Journal of Alzheimer's Disease xx (2022) x–xx
DOI 10.3233/JAD-220246
IOS Press

Salivary Proteomics Reveals Significant Changes in Relation to Alzheimer's Disease and Aging

Cristina Contini[a,1], Simone Serrao[a,1], Barbara Manconi[a,*], Alessandra Olianas[a], Federica Iavarone[b,c], Alessandra Bizzarro[c], Carlo Masullo[d], Massimo Castagnola[e], Irene Messana[f], Giacomo Diaz[g,2] and Tiziana Cabras[a,2]
[a]*Department of Life and Environmental Sciences, University of Cagliari, Cagliari, Italy*
[b]*Department of Basic Biotechnological Sciences, Intensive and Perioperative Clinics, Catholic University of the Sacred Heart, Rome, Italy*
[c]*Policlinico Universitario "A. Gemelli" Foundation – IRCCS, Rome, Italy*
[d]*Department of Neuroscience, Section Neurology, Catholic University of the Sacred Heart, Rome, Italy*
[e]*Proteomics laboratory, European Centre for Research on the Brain, "Santa Lucia" Foundation – IRCCS, Rome, Italy*
[f]*Institute of Chemical Sciences and Technologies "Giulio Natta", National Research Council, Rome, Italy*
[g]*Department of Biomedical Sciences University of Cagliari, Cagliari, Italy*

Accepted 1 July 2022
Pre-press 26 July 2022

Abstract.
Background: Aging is a risk factor for several pathologies as Alzheimer's disease (AD). Great interest exists, therefore, in discovering diagnostic biomarkers and indicators discriminating biological aging and health status. To this aim, omic investigations of biological matrices, as saliva, whose sampling is easy and non-invasive, offer great potential.
Objective: Investigate the salivary proteome through a statistical comparison of the proteomic data by several approaches to highlight quali-/quantitative variations associated specifically either to aging or to AD occurrence, and, thus, able to classify the subjects.
Methods: Salivary proteomic data of healthy controls under-70 (adults) and over-70 (elderly) years old, and over-70 AD patients, obtained by liquid chromatography/mass spectrometry, were analyzed by multiple Mann-Whitney test, Kendall correlation, and Random-Forest (RF) analysis.
Results: Almost all the investigated proteins/peptides significantly decreased in relation to aging in elderly subjects, with or without AD, in comparison with adults. AD subjects exhibited the highest levels of α-defensins, thymosin β4, cystatin B, S100A8 and A9. Correlation tests also highlighted age/disease associated differences. RF analysis individuated quali-/quantitative variations in 20 components, as oxidized S100A8 and S100A9, α-defensin 3, P-B peptide, able to classify with great accuracy the subjects into the three groups.
Conclusion: The findings demonstrated a strong change of the salivary protein profile in relation to the aging. Potential biomarkers candidates of AD were individuated in peptides/proteins involved in antimicrobial defense, innate immune system, inflammation, and in oxidative stress. RF analysis revealed the feasibility of the salivary proteome to discriminate groups of subjects based on age and health status.

Keywords: Aging, α-defensins, Alzheimer's disease, RF analysis, salivary proteome, S100A8, S100A9, thymosin β4

[1]These authors contributed equally to this work and share first authorship.

[2]These authors contributed equally to this work and share last authorship.

*Correspondence to: Prof. Barbara Manconi, Department of Life and Environmental Sciences, University of Cagliari, Monserrato Campus, ss 554, 09042 Monserrato, CA, Italy. Tel.: +39 0706754508; E-mail: bmanconi@unica.it.

INTRODUCTION

Aging is a risk factor for many pathologies, including neurodegeneration, cancer, osteoarthritis, and many others, so the availability of biomarkers of biological aging has been pointed as highly relevant to succeed in the early identification of patients at high age-related risk [1]. The potential of proteomic and metabolomic studies has recently gained attention to understand the molecular mechanisms influencing aging and longevity [1, 2], and in the definition of the aging clock [3]. Johnson et al. [4] performed a systematic review of 36 studies involving a total of 3,301 subjects aged 18–76 years old highlighting the upregulation of 23 proteins, e.g., vascular endothelial growth factor, pleiotrophin, and fibrinogen alpha, in the oldest subjects. Proteomic studies also highlighted a correlation between post-translational modifications and age or age-related diseases, in particular acetylation, oxidation, nitrosylation, and chlorination of several proteins [5]. Among the tissues and body fluids investigated for biomarker discovery by proteomics, saliva represents one of the most advantageous due to the painless, non-invasive, and safe collection [6]. Furthermore, human saliva includes both specific proteins of the oral cavity and proteins common to other tissues and body fluids. For this reason, the interest in its prognostic and diagnostic employment is increasing [7–9]. Human saliva composition varies between individuals depending on a multitude of factors, including sex, health status, circadian rhythms, habits, nutritional factors, and age [10]. Age-related changes in salivary proteome were highlighted by several studies focused on individuals from 180 days after birth to adulthood [11–15]. Anyway, to date, few studies on the changes of the salivary protein profile in individuals older than 60 years have been performed, and in some cases with opposite results. Significant decreased levels related to age of histatins [16], mucins 1–2 [17, 18], lactoferrin and secretory immunoglobulin A [19], and peroxidase activity [20] were highlighted by different methodologies. Conversely, Nagler et al. [21] found increased levels of secretory immunoglobulin A, lysozyme, amylase, and albumin. High levels of cystatins A and B and small proline-rich protein 3 in old totally edentulous subjects have been determined by a mass spectrometry-based approach [22], as well as lower levels of α-defensins [23]. In a recent review, telomere length, DNA methylation, mucin 1 and protein carbonylation have been indicated as the main biological hallmarks aging-related measurable in saliva [24]. Alzheimer's disease (AD) represents the major form of dementia in people over 65 years old [25]. AD diagnosis is based on presence and combination of three different parameters: clinical and neuropsychological evidence of cognitive impairment, levels of amyloid-β (Aβ) and tau proteins measurable in the brain either by positron emission tomography or by analysis of the cerebrospinal fluid [26]. To overcome the invasiveness of cerebrospinal fluid tests, peripheral biomarkers on other tissues and biofluids, like blood cells, plasma, eyes, saliva, and skin, have been studied [27, 28]. Recently, our group revealed significant changes in the salivary proteome of AD patients in comparison with a healthy control group by high-performance liquid chromatography (HPLC) coupled to electrospray ionization ion trap mass spectrometry (ESI-IT-MS) [29]. In the present study, we performed a statistical investigation based on exact Mann-Whitney test, Random Forest (RF), Multidimensional Scaling and Hierarchical Cluster Analysis, to compare the salivary protein profiles, detectable by HPLC-ESI-IT-MS, of healthy adult (under 70 years old) and elderly (over 70 years old) subjects, both healthy and affected by AD. The aim of the study was, therefore, to individuate salivary biomarkers related to aging or AD and useful to classify accurately the subjects.

MATERIALS AND METHODS

Demographic and clinical features of subjects included into the study

We used proteomic salivary data of thirty-five adult healthy controls (aHC, 18 females, 17 males, 46 ± 12 mean age $\pm$ standard deviation) selected among those included in the study of Serrao et al. [30]. Proteomic data of the elderly healthy controls (eHC, 18 females and 16 males, 78 ± 5) and AD patients (23 females and 12 males; 80 ± 6), reported in the study of Contini et al. [29], were also utilized. Table 1 reports demographic features of all the subjects included. The informed consent process for sample's collection agreed with the latest stipulations established by the Declaration of Helsinki. The study approval was obtained by the formal ethical committees of the Catholic University of Rome and of the University of Cagliari. No subjects included were affected by any major oral disease (periodontitis, caries, or dry mouth), moreover, they had not history of radiotherapy or chemotherapy and were carefully selected as non-smokers. The elderly subjects enrolled as

