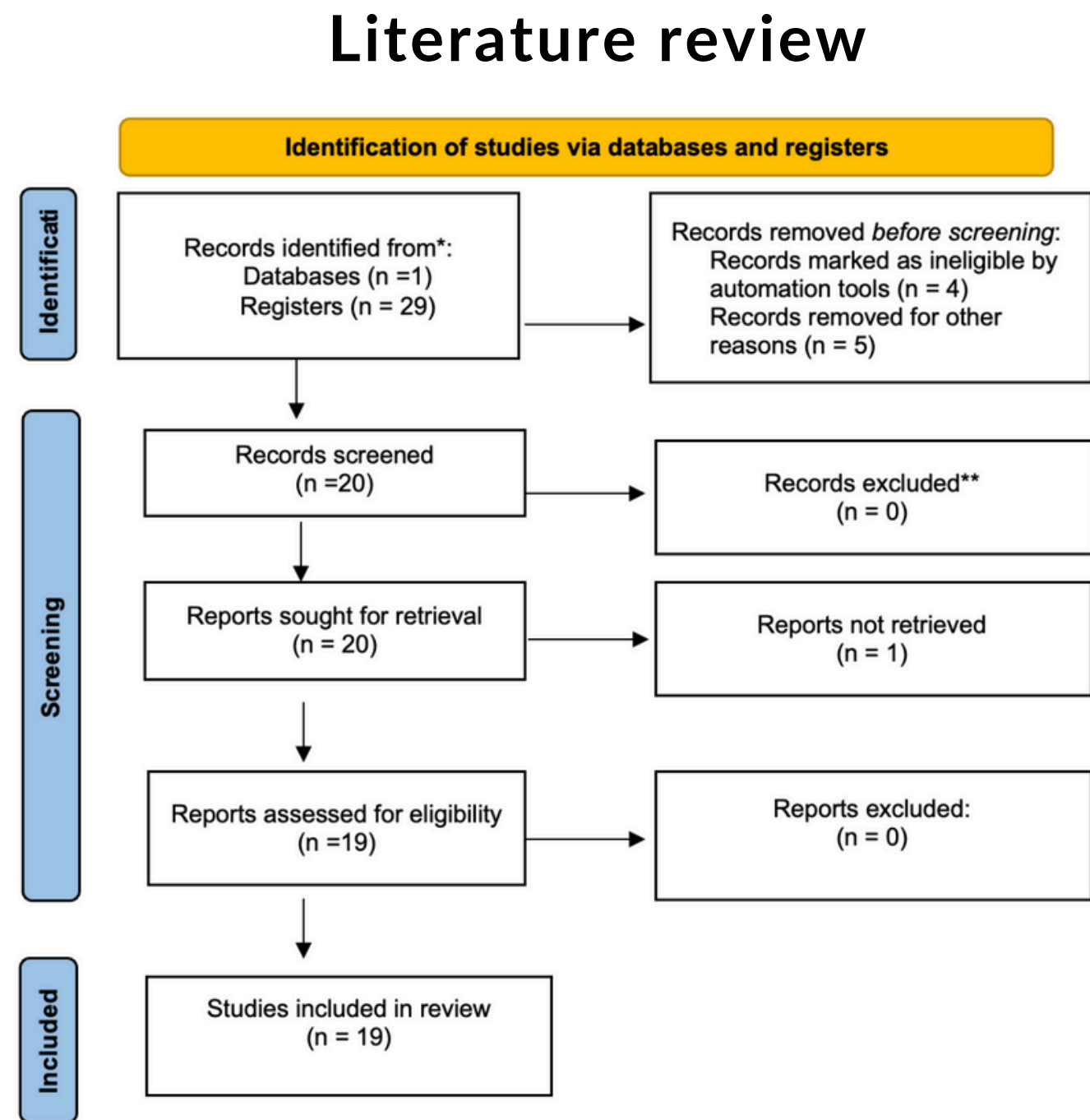
Foram extraídos descritores relacionados com a carga positiva das sequências, como a frequência de resíduos catiónicos (K, R, H) e a sua distribuição ao longo da cadeia peptídica.

