

Abstract

This thesis focuses on an innovative approach combining two high-throughput methods, DNA metabarcoding and bottom-up proteomic analysis, for the detection and identification of allergens in dust, including risk organisms.

The theoretical part focuses on the composition of dust, its physicochemical properties and the main groups of allergens and the main agents such as dust mites, fungi, bacteria and various environmental particles.

The methodological part deals with the possibilities of dust particle sampling, detection methods and identification of allergens. As a more effective alternative to immunochemical methods, proteomic analysis is presented, which allows comprehensive identification of a wide range of proteins and allergens, including potential uncharacterized allergens and proteins indicating the presence of risk microorganisms. A modern approach involving DNA sequencing is described as a complementary method that provides detailed information on the taxonomic composition of samples, which significantly improves the accuracy of identification at the protein level.

The practical part of the work combines proteomic and genomic analysis for detailed characterization of the dust samples. Optimization of the DNA isolation protocol was performed, followed by successful sequencing of the amplified 16S rRNA and 18S rRNA genes, which provided information on prokaryotic and eukaryotic organisms. Information from DNA metabarcoding was used to create a database for the evaluation of proteomic data from nanoLC-MS/MS. This combined approach enabled the identification of a wide range of allergens from grasses, animal allergens, human proteins or common indoor and outdoor microorganisms such as *Cladosporium*, *Aspergillus* and *Alternaria*. An approach based on protein splicing using feces of the mite *D. farinae* was also tested as a standard for assessing the presence of mite allergens in dust samples. Samples from a variety of environments were tested, including public places with high concentrations of people and CARC areas that are rich in allergen producers. Samples were collected by two methods, by swabbing and by using an air purifier, and the results were compared.

Overall, this work demonstrates that the combination of proteomic and genomic methods and the integration of data from them can provide a comprehensive picture of the biological composition of dust and its potential health risks, and highlights the importance of a proper database setup.

Keywords: dust, allergens, proteomic analysis, metabarcoding, DNA, proteins