Table 1
Demographic data of aHC, eHC, and AD patients involved in the study

aHC	Sex and age	eHC	Sex and age	AD	Sex and age
#1	F, 44	#1	M, 70	#1	M, 82
#2	M, 53	#2	M, 85	#2	F, 80
#3	F, 60	#3	F, 84	#3	M, 85
#4	F, 43	#4	M, 82	#5	F, 63
#5	F, 38	#5	F, 81	#6	M, 78
#6	F, 39	#6	F, 81	#7	M, 85
#7	F, 23	#7	F, 79	#8	F, 81
#8	F, 49	#8	M, 74	#9	F, 78
#9	M, 55	#9	M, 71	#10	M, 85
#10	M, 54	#10	M, 78	#11	F, 80
#11	M, 36	#11	M, 76	#13	F, 79
#12	M, 24	#12	F, 77	#14	F, 82
#13	M, 27	#13	M, 74	#16	F, 83
#14	M, 53	#14	M, 87	#17	F, 63
#15	M, 53	#15	M, 73	#18	F, 80
#16	M, 58	#16	F, 81	#19	F, 80
#17	F, 43	#17	F, 82	#20	M, 87
#18	M, 45	#18	F, 72	#21	M, 81
#19	F, 64	#19	F, 86	#22	M, 87
#20	M, 38	#20	F, 73	#23	F, 75
#21	F, 52	#22	F, 78	#24	F, 75
#22	M, 36	#23	F, 79	#25	F, 83
#23	F, 57	#25	F, 75	#26	F, 84
#24	F, 60	#24	F, 78	#27	F, 81
#25	F, 59	#26	M, 75	#28	F, 84
#26	F, 40	#27	F, 89	#29	F, 92
#27	F, 27	#28	M, 78	#30	M, 86
#28	M, 33	#29	M, 73	#31	M, 77
#29	F, 54	#30	F, 76	#32	F, 88
#30	F, 67	#31	F, 81	#33	F, 81
#31	M, 46	#32	M, 81	#34	F, 77
#32	M, 55	#33	M, 84	#35	M, 87
#33	M, 56	#34	F, 72	#36	M, 84
#34	F, 30	#35	M, 80	#37	F, 77
#35	M, 62			#38	F, 78

controls suffered from common age-related illness, such as hypertension, and were treated with standard drugs. However, none control subject used antidepressants or anticholinergic drugs.

Among the elderly subjects with or without AD, 50% carried a dental prosthesis. The diagnosis of AD, made according to standardized criteria [25], classified thirteen patients as moderate AD and the remaining twenty-two as mild AD.

Proteomic data

For sample collection procedure and treatment, and the experimental conditions of the HPLC-low resolution-ESI-IT-MS, and HPLC-high resolution-MS/MS (by LTQ-Orbitrap Elite or LTQ-Orbitrap XL instruments) analysis, which were applied for quantification and identification of peptides and proteins, see references [29, 30]. Briefly, the label-free quantitation of peptides and proteins was performed by measuring the area of the eXtracted Ion Current (XIC) peaks revealed by HPLC-low resolution-MS analysis. XIC peaks were generated by extracting ion current produced by specific m/z ions selected for each protein/peptide, the following parameters have been applied: baseline window 15, area noise factor 50, peak noise factor 50, peak height 15%, and tailing factor 1.5. The estimated percentage error of the XIC analysis was < 8%. Area of the XIC peaks, expressed by arbitrary units, is proportional to the protein concentration, and, under constant analytical conditions, it allows performing relative quantification of the same protein in different samples and quantify an indefinite number of proteins/peptides in a unique analysis [31, 32]. Eventual dilution errors occurring during sample collection were adjusted by correcting XIC peak areas of each peptide/protein with the XIC peak area of the leu-enkephalin used as internal standard in the aqueous 0.1% trifluoroacetic solution added to whole saliva in ratio 1:1(vol/vol) at the collection time [29]. The following equation was applied: Corrected Area of protein = Measured Area of protein * (Expected Area of Leu-enkephalin 25 μM/Measured Area of Leu-enkephalin). Supplementary Table 1 reports UniProt-KB codes, elution times, experimental and theoretical average mass values, multiply-charged ions used for the XIC procedure, and the detected post-translational modifications of the 61 peptides and proteins considered in this study. Moreover, we determined, in duplicate, the total protein concentration in μg/μl in the acid soluble fractions of each salivary sample by the bicinconinic acid assay (MicroBCA™ protein assay kit, 0.5–20 μg/mL, Thermo Fisher Scientific). The total protein concentration of every single sample was used to normalize the XIC peak areas of each peptide/protein detected in that sample as it follows: the value of the XIC peak area corrected with Leu-enkephalin was divided by the total protein concentration [29].

Statistical analysis

Non-parametric tests were applied to analyze the difference of total protein concentration among the three groups (Mann-Whitney and Kruskal-Wallis followed by Dunn's *post-hoc* test) using GraphPad Prism software (version 5.0). MS data were analyzed using three statistical methods: 1) multiple exact Mann-Whitney tests [33] to identify

proteins/peptides with different abundance between pair of groups, 2) multiple Kendall correlations [34] to identify correlated proteins/peptides within groups, and 3) RF analysis [35] to provide a classification of subjects into different groups. Statistical analysis considered both single proteoforms and the sum of proteoforms of the same protein, to simplify we call both "components" in the text. The number of components examined in this study is 76. Distribution of XIC peak areas of every protein/peptide showed a considerable deviation from normality using Kolmogorov-Smirnov test and several goodness-of-fit tests (Shapiro-Wilk, Anderson-Darling, Lilliefors, with p-values < 0.0001 in almost all tests, data not shown). Thus, non-parametric exact Mann-Whitney tests (between groups) and Kendall correlations (within groups) were adopted. Significant p-values of simultaneous multiple tests were verified by the Benjamini-Hochberg procedure [36] to keep a cumulative false discovery ratio among all the tests lower than 5%. Multidimensional scaling analysis was applied to Kendall correlations to obtain a dimensionally reduced diagram of co-expressed proteins. The classification of subjects was obtained using RF analysis. Algorithm parameters, such as the number of trees to grow and the number of features randomly sampled for each split, were preliminarily tuned to minimize the classification error. RF was applied to three data set combinations: 1) aHC and eHC, characterized by differences in age; 2) eHC and AD, characterized by differences in normal or pathological conditions; 3) aHC, eHC, and AD, characterized by differences both in age and normal or pathological conditions. Classification accuracy was calculated as the proportion of correct assessments (both true positive and true negative) to the total number of assessments. The Boruta method [37] was used to individuate a subset of proteins and peptides to use for an accurate RF analysis. Then, the significance of each protein/peptide for classification was expressed by a score indicating the mean decrease of the Gini index (MDG), that is a measure of impurity. For a single decision tree, the Gini index ranges from 0 (no impurities, 100% correct classification) to 1 (total impurity, elements are randomly distributed across classes). MDG averages the decrease in impurity for each tree of the whole 'forest', produced by each protein. Dimensionally reduced diagrams of RF classifications were obtained by multidimensional scaling and hierarchical cluster analysis using the RF proximity values (the normalized frequency of trees that contain the two samples in the same end node). For the hierarchical cluster analysis, we used the Ward's agglomerative method and 1-proximity as distance between each pair of samples. Multidimensional scaling analysis was computed using the singular value decomposition method, which ensures a matrix factorization numerically accurate even in the presence of a high degree of multicollinearity (i.e., multiple correlations). Multivariate analyses were made using the software "R" (RCoreTeam. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. https://www.R-project.org/).

Functional pathway analysis

Functional pathway analysis of the most discriminating components evaluated by RF in the classification of elderly and adult subjects, as well as elderly and AD subjects, was performed via ClueGO plugin (v.2.5.8) from Cytoscape software (v. 3.9.1) [38] by using the Kyoto Encyclopedia of Genes and Genomes and Reactome Pathway database. Enrichment right-sided hypergeometric test and Bonferroni step down statistical options were included and only pathways with $p \leq 0.05$ were accepted. Minimum and maximum tree interval values were 3–7, evidence code decision tree was set at "all", minimum number of 2 genes and 4% of genes selected for GO terms, and the kappa score set at 0.4.

