

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PH.D.)

CONSTRUCTION OF ANTIMICROBIAL AND IMMUNOMODULATORY PEPTIDE
DATABASE AND *IN SILICO* ANALYSIS OF THE ANTIMICROBIAL PEPTIDES IN
DIFFERENT PATHOLOGICAL CONDITIONS

By Ajneesh Kumar

Supervisor: Dr. Éva Csősz



UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF MOLECULAR, CELLULAR AND IMMUNE BIOLOGY

DEBRECEN, 2022

Table of Contents

Abbreviations	4
1. INTRODUCTION.....	5
1.1 Antimicrobial peptides as a part of the innate immune system.....	5
1.2 Mechanisms of AMP action.....	6
1.3 Structural classification of AMPs	6
1.4 AMPs present in the bodily fluids.....	7
1.4.1 Tear AMPs.....	7
1.4.2 Salivary AMPs.....	7
1.4.3 Blood AMPs	8
1.4.4 Urine AMPs.....	8
1.4.5 Cerebrospinal fluid AMPs	9
1.5 Role of AMPs in Alzheimer's disease	9
1.6. Role of AMPs in oral squamous cell carcinoma	10
1.7 Network analysis and visualization tools	11
1.7.1 String-DB database.....	11
1.7.2 Cytoscape network analysis tool	11
2. AIMS	13
3. METHODS	14
3.1 Generation of Antimicrobial and Immunomodulatory Peptide database (UDAMP Database).....	14
3.2 Examination of AMPs in proteomic datasets related to Alzheimer's disease	15
3.3 Network analysis of datasets related to Alzheimer's disease.....	16
3.4. Network analysis of data related to oral squamous cell carcinoma	16
4. RESULTS	18
4.1 Generation of the comprehensive human antimicrobial and immunomodulatory peptide database (UDAMP Database)	18
4.2 Functional examination of the human AMPs.....	19

4.2.1 Gene ontology (GO) of the human AMPs.....	19
4.2.2 Protein-Protein interaction network analysis of the human AMPs	19
4.2.3 Gene-gene interaction analysis of the human AMPs.....	20
4.2.4 Top GO terms enriched for AMPs in the brain, CSF, and blood	22
4.2.5 Gene-gene interaction network for the AMPs found in the brain, CSF, and blood	23
4.3 AMPs characteristic of Alzheimer's disease in the brain, CSF, and blood	24
4.3.1. Functional analysis of the network of AMPs characteristic of Alzheimer’s disease in brain, CSF, and blood.....	25
4.3.2 Network analysis of the AMPs characteristic of the AD in brain, CSF and blood .	29
4.4 Examination of antimicrobial peptides in oral squamous cell carcinoma.....	31
4.4.1 Examination of selected salivary proteins	31
4.4.2 Differentially expressed proteins in different oral disease groups	35
4.4.3 Network analysis of proteins and AMPs with statistically significant changes among the disease groups	37
4.4.4 Gene-gene interaction network analysis of differentially expressed proteins between the studied groups.....	41
5. DISCUSSION	44
6. NOVEL FINDINGS.....	49
7. SUMMARY	50
8. ACKNOWLEDGMENT	51
9. REFERENCES.....	52

Abbreviations

AD – Alzheimer’s disease	IL6 – Interleukin 6
AMPs – Antimicrobial and Immunomodulatory Peptides	NPY – Neuropeptide Y
ANXA1 – Annexin-1	OLK – Oral leukoplakia
APP – Amyloid-beta precursor protein	OLP – Oral lichen planus
B2M – Beta-2-microglobulin	OSCC – Oral squamous cell carcinoma
CCL11 – Eotaxin	PEA – Proximity Extension Assay
CCL – C-C motif chemokine 2, 4, 11, 17, 20, 21, 25, 28	PPBP – Platelet precursor basic protein
CCR - C-C chemokine receptor type 1,5 ,7	PPI – Protein-protein interaction
CLU – Clusterin	RETN – Resistin
CSF – Cerebrospinal fluid	S100A7 – Protein S100A7
CXCL1 – Growth regulated alpha protein	S100A9 – Protein S100A9
CXCL – C-X-C motif chemokine 1, 5, 8, 9, 10, 11, 12	SERPINA1 – Alpha-1-antitrypsin
CYR61 – CCN family member 1	SERPIND1 – Heparin Cofactor 2
DEFA4 – Defensin alpha 4	SERPINE1 – Plasminogen activator inhibitor 1
EGF – Epidermal growth factor	SLPI – Antileukoprotinase
F2 – Prothrombin	TAC1 – Protachykinin 11L
FAU – FAU ubiquitin-like and ribosomal protein S30	TLR3 – Toll-like receptor 3
FGA– Fibrinogen alpha chain	TNF – Tumor necrosis factor
GIN – Gene-gene interaction networks	TNFRSF – Tumor necrosis factor receptor superfamily members 10A, 10C, 10D, 10B
GO – Gene Ontology	UDAMP – University of Debrecen Antimicrobial and Immunomodulatory Peptide Database
GPNMB – Transmembrane glycoprotein NMB	UTIs – Urinary tract infections
HP – Haptoglobin	VGF – Neurosecretory VGF
	VIP – Vimentin peptides

1. INTRODUCTION

Multicellular organisms protect themselves from harmful microbes. The skin, conjunctiva, respiratory tract and lungs, mouth, digestive tract, urinary tract, and reproductive tract are the best-known places for the first encounter with microbes (1). Antimicrobial and immunomodulatory peptides (AMPs) are the first line of defence in various tissues and organs, protecting our bodies from potential pathogens (1,2). AMPs are present in all body fluids; they can be found in saliva, sweat, blood, tear, and urine. AMPs have a broad spectrum of action and can be anionic, cationic, and amphipathic (3). They can be divided into three categories based on their origin: (i) natural, ribosomally synthesized peptides, (ii) non-ribosomally produced natural peptides, or (iii) non-natural synthetic peptides (4).

1.1 Antimicrobial peptides as a part of the innate immune system

AMPs are the component of the first line of defence (5). Several AMPs are essential in the innate immune system in different organisms from lower eukaryotes to humans (2). They can induce various responses, i.e., cell differentiation, cell recruitments, antibacterial, antiviral and antifungal responses, etc. (Figure1) (6).

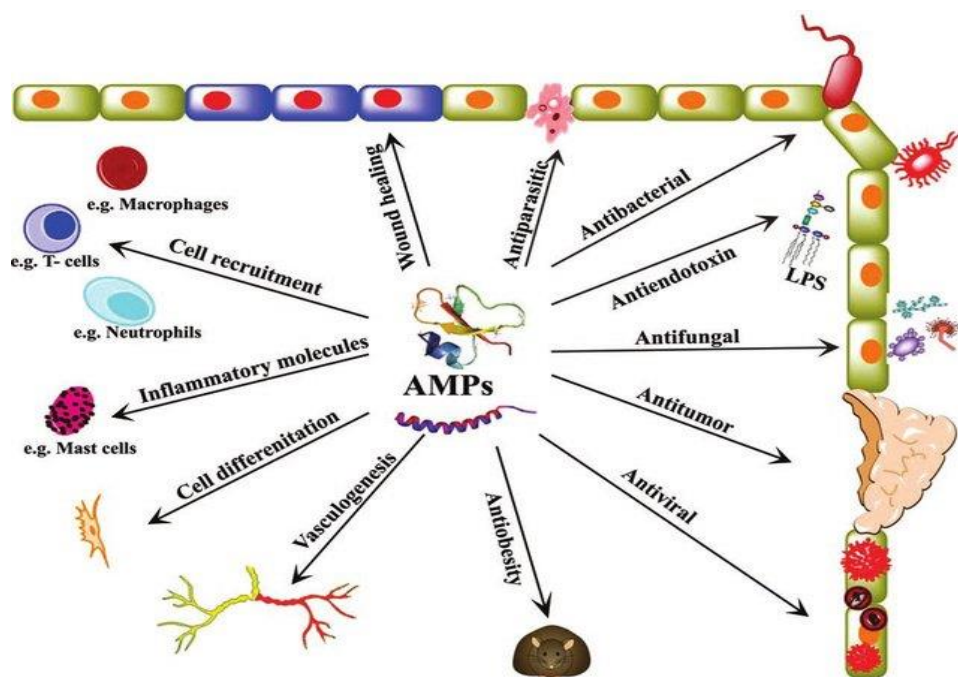


Figure 1: Responses produced by AMPs at different sites of infection (6).

1.2 Mechanisms of AMP action

Understanding the AMP's mechanism of action will drive their development as potential therapeutic agents against pathogens. Initially, targeting the membrane was thought to be the only mechanism of action, but now there is more information that AMPs have different mode of actions (7). There are two types of the mechanism of action of AMPs: direct killing and immunomodulation (Figure 2).

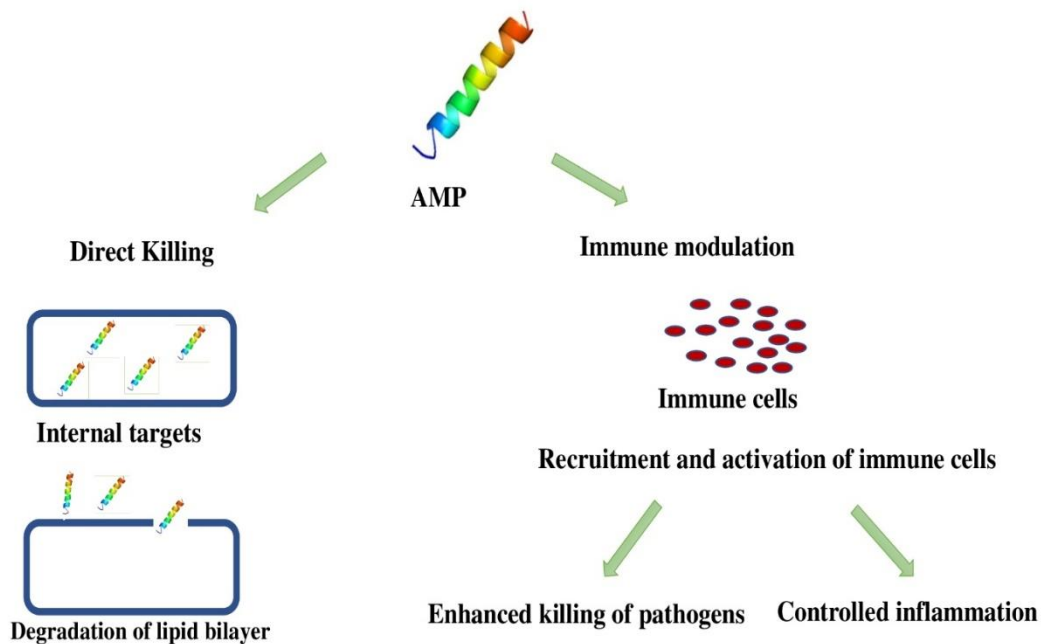


Figure 2: Mechanisms of action of AMPs: direct killing and immune modulation (7).

The cationic AMPs are direct killers. They interact with the negative charge of the bacterial cell wall, form a pore-like structure, and disrupt the cell wall killing the bacteria (7). On the other hand, AMPs can do immunomodulation by recruiting the immune cells to kill the pathogens and control the inflammation at the site of infection (7).

1.3 Structural classification of AMPs

There are many naturally occurring AMPs, and their classification can be challenging. AMPs have a broad range of lengths; usually, they can vary from 10 to 150 amino acids (8,9). The most common way to classify AMPs is the structural classification. Based on the primary, secondary and tertiary structures, the AMPs are of four types: (i) linear alpha-helical peptides, e.g., cathelicidin LL-37, magainin-2, (ii) linear peptides rich in particular amino acids, e.g.,

indolicidin, antibacterial protein PR-39, (iii) beta-sheets containing peptides, e.g., protegrin, human defensin-5 dimer, and (iv) peptides with alpha and beta structural elements, e.g., phorimicin, beta defensin-1 (9).

AMPs typically have distinct structural domains, such as a defensin fold in defensins and a cathelicidin domain in cathelicidins providing with extra stability even in harsh environment, such as high salt concentration, low pH, etc. (10–12), but some, such as cationic intrinsically disordered peptides lack a well-defined three dimensional structure (13).

1.4 AMPs present in the bodily fluids

Endogenous AMPs are present in human body fluids. Many AMPs have been identified in the blood, tear, saliva, urine, and cerebrospinal fluid (CSF) (14).

1.4.1 Tear AMPs

The tear film that coats the cornea and conjunctiva serves vital functions. It lubricates the ocular surface epithelia, helps to provide a smooth surface for refracting light, supplies oxygen, and is an important component of the eye's innate defence system, protecting against various potential pathogens (15). A variety of microorganisms can infect the cornea and conjunctiva. *Pseudomonas aeruginosa*, *Serratia marcescens*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, and *Streptococcus pneumoniae* are the most common bacterial pathogens of the ocular surface (16). Because it is in direct contact with the external environment, the tear film serves as the first line of defense against pathogens attempting to breach the physical barrier created by the ocular surface epithelia. The tear film helps washing out invading pathogens in conjunction with blinking and reflex tearing. It contains a plethora of AMPs, such as lysozymes, lipocalins, lacritins, and lactoferrins produced by lacrimal gland, or dermcidin produced by epithelial cells (15) to facilitate pathogen killing or to inhibit microorganism replication (17–19).

1.4.2 Salivary AMPs

Microorganisms and pathogens easily access the oral cavity and the rest of the body through the gastrointestinal tract. The oral cavity has a high microbial load, which can lead to a variety of oral diseases (20). To combat the various types of oral diseases, saliva contains a number of AMPs having a role in the host-defence. Because of its antibacterial, antioxidant, and antifungal properties, saliva, and oral mucosa, both serve as a potential line of defence to

encounter pathogens (20). As a response of host-defence mechanism, many AMPs are produced by salivary glands and found in the saliva (14). The most abundant ones are alpha-amylase, mucins, cystatins, proline-rich peptides, and serum albumin (21). Some other AMPs were identified in the saliva, such as histatins, defensins, LL-37 cathelicidins, and saliv (14,22,23).

1.4.3 Blood AMPs

Blood is an essential component of the human body. It is also a crucial component of the innate immune system. Blood is thought to contain vital components that mediate the immune response to infections and pathogens. Blood cells help blood-borne immune components to circulate freely in the human body. Some of these so-called blood-borne immune components are the AMPs (24).

From the approximately 12,000 identified serum proteins, more than 400 can act in the first line of host defense. Various types of AMPs are deployed by the blood cells at the site of infection. Antimicrobial peptides can be present in blood cells (25), the neutrophil granules for example, contain AMPs such as lactoferrin, cathelicidins and lysozymes (24). AMPs derived from bigger proteins can also be found in blood such as haemoglobin-derived peptides, the hemocidins (26).

1.4.4 Urine AMPs

Urine is a significant bodily fluid with an important excretory function. It can also be the ground for a variety of microbial infections present in urinary tract. These types of infections are known as urinary tract infection (UTIs). UTIs are the most common type of the human infections. To protect the human body from these UTIs, AMPs are produced in the urinary tract. Urine contains more than 7000 proteins and many of the urine proteins are part of the defense system (27). The AMPs present in urine are mostly produced by the epithelial cells in the urinary tract (28). The most abundant AMPs in urine were found to be serum albumin, uromodulin, alpha-1-microglobulin, kininogen and various immunoglobulin chains (29). AMPs such as ribonuclease 6 and 7 were also found in the urine and aid in urinary tract sterility (30). Other AMPs found in urine are neuropeptide Y, protachykinin-1, and vasoactive intestinal peptides (14). Cathelicidin is another AMP found in lower concentrations in adult human urine and higher concentrations in children with urinary tract infections (31).

1.4.5 Cerebrospinal fluid AMPs

The CSF is found in ventricles, in the subarachnoid spaces of the cranium and spine. It helps in supplying nutrients, in waste elimination, and brain protection being a fluid coat and protects the central nervous system (32). CSF is an essential fluid in the diagnosis of many neurodegenerative conditions, such as Alzheimer's disease and Parkinson disease. The progression of certain neurodegenerative diseases can be monitored by examining CSF protein amounts (32,33).

Until now, more than 4000 proteins have been identified in the human CSF (27). Approximately 8% of its proteins take part in host defense, for e.g., amyloid-beta precursor protein (APP), clusterin (CLU), and protein S100A7 (S100A7) (34,35). It was also discovered that the level of LL-37 cathelicidin and defensins' in CSF is increased upon bacterial challenge (36–38).

1.5 Role of AMPs in Alzheimer's disease

Alzheimer's disease (AD) is a complex disease that, despite medical advances, is extremely difficult to diagnose. It is very critical to detect this disease in its early stages. The main pathological features of AD are amyloid-beta depositions in the brain and tau protein hyperphosphorylation, which results in neuropathological lesions (39). These neuropathological lesions cause/are the consequence of inflammation and can lead to synaptic loss, and neurodegeneration, which results in cognitive impairment, characteristic to AD (39). Several studies have identified some potential AMP biomarkers that might play a role in the pathological conditions related to AD, in addition to amyloid beta and tau protein. Presenillin 1, presenilin 2, β -site APP-cleaving enzyme 1, alpha-1-antitrypsin (SERPINA1), apolipoprotein E, neurosecretory VGF (VGF), and complement components have been identified as potential biomarkers for AD (40).

Several scientific research have revealed that the APP plays a central role in AD (41). Despite significant data indicating the presence of A β and A β deposits, the so-called amyloid hypothesis of AD is being challenged (41,42). Current evidence suggests that A β has antimicrobial and tumor suppressor activity, as well as helps brain to recover from injury, which aids in the maintenance of the blood-brain barrier and regulates synaptic functions (43). Given these findings, it is not surprising that clinical trials aimed at preventing A β accumulation and the formation of A β deposits did not improve patients' clinical outcomes, and in many cases worsened them (43).

Because no drug was able to prevent A β deposit formation in clinical trials, research data on A β and APP as antimicrobial peptides has led to the formulation of other models and hypotheses, such as the antimicrobial protection hypothesis and pathogenic model of AD. According to the antimicrobial protection hypothesis, A β is a part of the immune system's defense mechanism against infections, but the chronic immune system activation, combined with A β and tau pathology leads to neurodegeneration and dementia (44).

According to the pathogenic model of Alzheimer's disease, the presence of pathogens can cause the appearance of the disease (45). A β peptide has been found to activate microglia, resulting in increased release of pro-inflammatory cytokines such as interleukins, prostaglandins, and leukotrienes as part of an inflammatory response (46). Several studies have shown that A β is active against at least eight different pathogens and induces the action of LL-37 cathelicidin and other AMPs (43,47).

1.6. Role of AMPs in oral squamous cell carcinoma

Oral cancer refers to a category of malignant tumours that affect the oral cavity, pharyngeal areas, and salivary glands. All of these malignancies are classified as oral squamous cell carcinoma (OSCC), accounting for up to 80-90 percent of all cancers of the head and neck region. Smoking, alcohol, UV radiation, human papillomavirus infection, *Candida albicans*, and dietary inadequacies are major risk factors of OSCC (48). OSCC is more common in elderly people and is diagnosed at an average age of 60 years. The OSCC's development takes several years.