4. COMPARAÇÃO ADICIONAL VIA BLAST

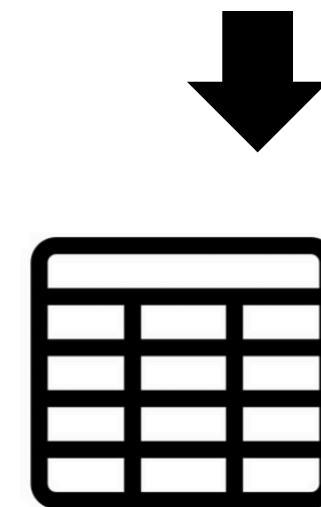
BLAST com um dataset negativo (non-AMPs) para reforçar a especificidade da seleção.

*DRAMP: <http://dramp.cpu-bioinfor.org>

METODOLOGIA

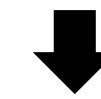


AMPs com dupla
atividade

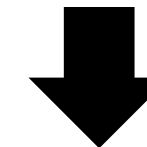
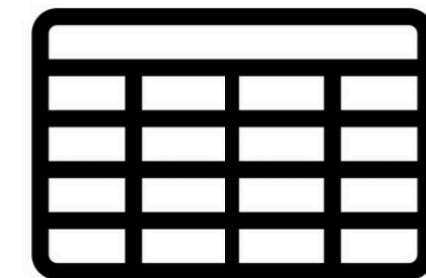
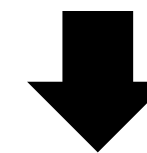


+

DRAMP



filtrar “antimicrobial”
“anticancer”



- Remover duplicados

Guidelines PRISMA e keywords como: “AMPs”,
“antimicrobial”, “anticancer”



Alinhamento “pairwise”:

AMPs filtrados com dupla actividade e AMPs com atividade "antimicrobial"

Cálculo e categorização da identidade global de cada par de péptidos, através função para cálculo da percentagem de identidade:

Alto grau de identidade ($\geq 90\%$) – strict.tsv

Identidade moderada ($< 90\%$) – relaxed.tsv

Alinharam 4009 AMPs com $\geq 50\%$ identidade que não estão no dataset

- A maioria dos peptídeos de função dupla tem similaridade parcial com os peptídeos antimicrobianos simples.