RESULTS

Figure 1 represents the workflow applied in this study in which we analyzed, by different statistical approaches, the abundances of 61 selected salivary proteins and peptides measured in our previous studies [29, 30]. The XIC peak areas are the quantitative data obtained by MS analysis used to compare the abundances of every peptide and protein among the three groups. Proteins and peptides of our interest belonged to the following families: acidic proline-rich proteins (aPRPs), statherin, histatin (Hst) 1, 3, 5 and 6, P-B peptide, cystatins A, B, C, D, and salivary (S-type), α-defensins 1–4, thymosin β4 (Tβ4), antileukoproteinase (Secretory Leukocyte Protease Inhibitor, SLPI), S100A7, A8, A9, and A12 proteins (Supplementary Table 1), including modified proteoforms generated by phosphorylation, proteolysis, N-terminal acetylation, methionine or tryptophan oxidation, and cysteine oxidation (formation of disulphide bridges, glutathionylation, cysteinylation, and nitrosylation). In the present study we implemented

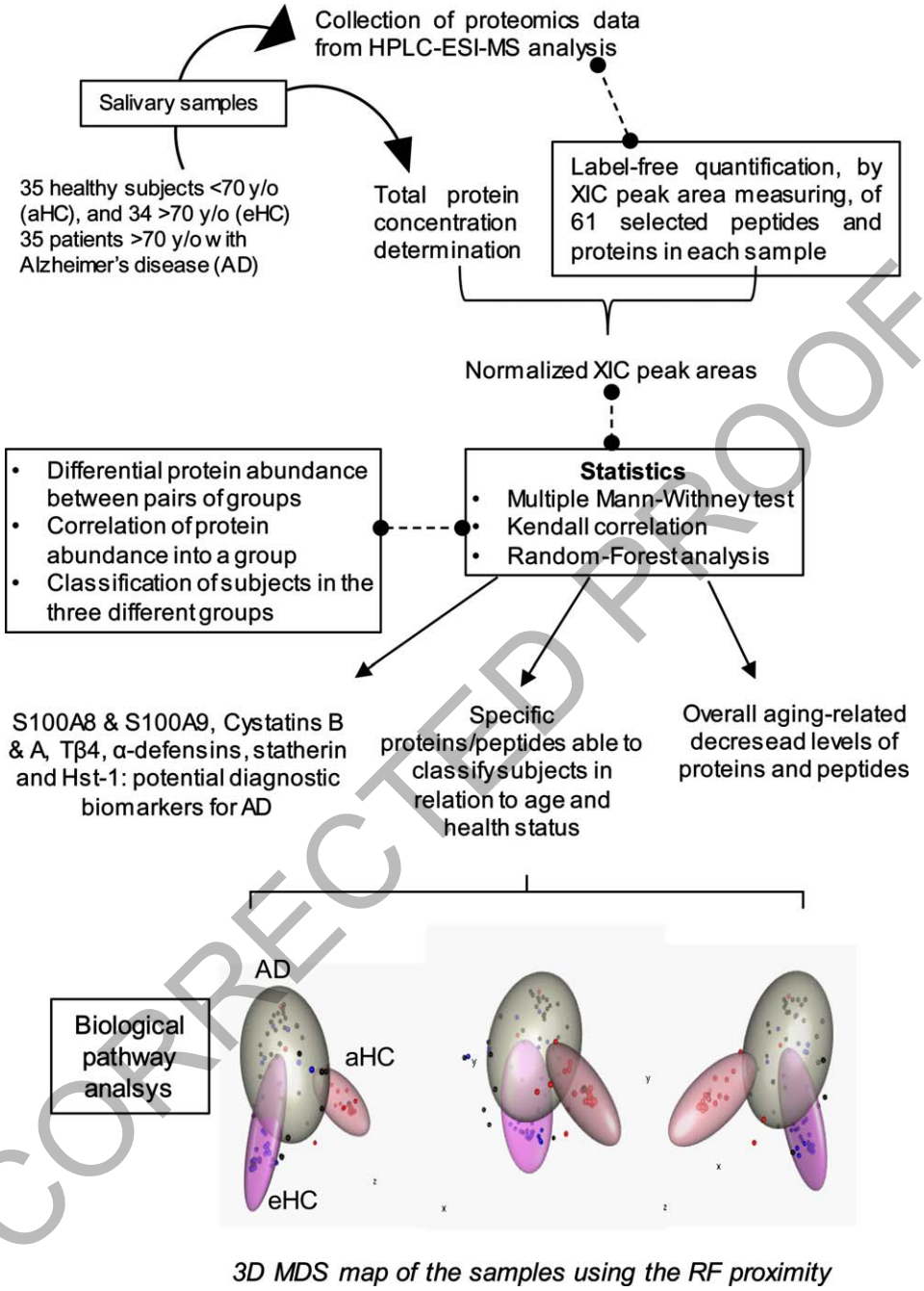

Fig. 1. Workflow drafting the different steps performed in the data analysis of this study.

the number of investigated proteoforms adding cystatin S2 mono-oxidized, S100A12, S100A9 long glutathionylated (L-SSG) and its phosphorylated and oxidized proteoforms, which were not included in the two previous studies [29, 30]. The typical total ion current chromatographic profiles of the acidic soluble fractions of whole saliva from aHC (Fig. 2A), eHC (Fig. 2B), and AD subjects (Fig. 2C), analyzed by HPLC-ESI-MS, is shown with the elution ranges of the protein families considered.

All the XIC peak areas, utilized for the several statistical analyses, have been normalized on the

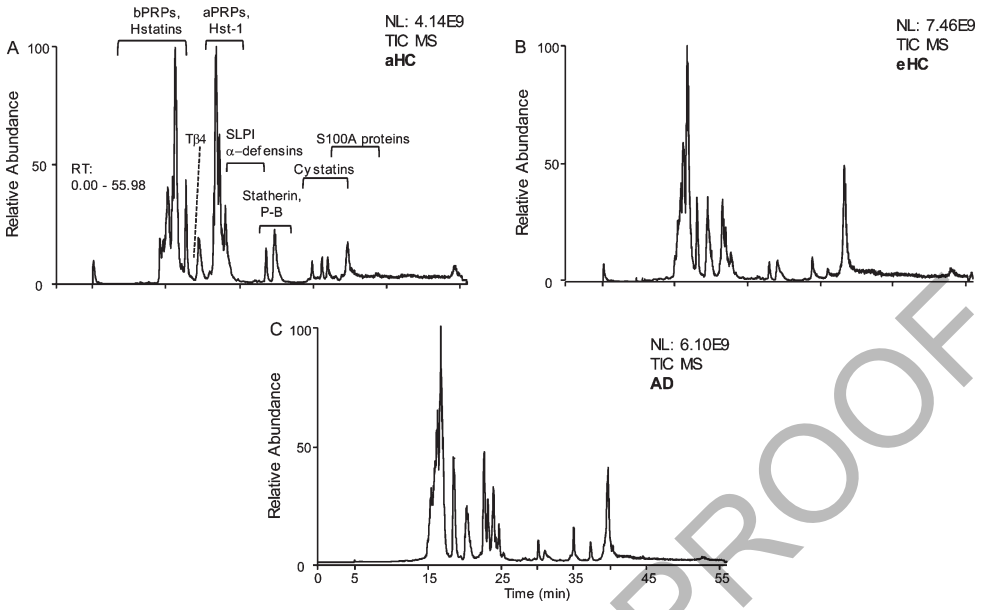

Fig. 2. Total ion current chromatographic profile of acidic-soluble fraction of saliva from aHC (A), eHC (B), and AD patient (C) obtained by RP-HPLC-ESI-low-resolution MS analysis.

total protein concentration measured in every sample. This parameter resulted to be significantly higher in eHC compared to aHC group (p-value < 0.01) and to AD patients (p-value < 0.05) (Supplementary Figure 1). Preliminary statistical analysis performed on un-normalized data produced results analogous to those ones obtained after normalization. To improve the statistical validity and accuracy of our study we preferred to use normalized XIC peak areas.

Statistical analysis of the proteins/peptides abundances by multiple Mann-Whitney tests

The XIC peak areas, the frequencies, and the results of the Mann-Whitney tests of all the components measured in aHC, eHC, and AD patients, are shown in Table 2. In the case of aPRPs, statherin, P-B peptide, histatins, cystatin A, B, S1, S2, SN, S100A8, S100A9, and α-defensins 1–4 the sum of the XIC peak areas of all their proteoforms (Supplementary Table 1) was also considered, and thus Table 2 reports overall 76 components including 61 proteins/peptides and 15 proteoform sums. Results of Mann-Whitney tests are also graphically shown in Supplementary Figure 2.

The comparison between aHC and eHC (Table 2 and Supplementary Figure 2A) showed a significant decrease of the abundance of 69 out of the 76 analyzed components (91%) in eHC group. The lowest abundance concerned the P-B peptide and particularly its fragment des1–7, the sum of PRP-3 and PRP-1 proteoforms and the oxidized proteoforms of S-type cystatins (indicated as "ox"). Some components were often not detectable in the two control groups, probably since their abundance was under the sensitivity of our MS apparatus, such as S100A7, the form of PRP3 missing for Arginine at position 106 ($desR_{106}$), cystatin B cysteinylated (SSC), Hst-3, SLPI, the non-phosphorylated (0P) proteoforms of PRP-1 and Hst-1. The comparison between eHC and AD (Table 2 and Supplementary Figure 2C) showed significant higher abundance of 37 out of the 76 analyzed components (49%) in AD group. These were statherin and its proteoforms des1–9, des1–13 and, as consequence, the sum of all proteoforms; Hst-1, both phosphorylated and non-phosphorylated; P-C peptide, the several proteoforms of S100A8, S100A9, cystatins A, B, and SA, the oxidized form of cystatins S2 and SN, α-defensins 1–4 and Tβ4. AD patients with respect to aHC (Table 2 and Supplementary Figure 2B) showed significant lower abundance of cystatin A, S-type cystatins, Hst-5 and Hst-6, PRP-1, PRP-3, except for its non-phosphorylated form, P-C peptide, statherin and its truncated forms $desT_{42}F_{43}$ and $desD_1$, and P-B peptide. Only the sum of the oxidized forms of S100A8 showed an opposite trend, being increased in AD group with respect to aHC.