The buccal mucosa, the ventral and lateral sides of the tongue, the floor of the mouth, and the retromolar area are common sites for OSCC in the oral cavity. OSCC can also be caused by precancerous abnormalities such as oral leukoplakia (OLK) or oral lichen planus (OLP) (49,50).

OLP is a disease characterized by inflammation affecting the mucous membranes of the mouth. OLP develops after viral infections such as herpes simplex, Epstein-Barr virus, human papillomavirus, and hepatitis C virus (51).

OLK is described as firm, adhesive white patches on the oral mucosa that cannot be scratched. It can be of two kinds: white patches on the tongue or white patches on the tongue base (52,53).

OLK and OLP are considered precancerous conditions, as in some circumstances, they can progress to OSCC (54).

Regarding AMPs, many of them have been identified in the saliva (55), and some of the previously identified biomarkers for oral cancer and precancerous lesions belong to the AMP family. Such examples are protein S100A8, SLPI, lactotransferrin and CLU.

1.7 Network analysis and visualization tools

In the biological aspect, the networks are used to measure the interactions between different types of biological data, whether protein-protein interactions or gene-gene interactions. Network analysis and visualization are used to demonstrate the relationship between different types of data generated by high throughput experiments (56–58).

Currently, many tools are available for network interpretation, analysis, and visualization. Some tools can even build a network from user-defined data retrieved from the high throughput experiments (59). The two very often used tools for network analysis are the String-DB (60) and Cytoscape (61).

1.7.1 String-DB database

The String-DB contains the interaction data between two proteins contributing to a particular biological function (62,63). Proteins are represented by nodes, whereas the edges represent the functional associations. The thickness of the line represents the degree of confidence prediction of the interaction. All association evidence in the String-DB is divided into seven evidence channels: Gene neighbourhoods, Gene fusions, Gene co-occurrence, Co-expression, Experiments, Databases, and Text mining. The user can enable or disable an evidence channel according to their needs. In String-DB, the individual scores are calculated for every evidence channel. 0.90 is the highest confidence score for evidence in String-DB. String-DB can process single protein queries, multiple proteins lists, and proteins with values and ranks to produce a highly interconnected network. Pathway interactions and molecular functions can also be studied using the String-DB. It is possible to process the String-DB network in the Cytoscape network analysis tool for extended network analysis (60).

1.7.2 Cytoscape network analysis tool

Cytoscape is open-source software used for the network analysis of the high-throughput experimental data and the biomolecular interaction networks (64). The central organizational metaphor of Cytoscape is a network diagram in which molecular species are represented as nodes, and intermolecular interactions are represented as connections, i.e., edges, between the

nodes. The core software component of Cytoscape provides essential functions for integrating any data into a graph: visual representation of the chart and integrated data, selection and filtering tools, and an interface for external methods implemented as plugins. Some of such plugins, like ClueGO (65) and CluePedia (66) can help us to perform extensive network analysis.

1.7.2.1 ClueGO

ClueGO is a tool for the biological interpretation of large protein lists (65). ClueGO integrates the GO ontology terms and also KEGG/Biocarta database's data and creates a functionally organized GO/pathway network. It can analyze a list of genes or compare two lists and comprehensively displays functionally grouped terms. With a one-click update option, ClueGO can automatically download the latest GO/KEGG version. Gene IDs are directly imported to the ClueGO platform in a simple text format. ClueGO uses a Cytoscape environment to show the GO terms network (65).

1.7.2.2 CluePedia

The CluePedia Cytoscape plugin is a search tool for new markers potentially related to a pathway (66). CluePedia calculates linear and non-linear statistical dependencies from experimental data. Genes, proteins, and miRNAs can be linked based on *in silico* or experimental information and integrated into a ClueGO network of terms/pathways. CluePedia Cytoscape plugin is user-friendly and easy to use. Users can directly input their experimental data or gene data and compare the data with the available interaction and public information in the String-DB (60), Inact (67), KEGG (68), and GO (69) databases. CluePedia is very helpful in enriching pathways because it can functionally annotate the gene that participates in a particular pathway (70).

2. AIMS

Our main goal was to generate a unified human AMP database and to acquire information on the involvement of AMPs in different pathological conditions.

As far as AMPs have a role in innate immune response, which is involved in the pathophysiology of AD, and that some AMPs have been explicitly related to AD, our goal was to determine the extent to which AMPs may be involved in the pathogenesis of this neurodegenerative disease. We were excited to learn more about the AMPs discovered in the brain, cerebrospinal fluid, and blood in association with AD, and we wanted to investigate how AD affects the network of AMPs.

At the same time, we wanted to enhance data reutilization, as the reanalysis of genomics and proteomics datasets by bioinformatics approaches can be an important tool to examine large amounts of reliable data deposited in publicly accessible datasets. In our study we aimed to use high-quality proteomics data from publicly accessible sources to examine the presence and function of AMPs in AD.

In the case of the OSCC, OLP and OLK our main goal was to do the extensive network analysis of the salivary proteins and AMPs to obtain disease-related information. We also intended to do an extensive pathway analysis to find the biologically relevant information related to salivary proteins and AMPs.

3. METHODS

3.1 Generation of Antimicrobial and Immunomodulatory Peptide database (UDAMP Database)

We have collected all human AMPs listed in Collection of Antimicrobial peptides CAMP (71), Antimicrobial Peptide Database APD (72), Database of Antimicrobial Peptides dbAMP (73), Linking Antimicrobial Peptides LAMP (74), and Database of Antimicrobial Activity and Structures of Peptides DBAASP (4). We made a PubMed search between May 2020 and June 2020 using the keywords "human antimicrobial peptide" and listed all hits not present in online AMP databases. We have curated the retrieved data, removed the redundancies, and compiled a unified comprehensive human antimicrobial and immunomodulatory peptide database. We named our database as University of Debrecen Antimicrobial and Immunomodulatory Peptide (UDAMP) Database. UniProt IDs of the AMPs were used to generate the Gene IDs from the UniProt (75) using its Applications Programming Interface. NCBI was used to find cross annotations and links between gene IDs and PubMed IDs. When we got our gene IDs, the "gene2pubmed" annotation file was used to retrieve the PubMed IDs of the genes through their gene IDs. E-utilities (47), a programming interface for the NCBI Entrez databases, was used to retrieve the article metadata based on PubMed IDs. We run a home-written script to retrieve an offline keyword search on the article metadata.

As we got the high data redundancy in the AMP sequence, we crosschecked the individual protein/peptide sequences. They were matched against the UniProtKB/SwissProt proteome database using BLASTP with the following settings: the target organism was "Homo sapiens", the E threshold was set to 1e-10; scoring matrix: BLOSUM62 (default), filtering: none, gap cost: existence 11, extension: 1 (default), limit of hits: 250. For each AMP of UDAMP: protein name, gene name, UniProt ID, GI number, peptide sequence, length of peptide, antiviral, antibacterial or antifungal activity, Gram type in case of AMPs having antibacterial activity, method of validation, database ID from where the AMP was downloaded, and the Protein Data Bank (PDB) ID indicating the structure of the AMP was retrieved. The online source databases of AMPs provided various identifiers for the listed AMPs. The UniProt and NCBI GenBank databases were queried to retrieve or verify the protein name, gene name, UniProt ID, and GI number. The Protein Databank ID was retrieved from the RCSB Protein Database; all other data were obtained from the online database of AMPs or a scientific article deposited in PubMed.

3.2 Examination of AMPs in proteomic datasets related to Alzheimer's disease

We have retrieved the Alzheimer's disease datasets, including human donors, from the ProteomeXchange repository (<http://www.proteomexchange.org/>). Datasets deposited before 2019 were selected, where the sample from AD and matched non-AD controls were investigated. We omitted the datasets involving only patients with AD and searched for comparisons between AD and healthy/control data. We have retrieved datasets originating from the analysis of brain, CSF, or blood samples (Table1).

Dataset Identifier	Source Database	Sample Type	References
PXD009199	ProteomXchange	Brain	(76)
PXD008807	ProteomXchange	Brain	(77)
PXD012203	ProteomXchange	Brain	(78)
PXD004510	ProteomXchange	Brain	(79)
PXD005321	ProteomXchange	Brain	(80)
PXD000067	ProteomXchange	Brain	(81)
PXD014376	ProteomXchange	Brain	(82)
PXD006122	ProteomXchange	Brain	(83)
PXD014557	ProteomXchange	Brain	(84)
PMID:30137212	PubMed	Brain	(85)
PXD007694	ProteomXchange	Brain	(86)
PXD012851	ProteomXchange	CSF	(87)
PXD008098	ProteomXchange	CSF	(35)
PXD011482	ProteomXchange	Blood	(67)
PMID:27911324	PubMed	Blood	(88)

Table 1. The list of datasets used in our study

We checked which of the AMPs listed in the UDAMP database were present in the ProteomeXchange datasets and evaluated the changes in their levels. Our results were based on the statistical analysis performed by the authors. AMPs whose level showed a statistically significant change in the AD samples compared to control were defined with labels "increase"

and "decrease", respectively. In contrast, AMPs without any statistically significant change were labelled as "no significant change". "0" was assigned to those AMPs absent from the ProteomeXchange datasets.

3.3 Network analysis of datasets related to Alzheimer's disease

AMPs that were present in the UDAMP database and showed a statistically significant change in the downloaded datasets from the ProteomeXchange datasets were analyzed through network analysis. We have used Cytoscape (61) with String-DB v11 (60) and ClueGO v2.5.7 (65), plus CluePedia v1.5.7 (70) plugins. Query proteins were imported in String-DB, where networks were generated at a confidence score of 0.90. Networks containing AMPs with statistically significant change in AD compared to controls and their 50 first shell interactors were retrieved. The networks were imported to the Cytoscape and analyzed using the "Analyze Network" tool available under the "Tools" menu. Pathway analysis was performed using the ClueGO tool using default settings, and the number of gene visualization thresholds was set to the highest number, i.e. 1000 in CluePedia. The betweenness centrality and the degree of distribution were obtained for the networks. The Cytoscape network was imported to ClueGO and CluePedia. The gene function analysis was performed using Cytoscape network in CluePedia by using default settings. The used interactions for the network analysis was, activation, binding, co-expression, and inhibition selected from the Sting-DB. Pathway analysis was also performed using DAVID (89). DAVID is a web-based gene ontology tool that includes functional genomic annotations as well as a good graphical summary. It takes a gene and protein list as input and annotates it using gene ontology data. Its current design elements provide automated solutions that allow researchers to quickly discover the biological significance of gene and protein list data (89). AMPs were further submitted for gene ontology analysis in the DAVID and top 10 enriched pathway were selected by investigating the gene counts of the pathways.

3.4. Network analysis of data related to oral squamous cell carcinoma

Saliva sample were collected from OSCC, OLP, OLK, and control groups (3 donor/group, mean age 62 years). The sample collection was approved by the Ethical Committee of the University of Debrecen (approval number: 45038/2012/EKU) and was done in accordance with the Declaration of Helsinki. The samples were analysed by proximity extension assay

(PEA) at the Olink proteomics facility (Uppsala, Sweden) by applying the Oncology II and Inflammation panels (www.olink.com). It was checked if the detected salivary proteins were AMPs, namely, if they were present in the UDAMP database. String-DB (60) and CluePedia v1.5.7 (70) were used to create the networks of differentially expressed proteins and AMPs. The network generation and examination of the generated networks was done as it was described in the case of AD (section 3.3), with one exception: in case of the String network generation the medium stringency (0.7) was applied.

4. RESULTS

4.1 Generation of the comprehensive human antimicrobial and immunomodulatory peptide database (UDAMP Database)

We created a large unified human AMP database and named it University of Debrecen Antimicrobial and Immunomodulatory Peptide database (UDAMP). UDAMP contains data from online databases as well as the scientific literature.

In the year 2019, when the database was originally made, the UDAMP had 186 AMPs. Each AMP had a unique protein and gene identifier (gene and protein name, UniProt ID, GI number), as well as the database name from which it was obtained. Every AMP also had a peptide sequence, Protein Databank ID (PDB ID), where the 3D structure was available, and the type of antimicrobial activity. In the case of the antibacterial peptides, it was indicated if they are effective against Gram positive or Gram negative bacteria. The validation cell was used to categorize antimicrobial activity as experimentally validated or predicted (Figure 3).

UDAMP ID	UDAMP89
Gene Name	DEFA5
UniProt ID	Q01523
Protein Name	Defensin-5
Sequence	ATCYCRTGRCATRESLSGVCEISGRLYRLCCR
Length	32
PubMed ID	18191790
Gi Number	109156995
Activity	Antibacterial, Antifungal
Gram Activity	Gram Positive, Gram Negative
Validation	Experimentally Validated
Database ID	CAMPSQ730
PDBID (If available)	1ZMP

Figure 3: Architecture of the UDAMP database.

Meanwhile the database was updated, currently containing 670 entries. The database was uploaded to the Figshare and can be accessed freely through the weblink https://figshare.com/articles/dataset/UDAMP_Database/20926873.

4.2 Functional examination of the human AMPs

In order to get more information on the human AMPs, a functional analysis and pathway analysis, along with gene-gene interaction network analysis was carried out.

4.2.1 Gene ontology (GO) of the human AMPs

ClueGO and DAVID was used to analyze the significantly enriched pathways among the human AMPs (Figure 4).

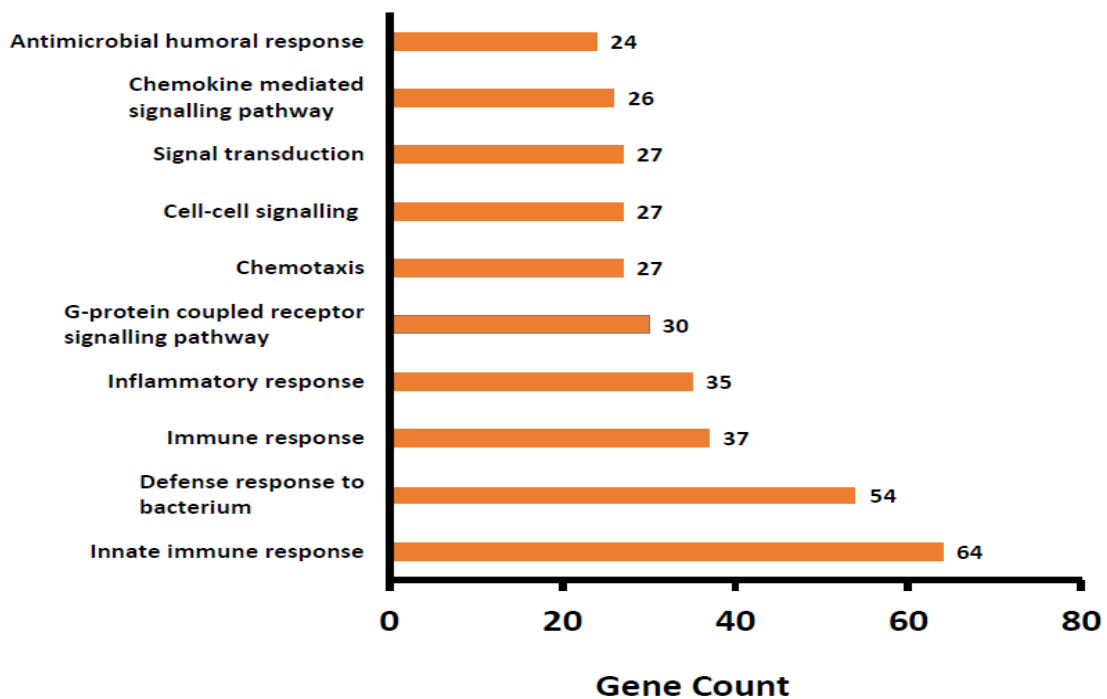


Figure 4. Top 10 enriched gene ontology (GO) terms for AMPs present in the human antimicrobial and immunomodulatory peptide (UDAMP) Database. The enriched GO terms are shown on the "y" axis, while the gene counts are shown on the "x" axis.

As it was expected, the extensive GO analysis reveals pathways such as antimicrobial response, immune system-related functions, and signalling.

4.2.2 Protein-Protein interaction network analysis of the human AMPs

To better understand the subtle relations among AMPs, a protein-protein interaction analysis using the Sting-DB and Cytoscape was carried out (Figure 5).

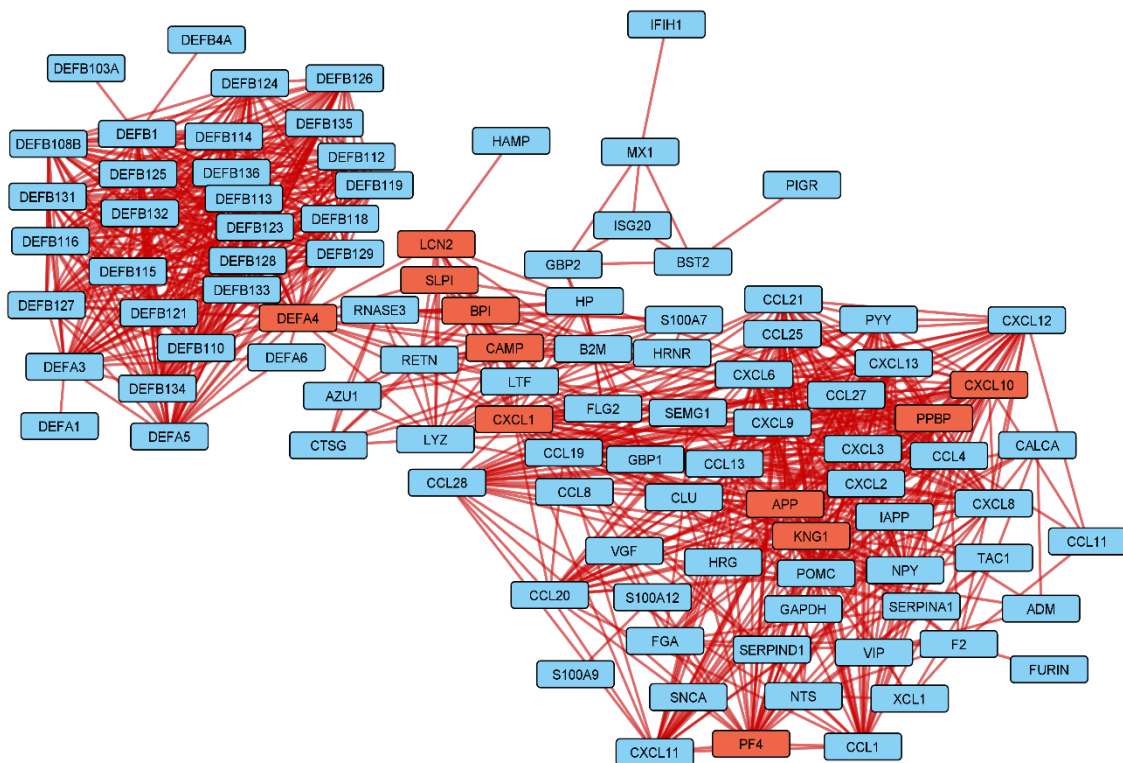


Figure 5: Protein–protein interaction (PPI) network. Each rectangle represents a protein, and the line represents the interaction. Proteins with central role are highlighted with red.