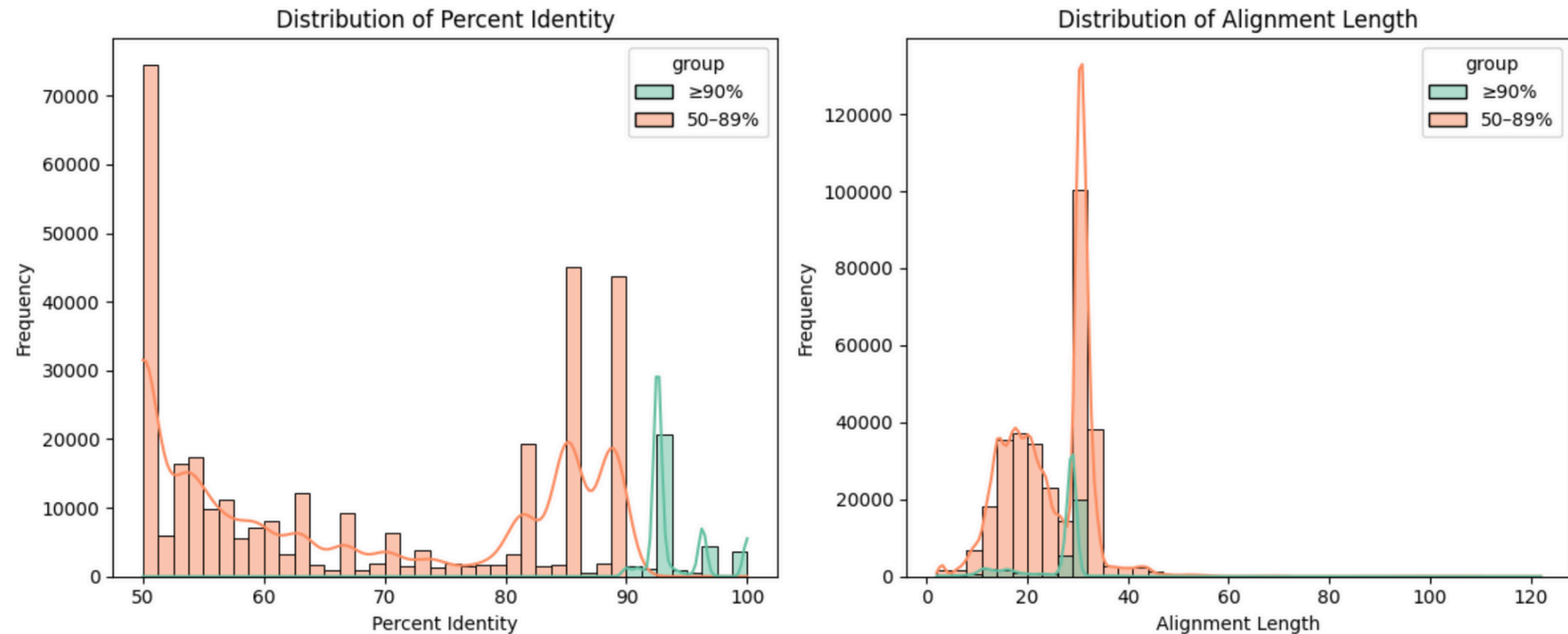


Figura 1. Gráficos obtidos dos resultados do alinhamento “pairwise”



Alinhamento “pairwise”: Alinhamentos mais significativos

- Foram identificadas múltiplas variantes da família peptídica Bim BH3. As diferentes mutações introduzidas não influenciaram essas funções, sugerindo robustez estrutural e funcional. Esta família representa um alvo promissor para engenharia racional de peptídeos terapêuticos de função dupla. Neste contexto, os derivados de Bim BH3 destacam-se como candidatos particularmente relevantes.

	Name	Activity
9701	Bim BH3 (E158S)	Antimicrobial, Anticancer
9698	Bim BH3 (I155R, E158A)	Antimicrobial, Anticancer
9710	Bim	Antimicrobial, Anticancer
9700	p-Bim BH3 (R154S, I155R, E158S)	Antimicrobial, Anticancer
9699	Bim BH3 (R154S, I155R, E158S)	Antimicrobial, Anticancer
9697	p-Bim BH3 (I155R, E158S)	Antimicrobial, Anticancer
9696	Bim BH3 (I155R, E158S)	Antimicrobial, Anticancer
9713	Bim-A2eT-E2gG	Antimicrobial, Anticancer
9711	Bim-A2eT	Antimicrobial, Anticancer
9712	Bim-A2eT-I2dM	Antimicrobial, Anticancer

Figura 2. Resultado obtido dos dez alinhamentos mais significativos



Descritor Baseado na Distribuição de Carga

A categorização dos resíduos com base na carga elétrica é uma abordagem bem estabelecida na análise de péptidos antimicrobianos .

Definiram-se classes de carga para os aminoácidos: positivos, negativos e neutros.

Implementou-se uma função para calcular a distribuição dos resíduos carregados ao longo das sequências peptídicas. As características extraídas da distribuição de carga foram integradas no dataset original.