Table 2
XIC peak areas (median and interquartile range) normalized on total protein concentration, and frequencies (F) of the salivary proteins/peptides in aHC, AD patients and eHC

	aHC				AD				eHC				aHC versus eHC		aHC versus AD		eHC versus AD	
	XIC Peak Area			F	XIC Peak Area			F	XIC Peak Area			F						
Components	25th perc	median	75th perc		25th perc	median	75th perc		25th perc	median	75th perc		*p*	change	*p*	change	*p*	change
S100A12	1.2E05	1.9E05	3.9E05	8/35	1.0E05	1.7E05	3.1E05	3/35	8.1E04	9.7E04	1.6E05	5/34	<0.001	↓eHC	ns		ns	
S100A8	1.4E05	2.3E05	7.8E07	11/35	1.1E05	2.0E05	1.3E08	9/35	7.4E04	9.4E04	1.5E05	2/34	<0.00001	↓eHC	ns		<0.01	↑AD
S100A8-Hyperox	1.0E05	1.6E05	2.5E05	2/35	1.1E05	1.8E05	3.3E05	5/35	7.7E04	9.5E04	1.5E05	3/34	<0.01	↓eHC	ns		<0.01	↑AD
S100A8-SNO	1.1E05	1.6E05	2.7E05	2/35	1.3E05	2.8E05	6.0E07	12/35	7.4E04	9.4E04	1.4E05	0/34	<0.0001	↓eHC	ns		<0.00001	↑AD
Sum_S100A8-ox	2.2E05	3.3E05	5.4E05	5/35	3.5E05	1.1E06	1.1E08	18/35	1.5E05	1.9E05	3.0E05	3/34	<0.001	↓eHC	<0.01	↑AD	<0.00001	↑AD
Sum_S100A8	4.1E05	6.8E05	9.4E07	14/35	5.8E05	1.1E08	2.208	24/35	2.3E05	2.9E05	4.6E05	5/34	<0.0001	↓eHC	ns		<0.00001	↑AD
S100A7	1.4E05	2.0E05	1.2E07	10/35	1.1E05	1.9E05	9.2E06	9/35	9.0E04	1.5E05	3.5E07	11/34	ns		ns		ns	
S100A9S	1.7E05	4.0E05	4.4E08	17/35	2.0E05	9.3E07	2.2E08	20/35	8.8E04	1.5E05	5.9E07	15/34	<0.01	↓eHC	ns		<0.001	↑AD
S100A9S-1P	1.4E05	1.9E05	5.2E07	9/35	1.2E05	3.0E05	9.3E07	13/35	7.7E04	9.7E04	1.6E05	4/34	<0.001	↓eHC	ns		<0.01	↑AD
S100A9S-ox	1.7E05	4.7E07	2.1E08	19/35	1.2E05	3.0E05	1.1E08	14/35	8.1E04	1.1E05	1.7E07	10/34	<0.0001	↓eHC	ns		<0.05	↑AD
S100A9S-1Pox	1.0E05	1.7E05	3.2E05	5/35	1.1E05	2.0E05	3.5E05	7/35	7.7E04	9.5E04	1.5E05	1/34	<0.001	↓eHC	ns		<0.01	↑AD
Sum_S100A9S_and_S-1P	3.3E05	8.0E05	5.6E08	18/35	4.4E05	1.0E08	3.2E08	23/35	1.8E05	3.1E05	7.1E07	15/34	<0.01	↓eHC	ns		<0.001	↑AD
Sum_S100A9S-1P_and_S-1Pox	2.9E05	5.4E05	1.5E08	14/35	3.9E05	4.8E07	1.0E08	20/35	1.5E05	1.9E05	3.2E05	4/34	<0.0001	↓eHC	ns		<0.001	↑AD
Sum_S100A9S-ox_and_S-1Pox	3.3E05	7.1E07	2.6E08	19/35	2.5E05	6.1E05	1.5E08	15/35	1.6E05	2.1E05	2.9E07	10/34	<0.0001	↓eHC	ns		<0.05	↑AD
Sum_S100A9S	7.1E07	3.0E08	7.6E08	32/35	1.1E08	2.0E08	3.5E08	33/35	3.5E05	6.2E05	9.6E07	16/34	<0.00001	↓eHC	ns		<0.00001	↑AD
Sum_S100A9L-SSG	5.8E05	1.5E06	2.4E08	20/35	9.0E05	4.9E07	1.8E08	26/35	3.5E05	6.2E05	5.0E07	16/34	<0.05	↓eHC	ns		<0.01	↑AD
Cystatin_A	1.7E08	3.3E08	5.2E08	33/35	8.8E07	1.5E08	2.6E08	32/35	4.0E07	6.5E07	1.6E08	29/35	<0.00001	↓eHC	<0.001	↓AD	<0.01	↑AD
Cyst_A-NAcetyl	2.4E07	5.9E07	1.1E08	29/35	1.8E06	2.2E07	3.6E07	26/35	1.7E05	1.1E07	2.4E07	23/35	<0.0001	↓eHC	<0.001	↓AD	ns	
Sum_cyst_A	2.1E08	3.9E08	5.6E08	34/35	9.4E07	1.7E08	3.0E08	32/35	4.7E07	7.9E07	1.9E08	30/34	<0.00001	↓eHC	<0.001	↓AD	<0.05	↑AD
Cystatin_B-SSG	4.3E07	7.2E07	1.3E08	29/35	2.1E07	4.1E07	1.2E08	30/35	1.1E07	1.9E07	3.1E07	27/34	<0.0001	↓eHC	ns		<0.01	↑AD
Cyst_B-SSC	1.5E05	1.5E07	4.1E07	18/35	2.6E05	1.5E07	2.8E07	21/35	1.1E05	4.0E06	1.3E07	18/34	ns		ns		<0.05	↑AD
Cyst_B-S-S_dimer	8.9E06	5.5E07	1.2E08	26/35	1.7E07	5.4E07	7.9E07	31/35	1.5E05	1.7E07	3.3E07	22/34	<0.01	↓eHC	ns		<0.001	↑AD
Sum_cyst_B	7.5E07	1.5E08	3.0E08	30/35	4.5E07	1.1E08	2.4E08	33/35	1.6E07	4.5E07	7.3E07	27/34	<0.0001	↓eHC	ns		<0.01	↑AD
Cystatin_C	1.6E05	3.8E05	6.8E07	16/35	1.2E05	1.8E05	3.2E05	6/35	8.1E04	9.7E04	2.4E05	6/34	<0.00001	↓eHC	<0.01	↓AD	ns	
Cystatin_D_R_{26}des1-5	1.7E05	4.0E07	1.7E08	20/35	1.3E05	2.8E05	5.3E07	13/35	9.0E04	3.6E05	5.8E07	16/34	<0.05	↓eHC	ns		ns	
Cystatin_S	1.6E05	4.0E05	1.4E08	17/35	1.2E05	1.9E05	3.2E05	5/35	8.1E04	1.2E05	2.1E07	9/34	<0.001	↓eHC	<0.01	↓AD	ns	
Cyst_S1	2.1E08	6.8E08	1.3E09	33/35	1.1E08	3.2E08	8.1E08	29/35	1.0E08	3.3E08	6.6E08	30/34	<0.05	↓eHC	ns		ns	
Cyst_S2	3.6E07	2.3E08	4.0E08	31/35	2.2E05	8.0E07	2.6E08	20/35	3.4E07	8.4E07	2.1E08	26/34	<0.01	↓eHC	<0.01	↓AD	ns	
Cyst_SN	5.7E08	1.6E09	2.4E09	34/35	1.6E08	4.8E08	1.6E09	30/35	2.6E08	5.4E08	8.8E08	29/34	<0.001	↓eHC	<0.05	↓AD	ns	
Cyst_SA	2.5E05	2.0E08	4.3E08	24/35	1.2E05	2.1E05	5.9E05	8/35	7.7E04	9.7E04	2.4E05	7/34	<0.00001	↓eHC	<0.01	↓AD	ns	
Cyst_S1-ox	3.3E05	1.9E08	4.9E08	24/35	1.1E05	2.0E05	3.5E07	10/35	7.7E04	1.1E05	2.0E05	7/34	<0.00001	↓eHC	<0.0001	↓AD	ns	
Cyst_S2-ox	1.9E05	2.3E07	1.4E08	18/35	1.0E05	1.7E05	2.9E05	3/35	7.4E04	9.4E04	1.5E05	1/34	<0.00001	↓eHC	<0.0001	↓AD	<0.05	↑AD
Cyst_SN-ox	8.0E05	2.3E08	5.7E08	25/35	1.2E05	2.6E05	5.2E07	12/35	7.7E04	1.1E05	1.4E07	9/34	<0.00001	↓eHC	<0.0001	↓AD	<0.05	↑AD
Sum_cyst_S1	4.4E08	8.0E08	1.8E09	33/35	1.5E08	3.4E08	8.1E08	31/35	1.2E08	3.3E08	6.6E08	30/34	<0.001	↓eHC	<0.01	↓AD	ns	
Sum_cyst_S2	1.4E08	2.5E08	6.8E08	32/35	4.4E05	9.6E07	2.6E08	21/35	3.5E07	8.4E07	2.1E08	26/34	<0.0001	↓eHC	<0.001	↓AD	ns	
Sum_cyst_SN	8.1E08	1.9E09	2.8E09	34/35	1.7E08	4.8E08	1.6E09	32/35	2.6E08	5.4E08	8.8E08	31/34	<0.00001	↓eHC	<0.01	↓AD	ns	
Hst-1-1P	3.7E05	2.4E08	5.7E08	24/35	3.2E07	1.3E08	2.3E08	33/35	1.5E07	6.5E07	1.1E08	29/34	<0.01	↓eHC	ns		<0.05	↑AD
Hst-1-0P	1.4E05	2.7E05	6.4E07	14/35	2.1E05	2.7E07	4.2E07	24/35	1.1E05	3.4E06	2.1E07	17/34	ns		ns		<0.05	↑AD