A dense cluster of proteins was observed containing defensins linked by defensin alpha 4 (DEFA4) to the cluster of proteins containing AMPs like antileukoproteinase (SLPI), bacterial permeability increasing protein, lipocalin-2 and cathelicidin. The other cluster contained mainly cytokines, cystatins, APP and other AMPs.

Based on the betweenness centrality and degree of distribution, proteins such as amyloid beta precursor proteins (APP), DEFA4, growth regulated alpha protein (CXCL1), C-X-C motif chemokine 10 (CXCL10), kininogen 1, platelet precursor basic protein (PPBP), and platelet factor 4 were playing a central role in the network.

4.2.3 Gene-gene interaction analysis of the human AMPs

To have a functional overview, the gene-gene interaction analysis was performed using the Cytoscape plugin CluePedia. Gene-gene interaction analysis revealed many subnetworks and proteins which were playing significant role in the human AMPs network (Figure 6).

and SLPI, which inhibited the polymeric immunoglobulin receptor. Interferon-induced with helicase C domain 1 and interferon alpha-2 were mutually activating each other, while interferon alpha-2 was activating the CXCL10, and Interferon-induced with helicase C domain 1 was inhibiting it. Apart from the main network, there were some independent pairs found such as alpha-antitrypsin (SERPINA1) was activating hepcidin, DEFB4A the apolipoprotein B mRNA editing enzyme, catalytic subunit 3G, and guanylate-binding protein 1 the interferon-induced GTP-binding protein Mx1. APP and glyceraldehyde -3- phosphate dehydrogenase were reciprocally inhibiting each other.

4.2.4 Top GO terms enriched for AMPs in the brain, CSF, and blood

The DAVID and ClueGO was used to analyse GO functions. The top 10 GO functions were obtained in order to examine the regulated pathways in the network of AMPs characteristic to the brain, CSF and blood (Figure 7).

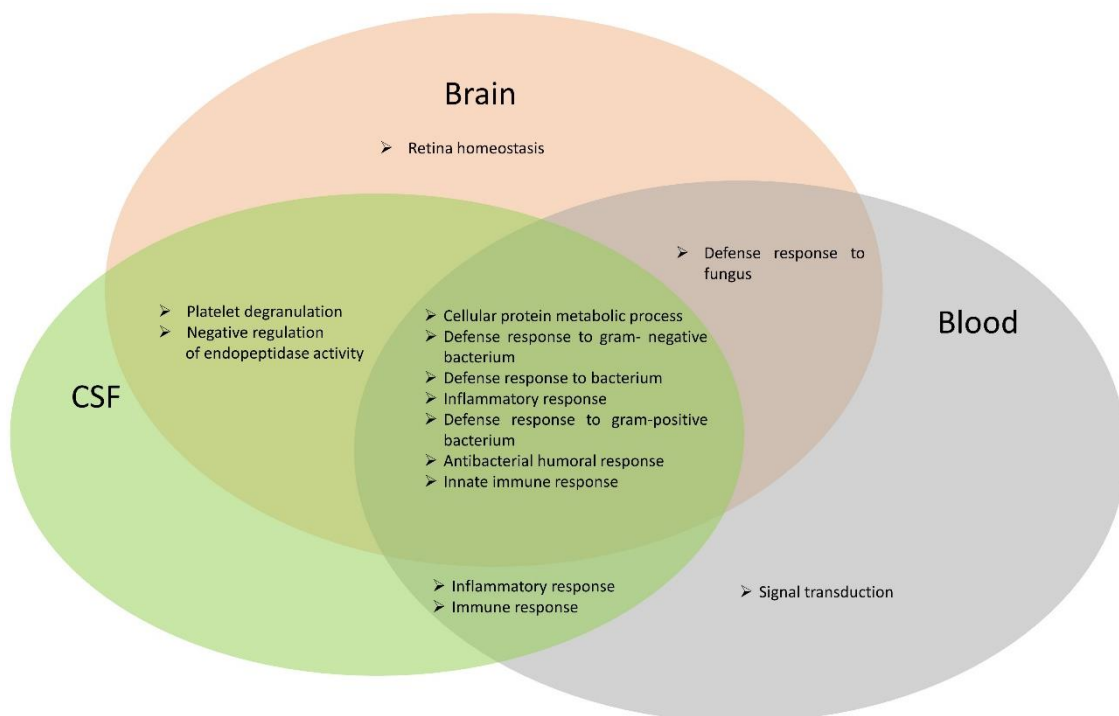


Figure 7: Venn diagram of the top 10 enriched GO terms for AMPs identified in the brain, CSF, or blood by DAVID.

Bacterial defense, inflammation, and innate immune response were the primarily enriched GO functions in the AMPs characteristic to the brain, CSF, and blood. Some of the

GO functions were sample specific. The cellular metabolic process was enriched in all three sample types indicating a basic function of metabolism control, but retinal homeostasis was only found in the brain. Platelet degranulation and negative regulation of endopeptidase activity were found in the brain and CSF, while the defense response to fungi was found in the brain and the blood.

4.2.5 Gene-gene interaction network for the AMPs found in the brain, CSF, and blood

Gene-gene interaction analysis was performed by using the CluePedia plugin in the Cytoscape. This analysis revealed the presence of the different regulatory circuits among the AMPs present in all sample types (Figure 8).

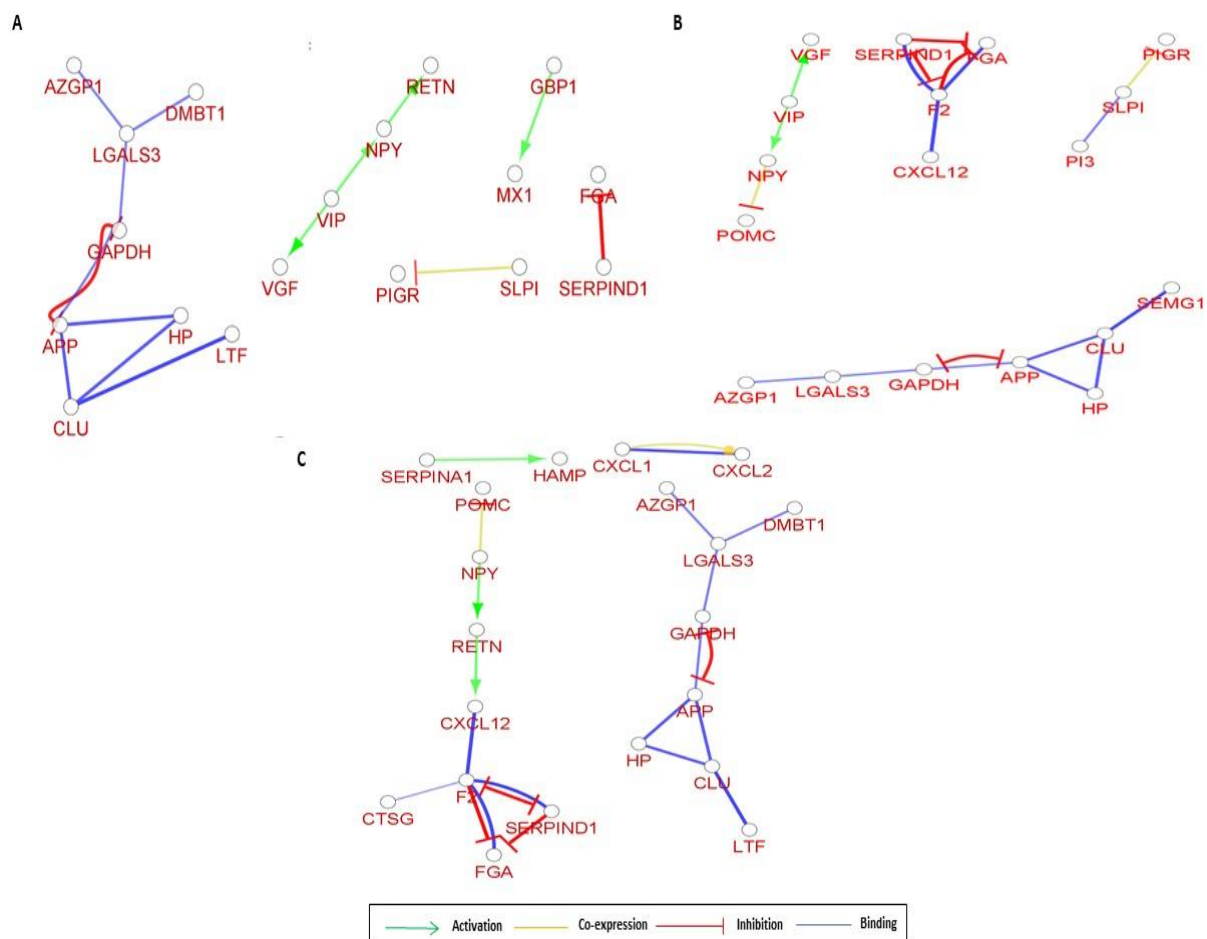


Figure 8: Gene-gene interactions networks characteristic for AMPs identified in brain (A), in CSF (B), and in blood (C). The circles represent a gene/protein and the lines indicate interactions. The lines with arrow represent activation, blocking lines represent inhibition, simple lines represent protein–protein interaction. Line color indicates the type of

interaction: green color refers to activation, red color to inhibition, blue color to binding, whereas yellow color shows co-expression. On all panels the proteins are labelled with their gene name.

The network of AMPs observed in the brain, CSF and blood shared some similarities. For example, the small subnetwork containing the reciprocal inhibition between APP and glyceraldehyde -3- phosphate dehydrogenase was found in the all sample types. The inhibitory loop involving F2, FGA and SERPIND1 was present in CSF and blood and was only partially present in the brain, while the VIP—VGF—NPY regulatory subnetwork was present in brain and CSF and only partially in blood.

4.3 AMPs characteristic of Alzheimer's disease in the brain, CSF, and blood

To retrieve the AMPs in Alzheimer's disease, we examined AD datasets obtained from the ProteomXchange repository. For the data analysis, 11 datasets from brain tissue, 2 from CSF, and 2 from blood were used. Identified AMPs were used for the GO analysis and network analysis. We examined the AMPs that showed a statistically significant difference between the AD and control groups (Figure 9).

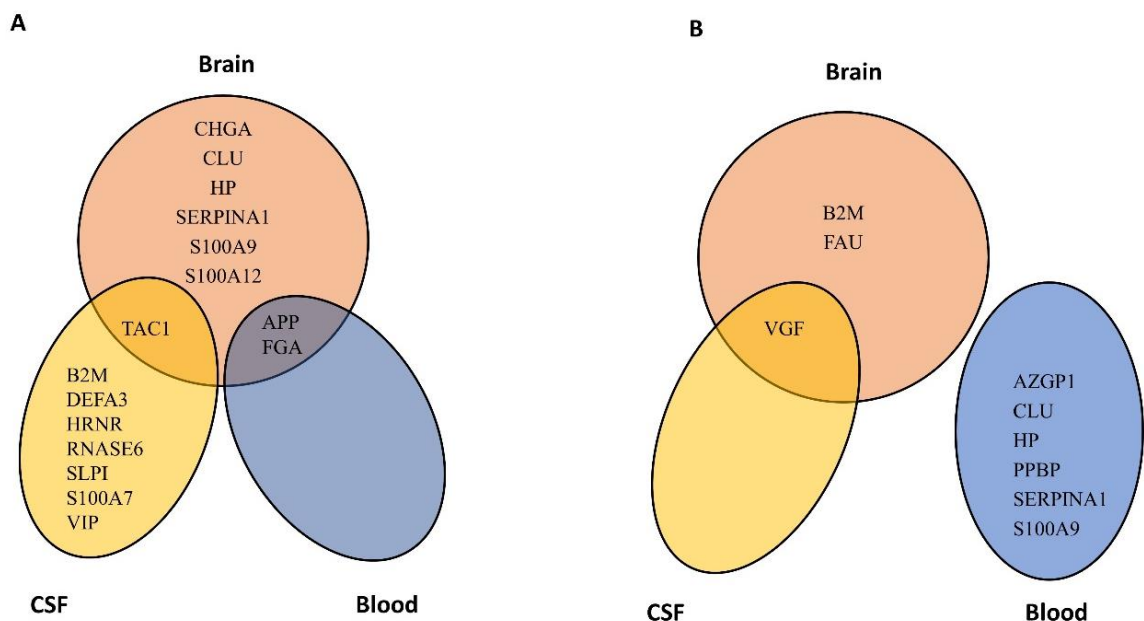


Figure 9. (A) AMPs whose amount increased statistically significantly in AD compared to control. (B) AMPs whose amount dropped statistically significantly in AD compared to control.

control. The Venn diagrams show the proteins corresponding to each group. Orange color shows AMPs found in the brain, yellow refers to AMPs identified in CSF, while blue shows AMPs found in the blood. The proteins are shown according to their gene names.

There were nine AMPs increased in the brain characteristic to the AD. The increased AMPs were APP, SERPINA1, chromogranin-A, CLU, FGA, protein S100A9 (S100A9), protein S100A12 (S100A12), haptoglobin (HP), and TAC1 (Figure 9A).

The amount of three AMPs, FAU ubiquitin-like and ribosomal protein S30 (FAU), VGF, and beta-2-microglobulin (B2M), decreased in AD. In the CSF, SLPI, B2M, neutrophil defensin 3, hornerin, TAC1, protein S100A7 (S100A7), ribonuclease 6, and VIP were found in increased, whereas VGF in decreased amount in AD.

The amount of APP and FGA in the blood samples was found to be increased in AD. On the other hand, the amount of SERPINA1, CLU, HP, PPBP, S100A9, and zinc-alpha-2-glycoprotein was decreased (Figure 9B).

4.3.1. Functional analysis of the network of AMPs characteristic of Alzheimer's disease in brain, CSF, and blood

ClueGo and DAVID were used to perform functional analysis of the AMPs characteristic of AD in different sample types. The GO terms enriched in different sample types were identified using functional analysis. Using Venn diagrams, the top enriched GO terms for AMPs increased or decreased in the brain were listed (Figure 10).

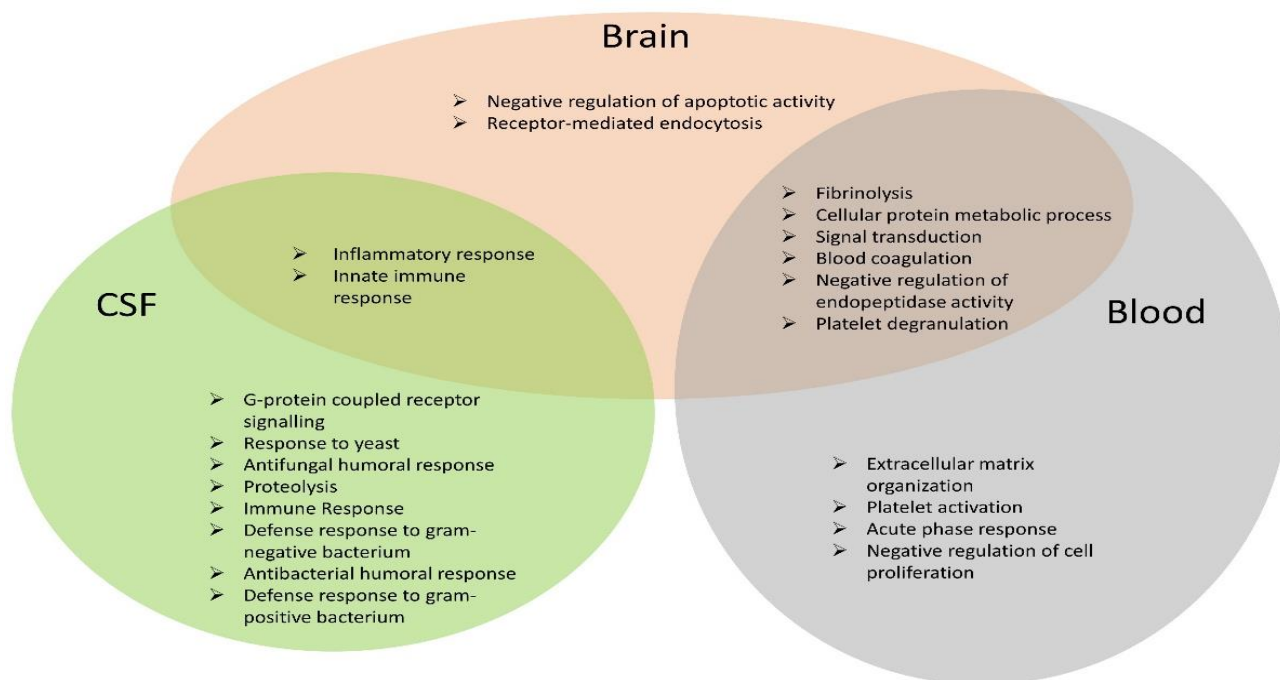


Figure 10: Venn diagram of the top 10 enriched GO terms according to DAVID for AMPs increased in the brain, CSF, or blood.

Several of the GO functions enriched in the brain for AMPs increased in AD were similar to those enriched in the overall AMP network (Figure 4). Beside the brain, the inflammatory response and innate immune response were enriched also in CSF, and signal transduction in the blood. Blood coagulation (fibrinolysis, platelet degranulation, and blood coagulation), cellular metabolic processes, endocytosis, and negative regulation of endopeptidase activity, and apoptotic activity were newly enriched functions (Figure 10).

Majority of the enriched GO terms in the CSF were related to antimicrobial defense, immune response, inflammation, and G-protein receptor signaling pathway, and also present in the main AMP network. Proteolysis and the response to yeast was the newly enriched functions. Blood coagulation (fibrinolysis, platelet activation, and degranulation), acute phase response, extracellular matrix organization, signal transduction, cellular protein metabolic process, and negative regulation of endopeptidase activity and of cell proliferation were the enriched functions in the blood. Inflammatory and innate immune response was shared by the brain and CSF. In contrast, blood clotting, negative regulation of endopeptidase activity, signal transduction, and cellular protein metabolism were shared by the brain and the blood (Figure 10).

Regarding the AMPs found to be decreased in AD, the RNA metabolism, translation, viral transcription, cell adhesion were newly discovered enriched functions (Figure 11).