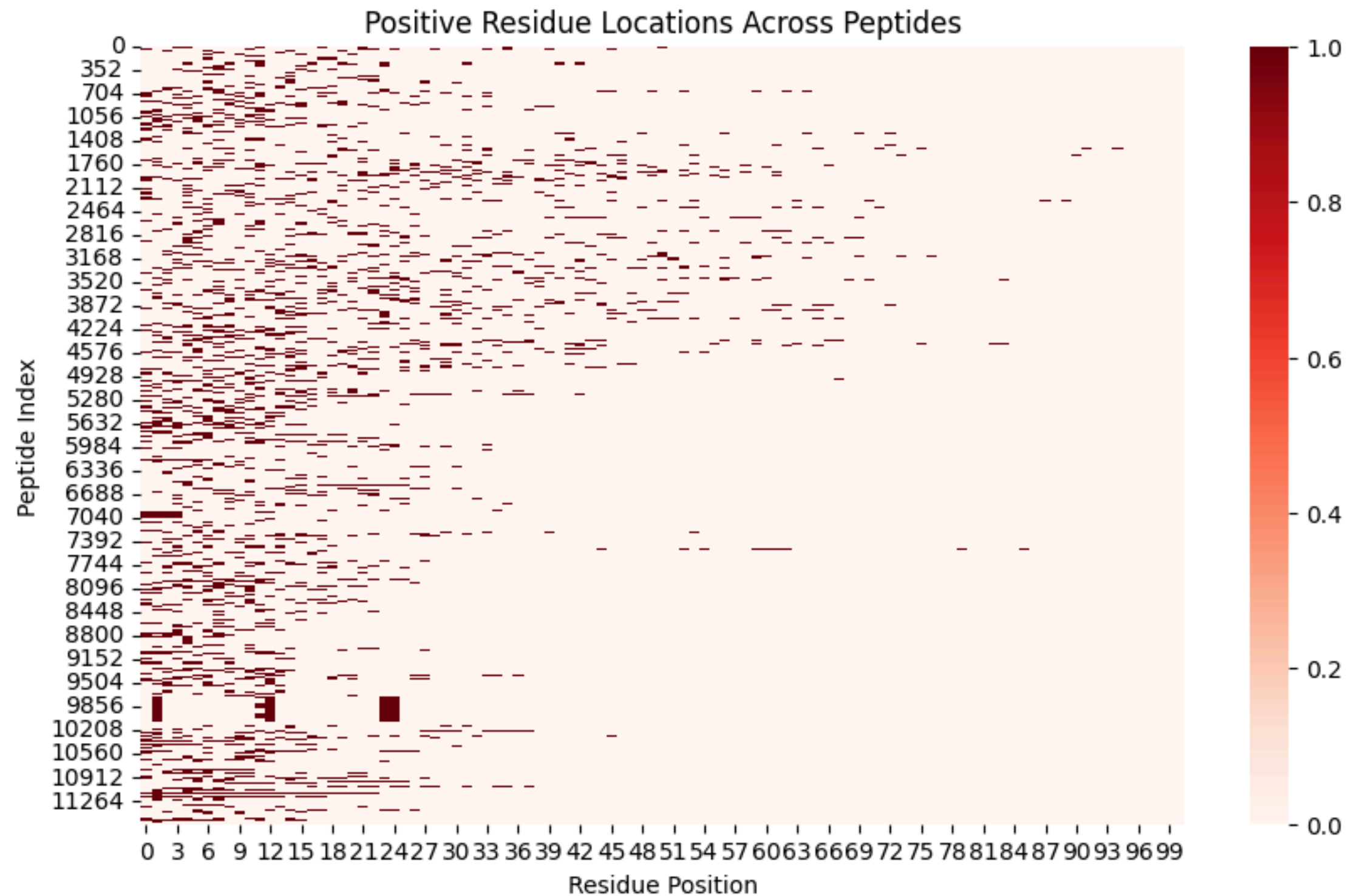
Feature que queremos extrair: Distribuição de resíduos carregados positivamente na sequência

```
charge_classes = {  
    'positive': ['K', 'R', 'H'],  
    'negative': ['D', 'E'],  
    'neutral': ['A', 'N', 'C', 'Q', 'G', 'I', 'L', 'M', 'F', 'P', 'S', 'T', 'W', 'Y', 'V']  
}
```

Figura 3. Representação da definição das classes de carga para aminoácidos positivos, negativos e neutros



Distribuição de carga positiva: Visualização



- Resíduos carregados positivamente (lisina - K, arginina - R e histidina - H).
- Distribuição de resíduos positivos ao longo das sequências, com maior densidade na região inicial;
- Este padrão é compatível com mecanismos de interação com membranas celulares negativamente carregadas.

Figura 4. Distribuição da carga positiva ao longo dos peptídeos. Cada linha representa um peptídeo e cada coluna uma posição relativa (0% a 100%) ao longo da sequência.