(Continued)

Table 2
(Continued)

	aHC				AD				eHC				aHC versus eHC		aHC versus AD		eHC versus AD	
	XIC Peak Area			F	XIC Peak Area			F	XIC Peak Area			F						
Components	25th perc	median	75th perc		25th perc	median	75th perc		25th perc	median	75th perc		*p*	change	*p*	change	*p*	change
Sum_Hst-1	7.5E05	2.7E08	6.8E08	25/35	5.1E07	1.4E08	2.9E08	33/35	1.6E07	6.9E07	1.4E08	28/34	<0.01	↓eHC	ns		<0.05	↑AD
Hst-3_1-24_Hst-5	1.9E08	4.4E08	1.1E09	30/35	1.8E05	9.2E07	1.5E08	20/35	9.6E04	7.9E06	9.1E07	17/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
Hst-3_1-25_Hst-6	1.7E05	3.5E07	2.1E08	18/35	1.2E05	2.0E05	2.3E06	8/35	9.1E04	1.5E05	3.9E07	10/34	<0.001	↓eHC	<0.01	↓AD	ns	
Hst-3	1.5E05	3.8E05	2.1E08	17/35	1.5E05	6.3E05	5.8E07	17/35	1.1E05	1.6E07	5.6E07	18/34	ns		ns		ns	
Sum_Hst-3	2.8E08	5.9E08	1.5E09	33/35	6.1E05	1.4E08	2.7E08	25/35	4.0E05	3.8E07	2.1E08	22/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
Tβ4	1.6E05	6.6E07	1.7E08	22/35	3.3E05	3.5E07	8.0E07	25/35	9.0E04	4.8E06	5.0E07	17/34	<0.001	↓eHC	ns		<0.01	↑AD
α-defensin 1	5.1E07	1.8E08	3.2E08	32/35	6.2E07	1.1E08	2.3E08	32/35	2.7E06	3.4E07	7.5E07	25/34	<0.0001	↓eHC	ns		<0.0001	↑AD
α-defensin 2	6.4E07	1.3E08	2.3E08	30/35	3.8E07	1.1E08	1.6E08	30/35	1.3E07	2.6E07	5.5E07	27/34	<0.0001	↓eHC	ns		<0.001	↑AD
α-defensin 3	2.5E05	4.1E07	1.3E08	24/35	4.6E05	5.4E07	1.0E08	25/35	1.3E05	2.9E06	1.9E07	14/34	<0.001	↓eHC	ns		<0.0001	↑AD
α-defensin 4	1.6E05	3.1E07	6.1E07	20/35	2.1E05	1.3E07	3.4E07	18/35	8.3E04	1.2E05	8.4E06	11/34	<0.0001	↓eHC	ns		<0.001	↑AD
Sum_ α -defensins	2.0E08	4.8E08	8.5E08	34/35	1.2E08	2.9E08	4.7E08	34/35	2.4E07	6.5E07	1.8E08	27/34	<0.00001	↓eHC	ns		<0.0001	↑AD
PRP-1-2P	6.7E09	1.0E10	2.0E10	35/35	1.4E09	3.2E09	5.6E09	34/35	1.1E09	2.5E09	4.6E09	33/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
PRP-1-1P	6.0E08	1.3E09	2.1E09	34/35	1.6E08	3.6E08	7.5E08	34/35	8.0E07	2.9E08	5.5E08	31/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
PRP-1-0P	1.4E05	1.1E07	1.1E08	17/35	1.2E05	1.9E05	4.4E05	7/35	9.4E04	7.4E06	3.8E07	17/34	ns		<0.05	↓AD	ns	
PRP-1-3P	6.4E07	2.2E08	3.4E08	33/35	1.2E05	2.1E05	5.9E05	8/35	1.2E05	4.0E05	4.0E07	15/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
Sum_PRP-1	8.2E09	1.2E10	2.3E10	35/35	1.8E09	3.7E09	6.1E09	35/35	1.3E09	2.9E09	5.2E09	33/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
PRP-3-2P	1.8E09	3.4E09	7.6E09	35/35	5.2E08	9.2E08	2.0E09	35/35	2.9E08	8.0E08	1.3E09	33/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
PRP-3-1P	2.5E08	5.2E08	8.9E08	35/35	1.1E08	1.5E08	2.9E08	35/35	3.7E07	1.5E08	1.9E08	31/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
PRP-3-0P	1.1E05	1.9E05	3.0E07	12/35	1.3E05	2.1E05	9.9E06	10/35	8.8E04	1.3E05	2.3E06	9/34	<0.05	↓eHC	ns		ns	
PRP-3-2P_desR$_{106}$	1.2E08	2.6E08	8.4E08	30/35	6.6E07	1.4E08	2.5E08	31/35	2.6E07	2.0E08	3.2E08	27/34	ns		<0.01	↓AD	ns	
Sum_PRP-3	2.5E09	4.7E09	9.6E09	35/35	8.2E08	1.4E09	2.4E09	35/35	5.8E08	1.1E09	1.8E09	33/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
P-C_peptide	1.1E09	2.3E09	4.3E09	35/35	4.7E08	8.4E08	1.3E09	34/35	2.0E08	4.9E08	8.2E08	34/35	<0.00001	↓eHC	<0.0001	↓AD	<0.01	↑AD
Statherin-2P	9.7E08	1.6E09	2.7E09	35/35	3.9E08	6.5E08	1.1E09	35/35	2.5E08	4.5E08	6.0E08	33/34	<0.00001	↓eHC	<0.001	↓AD	<0.05	↑AD
Stath-1P	3.3E05	3.0E07	4.5E07	25/35	4.4E06	1.3E07	3.2E07	27/35	2.8E05	8.5E06	1.5E07	25/35	<0.01	↓eHC	ns		ns	
Stath_desF$_{43}$	9.6E07	1.9E08	4.2E08	34/35	6.5E07	1.1E08	2.1E08	34/35	3.6E07	7.9E07	1.7E08	33/34	<0.01	↓eHC	ns		ns	
Stath_desT$_{42}$F$_{43}$	2.7E07	5.5E07	1.1E08	34/35	1.3E07	2.5E07	5.3E07	33/35	9.5E06	1.8E07	4.0E07	30/34	<0.001	↓eHC	<0.01	↓AD	ns	
Stath_desD$_{1}$	4.8E07	9.7E07	1.5E08	35/35	2.9E05	1.9E07	6.1E07	24/35	9.7E06	2.9E07	7.0E07	32/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
Stath_des1-9	3.4E07	7.8E07	1.5E08	30/35	4.6E06	5.8E07	8.7E07	26/35	2.8E05	1.8E07	3.4E07	24/34	<0.0001	↓eHC	ns		<0.01	↑AD
Stath_des1-10	2.7E07	5.0E07	8.8E07	32/35	4.9E05	2.9E07	4.4E07	25/35	7.2E06	1.6E07	3.0E07	27/34	<0.00001	↓eHC	<0.01	↓AD	ns	
Stath_des1-13	1.5E07	3.5E07	5.9E07	33/35	1.3E05	1.8E07	2.7E07	23/35	9.3E04	3.5E05	1.4E07	16/34	<0.00001	↓eHC	<0.001	↓AD	<0.01	↑AD
Sum_statherin	1.5E09	2.6E09	3.7E09	35/35	5.6E08	1.0E09	1.7E09	35/35	4.3E08	7.1E08	9.6E08	34/34	<0.00001	↓eHC	<0.001	↓AD	<0.05	↑AD
P-B_peptide	1.4E09	3.0E09	3.9E09	35/35	3.7E08	5.5E08	9.8E08	35/35	3.0E08	5.0E08	9.8E08	33/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
P-B_des1-5	6.8E07	1.3E08	4.4E08	32/35	2.2E07	4.5E07	9.8E07	34/35	2.8E07	5.8E07	1.6E08	34/34	<0.01	↓eHC	<0.001	↓AD	ns	
P-B_des1-7	3.0E08	4.8E08	7.0E08	34/35	4.9E07	9.1E07	1.4E08	34/35	4.2E07	7.1E07	1.2E08	34/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
P-B_des1-4	5.9E07	1.1E08	2.7E08	30/35	4.8E05	3.8E07	8.6E07	25/35	2.2E07	5.1E07	1.0E08	30/34	<0.01	↓eHC	<0.001	↓AD	ns	
P-B_des1-12	5.0E07	9.4E07	1.7E08	34/35	2.3E07	5.1E07	8.0E07	34/35	2.3E07	5.6E07	8.7E07	31/34	<0.05	↓eHC	<0.01	↓AD	ns	
Sum_P-B_peptide	2.2E09	4.2E09	5.2E09	35/35	5.7E08	9.4E08	1.4E09	35/35	4.8E08	8.3E08	1.2E09	35/35	<0.00001	↓eHC	<0.00001	↓AD	ns	
SLPI	1.1E05	1.6E05	2.7E05	5/35	1.4E05	3.2E05	1.6E07	14/35	8.3E04	1.9E05	8.7E06	14/34	ns		ns		ns	