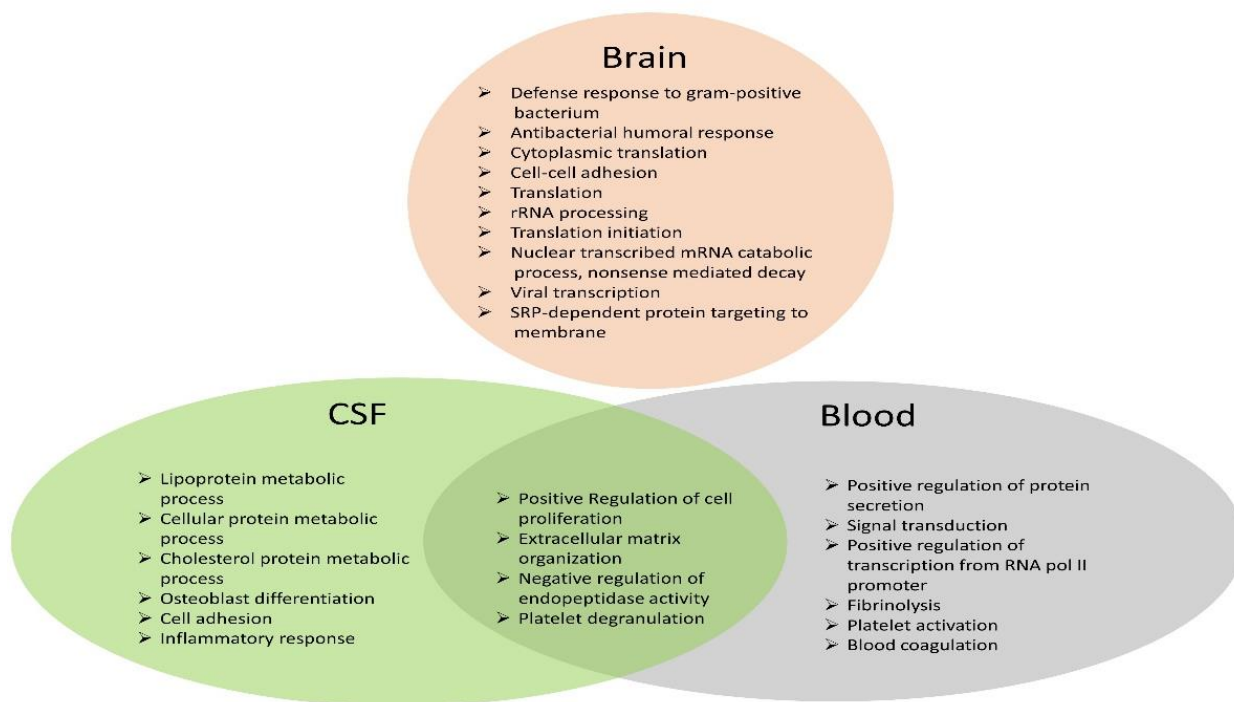


Figure 11: Venn diagram of the top 10 enriched GO terms according to DAVID in case of AMPs decreased in the brain, CSF, or blood.

Out of top 10 enriched GO functions, the six characteristic to CSF were related to lipid (lipoprotein and cholesterol) metabolism, cell adhesion, osteoblast differentiation, endopeptidase activity, platelet degranulation, inflammation, and cell adhesion. Blood coagulation, signal transduction, protein secretion, endopeptidase activity regulation, cell proliferation, transcription, and extracellular matrix organization were more prevalent in the blood (Figure 11). Cell proliferation regulation and endopeptidase activity, the extracellular matrix organization, and the platelet degranulation functions were general in both blood and CSF.

ClueGO's GO analysis revealed additional information to DAVID about functional enrichment (Figure 12). Aside from the enriched GO functions in DAVID in the case of AMPs increased in AD, the amyloid precursor protein metabolic process, neuropeptide and tachykinin signaling pathway, and antioxidant activity have been found only in brain samples.

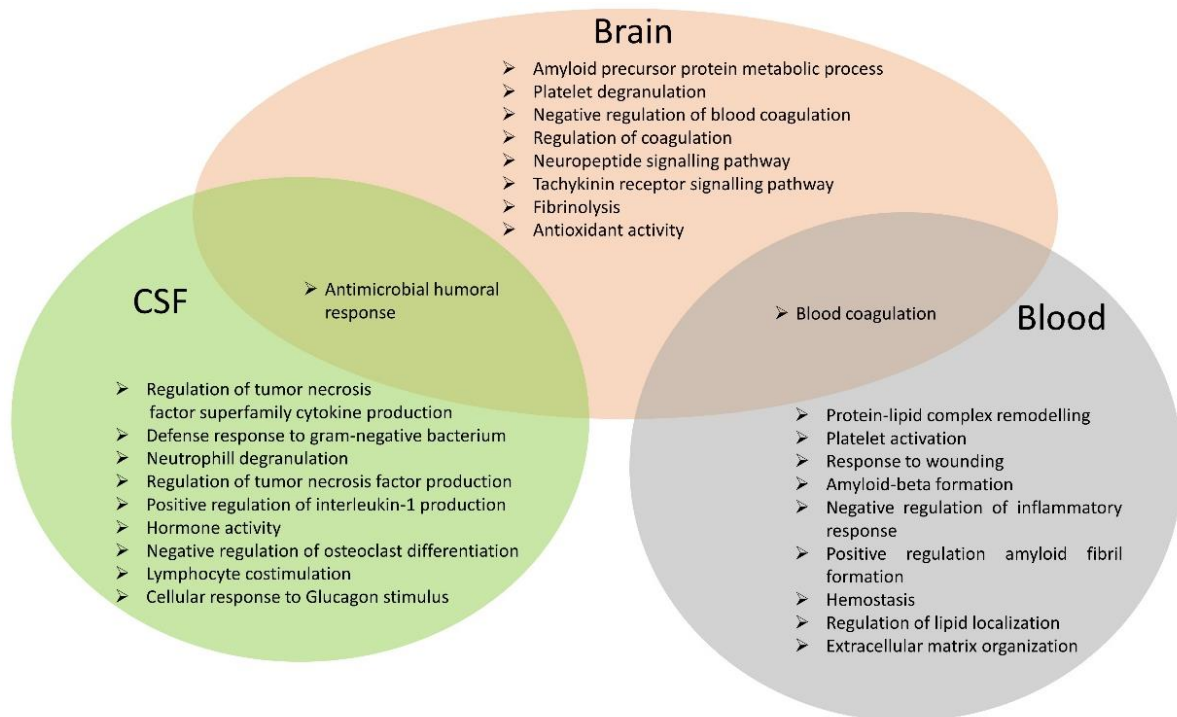


Figure 12: Venn diagram of the top 10 enriched GO terms according to ClueGO for AMPs increased in the brain, CSF, or blood.

Regulation of lipid localization, protein–lipid complex remodelling, amyloid beta formation, and response to wounding were newly enriched GO functions in the blood in the case of AMPs increased in AD. Along with this, the cellular response to glucagon stimulus and regulation of osteoclast differentiation were considered new enriched GO functions in the CSF (Figure 12).

When running ClueGO, the additional functions to those identified by DAVID in the case of AMPs decreased in AD, were antigen processing and DNA damage response in the brain, regulation of amyloid fibril formation in the blood, and response to axon injury, regulation of insulin-like growth factor receptor signaling, cellular response to cAMP, post-translational protein modification, and membrane protein proteolysis in the CSF (Figure 13).

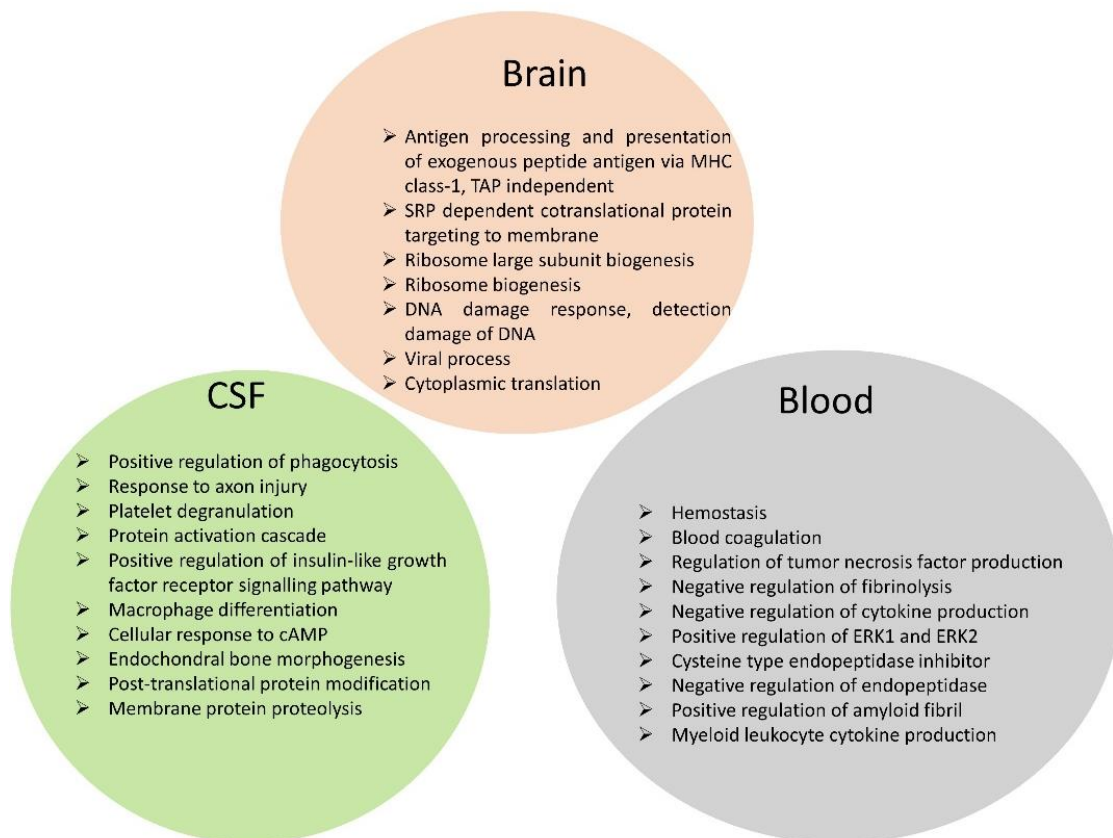


Figure 13: Venn diagram of the top 10 enriched GO terms according to ClueGO for AMPs decreased in the brain, CSF, or blood.

4.3.2 Network analysis of the AMPs characteristic of the AD in brain, CSF and blood

Given that there were no considerable differences between the gene interaction networks of AMPs in the UDAMP Database and those reported in the brain, CSF, and blood, we were keen to study the changes in gene-gene interaction networks in the case of AMPs related with AD in the next step. We wanted to find out which proteins could be the most important in the AMP network by identifying the key hub proteins. As a result, we investigated the protein–protein interaction networks and computed the betweenness centrality and degree of distribution. In line with earlier research, we discovered APP to be the primary hub protein in the AMP network observed in the brain.

Next we generated the gene-gene interaction network of all proteins altered in AD, independent of the direction of change (increase or decrease) or the location of identification (brain, CSF, or blood) (Figure 14).

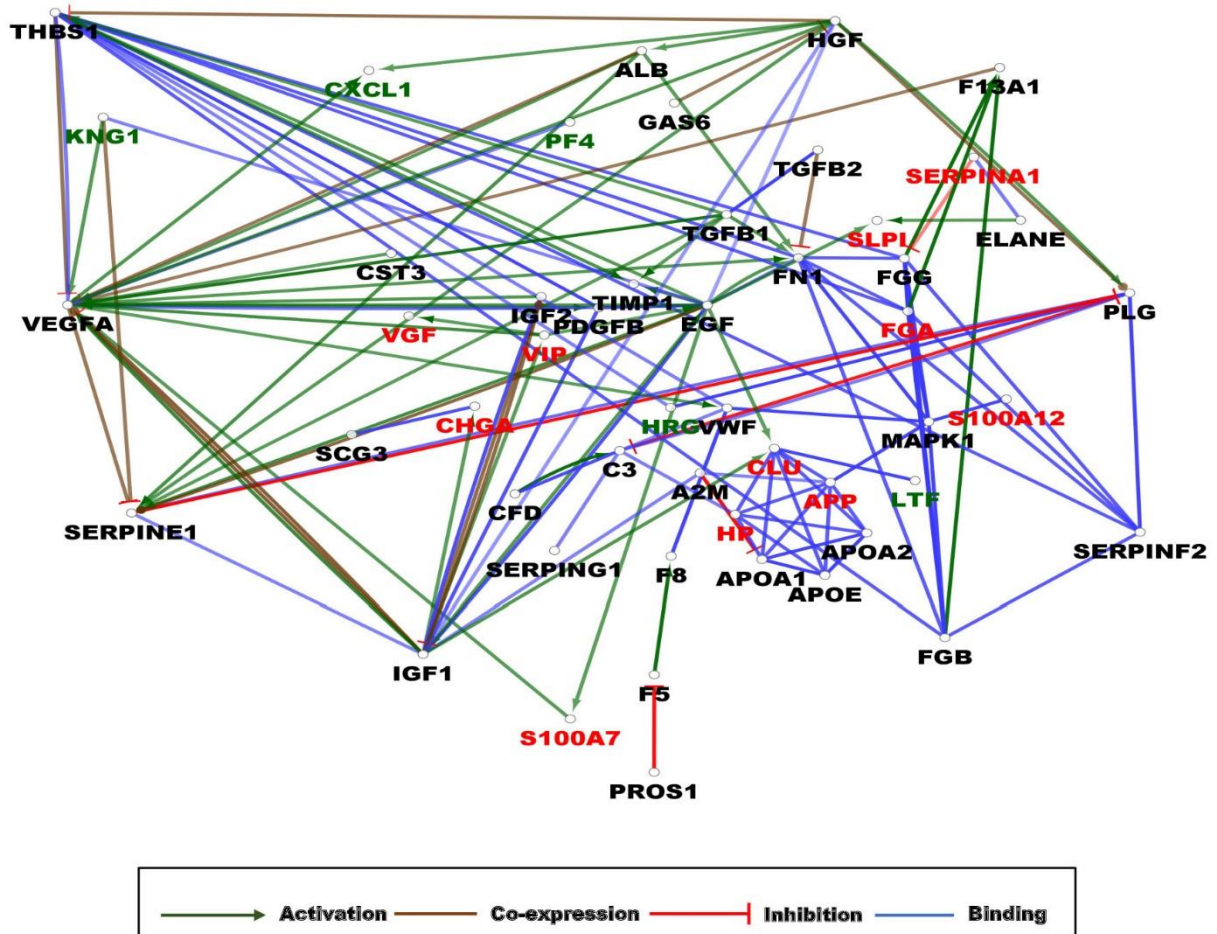


Figure 14: Gene-gene interaction network of AMPs characteristic of AD. The circles represent a gene/protein, and the lines indicate interactions. The lines with an arrow represent activation, blocking lines represent inhibition, and simple lines represent protein-protein interaction. Line color indicates the type of interaction: green color refers to activation, red color to inhibition, blue color to binding, whereas brown color shows co-expression. The red color of the protein label indicates AMPs whose amount changed in a statistically significant manner between AD and control samples, the black color shows their interacting partners with no AMP activity, while the green color indicates interactors with AMP activity. The proteins are labeled with their gene name.

We could observe that the main sites of activations were vascular endothelial growth factor A, and plasminogen activator inhibitor 1 (SERPINE1). A cluster with multiple interactions among its members containing the apolipoproteins APOA1, APOA2, and apolipoprotein E, as well as CLU, HP, mitogen-activated protein kinase 1, and APP was present, whereas proteins implicated in blood clotting such as FGA, fibrinogen chain beta and gamma, fibronectin, alpha-2-antiplasmin, and thrombospondin-1 were found in another cluster.

Vascular endothelial growth factor A was the major hub protein, followed by epidermal growth factor (EGF), insulin-like growth factor 1, SERPINE1, and fibronectin, as proteins activated or inhibited by many other proteins (Figure 14).

4.4 Examination of antimicrobial peptides in oral squamous cell carcinoma

Saliva is a bodily fluid with great diagnostic potential. Besides being used for diagnosis, it can also be used for screening, further increasing the diagnostic value of this body fluid. Considering the presence of AMPs in saliva, our aim was to evaluate their role in OSCC and precancerous conditions such as OLP and OLK. A study was carried out by our group to observe the altered proteins characteristic of the OSCC and the other precancerous conditions: OLK and OLP. Saliva sample taken from 4 groups of donors (OSCC, OLP, OLK, control) were analysed by proximity extension assay (PEA). Three donors were recruited to each group and the donors were age matched. The PEA is a combination of immunoassay and qPCR for the quantification of proteins in liquid samples with high sensitivity. A matched pair of antibodies linked to unique oligonucleotides (proximity probes) recognize the protein. As a result, the probes become close to each other, hybridize and using quantitative real-time PCR quantification is possible. The sensitivity and specificity is provided by the antibodies, while the large dynamic range by the qPCR.

4.4.1 Examination of selected salivary proteins

176 out of examined 184 proteins were detected in saliva and it was checked if they were present in the UDAMP database. 15 proteins were present in the UDAMP database being AMPs. The AMPs and their 50 first shell of interactors were used to generate gene-gene interaction networks (Figure 15A).

In order to better examine the activation and the binding events among the proteins in the network, from the composite network containing all interactions only the activation (Figure 15B) and the binding (Figure 15C) events, respectively, were extracted. The network analysis indicated activation hubs, showing proteins which were activated by many other proteins in the network. A highly interconnected cluster with multiple activations containing chemokines and chemokine receptors could be observed. Interleukin (IL) 6 has functioned as the activation hub for numerous proteins, as evidenced by the fact that it was activated by several proteins and vice versa. IL6 activated interleukin 4, interleukin1beta, and interleukin 17A, while tumor necrosis factor (TNF) activated IL6.

and simple lines represent protein-protein interaction. Line color indicates the type of interaction: green color refers to activation, red color to inhibition, blue color binding, whereas yellow color shows co-expression. The proteins were listed according to their gene names.

The activations and interactions among the FAS-associated death domain protein and members of the TNF receptor family (TNFRSF10A, TNFRSF10C, TNFRSF10D, and TNFRSF10B) showed another highly interconnected cluster. The degree of distribution analysis regarding the protein-protein interactions revealed C-X-C chemokine receptor type 1, C-X-C chemokine receptor type 2 and C-C chemokine receptor type (CCR1, CCR5, CCR7) as hub proteins in the binding network (Figure 15B), while taking into account the activation, the main hub proteins were CCL2, IL6, CXCL8, CCR5 and CCR7 (Figure 15C).

4.4.2 Differentially expressed proteins in different oral disease groups

The amount of proteins that were present in the more than 80% of the samples were examined. Out of 146 proteins 37 showed the statistically significant change at least in 2 groups (Figure 16).