Distribuição de carga positiva: Coorelação

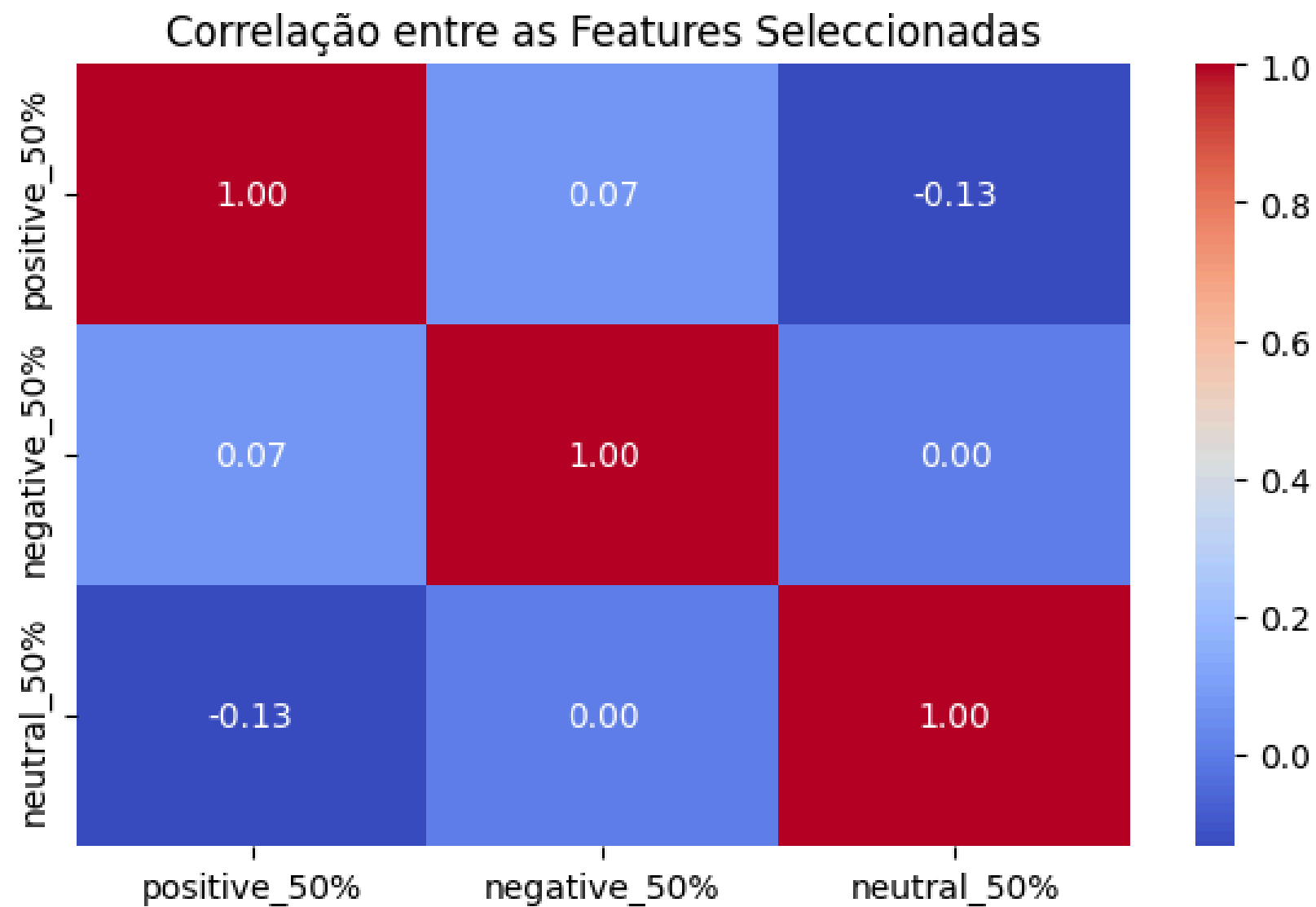


Figura 5. Visualização da correlação de Pearson entre as três variáveis selecionadas: positive_50%, negative_50% e neutral_50%.

Foram avaliadas três variáveis representativas da distribuição de carga ao longo da sequência: positive_50%, negative_50% e neutral_50%.

- Os resultados mostram correlações fracas ou nulas entre estas variáveis, evidenciando a sua independência estatística isto garante que cada feature contribui de forma única para a representação da sequência.



Análise de Homologia via BLAST: Dual-Activity AMPs vs. Non-AMPs

O objetivo é identificar sequências com alta similaridade entre os 2 grupos, o que poderia indicar falsos positivos ou péptidos redundantes. Assegura que os péptidos selecionados são aptos para downstream machine learning.

A presença destas sequências poderia numa fase futura comprometer a eficácia dos modelos de previsão ao introduzir ruído e ambiguidade.

Etapa inicial de limpeza das sequências para manter apenas os 20 aminoácidos standard (A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y). As sequências que contêm aminoácidos ambíguos ou raros podem introduzir ruído ou erros.