On the right, results of Mann-Whitney exact tests for group comparisons, with a false discovery rate < 5%. The range of color tones from yellow to red denotes the magnitude of significant *p*-values. Color tones are continuous and more accurate over significance thresholds.

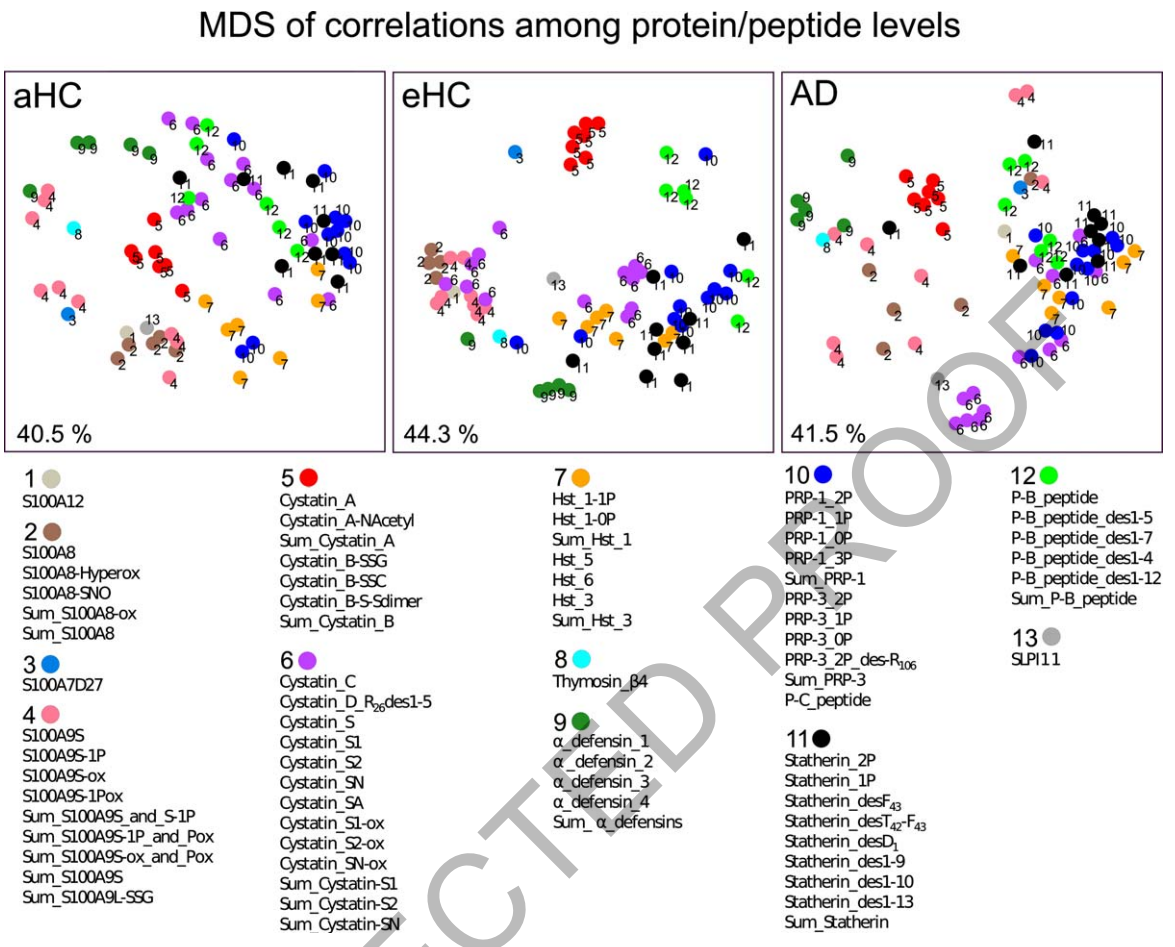

Fig. 3. Multidimensional Scaling (MDS) diagrams of Kendall correlations between components. To facilitate the understanding of the diagrams, the 76 components are grouped into different categories, numerically and color encoded, based on their structural/functional similitudes and secretory origin. The degree of clustering of points accounts for the degree of proteins/peptides co-expression. Percent values indicate the percent of information contained in bi-dimensional multidimensional scaling diagrams, relative to all the information contained in the whole multi-dimensional structure.

Correlation analysis of proteins/peptides within the groups

A diagram of correlated proteins/peptides within each group was obtained by multidimensional scaling analysis applied to Kendall correlations, as showed in Fig. 3, where clusters generated by components with correlated levels are evident. To facilitate the understanding of multidimensional scaling diagrams, the 76 components were subdivided into 13 categories based on their structural/functional analogies and secretory origin. The most compact cluster, in all the groups, was that of cystatins A and B (category 5 in Fig. 3). Less compact clusters were formed by histatins (category 7), α-defensins 1–4 (category 9), aPRPs (category 10), statherin family (category 11), and P-B peptide family (category 12), without appreciable differences between groups. Conversely, differences were found in the degree of clustering of three categories: category 2, including S100A8 and its proteoforms, which was relatively more compact in aHC and eHC than in AD; category 4, including S100A9 and its proteoforms, which was relatively more compact in eHC than in aHC and AD; category 6, including cystatins C, D and S, and category 9 (α-defensins 1–4) relatively more compact in eHC and AD than in aHC. In addition, the clustering of different categories was evaluated: a strong proximity between categories 2 and 4, including S100A8 and S100A9 respectively, was present in eHC. Moreover, S100A12 was observed to cluster with S100A8 in aHC and with S100A8 and A9 in eHC. In AD group a good relationship between Tβ4 and α-defensins 1–4 was found.