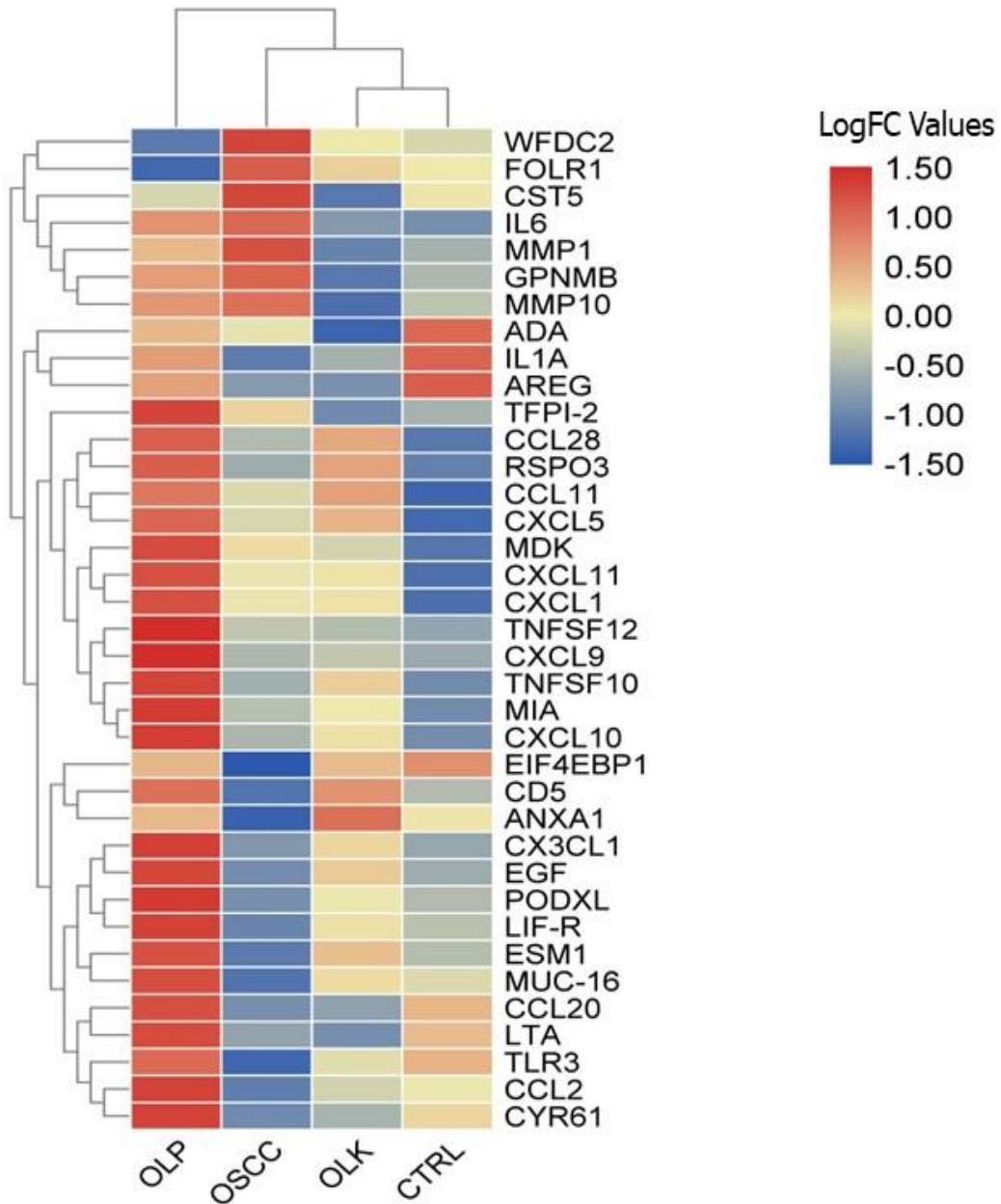


Figure 16: Heat map and hierarchical clustering of differentially expressed proteins in the studied groups. The proteins are labelled according to their gene symbols. The values on the scale demonstrate log fold change (FC) of proteins between different groups.

Most of the differentially expressed proteins between OLP and control group were chemokines, such as CCL28, CCL11, CXCL5, CXCL11, fractalkine, CXCL10, MCP1/CCL2, and TNF ligand superfamily member 12, melanoma- derived growth regulatory protein, midkine, r-spondin-3, tissue factor pathway inhibitor 2, and TNF ligand superfamily member 10.

Adenosine deaminase, amphiregulin, eotaxin (CCL11), CXCL11, transmembrane glycoprotein NMB (GPNMB), IL1, IL6, matrix metalloproteases MMP1 and MMP10, and toll like receptor 3 (TLR3) were found to be differentially expressed between the OSCC and control groups.

Only two proteins were discovered to be changed in a statistically significant manner between the OLK and control groups, the chemokines CXCL11 and CCL11.

Many chemokines such as CCL20, CX3CL1, CXCL1, CXCL9, CXCL10, CXCL11, and MCP1/CCL2, mucin-16, t-cell surface glycoprotein-5, CCN family member 1 (CYR61), epidermal growth factor (EGF), endothelial cell-specific molecule 1, leukemia inhibitory factor receptor, podocalyxin, r-spondin-3, band 3 anion transport protein 1, TLR3, TNF ligand superfamily member 12, and WAP four-disulfide core domain protein 2 were the proteins that differed between OSCC and OLP.

Annexin-1 (A), cystatin-5, eukaryotic translation initiation factor 4E-binding protein 1, GPNMB, IL6, MMP1, MMP10, and WAP four-disulfide core domain protein 2 were the proteins that were altered differently between OLK and OSCC.

On the other hand, only four proteins were found to be differently expressed between OLK and OLP: CYR61, MCP1/CCL2, melanoma- derived growth regulatory protein, and midkine.

4.4.3 Network analysis of proteins and AMPs with statistically significant changes among the disease groups

Proteins and AMPs with statistically significant differences between two groups were examined for protein-protein interactions in order to gain a better understanding of the OSCC and precancerous conditions. The protein-protein interaction and GO analysis of OSCC vs control group were performed using string-DB (Figure 17).

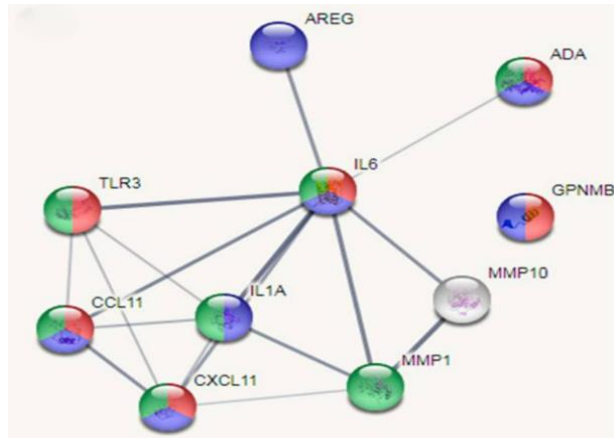


Figure 17: Protein–protein interaction (PPI) network of proteins generated by String-DB, with significantly different expression between the OSCC vs control. Within the network each circle represents a protein, and the lines between the circles represent protein–protein interactions. The line width is correlated with the strength of the supporting evidence for the interaction; the wider the line, the more experimental evidence supports the interaction. The circle color indicates enriched GO functions. The green color of the circle represents immune system functions, the red color represents positive regulation of intracellular signal transduction, and the blue color represents regulation of cell proliferation.

The proteins and AMPs that were differentially expressed between the OSCC and controls were associated with functions such as inflammatory response, regulation of growth factor production, collagen catabolism, positive regulation of intracellular signaling, and cellular proliferation.

The proteins from the OLK and control groups were then examined for protein–protein interaction (PPI), and GO function analysis was performed (Figure 18).



Figure 18: Protein–protein interaction network of proteins generated by String-DB, with significantly different expression between the OLK vs control. Within the network each circle represents a protein, and the line between the circles represents protein–protein interaction. The circle color indicates enriched GO functions. The red color represents positive regulation of intracellular signal transduction, the blue color the regulation of cell proliferation, and yellow color represents the regulation of signalling receptor activity.

In the case of OLK vs control the CXCL11 and protein CCL11 was differentially expressed. These proteins were involved in biological processes like chemokine-signaling, antimicrobial humoral immune response, and lymphocyte chemotaxis.

The PPI and GO function analysis for the OLP vs control group was also performed (Figure 19).

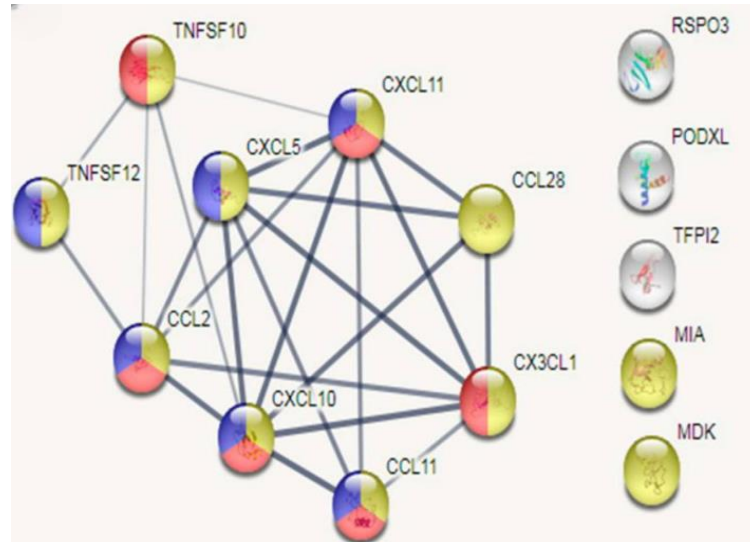


Figure 19: Protein–protein interaction network of proteins generated by String-DB, with significantly different expression between the OLP vs control. Within the network each circle represents a protein, and the lines between the circles represent protein–protein interactions. The line width is correlated with the strength of the supporting evidence for the interaction; the wider the line, the more experimental evidence supports the interaction. The circle color indicates enriched GO functions. The red color represents positive regulation of intracellular signal transduction, the blue color regulation of cell proliferation, and yellow represents the regulation of signalling receptor activity.

In the case of OLP vs control groups, the enriched GO terms were related to the signalling pathways and immune response. Ten of the 14 proteins were shown to be involved in cell migration, which corresponds to histological observations of T-cell and lymphocyte infiltration in the epithelium and lamina propria. Proteins involved in positive stimulus response, stress response, and signalling may indicate the link between viral infection and OLP induction, as well as the link between autoimmune disorders and OLP development.

The PPI network and GO analyses of OSCC vs OLP groups was generated (Figure 20).

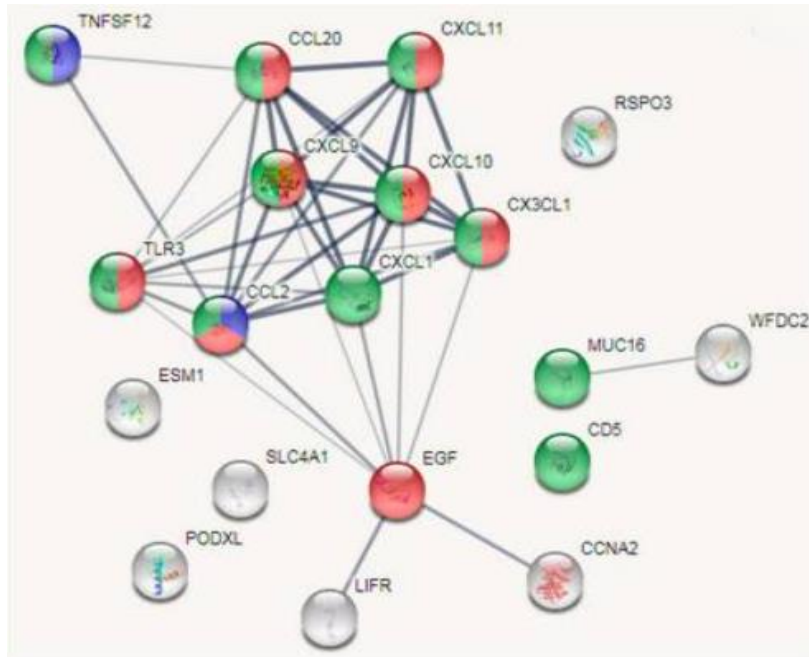


Figure 20: PPI network of proteins generated by String-DB, with significantly different expression between the OSCC vs OLP. Within the network each circle represents a protein, and the lines between the circles represent protein–protein interactions. The line width is correlated with the strength of the supporting evidence for the interaction; the wider the line, the more experimental evidence supports the interaction. The circle color indicates enriched GO functions. The green color represents immune system processes, the red represents positive regulation of intracellular signal transduction, and the blue color represents regulation of cell proliferation.

The bulk of the proteins in the network are members of the cytokine-cytokine receptor interaction KEGG pathway. The biological processes of the leukocyte chemotaxis and inflammatory response were the enriched functions between the OSCC and OLP.

For the examination of the differences between the OSCC and OLK groups, the PPI and GO analysis was performed (Figure 21).

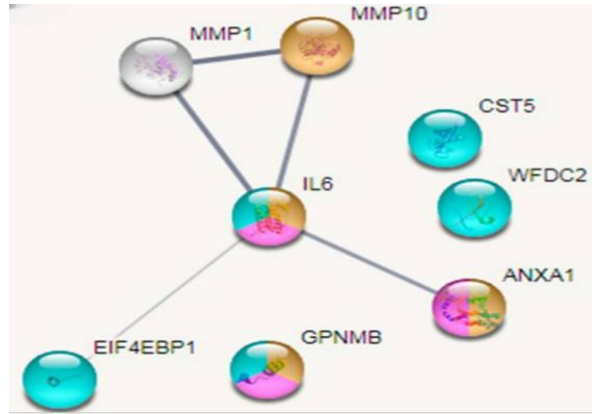


Figure 21: Protein–protein interaction network of proteins generated by String-DB, with significantly different expression between the OSCC vs OLK. Within the network each circle represents a protein, and the lines between the circles represent protein–protein interactions. The line width is correlated with the strength of the supporting evidence for the interaction; the wider the line, the more experimental evidence supports the interaction. The circle color indicates enriched GO functions. The pink color represents the negative regulation of cell death, aqua blue color the regulation of cellular metabolic processes, and brown color represents the regulation of cell migration.

The enriched functions between OSCC and OLK were immune response, cell migration, negative regulation of cell death, collagen catabolism, and the regulation of endopeptidase activity.

There were no interactions found in the proteins differentially expressed between OLP and OLK.

4.4.4 Gene-gene interaction network analysis of differentially expressed proteins between the studied groups

A CluePedia search was undertaken to acquire further information on the proteins that showed statistically significant changes between the groups, and the types of interactions were investigated (Figure 22).

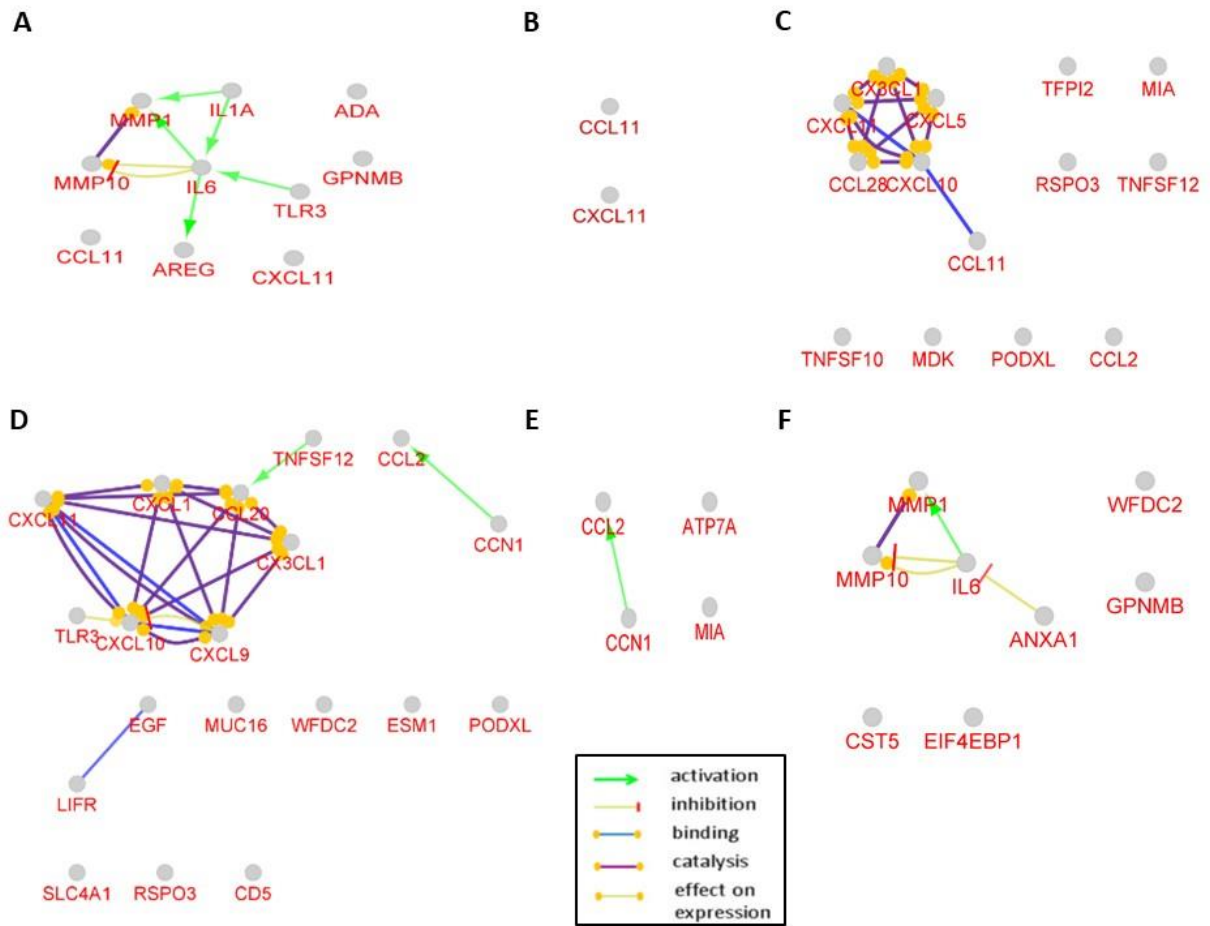


Figure 22: Gene-gene interaction networks (GIN) of the genes corresponding to proteins, with significantly different expression between the patient groups. Each circle represents a gene and the lines represent the different types of interaction between the genes. The panels show the GIN of the genes of the significantly different proteins between (A) OSCC and control, (B) OLK and control, (C) OLP and control group, (D) OSCC and OLP, (E) OLK and OLP and (F) OLK and OSCC. Line color indicates the type of interaction: green color refers to activation, red color to inhibition, blue color to binding, whereas brown color shows co-expression. The proteins are labeled with their gene name.

The relationships found with String analyses were validated in the GINs. The study of GINs added more information to the data collected by the examination of the network using only String-DB.

In the OSCC vs control, it was observed that AMP IL6 was activated by TLR3 and IL1A, while IL6 further activated the amphiregulin and MMP1 and inhibited MMP10 (Figure 22A). In addition to the previously known information, activation of CCL20 by TNFSF12 and

of CCL2 by CYR61/CCN1 was identified in the networks of proteins with statistically significant differences between OSCC and OLP (Figure 22D). The network between OLK and OSCC, on the other hand, revealed that ANXA1 inhibits IL6, which then activates MMP1 and inhibits MMP10 (Figure 22F).

5. DISCUSSION

The purpose of this study was to look into the potential involvement of AMPs in different pathological conditions. AMPs, as part of the innate immune system, control immune responses, while also directly reducing infections, either by killing pathogens or inhibiting their proliferation (11).

Our aim was to discover which previously identified AMPs by proteomics methods in the brain, blood, and CSF, could play a role in the pathogenesis of Alzheimer's disease (AD). The pivotal role of APP in the pathogenesis of AD has been extensively investigated, and on the basis of the amyloid hypothesis, the A β produced by APP cleavage plays a very important role in the disease's progression (90). Along with the amyloid hypothesis, research data have shown that bacterial or viral infections might cause the disease to manifest, and immune system activation is essential for the progression of AD, pointing to a complicated phenomenon lurking beneath the symptoms' appearance (91–94). Various groups discovered a direct link between A β and antimicrobial defense, revealing that A β peptides can lead to bacterial death and aggregated A β can assist in pathogen immobilization (91). It was demonstrated that A β can interact with viral coat proteins when acting against influenza strains and herpes simplex virus, further highlighting its AMP feature (95–97).

