

Facilitating the Screening of AMP Literature via Text Mining Strategies

Paula Jorge, Maria Olívia Pereira and Anália Lourenço*

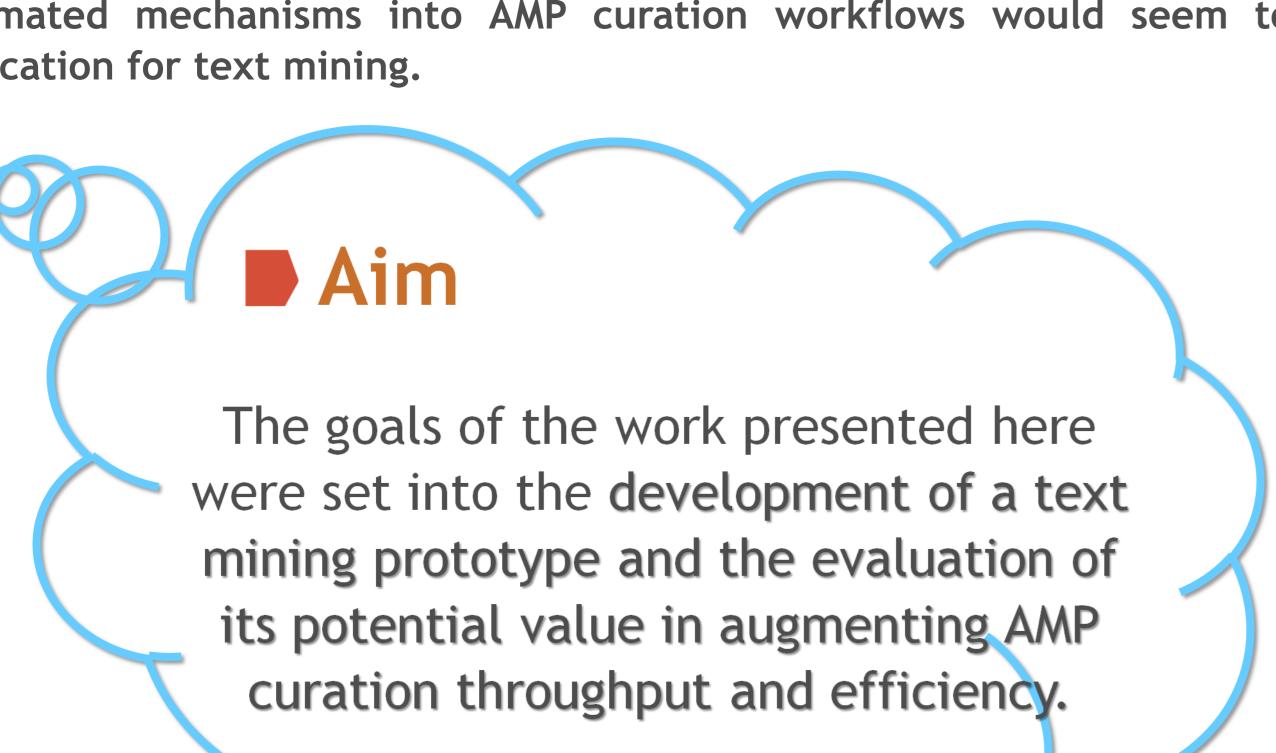
*IBB-Institute for Biotechnology and Bioengineering, Centre of Biological Engineering, Universidade do Minho, Campus de Gualtar 4710-057, Braga, Portugal

Introduction

Antimicrobial peptides (AMPs) are nowadays importantly recognized, mostly due to the raise of infections caused by multi-resistant pathogens, and have been hailed as a potential solution to the shortage of novel antibiotics, due to their general unspecific mechanism of action which impairs the acquiring of resistance.

Considerable efforts have been called to the study of antimicrobials with new mechanisms of actions, leading to a proficiency of scientific publications and an emergence of specialized databases. As the volume of related literature increases, it becomes urgent to help researchers and curators keeping up with it, especially by providing means for the systematic screening and prioritization of literature of interest.

Given the success reported in other biomedical domains, the integration of semiautomated mechanisms into AMP curation workflows would seem to be an ideal application for text mining.