Table 3
Mean decrease of the Gini (MDG) scores of the 20 most important proteins/peptides or their sum, generically indicated as components, identified by RF

aHC versus eHC		eHC versus AD		aHC versus eHC versus AD	
Component	MDG	Component	MDG	Component	MDG
P-B_des1-7	11.82	Sum_S100A8	5.92	P-B_des1-7	14.60
PRP-1-2P	4.29	α-defensin3	4.14	Sum_S100A8	5.17
Sum_PRP-1	3.70	Sum_S100A8-ox	3.92	PRP-1-3P	4.93
Hst-3_1-24_(Hst-5)	1.19	Sum_S100A9(S)	2.81	Sum_PRP-1	3.56
Sum_P-B_peptide	1.10	S100A8-SNO	2.01	α-defensin3	3.37
Sum_S100A9(S)	1.00	Sum_S100A9(S)-1P_and_(S)-1Pox	1.80	Sum_S100A8-ox	3.08
P-B_peptide	0.94	Stath_des-F43	1.74	PRP-1-2P	3.03
PRP-3-2P	0.91	PRP-1-2P	1.50	Sum_S100A9(S)	2.94
Cyst_S2-ox	0.87	Stath_des1-13	1.17	S100A8-SNO	1.81
Cyst_SN-ox	0.72	Cystatin_A	1.14	Sum_S100A9(S)-1P_and_(S)-1Pox	1.80
PRP-1-3P	0.51	Stath_des1-9	1.08	Sum_α-defensins	1.64
Sum_cyst_SN	0.50	Sum_α-defensins	1.02	Hst-3_1-24_(Hst-5)	1.48
Sum_α-defensins	0.44	Cyst_B-S-S_dimer	0.99	α-defensin1	1.44
PRP-3-1P	0.43	α-defensin4	0.97	P-B_peptide	1.36
α-defensin4	0.42	α-defensin1	0.86	Sum_P-B_peptide	1.23
Cyst_S1-ox	0.40	Sum_S100A9(S)_and_(S)-1P	0.74	Sum_cyst_SN	1.23
Sum_PRP-3	0.35	α-defensin2	0.68	α-defensin2	1.06
Cystatin_A	0.31	S100A9(S)-1P	0.59	Cyst_B-S-S_dimer	1.06
P-C_peptide	0.30	Tβ4	0.54	α-defensin4	1.05
Sum_Hst-3	0.30	S100A8	0.37	Stath_des1-13	1.02

Classification of subjects by RF analysis

RF was applied to a subset of components selected according to the Boruta method (Supplementary Figure 3), indeed, the use of a subset of proteins/peptides individuated with the Boruta method, compared to the use of all the proteins/peptides analyzed, resulted in a consistent increase of classification accuracy [37]. The number of selected components varied in the three analyses, they were 47 for aHC-eHC, 22 for eHC-AD, and 38 for aHC-eHC-AD, for this reason a limit of 20 components with the highest MDG scores of each analysis was chosen (Table 3), being this parameter fundamental to individuate proteins/peptides important for classification.

Confusion matrices and sensitivity/specificity of classifications are shown in Fig. 4A. Classification of samples in either aHC or eHC group showed the highest accuracy (97.1%), followed by classification in either aHC, eHC or AD (82.7%), and in either eHC or AD (79.8%). It should be noted that these findings were validated by 'out of bag' samples, a method that consists in creating separate sets of training and test samples, composed by 72% and 36% of the entire set of data, respectively. According to MDG scores, the most important components for the classification of samples in either aHC or eHC groups were the des1–7 fragment of P-B peptide and aPRPs, especially PRP1 proteoforms. The most important components for the classification of the samples in either eHC or AD group were all the S100A8 proteoforms, mainly the oxidized and nitrosylated forms, all the S100A9 proteoforms and α-defensins, mainly α-defensin 3. The classification of samples mixing the three groups together (aHC-eHC-AD) shared the components already identified in the previous two analyses. It is to note that MDG scores were in close accordance with the Boruta selection, while, both MDG and Boruta scores were not consistent with Mann-Whitney tests. For example, in the classification of eHC-AD mixed samples, PRP1-2P was the 8th most 'important' component in the MDG ranking (Table 3) and the 10th in the Boruta ranking (not shown), but the same protein did not appear to be differentially expressed by the Mann-Whitney test (Table 2 and Supplementary Figure 2C). This apparent contrast represents an essential difference between multivariate RF classification and univariate comparisons and will be discussed in detail in the next section. Diagrams of RF classifications were obtained by multidimensional scaling analysis, using the proximity between each pair of samples as distance (Fig. 4B). A 3D movie showing the clustering of the three groups using the first three multidimensional scaling axes is reported in the Supplementary Material. RF classification of aHC-eHC-AD mixed samples was also shown by hierarchical cluster analysis (Fig. 5), which

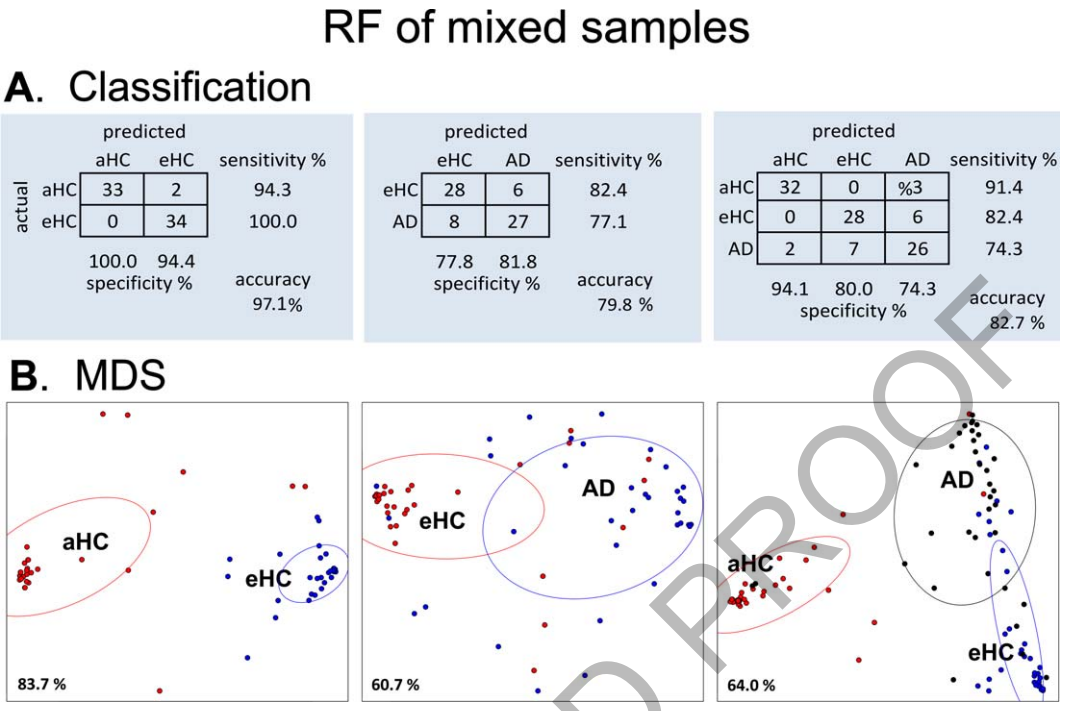

Fig. 4. RF applied to the three mixed data sets. Confusion matrices of RF classifications (A), validated by out-of-bag samples. Matrix rows represent the actual classes, while columns represent the predicted classes. Marginal columns show the frequency of false negatives, while marginal rows show the frequency of false positives. Multidimensional Scaling (MDS) diagrams (B) showing the relationships among subjects, using the proximity values calculated by RF. Each group is delimited by a dispersion ellipse with a confidence of 1.6 standard deviations. Percent values indicate the percent of information contained in bi-dimensional multidimensional scaling diagrams, relative to all the information contained in the whole multi-dimensional structure.

represents the hierarchical clustering of subjects of the aHC-eHC-AD mixed data set. aHC subjects are shown in red, eHC in blue and AD in black. Clustering was obtained using RF proximity values as distances and the Ward method as clustering criterion. The dendrogram shows three main clusters, each composed primarily of aHC or eHC or AD. The relative frequencies of aHC, eHC, and AD in each of the three main clusters are represented by the pie charts.

Pathway analysis

The biological pathway analysis was performed considering the 20 components with the highest MDG scores that allowed the RF classification of the samples in the three groups (Table 3). As it concerns the components differentiating the eHC from the aHC group (Fig. 6A), the Cytoscape ClueGo plugin individuated a functional pathway among antimicrobial components such as α-defensins 1–4, histatins 3 and 5, and S100A9 protein. Among the components differentiating the eHC from the AD group (Fig. 6B), the analysis placed the α-defensins in the antimicrobial pathway and revealed functional pathways around the several biological roles of S100A8 and S100A9 proteins, such as the immune system disease, the activation of NADPH oxidases and various pathways associated to the signaling cascade of the toll-like receptor.