To examine this subject, we first gathered all accessible human AMPs by creating the UDAMP Database. The database presently has 670 AMPs and is accessible at Figshare https://figshare.com/articles/dataset/UDAMP_Database/20926873, but at the time of the analysis, it contained 186 AMPs.

The network analysis of all human AMPs identified three clusters of interacting proteins. Cluster 1 featured a densely linked network of defensins, with DEFA4 acting as a hub connecting it to Cluster 2. Cluster 3 contained cytokines as well as other AMPs such as neuropeptides (VIP, NPY, VGF, etc.), APP, blood coagulation factors (F2, FGA, SERPINA1), and so on. The link between Clusters 1 and 3 was established via Cluster 2, which had fewer AMPs than the other two clusters. CLU and CCL25, CCL21, and B2M were the hub proteins that connected Clusters 2 and 3. APP, DEFA4, CXCL1, CXCL10, kininogen 1, PPBP, POMC, PF4, and other AMPs were hub proteins, with the most important biological roles relating to defense, innate immune system, and signal transduction. As expected, the examination of the network of human AMPs recapitulated knowledge from the scientific literature regarding their activities and roles in human beings (11,98,99).

The UDAMP Database listed 46 human AMPs discovered in the brain, 46 in the CSF, and 75 in the blood. The collection of AMPs found in each sample type overlapped but was not identical. In terms of network analysis, each sample type had certain distinct characteristics; nevertheless, in terms of enriched GO functions, AMPs discovered in the brain, CSF, or blood exhibited general AMP functions linked to antimicrobial defense, immune system activation, and blood coagulation.

The situation was different when it came to the proteins that showed a statistically significant alteration in AD. The levels of nine AMPs were discovered to be increased in a statistically significant manner in the brain, and two of the AMPs were also found in the blood. These two proteins (FGA and APP) may have diagnostic potential in AD from blood, and there are already several diagnostic techniques for using FGA and APP as biomarkers (100–102). According to our findings and those presented in the scientific literature, APP is the key hub protein with the greatest number of interactions in AD (103,104). In terms of CSF, only TAC1 overlapped with AMPs elevated in the brain and no overlap with AMPs increased in the blood was observed. When the AMPs that showed a statistically significant decrease in AD were analysed, four AMPs that were increased in the brain were found to be decreased in the blood (CLU, HP, S100A9, SERPINA1). A high level of CLU (or ApoJ) and HP was already linked to AD, but their higher levels were found not only in the brain but also in the blood (105–108). Given that the blood collection period in the different tests varied in relation to the beginning of AD, additional study with consistent sample collection times is likely needed to clarify the situation regarding the dynamics of the change in the amount of these proteins in the blood. B2M levels rose in the CSF while decreased in the brain. FAU levels in the brain decreased, as the level of VGF in both the brain and the CSF. Data published in peer-reviewed journals show higher B2M and FAU levels in the brain, suggesting them as viable targets for AD treatment (109–111). FAU, or 40S ribosomal protein S30, plays a role in DNA repair and function, and its levels in the AD brain appeared to be decreased according to another study (112).

Our results on TAC1 and VGF are consistent with previously reported data: TAC1-derived neuropeptides have a neuroprotective role in AD (113), and VGF was found as a major driving gene/protein of AD in a mouse model. It was discovered that increasing VGF level helped by reducing the memory impairment symptoms in mice (114). A network study revealed that a gradual decrease in the amount of VGF, along with other proteins, was diagnostic of asymptomatic AD and AD (115).

Aside from the AMPs stated above, there are other AMPs that either considerably increased or decreased in AD patients. Chromogranin-A has been demonstrated to play a role

in activating microglia, which can emit neurotoxins and cause neuronal death (116). Some of the other AMPs have already been related to AD. S100A7 is a recently identified possible biomarker for AD (34,117) and PPBP and zinc-alpha-2-glycoprotein were discovered to be downregulated in the serum of patients with AD (88). Our findings support their downregulation in CSF. The rise in VIP can be viewed as an attempt to slow the progression of neurodegeneration since VIP inhibits the activation of microglial cells, preventing the production of additional inflammatory cytokines (118).

When the activities of the differentially expressed AMPs that are not connected to the innate immune system were investigated, it was revealed that the majority of the proteins had protective effects, such as the antioxidant HP, which is responsible for heme binding, and CLU, which inhibits the development of amyloid fibrils and reduces protein aggregation caused by stress (119). FGA, SERPINA1, and PPBP all play a role in hemostasis (120–122) preventing blood loss.

In terms of the network analysis of AMPs changed in AD, our findings support the role of inflammation and innate defense systems in the illness's progression. Aside from these broad roles, changes influencing protein post-translational modification may be associated to tau hyperphosphorylation (123), ubiquitination (76), glycosylation (79), or citrullination (124) of brain proteins, all characteristic to AD. Dietary excess of lipids and lipid metabolism impairment, as well as insulin intolerance along with a lack of neurotrophic factors, can all contribute to the cognitive deterioration seen in AD (125–131). Our findings add to the evidence for changes in lipoprotein metabolism and cholesterol metabolism seen in CSF samples obtained from patients with AD. We could also highlight the importance of hemostasis via network analysis. The function of blood clotting in the development of AD was investigated, and it was discovered to be damaged (132). Some clotting factors may damage synapses, providing new possible candidates for diagnostic and therapeutic intervention (133).

Changes in GO functions such as APP metabolism, amyloid fibril formation, inflammation, post-translational protein modification, translation, antimicrobial defense, hemostasis, response to axonal injury, regulation of IGFR signaling, cellular response to cAMP, and others suggest that AMPs play a central role in the disease's development.

Studies aimed at understanding the pathophysiological events that contribute to the formation of AD and its progression have revealed protein modules that play a role in the illness's pathogenesis. Protein modules and proteins involved in synapse function, mitochondrial function, glucose and carbohydrate metabolism, extracellular matrix, cytoskeleton, and RNA binding/splicing have all been associated to AD (134,135) being

consistent with our findings regarding changed activities linked to extracellular matrix architecture observed in CSF and blood, as well as altered RNA metabolism observed in the brain and blood. RNA-binding proteins, alterations linked to RNA metabolism have been observed in AD, and RNA splicing has been shown to be significant in the exacerbation of AD symptoms (135). Other studies have identified proteostasis, RNA homeostasis, immune response, neuroinflammation, synaptic transmission, vesicular transport, cell signaling, cellular metabolism, lipid homeostasis, mitochondrial function, cytoskeleton organization, and myelin-axon interactions as key players in the pathology of AD (136), and the majority of these functions have been demonstrated in our study as well.

On the basis of composite network analysis, VEGF was identified as a functional hub protein modulated by the greatest number of proteins, and it has been shown to have a role in AD by other groups (137). Similarly, the other hub proteins identified: insulin-like growth factor 1 (138), EGF (139,140) SERPINE1 (141), fibronectin (142) were linked earlier to aging, neurodegeneration, and AD.

Despite their small quantity (0.5-1% of all discovered proteins), the network analysis of AMPs might indicate the crucial activities mentioned previously in AD. Our findings highlight the necessity of AMP testing and the function of the innate immune system and blood clotting in AD. At the same time, we could find further evidence indicating that AMPs play a significant role in AD. As a result, AMPs may be suitable candidates for additional mechanistic research aimed at understanding their precise function in the complicated pathophysiology of AD (143) and these may be used as therapeutic targets in the future.

The other pathological condition where the importance of AMPs was examined during our study was related to the oral cancer and precancerous lesions. Oral squamous cell carcinoma (OSCC), along with the precancerous conditions: oral leucoplakia (OLK) and oral lichen planus (OLP) are responsible for a high number of pathological conditions appearing in patients (144,145). Saliva can be a sample of choice when we would like to examine patients efficiently and possibly to screen the patients with high risk for OSCC or precancerous conditions (146). In order to be able to perform screening, first proper biomarkers are needed. Our group applied proximity extension assay (PEA) for the thorough examination of saliva and could demonstrate that the PEA technique is suitable for saliva analysis as well and it can be the technique of choice when extensive saliva analysis is required. Our aim during this study was to examine the AMPs showing differential expression between the OSCC, OLK, OLP and control groups, respectively.

The differentially expressed proteins between the OSCC and the control groups were mostly involved in biological functions having role in carcinogenesis. The inflammatory response, regulation of growth factor production, collagen catabolism, positive regulation of intracellular signalling, and regulation of cellular proliferation are all well-known to be associated with tumor formation and progression (147).

Between the OLK and control groups, only two proteins were seen to be differentially expressed. Both of them were chemokines, involved in many general processes like chemokine-signaling, antimicrobial humoral immune response, and lymphocyte chemotaxis (148).

The 14 proteins differentially expressed between in OLP compared to the control group were involved in signaling pathway and immune response, being possibly related to the chronic inflammatory background observed in case of OLP (149). Ten of the proteins were responsible for cell migration, being in line with the histopathological findings of T-cell and lymphocyte infiltration in the epithelium and lamina propria observed in OLP (148). Proteins responsible for positive response to stimulus, stress response, and signaling may reflect the association between viral infection and OLP induction, and OLP and autoimmune conditions (150).

Further evidence highlights the inflammatory background seen in OLP. Among proteins differentially expressed between OSCC and OLP, many were involved in leukocyte chemotaxis and inflammatory response (151), correlating with the data observed by histopathology, namely, the infiltration of various leukocyte types in the epithelium and lamina propria (149).

Considering the difference between OSCC and OLK the differentially expressed proteins indicate the presence of the malignant transformation. The enriched functions were related to immune response, cell migration, negative regulation of cell death, collagen catabolism, and the regulation of endopeptidase activity. The increase in MMPs expression could also be observed. The MMPs are responsible for the catabolism of collagen in the basement membrane and the disassembly of extracellular matrix (152). Another important differentially expressed protein was ANXA1, which is involved in a large number of biological functions (153). ANXA1 and GPNMB are important in malignant transformation, being involved in procarcinogenic biological events, such as negative regulation of cell death and regulation of G1/S transition of mitotic cell cycle. It was also shown, that GPNMB is overexpressed in many cancer types, such as brain and breast cancers (153,154). The IL6, was shown earlier by our and other groups as characteristic to OSCC, being involved in inflammatory- and carcinogenesis-related biological functions (155–157).

The proteins differentially expressed between the OLP and OLK groups had a role in signalling, but did not show interactions among each other. No specific common biological functions could be drawn from these unrelated proteins. This further confirmed the understanding that OLK and OLP are two unrelated conditions that separately progress to OSCC (158,159).

As far as most of these proteins belong to the AMP family, it is worth examining the AMPs in relation to oral cancers and premalignant conditions as well.

Our study indicates that AMPs constitute a functionally important class of proteins and their analysis can give valuable insights on the pathogenesis of Alzheimer's disease and oral cancer and precancerous conditions. Their utilization potential should be tested on other diseases as well, but according to our data, their analysis can provide with useful information about the disease and potential biomarkers able to help the diagnosis or the progression of the diseases.

6. NOVEL FINDINGS

- A comprehensive database containing the human AMPs was generated. The UDAMP database currently has 670 entries.
- We have successfully reutilized proteomics data deposited in publicly available databases such as ProteomeXchange (www.proteomexchange.org).
- We have examined the AMPs characteristic to AD in brain, CSF and blood. According to our data the AMPs make up 0.5–1% of all identified proteins, but their network analysis could give information on the functions related to AD, previously reported in the scientific literature.
- Our data emphasize the key role of the innate immune system and blood clotting in the development of AD and suggests the deep involvement of AMPs in AD. They can be candidates for further mechanistic studies aiming to understand their exact role in the complex role in the pathophysiology of AD and may serve as targets for future therapies.
- We thoroughly examined the proteomics data arising from the analysis of saliva samples originating from patients with OSCC, OLK or OLP and discovered that many of the expressed proteins, such as CCL20, IL6, and CCL2, were AMPs.
- We discovered that inflammation and malignant transformation are two major pathways expressed by AMPs. Our findings shed light on the central role of IL6 in the OSCC, OLK,

and OLP.

- We found no interaction between the proteins characteristic to OLP or OLK groups, indicating that these two conditions are independent of one another and progress to OSCC.

7. SUMMARY

This thesis focuses on the examination of the role of AMPs in two different pathological conditions: namely Alzheimer's disease (AD) and oral cancer along with two precancerous conditions.

Given the inflammatory nature of AD, we aimed the examination of antimicrobial and immunomodulatory peptide (AMP) family in human proteome datasets deposited in the publicly available proteomics repositories. First, a unified, complete human antimicrobial and immunomodulatory peptide database (UDAMP) was created, coupled with the examination of datasets downloaded from publicly available sources containing high-quality proteomics data derived from AD and control groups. Extensive network analysis was used to study the possible implications of AMPs in the pathogenesis of AD. AMPs made up less than 1% of all detected proteins in the brain, but their study could replicate the abnormalities previously identified in AD. Our findings emphasize the significance of the innate immune system and blood clotting in the progression of AD and can provide with new potential therapeutic targets.

Saliva is a bodily fluid with great diagnostic potential aiming to help diagnose oral malignancies. Apart from oral squamous cell carcinoma (OSCC), it is critical to examine precancerous conditions such as oral lichen planus (OLP) and leukoplakia (OLK) as well. Our aim was to assess the changes in salivary AMPs in case of patients with OSCC, OLK or OLP. Our research revealed pathways and proteins characteristic to each condition and highlighted the distinct natures of the OLK and OLP.

Our study indicates that AMPs constitute a functionally important class of proteins and their analysis can give valuable insights on the pathogenesis of AD, OSCC, OLK and OLP. The examination of AMPs can provide with useful information about the disease and potential biomarkers able to help the diagnosis or the progression of the diseases.

8. ACKNOWLEDGMENT

Among all those who contributed to my Ph.D. thesis, I would like to thank my supervisor, Dr. Éva Csősz, for her continual support, scholarly conversations, and the chances she provided me during my studies.

I am thankful to Prof. József Tőzsér, head of the Department of Biochemistry and Molecular Biology, for the opportunity to conduct research in a well-equipped environment.

I am grateful to my previous and recent colleagues at the Proteomics Core Facility, notably Dr. Gergő Kalló, Erdenetsetseg Nokhoijav, Andrea Guba, Azhar Orynbek, Renáta Kovács, Julianna Kökényesiné Csáki and VoMinh Doan.

I also want to thank the Tempus Public Foundation for giving me the Stipendium Hungaricum scholarship, which allowed me to study in Hungary.

Last but not least, I want to thank my wife, parents, and grandmother for all their aid, love, and encouragement throughout my life.

This work was supported by the Stipendium Hungaricum Scholarship Program, by the GINOP-2.3.3-15-2016-00020, and the Hungarian Scientific Research Fund (OTKA) K105034.

9. REFERENCES

1. Ganz T. The role of antimicrobial peptides in innate immunity. *Integr Comp Biol.* 2003;43(2):300–4.
2. Jenssen H, Hamill P, Hancock REW. Peptide antimicrobial agents. *Clin Microbiol Rev.* 2006;19(3):491–511.
3. Huan Y, Kong Q, Mou H, Yi H. Antimicrobial Peptides: Classification, Design, Application and Research Progress in Multiple Fields. *Front Microbiol.* 2020;11(October):1–21.
4. Gogoladze G, Grigolava M, Vishnepolsky B, Chubinidze M, Duroux P, Lefranc MP, et al. DBAASP: Database of antimicrobial activity and structure of peptides. *FEMS Microbiol Lett.* 2014;357(1):63–8.
5. Bardan A, Nizet V, Gallo RL. Antimicrobial peptides and the skin. *Expert Opin Biol Ther.* 2004;4(4):543–9.
6. Pasupuleti M, Schmidtchen A, Malmsten M. Antimicrobial peptides: Key components of the innate immune system. Vol. 32, *Critical Reviews in Biotechnology.* 2012. p. 143–71.
7. Prashant K, Kizhakkedathu JN, Straus SK, Ala A. Antimicrobial peptides: Diversity, mechanism of action and strategies to improve the activity and biocompatibility in vivo. *Biomolecules.* 2018;8(1).
8. Guangshun W. Human antimicrobial peptides and proteins. *Pharmaceuticals.* 2014;7(5):545–94.
9. Koehbach J, Craik DJ. The Vast Structural Diversity of Antimicrobial Peptides. *Trends Pharmacol Sci [Internet].* 2019;40(7):517–28. Available from: <https://doi.org/10.1016/j.tips.2019.04.012>
10. Torres AM, Kuchel PW. The β -defensin-fold family of polypeptides. *Toxicon.* 2004;44(6):581–8.
11. Wiesner J, Vilcinskas A. Antimicrobial peptides: The ancient arm of the human immune system. *Virulence.* 2010;1(5):440–64.
12. Mojsoska B, Jenssen H. Peptides and peptidomimetics for antimicrobial drug design. *Pharmaceuticals.* 2015;8(3):366–415.
13. Latendorf T, Gerstel U, Wu Z, Bartels J, Becker A, Tholey A, et al. Cationic Intrinsically Disordered Antimicrobial Peptides (CIDAMPs) Represent a New Paradigm of Innate Defense with a Potential for Novel Anti-Infectives. *Sci Rep.*

- 2019;9(1):1–25.
14. Paulo Bastos, F´abio Trindade, Joao da Costa RF and RV. Human Antimicrobial Peptides in Bodily Fluids: Current Knowledge and Therapeutic Perspectives in the Postantibiotic Era. Wiley Online Libr. 2018;38(1):101–146.
 15. Alison M. McDermott. Antimicrobial Compounds in Tears Alison. Exp Eye Res. 2013;117.
 16. Karsten E. Diversity of Microbial Species Implicated in Keratitis: A Review. Open Ophthalmol J. 2012;6(1):110–24.
 17. Fleiszig SMJ, Kwong MSF, Evans DJ. Modification of *Pseudomonas aeruginosa* interactions with corneal epithelial cells by human tear fluid. Infect Immun. 2003;71(7):3866–74.
 18. Kwong MSF, Evans DJ, Ni M, Cowell BA, Fleiszig SMJ. Human tear fluid protects against *Pseudomonas aeruginosa* keratitis in a murine experimental model. Infect Immun. 2007;75(5):2325–32.
 19. McNamara NA, Andika R, Kwong M, Sack RA, Fleiszig SMJ. Interaction of *Pseudomonas aeruginosa* with human tear fluid components. Curr Eye Res. 2005;30(7):517–25.
 20. Khurshid Z, Naseem M, Sheikh Z, Najeeb S, Shahab S, Zafar MS. Oral antimicrobial peptides: Types and role in the oral cavity. Saudi Pharm J [Internet]. 2016;24(5):515–24. Available from: <http://dx.doi.org/10.1016/j.jsps.2015.02.015>
 21. Schulz BL, Cooper-White J, Punyadeera CK. Saliva proteome research: Current status and future outlook. Crit Rev Biotechnol. 2013;33(3):246–59.
 22. Güncü GN, Yilmaz D, Könönen E, Gürsoy UK. Salivary antimicrobial peptides in early detection of periodontitis. Front Cell Infect Microbiol. 2015;5(DEC):1–6.
 23. Grant M, Kilsgård O, Åkerman S, Klinge B, Demmer RT, Malmström J, et al. The Human Salivary Antimicrobial Peptide Profile according to the Oral Microbiota in Health, Periodontitis and Smoking. J Innate Immun. 2019;11(5):432–43.
 24. Levy O. Antimicrobial proteins and peptides of blood: Templates for novel antimicrobial agents. Blood. 2000;96(8):2664–72.
 25. Vasilchenko AS, Rogozhin EA, Vasilchenko A V., Kartashova OL, Sycheva M V. Novel haemoglobin-derived antimicrobial peptides from chicken (*Gallus gallus*) blood: purification, structural aspects and biological activity. J Appl Microbiol. 2016;121(6):1546–57.
 26. Mak P. Hemocidins in a functional and structural context of human antimicrobial