Filtragem baseada em homologia com conjunto negativo

```
# Define a threshold for significant matches
threshold_identity = 50

# Filter out the query sequences that have high identity alignments with the negative dataset
significant_matches = blast_results[blast_results['percent_identity'] >= threshold_identity]

# Get the list of query IDs to exclude
exclude_ids = significant_matches['query_id'].unique()

# Load the original peptides dataset
df = pd.read_excel("general_amps2.xlsx")

# Filter out the sequences from the dataset that have significant matches
filtered_df = combined_df[~combined_df['Sequence'].isin(exclude_ids)]

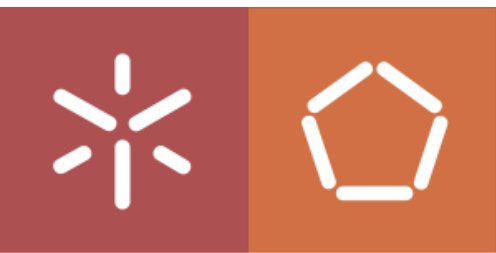
# Save the filtered dataset
filtered_df.to_excel("filtered_amps_without_negative_matches.xlsx", index=False)

print(f"Total peptides removed due to high-identity matches: {len(combined_df) - len(filtered_df)}")

Total peptides removed due to high-identity matches: 0
```

- Mesmo adotando um critério de identidade relativamente permissivo (50% de similaridade), não foram encontrados alinhamentos significativos com o conjunto negativo. Este resultado reforça que as sequências apresentam características distintivas em relação aos peptídeos inativos, reforçando a viabilidade do conjunto com atividade comprovada.
- Este conjunto negativo foi obtido a partir de um repositório online disponível <https://sourceforge.net/projects/axpep/>

Figura 6. Parte da implementação do pipeline de homologia com base em resultados BLAST



DATASET NEGATIVO

A definição de um conjunto negativo confiável é crítica para o desempenho e validade dos modelos de machine learning aplicados à previsão da atividade de AMPs. Neste trabalho foi utilizada um dataset negativo disponível publicamente de forma a permitir a aplicação de testes de similaridade e exclusão de redundâncias. A tabela seguinte apresenta alguns estudos da literatura que abordam a construção de conjuntos negativos para AMPs, destacando as fontes de dados utilizadas e os critérios aplicados para assegurar a ausência de atividade.

Tabela 1. Resumo das abordagens utilizadas na literatura para a definição de datasets negativos em modelos de previsão de AMPs

Nome do artigo	Fonte datasets negativos	Método
AmPEP: Sequence-based prediction of antimicrobial peptides using distribution patterns of amino acid properties and random forests	UniProt	Filtro NOT antimicrobial + CD-HIT
UniAMP: enhancing AMP prediction using deep neural networks with a large-scale dataset	UniProt	NOT antimicrobial + CD-HIT + ajuste comprimento
Benchmarks in antimicrobial peptide prediction are biased due to the selection of negative data	APD3 + EMEM	Anotação non-AMP + dataset anterior
AMP-RNNpro: a two-stage approach for identification of antimicrobial peptides and their functional activities	APD3 + artigos anteriores	Anotação non-AMP
SGAC: A Graph Neural Network Framework for Imbalanced and Structure-Aware AMP Classification	UniProt	NOT antimicrobial + clustering (pseudo-label)

RESULTADOS FUTUROS



Atualmente, está disponível um dataset robusto de AMPs com função dupla confirmada, devidamente caracterizados com descritores.

O desenvolvimento subsequente do projeto incidirá na implementação e validação de um modelo de machine learning de previsão robusto.

- Aplicação de algoritmos supervisionados;
- Avaliação com métricas padrão;
- Estratégia de validação cruzada para garantir generalização e robustez do modelo.





CONCLUSÃO:

A resistência antimicrobiana e a resistência no tratamento do cancro representam dois dos maiores desafios de saúde pública atuais. Os peptídeos antimicrobianos multifuncionais são explorados como alternativas terapêuticas.

O modelo preditivo resultante será uma ferramenta *in silico* para a identificação eficiente de AMPs multifuncionais, contribuindo para o avanço de terapias contra infeções resistentes e cancro.

A longo prazo, esta estratégia pode acelerar a descoberta racional de novos tratamentos e apoiar o desenvolvimento de agentes bioativos.

**OBRIGADA PELA
ATENÇÃO!**

BÁRBARA PINTO FREITAS, MARIA OLÍVIA PEREIRA E ANÁLIA LOURENÇO