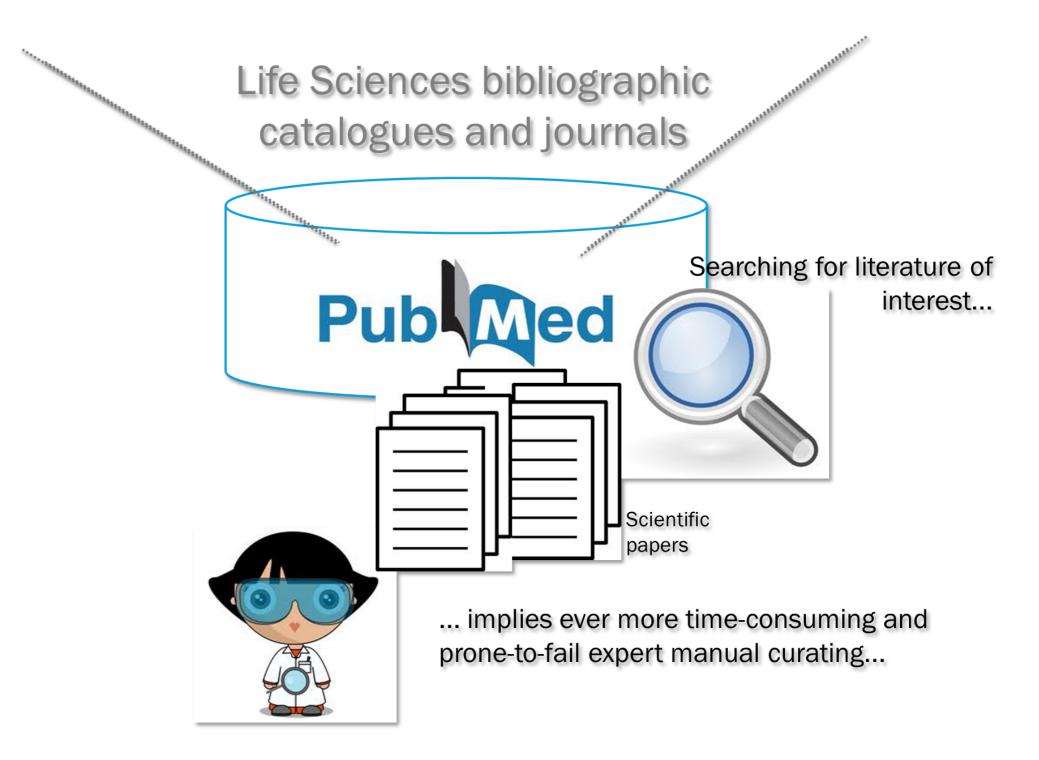


Methods

The text mining prototype was developed and evaluated using a corpus consisting of 1012 articles, published during 2011 and early 2012. The strategy to recognise biological entities of interest included:

- State-of-the-art bio-recognisers were used to tag organisms (LINNAEUS), chemical entities (OSCAR 4), and genes, proteins and other biological entities (ABNER).
- AMP terms were recognised using dictionary-based pattern matching. Vocabulary on AMPs derived from UniProt knowledge base, antibiotics lexicon from ARDB database and the antibiotics list in Wikipedia.
- An in-house lookup list of textual clues on AMP bio-activities and mechanisms of action was also included.

Article screening and ranking were based on statistical considerations. Especially, the number and diversity of unique concepts detected in the texts weighted by the relevance associate of the different biological categories and the degree of certainty associated to detection.



HOW CAN RESEARCHERS KEEP UP WITH NEW FINDINGS? HOW CAN DATABASES BE KEPT UPDATED?



Text Mining:

screening + relevance assessment + information extraction

Amps vial peptide hCAP-18/LL-37 protein and mRNA expressions in different periodontal disease **Proteins** nd mRNA expression in gingival tissues with different periodontal disease. MATERIALS AND nic periodontitis, and 10 healthy controls were included in this study. Periodontal parameters including pri index were assessed in study subjects. hCAP-18/LL-37 mRNA analysis by RT-PCR and immunohistoche concentration and integrity. RESULTS: This study demonstrated that hCAP-18/LL-37 was a product of neu ntly higher immunostaining of hCAP-18/LL-37 on neutrophils infiltrating in both epithelium and connective tissue samples of chronic periodontitis patients seemed to be upregulated compared with controls. While ulated hCAP-18/LL-37 mRNA expression levels, one generalized aggressive periodontitis patient showed CONCLUSIONS: hCAP-18/LL-37 has an important role in innate response during periodontal in Cells thogenesis of generalized aggressive periodontitis.

s in Yersinia ruckeri. We have previously demonstrated that Yersinia ruckeri resists cecropin B in a sible for the inducible cecropin B resistance of Y. ruckeri. Differences in gene expression associated kposed to a sublethal concentration of cecropin B and resultant changes in the messenger RNA pop iption polymerase chain reaction (DD-RT-PCR). A single band was consistently increased in intensit id used to screen a Y. ruckeri genomic DNA library. The DD-RT-PCR fragment shared 100% identity to ization defect (OLD) family of Y. ruckeri 29473. The genomic clone that was recovered was not identi (nucM) homologue. It was determined that transcription of the gene was upregulated following exposiof culture supernatants of Y. ruckeri following exposure to cecropin B was demonstrated. These find vo endonucleases in Y. ruckeri. The production and secretion of an endonuclease by Y. ruckeri in recellular and extracellular DNN in the toxic effects of cecropin B.

Genes







Examples of abstracts with highlighted and automatically annotated terms.

Organisms

Preliminary Results and Ongoing Work

- Preliminary prototype results are available at http://stardust.deb.uminho.pt/amps.
- A list of articles ranked by year, pmid, title and journal is presented as proof of concept. Users will be able to order articles by these categories.
- To retrieve descriptions of AMP sources or targets, users are able to search using as a query the taxonomic name or common name of the organism of interest.
- Also, search for commons drugs and antibiotics present in the papers is available, as well as for genes and proteins.

AMP Text Mining Computer-assisted compilation of articles This is a first proof of concept of the text mining antibiofilm data curation system. It encompasses articles published during 2011 and early in 2012. Use free text to retrieve articles mentioning AMPs, source or target organisms, or genes and proteins of interes Add another keyword | Search Eur J Clin Microbiol 21311938 2011 Bactericidal synergy of lysostaphin in combination with antimicrobial peptide Infect Dis Int J Antimicrob road-spectrum antimicrobial efficacy of peptide A3-APO in mouse models of multidrug-resistant wound and lung 21353493 2011 Agents Microbiology

Acknowledgments

The financial support from IBB-CEB, Fundação para a Ciência e Tecnologia (FCT) and European Community fund FEDER, through Program COMPETE, in the ambit of the FCT PTDC/SAU-SAP/113196/2009/FCOMP-01-0124-FEDER-016012 are gratefully acknowledged.