- peptides. *Front Biosci (Landmark Ed)*. 2008;6859–71.
27. Shao D, Huang L, Wang Y, Cui X, Li Y, Wang Y, et al. HBFP: a new repository for human body fluid proteome. *Database*. 2021;2021(April):1–14.
 28. Ali ASM, Townes CL, Hall J, Pickard RS. Maintaining a Sterile Urinary Tract: The Role of Antimicrobial Peptides. *J Urol [Internet]*. 2009;182(1):21–8. Available from: <http://dx.doi.org/10.1016/j.juro.2009.02.124>
 29. Zhao M, Li M, Yang Y, Guo Z, Sun Y, Shao C, et al. A comprehensive analysis and annotation of human normal urinary proteome. *Sci Rep*. 2017;7(1):1–13.
 30. Becknell B, Eichler T, Beceiro S, Li B, Easterling R, Ashley R, et al. Ribonucleases 6 and 7 have antimicrobial function in the human and murine urinary tract. 2015;87(1):151–61.
 31. Zasloff M. Antimicrobial peptides, innate immunity, and the normally sterile urinary tract. *J Am Soc Nephrol*. 2007;18(11):2810–6.
 32. Sakka L, Coll G, Chazal J. Anatomy and physiology of cerebrospinal fluid. *Eur Ann Otorhinolaryngol Head Neck Dis [Internet]*. 2011;128(6):309–16. Available from: <http://dx.doi.org/10.1016/j.anorl.2011.03.002>
 33. Spector R, Robert Snodgrass S, Johanson CE. A balanced view of the cerebrospinal fluid composition and functions: Focus on adult humans. *Exp Neurol [Internet]*. 2015;273:57–68. Available from: <http://dx.doi.org/10.1016/j.expneurol.2015.07.027>
 34. Qin W, Ho L, Wang J, Peskind E, Pasinetti GM. S100A7, a Novel Alzheimer's disease biomarker with non-amyloidogenic α -secretase activity acts via selective promotion of ADAM-10. *PLoS One*. 2009;4(1).
 35. Sathe G, Na CH, Renuse S, Madugundu AK, Albert M, Moghekar A, et al. Quantitative Proteomic Profiling of Cerebrospinal Fluid to Identify Candidate Biomarkers for Alzheimer's Disease. *Proteomics - Clin Appl*. 2019;13(4).
 36. Stuart BAR, Franitza AL, E L. Regulatory Roles of Antimicrobial Peptides in the Nervous System: Implications for Neuronal Aging. *Front Cell Neurosci*. 2022;16(March).
 37. Maffei FA, Heine RP, Whalen MJ, Mortimer LF, Carcillo JA. Levels of antimicrobial molecules defensin and lactoferrin are elevated in the cerebrospinal fluid of children with meningitis. *Pediatrics*. 1999;103(5 I):987–92.
 38. Brandenburg LO, Varoga D, Nicolaeva N, Leib SL, Wilms H, Podschun R, et al. Role of glial cells in the functional expression of LL-37/rat cathelin-related antimicrobial peptide in meningitis. *J Neuropathol Exp Neurol*. 2008;67(11):1041–54.

39. Drummond E, Wisniewski T. Using Proteomics to Understand Alzheimer's Disease Pathogenesis. *Alzheimer's Dis.* 2019;37–51.
40. Wattamwar P, Mathuranath PS. An overview of biomarkers in Alzheimer's disease. *Ann Indian Acad Neurol.* 2010;13(SUPPL. 2):1–15.
41. Hardy J, Selokoe D. The Amyloid Hypothesis of Alzheimer 's Disease. *Amyloid Int J Exp Clin Investig.* 2002;297(5580):353–7.
42. Potential fabrication in research images threatens key theory of Alzheimer's disease | Science | AAAS [Internet]. [cited 2022 Oct 20]. Available from: <https://www.science.org/content/article/potential-fabrication-research-images-threatens-key-theory-alzheimers-disease>
43. Gosztyla ML, Brothers HM, Robinson SR. Alzheimer's Amyloid- β is an Antimicrobial Peptide: A Review of the Evidence. *J Alzheimer's Dis.* 2018;62(4):1495–506.
44. Moir RD, Lathe R, Tanzi RE. The antimicrobial protection hypothesis of Alzheimer's disease. *Alzheimer's Dement [Internet].* 2018;14(12):1602–14. Available from: <https://doi.org/10.1016/j.jalz.2018.06.3040>
45. Vigasova D, Nemergut M, Liskova B, Damborsky J. Multi-pathogen infections and Alzheimer's disease. *Microb Cell Fact [Internet].* 2021;20(1):1–13. Available from: <https://doi.org/10.1186/s12934-021-01520-7>
46. Tuppo EE, Arias HR. The role of inflammation in Alzheimer's disease. *Int J Biochem Cell Biol.* 2005;37(2):289–305.
47. Soscia SJ, Kirby JE, Washicosky KJ, Tucker SM, Ingelsson M, Hyman B, et al. The Alzheimer's disease-associated amyloid β -protein is an antimicrobial peptide. *PLoS One.* 2010;5(3):1–10.
48. Pires FR, Ramos AB, de Oliveira JBC, Tavares AS, de Luz PSR, dos Santos TCRB. Oral squamous cell carcinoma: Clinicopathological features from 346 cases from a single oral pathology service during an 8-year period. *J Appl Oral Sci.* 2013;21(5):460–7.
49. Speight PM, Khurram SA, Kujan O. Oral potentially malignant disorders: risk of progression to malignancy. *Oral Surg Oral Med Oral Pathol Oral Radiol [Internet].* 2018;125(6):612–27. Available from: <https://doi.org/10.1016/j.oooo.2017.12.011>
50. Ganesh D, Sreenivasan P, Ohman J, Wallström M, Braz-Silva PH, Giglio D, et al. Potentially malignant oral disorders and cancer transformation. *Anticancer Res.* 2018;38(6):3223–9.

51. Lucchese A, Di Stasio D, Romano A, Fiori F, De Felice GP, Lajolo C, et al. Correlation between Oral Lichen Planus and Viral Infections Other Than HCV: A Systematic Review. *J Clin Med*. 2022;11(18).
52. Van Der Waal I, Schepman KP, Van Der Meij EH, Smeele LE. Oral leukoplakia: A clinicopathological review. *Oral Oncol*. 1997;33(5):291–301.
53. Villa A, Sonis S. Oral leukoplakia remains a challenging condition. *Oral Dis*. 2018;24(1–2):179–83.
54. Sun Y, Liu N, Guan X, Wu H, Sun Z, Zeng H. Immunosuppression Induced by Chronic Inflammation and the Progression to Oral Squamous Cell Carcinoma. *Mediators Inflamm*. 2016;2016.
55. Kalló G, Kumar A, Tőzsér J, Csősz É. Chemical Barrier Proteins in Human Body Fluids. *Biomedicines*. 2022;10(7):1472.
56. Pujana MA, Han JDJ, Starita LM, Stevens KN, Tewari M, Ahn JS, et al. Network modeling links breast cancer susceptibility and centrosome dysfunction. *Nat Genet*. 2007;39(11):1338–49.
57. Mummery-widmer JL, Yamazaki M, Stoeger T, Bhalerao S, Chen D, Dietzl G, et al. Genome-wide analysis of Notch signalling in *Drosophila* by transgenic RNAi. *Nat Genet*. 2010;458(7241):987–92.
58. Fraser AG, Marcotte EM. A probabilistic view of gene function. *Nat Genet*. 2004;36(6):559–64.
59. Daniele Merico, David Gfeller and GDB. How to visually interpret biological data using networks. *Nat Biotechnol*. 2009;27(10):921–4.
60. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47(D1):D607–13.
61. Shannon P, Markiel A, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang DR, Amin N, Benno Schwikowski and TI. Cytoscape: A Software Environment for Integrated Models. *Genome Res [Internet]*. 2003;13(22):426. Available from: <http://ci.nii.ac.jp/naid/110001910481/>
62. Snel B, Bork P, Huynen MA. The identification of functional modules from the genomic association of genes. *Proc Natl Acad Sci U S A*. 2002;99(9):5890–5.
63. Enright AJ, Ouzounis CA. Functional associations of proteins in entire genomes by means of exhaustive detection of gene fusions. *Genome Biol*. 2001;2(9):1–7.

64. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* [Internet]. 2003 Nov;13(11):2498–504. Available from: <http://ci.nii.ac.jp/naid/110001910481/>
65. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*. 2009;25(8):1091–3.
66. Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics* [Internet]. 2013 Mar;29(5):661–3. Available from: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btt019>
67. Dey KK, Wang H, Niu M, Bai B, Wang X, Li Y, et al. Deep undepleted human serum proteome profiling toward biomarker discovery for Alzheimer’s disease. *Clin Proteomics* [Internet]. 2019;16(1):1–12. Available from: <https://doi.org/10.1186/s12014-019-9237-1>
68. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res*. 2017;45(D1):D353–61.
69. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, et al. The Gene Oncology (GO) database and informatics resource. *Nucleic Acids Res*. 2004;32(DATABASE ISS.):258–61.
70. Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: Pathway insights using integrated experimental and in silico data. *Bioinformatics*. 2013;29(5):661–3.
71. Thomas S, Karnik S, Barai RS, Jayaraman VK, Idicula-Thomas S. CAMP: A useful resource for research on antimicrobial peptides. *Nucleic Acids Res*. 2009;38(SUPPL.1):774–80.
72. Zhe W, Guangshun W. APD: The antimicrobial peptide database. *Nucleic Acids Res*. 2004;32(DATABASE ISS.):590–2.
73. Jhong JH, Chi YH, Li WC, Lin TH, Huang KY, Lee TY. DbAMP: An integrated resource for exploring antimicrobial peptides with functional activities and physicochemical properties on transcriptome and proteome data. *Nucleic Acids Res*. 2019;47(D1):D285–97.
74. Zhao X, Wu H, Lu H, Li G, Huang Q. LAMP: A Database Linking Antimicrobial Peptides. *PLoS One*. 2013;8(6):6–11.

75. Bateman A, Martin MJ, O'Donovan C, Magrane M, Alpi E, Antunes R, et al. UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* 2017;45(D1):D158–69.
76. Measho HA, Eric BD, Lingyan P, Tian Z, M.Duong D, Gearing M, et al. Quantitative Analysis of the Brain Ubiquitylome in Alzheimer's Disease. *Proteomics.* 2018;18(20).
77. Xu J, Patassini S, Rustogi N, Riba-Garcia I, Hale BD, Phillips AM, et al. Regional protein expression in human Alzheimer's brain correlates with disease severity. *Commun Biol [Internet].* 2019;2(1). Available from: <http://dx.doi.org/10.1038/s42003-018-0254-9>
78. Adav SS, Park JE, Sze SK. Quantitative profiling brain proteomes revealed mitochondrial dysfunction in Alzheimer's disease. *Mol Brain.* 2019;12(1):1–12.
79. Wang S, Yang F, Petyuk VA, Shukla AK, Monroe ME, Gritsenko MA, et al. Quantitative proteomics identifies altered O-GlcNAcylation of structural, synaptic and memory-associated proteins in Alzheimer's disease. *J Pathol.* 2017;243(1):78–88.
80. Chadwick M. Hales^{2, 4}, Eric B. Dammer^{1, 4}, Qiudong Deng^{1, 4}, Duc M. Duong^{1, 4}, M, Gearing^{3, 4}, Juan C. Troncoso⁵, Madhav Thambisetty⁶, James J. Lah^{2, 4}, Joshua M. Shulman⁷, Allan I. Levey^{2, 4}, and Nicholas T. Seyfried^{1, 2 4}. Changes in the detergent-insoluble brain proteome linked to amyloid and tau in Alzheimer's Disease progression. 2016;16(23):3042–53.
81. Bai B, Hales CM, Chen PC, Gozal Y, Dammer EB, Fritz JJ, et al. U1 small nuclear ribonucleoprotein complex and RNA splicing alterations in Alzheimer's disease. *Proc Natl Acad Sci U S A.* 2013;110(41):16562–7.
82. Higginbotham L, Dammer E, Duong D, Modeste E, Montine T, Lah J, et al. Network Analysis of a Membrane-Enriched Brain Proteome across Stages of Alzheimer's Disease. *Proteomes.* 2019;7(3):30.
83. Bereczki E, Branca RM, Francis PT, Pereira JB, Baek JH, Hortobágyi T, et al. Synaptic markers of cognitive decline in neurodegenerative diseases: A proteomic approach. *Brain.* 2018;141(2):582–95.
84. Haytural H, Mermelekas G, Emre C, Nigam SM, Carroll SL, Winblad B, et al. The proteome of the dentate terminal zone of the perforant path indicates presynaptic impairment in Alzheimer disease. *Mol Cell Proteomics.* 2020;19(1):128–41.
85. Petyuk VA, Chang R, Ramirez-Restrepo M, Beckmann ND, Henrion MYR, Piehowski PD, et al. The human brainome: Network analysis identifies HSPA2 as a novel Alzheimer's disease target. *Brain.* 2018;141(9):2721–39.
86. Cherry JD, Zeineddin A, Dammer EB, Webster JA, Duong D, Seyfried NT, et al.

- Characterization of detergent insoluble proteome in chronic traumatic encephalopathy. *J Neuropathol Exp Neurol*. 2018;77(1):40–9.
87. Sjödin S, Brinkmalm G, Öhrfelt A, Parnetti L, Paciotti S, Hansson O, et al. Endo-lysosomal proteins and ubiquitin CSF concentrations in Alzheimer’s and Parkinson’s disease. *Alzheimer’s Res Ther*. 2019;11(1):1–16.
 88. Shen L, Liao L, Chen C, Guo Y, Song D, Wang Y, et al. Proteomics analysis of blood serums from Alzheimer’s disease patients using iTRAQ labeling technology. *J Alzheimer’s Dis*. 2017;56(1):361–78.
 89. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44–57.
 90. Yun-wu Z, Thompson R, Zhang H, Xu H. APP processing in Alzheimer’s disease. *Mol Brain* [Internet]. 2011;4(1):3. Available from: <http://www.molecularbrain.com/content/4/1/3>
 91. Kumar D, Kumar V, Choi SH, Washicosky KJ, Eimer WA, Tucker S, et al. Amyloid- β Peptide Protects Against Microbial Infection In Mouse and Worm Models of Alzheimer’s Disease. *Sci Transl Med*. 2017;8(340).
 92. Venegas C, Kumar S, Franklin BS, Dierkes T, Brinkschulte R, Tejera D, et al. Microglia-derived ASC specks crossseed amyloid- β in Alzheimer’s disease. *Nature*. 2017;552(7685):355–61.
 93. Dominy SS, Lynch C, Ermini F, Benedyk M, Marczyk A, Konradi A, et al. *Porphyromonas gingivalis* in Alzheimer’s disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. *Sci Adv*. 2019;5(1):1–22.
 94. András IE, Sewell BB, Toborek M. HIV-1 and amyloid beta remodel proteome of brain endothelial extracellular vesicles. *Int J Mol Sci*. 2020;21(8):1–39.
 95. WJ Lukiw¹, JG Cui¹, YY Li¹, PS Bhattacharjee², M Corkern¹, C Clement¹ EK, MJ Ball³, Y Zhao⁴ and JH. Acyclovir and A β 42 peptide attenuates HSV-1-induced miRNA-146a levels in human brain cells. *Neuroreport* [Internet]. 2010;21(14):922–7. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf>
 96. White MR, Kandel R, Tripathi S, Condon D, Qi L, Taubenberger J, et al. Alzheimer’s associated β -Amyloid protein inhibits influenza a virus and modulates viral interactions with phagocytes. *PLoS One*. 2014;9(7):1–9.
 97. Bourgade K, Le Page A, Bocti C, Witkowski JM, Dupuis G, Frost EH, et al. Protective Effect of Amyloid- β Peptides Against Herpes Simplex Virus-1 Infection in a Neuronal

- Cell Culture Model. *J Alzheimer's Dis*. 2016;50(4):1227–41.
98. Yuping L, Gallo RL. AMPed Up immunity: how antimicrobial peptides have multiple roles in immune defense. *Trends Immunol*. 2009;30(3):131–41.
 99. Chessa C, Bodet C, Jousselin C, Wehbe M, Lévêque N, Garcia M. Antiviral and Immunomodulatory Properties of Antimicrobial Peptides Produced by Human Keratinocytes. *Front Microbiol*. 2020;11(June):1–15.
 100. Irizarry MC. Biomarkers of Alzheimer Disease in Plasma. *NeuroRx*. 2004;1(2):226–34.
 101. Joung Wook L, Namkoong H, Kim HK, Kim S, Hwang DW, Na HR, et al. Fibrinogen gamma-A chain precursor in CSF: A candidate biomarker for Alzheimer's disease. *BMC Neurol*. 2007;7:1–6.
 102. Cortes-Canteli M, Zamolodchikov D, Ahn HJ, Strickland S, Norris EH. Fibrinogen and Altered Hemostasis in Alzheimer's Disease Marta. *J Alzheimers Dis* [Internet]. 2012;32(3):509–608. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf>
 103. Jiao L, Zhu X, Chen JY. Building disease-specific drug-protein connectivity maps from molecular interaction networks and PubMed abstracts. *PLoS Comput Biol*. 2009;5(7).
 104. Rao VS, Srinivas K, Kumar GS, Sujin G. Protein interaction network for Alzheimer's disease using computational approach. *Bioinformatics*. 2013;9(19):968–72.
 105. Spagnuolo MS, Maresca B, La Marca V, Carrizzo A, Veronesi C, Cupidi C, et al. Haptoglobin interacts with Apolipoprotein e and beta-amyloid and influences their crosstalk. *ACS Chem Neurosci*. 2014;5(9):837–47.
 106. Song IU, Kim Y Do, Chung SW, Cho HJ. Association between serum haptoglobin and the pathogenesis of alzheimer's disease. *Intern Med*. 2015;54(5):453–7.
 107. Gupta VB, Hone E, Pedrini S, Doecke J, O'Bryant S, James I, et al. Altered levels of blood proteins in Alzheimer's disease longitudinal study: Results from Australian Imaging Biomarkers Lifestyle Study of Ageing cohort. *Alzheimer's Dement Diagnosis, Assess Dis Monit* [Internet]. 2017;8:60–72. Available from: <https://doi.org/10.1016/j.dadm.2017.04.003>
 108. Foster EM, Dangla-Valls A, Lovestone S, Ribe EM, Buckley NJ. Clusterin in Alzheimer's disease: Mechanisms, genetics, and lessons from other pathologies. *Front Neurosci*. 2019;13(FEB):1–27.
 109. Smith LK, He Y, Park JS, Bieri G, Snethlage CE, Lin K, et al. B2-Microglobulin Is a

- Systemic Pro-Aging Factor That Impairs Cognitive Function and Neurogenesis. *Nat Med*. 2015;21(8):932–7.
110. Matthes F, Hettich MM, Schilling J, Flores-Dominguez D, Blank N, Wiglenda T, et al. Inhibition of the MID1 protein complex: a novel approach targeting APP protein synthesis. *Cell Death Discov* [Internet]. 2018;4(1). Available from: <http://dx.doi.org/10.1038/s41420-017-0003-8>
 111. Li Y, Li Y, Yu H, Yu H, Chen C, Chen C, et al. Proteomic Profile of Mouse Brain Aging Contributions to Mitochondrial Dysfunction, DNA Oxidative Damage, Loss of Neurotrophic Factor, and Synaptic and Ribosomal Proteins. *Oxid Med Cell Longev*. 2020;2020.
 112. Lee CH, Yoo KY, Choi JH, Hwang IK, Choi SY, Won MH. Ribosomal protein S3 immunoreactivity in the young, adult and aged gerbil hippocampus. *J Vet Med Sci*. 2011;73(3):361–5.
 113. Chen XY, Du YF, Chen L. Neuropeptides exert neuroprotective effects in alzheimer’s disease. *Front Mol Neurosci*. 2019;11(January).
 114. Beckmann ND, Lin WJ, Wang M, Cohain AT, Charney AW, Wang P, et al. Multiscale causal networks identify VGF as a key regulator of Alzheimer’s disease. *Nat Commun* [Internet]. 2020;11(1). Available from: <http://dx.doi.org/10.1038/s41467-020-17405-z>
 115. Seyfried NT, Dammer EB, Swarup V, Nandakumar D, Duc M, Yin L, et al. A Multi-Network Approach Identifies Protein-specific Co- expression in Asymptomatic and Symptomatic Alzheimer’s Disease. 2018;4(1):60–72. Available from: <https://www.synapse.org/#!/Synapse:syn2580853>
 116. Sathe G, Albert M, Darrow J, Saito A, Troncoso J, Pandey A, et al. Quantitative proteomic analysis of the frontal cortex in Alzheimer’s disease. *J Neurochem*. 2020;
 117. Fleming JM, Ginsburg E, Oliver SD, Goldsmith P, Vonderhaar BK. Hornerin, an S100 family protein, is functional in breast cells and aberrantly expressed in breast cancer. *BMC Cancer* [Internet]. 2012;12(1):1. Available from: ???
 118. Delgado M, Ganea D. Vasoactive intestinal peptide prevents activated microglia-induced neurodegeneration under inflammatory conditions: potential therapeutic role in brain trauma. *FASEB J*. 2003;17(13):1–17.
 119. Yerbury JJ, Poon S, Meehan S, Thompson B, Kumita JR, Dobson CM, et al. The extracellular chaperone clusterin influences amyloid formation and toxicity by interacting with prefibrillar structures. *FASEB J*. 2007;21(10):2312–22.
 120. J. C. RAU*, 1, L. M. BEAULIEU*, 1, J. A. HUNTINGTON† and FCC. Serpins in

- thrombosis, hemostasis and fibrinolysis. *J Thromb Haemost* [Internet]. 2007;5(1):102–15. Available from:
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf>
121. Smyth SS, Mcever RP, Weyrich AS, Morrell CN, Hoffman MR, Arepally GM, et al. Platelet functions beyond hemostasis. *J Thromb Haemost*. 2009;7(11):1759–66.
 122. Sørensen B, Tang M, Larsen OH, Laursen PN, Fenger-Eriksen C, Rea CJ. The role of fibrinogen: A new paradigm in the treatment of coagulopathic bleeding. *Thromb Res* [Internet]. 2011;128(SUPPL. 1):S13–6. Available from:
[http://dx.doi.org/10.1016/S0049-3848\(12\)70004-X](http://dx.doi.org/10.1016/S0049-3848(12)70004-X)
 123. Tan H, Wu Z, Wang H, Bai B, Li Y, Wang X, et al. Refined phosphopeptide enrichment by phosphate additive and the analysis of human brain phosphoproteome Haiyan. *Proteomics*. 2015;15(0):500–7.
 124. Deolankar SC, Patil AH, Koyangana SG, Subbannayya Y, Prasad TSK, Modi PK. Dissecting Alzheimer’s Disease Molecular Substrates by Proteomics and Discovery of Novel Post-translational Modifications. *Omi A J Integr Biol*. 2019;23(7):350–61.
 125. Schindowski K, Belarbi K, Buée L. Neurotrophic factors in Alzheimer’s disease: Role of axonal transport. *Genes, Brain Behav*. 2008;7(SUPPL. 1):43–56.
 126. Suzanne M de la M. Insulin resistance and Alzheimer’s disease Suzanne. *BMB Rep*. 2009;42(8):475–81.
 127. Allen SJ, J. Watson J, Dawbarn D. The Neurotrophins and Their Role in Alzheimers Disease. *Curr Neuropharmacol*. 2011;9(4):559–73.
 128. Hu N, Yu J, Tan L, Wang Y, Sun L, Tan L. Nutrition and the Risk of Alzheimer ’ s Disease. *Hindawi Publ Corp BioMed Res Int* [Internet]. 2013;2013:12 pages. Available from: <http://dx.doi.org/10.1155/2013/524820>
 129. Dharshini SAP, Taguchi YH, Gromiha MM. Investigating the energy crisis in Alzheimer disease using transcriptome study. *Sci Rep* [Internet]. 2019;9(1):1–13. Available from: <http://dx.doi.org/10.1038/s41598-019-54782-y>
 130. Mitra S, Behbahani H, Erikdotter M. Innovative therapy for Alzheimer’s disease-with focus on biodelivery of NGF. *Front Neurosci*. 2019;13(FEB):1–21.
 131. Kellar D, Craft S. Brain insulin resistance in Alzheimer’s disease and related disorders: mechanisms and therapeutic approaches. *Lancet Neurol* [Internet]. 2020;19(9):758–66. Available from: [http://dx.doi.org/10.1016/S1474-4422\(20\)30231-3](http://dx.doi.org/10.1016/S1474-4422(20)30231-3)
 132. Suidan GL, Singh PK, Patel-Hett S, Chen ZL, Volfson D, Yamamoto-Imoto H, et al. Abnormal clotting of the intrinsic/contact pathway in Alzheimer disease patients is

- related to cognitive ability. *Blood Adv.* 2018;2(9):954–63.
133. Merlini M, Rafalski VA, Rios Coronado PE, Gill TM, Ellisman M, Muthukumar G, et al. Fibrinogen Induces Microglia-Mediated Spine Elimination and Cognitive Impairment in an Alzheimer’s Disease Model. *Neuron* [Internet]. 2019;101(6):1099-1108.e6. Available from: <https://doi.org/10.1016/j.neuron.2019.01.014>
 134. Bader JM, Geyer PE, Müller JB, Strauss MT, Koch M, Leypoldt F, et al. Proteome profiling in cerebrospinal fluid reveals novel biomarkers of Alzheimer’s disease. *Mol Syst Biol.* 2020;16(6):1–17.
 135. Johnson ECB, Dammer EB, Duong DM, Ping L, Yin L, Higginbotham LA, et al. Large-scale Proteomic Analysis of Alzheimer’s Disease Brain and Cerebrospinal Fluid Reveals Early Changes in Energy Metabolism Associated with Microglia and Astrocyte Activation Erik. 2020;26(5):769–80.
 136. Zhang Q, Ma C, Gearing M, Wang PG, Chin LS, Li L. Integrated proteomics and network analysis identifies protein hubs and network alterations in Alzheimer’s disease. *Acta Neuropathol Commun.* 2018;6(1):19.
 137. Mahoney ER, Dumitrescu L, Moore AM, Cambrono FE, De Jager PL, Koran MEI, et al. Brain expression of the vascular endothelial growth factor gene family in cognitive aging and alzheimer’s disease. *Mol Psychiatry* [Internet]. 2019; Available from: <http://dx.doi.org/10.1038/s41380-019-0458-5>
 138. Westwo W, Beiser A, DeCarli C, Harris TB, Chen TC, He XM, et al. Insulin-like growth factor-1 and risk of Alzheimer dementia and brain atrophy. *Neurology.* 2014;82(18):1613–9.
 139. Zhou L, Zhao SZ, Koh SK, Chen L, Vaz C, Tanavde V, et al. In-depth analysis of the human tear proteome. *J Proteomics* [Internet]. 2012 Jul;75(13):3877–85. Available from: <http://dx.doi.org/10.1016/j.jprot.2012.04.053>
 140. Siddiqui S, Fang M, Ni B, Lu D, Martin B, Maudsley S. Central role of the EGF receptor in neurometabolic aging. *Int J Endocrinol.* 2012;2012.
 141. Akhter H, Huang WT, Van Groen T, Kuo HC, Miyata T, Liu RM. A Small Molecule Inhibitor of Plasminogen Activator Inhibitor-1 Reduces Brain Amyloid- β Load and Improves Memory in an Animal Model of Alzheimer’s Disease. *J Alzheimer’s Dis.* 2018;64(2):447–57.
 142. Anna L-P, Leszek J, Krzyanowska-Golab D, Radzik J, Kątnik-Prastowska MI. Molecular status of plasma fibronectin as an additional biomarker for assessment of alzheimer’s dementia risk. *Dement Geriatr Cogn Disord.* 2009;28(4):338–42.

143. Zhong Q, Zou Y, Liu H, Chen T, Zheng F, Huang Y, et al. Toll-like receptor 4 deficiency ameliorates β 2-microglobulin induced age-related cognition decline due to neuroinflammation in mice. *Mol Brain*. 2020;13(1):1–15.
144. Liu Y, Li Y, Fu Y, Liu T, Liu X, Zhang X, et al. Quantitative prediction of oral cancer risk in patients with oral leukoplakia. *Oncotarget*. 2017;8(28):46057–64.
145. Gandolfo S, Richiardi L, Carrozzo M, Broccoletti R, Carbone M, Pagano M, et al. Risk of oral squamous cell carcinoma in 402 patients with oral lichen planus: A follow-up study in an Italian population. *Oral Oncol*. 2004;40(1):77–83.
146. Csősz É, Márkus B, Darula Z, Medzihradzsky KF, Nemes J, Szabó E, et al. Salivary proteome profiling of oral squamous cell carcinoma in a Hungarian population. *FEBS Open Bio*. 2018;8(4):556–69.
147. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell* [Internet]. 2011;144(5):646–74. Available from: <http://dx.doi.org/10.1016/j.cell.2011.02.013>
148. Jin1 T, , Xuehua Xu2 and DH. Chemotaxis, chemokine receptors and human disease. *Cytokine*. 2008;44(1):1–8.
149. Kurago ZB. Etiology and pathogenesis of oral lichen planus: An overview. *Oral Surg Oral Med Oral Pathol Oral Radiol* [Internet]. 2016;122(1):72–80. Available from: <http://dx.doi.org/10.1016/j.oooo.2016.03.011>
150. Proost P, Struyf S, Van Damme J, Fiten P, Ugarte-Berzal E, Opdenakker G. Chemokine isoforms and processing in inflammation and immunity. *J Autoimmun*. 2017;85:45–57.
151. Diane E. Handy, PhD1, Rita Castro, PhD1,2, and Joseph Loscalzo, MD P. Epigenetic Modifications: Basic Mechanisms and Role in Cardiovascular Disease. *Circulation*. 2011;123(19):2145–56.
152. Yadav L, Puri N, Rastogi V, Satpute P, Ahmad R, Kaur G. Matrix metalloproteinases and cancer - Roles in threat and therapy. *Asian Pacific J Cancer Prev*. 2014;15(3):1085–91.
153. Wan YM, Tian J, Qi L, Liu LM, Xu N. ANXA1 affects cell proliferation, invasion and epithelial-mesenchymal transition of oral squamous cell carcinoma. *Exp Ther Med*. 2017;14(5):5214–8.
154. Maric G, Rose AAN, Annis MG, Siegel PM. Glycoprotein non-metastatic b (GPNMB): A metastatic mediator and emerging therapeutic target in cancer. *Onco Targets Ther*. 2013;6:839–52.
155. Csosz É, Lábiscsák P, Kalló G, Márkus B, Emri M, Szabó A, et al. Proteomics

- investigation of OSCC-specific salivary biomarkers in a Hungarian population highlights the importance of identification of population-tailored biomarkers. *PLoS One*. 2017;12(5):1–21.
156. Márton IJ, Horváth J, Lábiscsák P, Márkus B, Dezső B, Szabó A, et al. Salivary il-6 mRNA is a robust biomarker in oral squamous cell carcinoma. *J Clin Med*. 2019;8(11).
157. Kumari N, Dwarakanath BS, Das A, Bhatt AN. Role of interleukin-6 in cancer progression and therapeutic resistance. *Tumor Biol* [Internet]. 2016;37(9):11553–72. Available from: <http://dx.doi.org/10.1007/s13277-016-5098-7>
158. Bombeccari GP, Guzzi G, Tettamanti M, Giann AB, Baj A, Pallotti F, et al. Oral lichen planus and malignant transformation: A longitudinal cohort study. *Oral Surgery, Oral Med Oral Pathol Oral Radiol Endodontology* [Internet]. 2011;112(3):328–34. Available from: <http://dx.doi.org/10.1016/j.tripleo.2011.04.009>
159. Warnakulasuriya S, Ariyawardana A. Malignant transformation of oral leukoplakia: A systematic review of observational studies. *J Oral Pathol Med*. 2016;45(3):155–66.

LIST OF PUBLICATIONS PREPARED BY THE KENÉZY LIFE SCIENCE LIBRARY



UNIVERSITY of
DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN

H-4002 Egyetem tér 1, Debrecen

Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/422/2022.PL
Subject: PhD Publication List

Candidate: Ajneesh Kumar

Doctoral School: Doctoral School of Molecular Cellular and Immune Biology

List of publications related to the dissertation

1. **Kumar, A.**, Doan-Xuan, Q. M., Kunkli, B., Csósz, É.: Construction of unified human antimicrobial and immunomodulatory peptide database and examination of antimicrobial and immunomodulatory peptides in Alzheimer's disease using network analysis of proteomics datasets.
Front. Genet. 12, 1-15, 2021.
DOI: <http://dx.doi.org/10.3389/fgene.2021.633050>
IF: 4.772
2. Scholtz, B., Vo Minh, D., Kiss, C., Tar, I., **Kumar, A.**, Tózsér, J., Csósz, É., Márton, I.: Examination of Oral Squamous Cell Carcinoma and Precancerous Lesions Using Proximity Extension Assay and Salivary RNA Quantification.
Biomedicines. 8 (12), 1-16, 2020.
DOI: <http://dx.doi.org/10.3390/biomedicines8120610>
IF: 6.081

List of other publications

3. Kalló, G., **Kumar, A.**, Tózsér, J., Csósz, É.: Chemical Barrier Proteins in Human Body Fluids.
Biomedicines. 10 (7), 1-26, 2022.
DOI: <http://dx.doi.org/10.3390/biomedicines10071472>
IF: 4.757 (2021)
4. Nokhoijav, E., Guba, A., **Kumar, A.**, Kunkli, B., Kalló, G., Káplár, M., Somodi, S., **Garai, I.**, Csutak, A., Tóth, N., Emri, M., Tózsér, J., Csósz, É.: Metabolomic Analysis of Serum and Tear Samples from Patients with Obesity and Type 2 Diabetes Mellitus.
Int. J. Mol. Sci. 23, 1-19, 2022.
DOI: <http://dx.doi.org/10.3390/ijms23094534>
IF: 6.208 (2021)





5. Linkner, T. R., Ambrus, V. A., Kunkli, B., Szojka, Z., Kalló, G., Csősz, É., **Kumar, A.**, Emri, M.,
Tőzsér, J., Mahdi, M.: Cellular Proteo-Transcriptomic Changes in the Immediate Early-Phase
of Lentiviral Transduction.
Microorganisms. 9 (11), 1-21, 2021.
DOI: <http://dx.doi.org/10.3390/microorganisms9112207>
IF: 4.926

Total IF of journals (all publications): 26,744

Total IF of journals (publications related to the dissertation): 10,853

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on
the basis of the Journal Citation Report (Impact Factor) database.

08 September, 2022



CONFERENCE PARTICIPATION

Oral presentations:

Ajneesh Kumar: *Examination of antimicrobial and immunomodulatory peptides in Alzheimer's disease using network analysis of proteomics datasets and the DEAMP Database*, Winter Symposium of the Molecular and Cellular Immune Biology Doctoral School, Debrecen, Hungary, Jan 9 – 10, 2020.

Ajneesh Kumar: *Examination of antimicrobial and immunomodulatory peptides in Alzheimer's disease using network analysis of proteomics datasets and the DEAMP Database*, Winter Symposium of the Molecular and Cellular Immune Biology Doctoral School, Debrecen, Hungary, Jan 7 – 8, 2021.

Ajneesh Kumar: *Construction of Unified Human Antimicrobial and Immunomodulatory Peptide Database and Examination of Antimicrobial and Immunomodulatory Peptides in Alzheimer's Disease Using Network Analysis of Proteomics Datasets*, “Learn from each other”, meeting of the Proteomics Division of the Hungarian Biochemical Society, online, May 19, 2021

Ajneesh Kumar: *Network analysis of atheroma and complicated lesions in human atherosclerosis*, Winter Symposium of the Molecular and Cellular Immune Biology Doctoral School, Debrecen, Hungary, Jan 6 – 7, 2022.

Poster Presentations:

Ajneesh Kumar, Eva Csoz: *In-Silico Construction of Homo Sapiens Antimicrobial and Immunomodulatory Peptide (AMP) Database*, Network Tools and Applications in Biology NETTAB/BCC2019, Salerno, Italy, Nov 11 – 14, 2019.

Ajneesh Kumar, Eva Csoz: *In-Silico Construction of Homo Sapiens Antimicrobial and Immunomodulatory Peptide (AMP) Database DEAMP*, Human Proteome Organization HUPO Connect 2020 conference, online, Oct 19 – 22, 2020.

Ajneesh Kumar, Gergő Kalló, László Potor, Zoltán Hendrik, Gábor Méhes, Csaba Tóth, Péter Gergely, József Tózsér, József Balla, Éva Csósz: *Proteomics analysis followed by network analysis reveals proteins and functions characteristic to atheroma and complicated lesions in human atherosclerosis*, PROTEOMIC FORUM 2022 XIV. Annual Congress of the European Proteomics, Leipzig, Germany, Apr 3 – 7, 2022.

Ajneesh Kumar, Gergő Kalló, László Potor, Zoltán Hendrik, Gábor Méhes, Csaba Tóth, Péter Gergely, László Prókai, József Tózsér, József Balla, Éva Csósz: *Network analysis of proteomics data in human atherosclerosis*, Annual Meeting of the Hungarian Biochemical Society, Pecs, Hungary, Aug 25 – 27, 2022.

KEYWORDS

APP, Alzheimer's disease, UDAMP database, network analysis, oral cancer, OSCC, saliva