DISCUSSION

Overall, the results of the statistical analysis performed on salivary proteomic data from two groups of healthy subjects differing only in the age, namely adult (under 70 years old) and elderly subjects (over 70 years old), and a group of elderly affected by AD, clearly evidenced: 1) changes of the salivary proteome according to a physiological parameter, the age, and 2) changes of the salivary proteome in accordance the investigated pathology, AD. These results highlighted that some salivary proteins could classify subjects with great accuracy based on two parameters: age and disease, some proteins being potentially

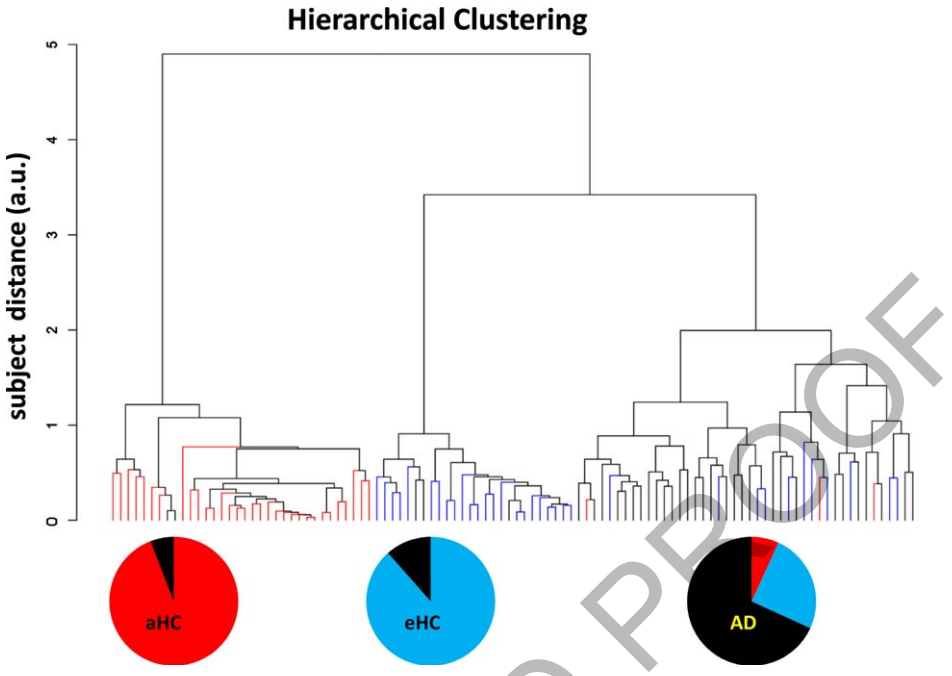

Fig. 5. Hierarchical clustering of subjects of the aHC-eHC-AD mixed data set. aHC subjects are shown in red, eHC in blue and AD in black. Clustering was obtained using RF proximity values as distances and the Ward method as clustering criterion. The dendrogram denotes three main clusters, each composed primarily of aHC or eHC or AD. The relative frequencies of aHC, eHC, and AD in each of the three main clusters are represented by the pie charts.

useful for diagnostic purposes. In particular, the aging-associated variations in the abundance of the salivary proteins were revealed between the adult group and the two elderly groups, with and without AD, reaching the 91% of the analyzed components with lower levels in healthy elderly than in adult controls. Elderly people, especially, showed a very noteworthy decreased level of components belonging to acidic proline-rich proteins, salivary cystatins, histatins, P-B peptide, and statherin families, which are proteins and peptides secreted by salivary glands [39]. Except for the P-B peptide, whose biological role is unknown, they are physiologically involved in the homeostasis, in the bacterial colonization and in antimicrobial defense of the oral cavity [7, 40]. The correlation analysis performed in this study evidenced that in all of the three groups their abundances were interrelated as expected for peptides/proteins with same origin and with harmonized functions. Moreover, these secretory components resulted to be among the most discriminant between elderly and adult subjects in the RF analysis of classification, in accordance with the Mann-Whitney test.

These findings appeared related to the physiological and histological changes occurring in the major salivary glands with aging. Indeed, the proportion of fat and fibrovascular tissues of salivary glands increases in elderly individuals [41], and the reduced volume of acinar cell secretion causes gland hypofunction [42]. However, our results were in accordance also with other studies, which suggested that age may influence the secretion of specific salivary components, indeed, different aging-related trends for diverse proteins have been observed highlighting increases of some and decreases of others [43]. It was demonstrated, for instance, that N-glycoproteins [44], amylase, and IgA [43], are more abundant in saliva of elderly with respect to younger subjects. This suggested that the highest total protein concentration measured in eHC group probably might have been affected by the concentration of amylase, IgA, or N-glycoproteins, which have not been analyzed in the present study. We observed a lower total protein concentration in elderly subjects with AD, whose pathological conditions could have interfered with the levels of these proteins. Indeed, de la Rubia et al. [45] observed a decreased level of IgA in AD subjects. It should be noted that we used the total protein concentration to normalize the quantitative data represented by the XIC peak areas of proteins/peptides, and that same results (not shown) were obtained without normalization. Thus, the nor-

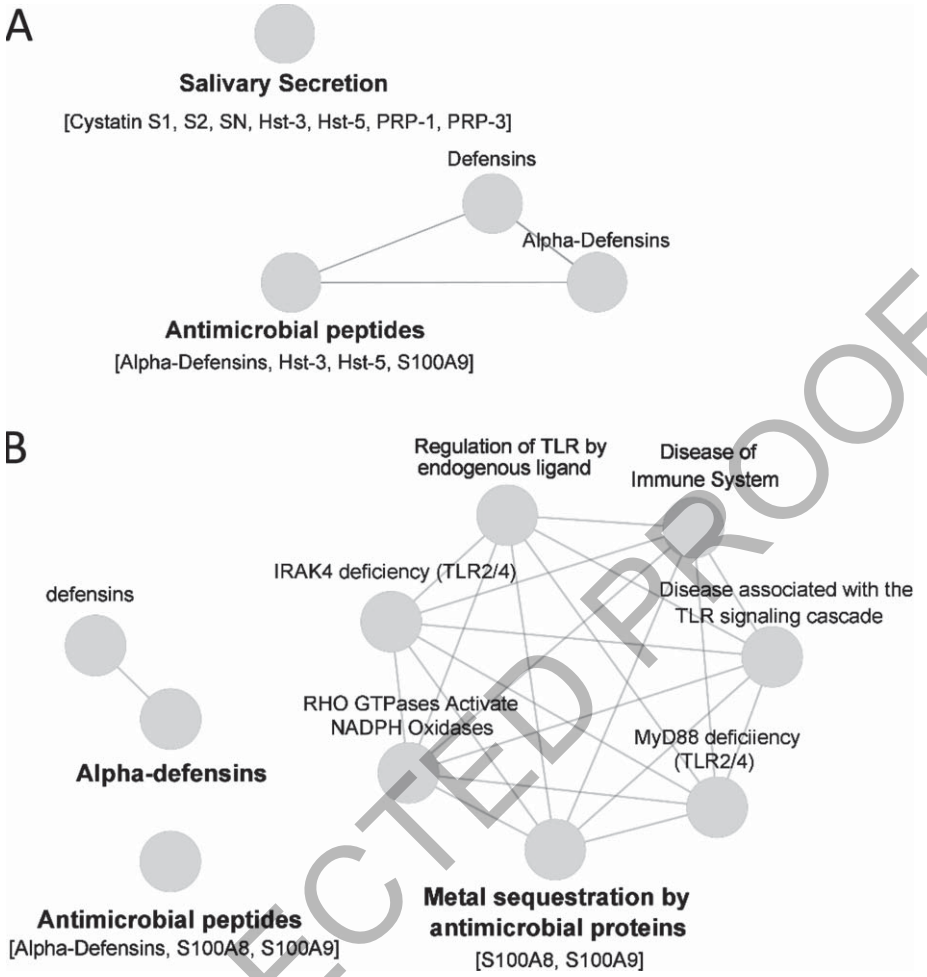

Fig. 6. Network representation of the functional pathway analysis considering the 20 components with the highest MDG scores used to classify the samples in the eHC, aHC, and AD groups. The analysis was performed by enriched KEGG and Reactome databases with the Cytoscape ClueGo plugin. Results obtained with components distinctive the eHC from aHC subjects in panel A, results obtained with components distinctive eHC from AD subjects in panel B.

malization on total protein concentration, even if affected by abundant proteins not considered in the analysis, did not produce an underestimation of the investigated proteins, and did not introduce errors in the statistical analysis.

Regarding peptides/proteins not secreted by salivary glands, aging-associated differences were observed between the two groups of healthy controls. In the younger subjects were measured higher levels of α-defensins 1, 2, 3, and 4, thymosin β4, cystatins A and B, and the several proteoforms of S100A9 and S100A8, all components largely expressed in several body fluids and tissues as modulator of the inflammatory processes and antimicrobial agents [40–48]. RF classification analysis individuated the α-defensins, cystatin A and the short S100A9 proteoforms among the 20 most discriminant between adult and elderly healthy controls. Moreover, the functional pathway analysis included α-defensins and S100A9 protein in the network of antimicrobial activities together with histatin 5. Overall, these results suggested that healthy subjects could be distinguished based on their salivary profiles that drastically changed with age, and that the variation involved especially components important to regulate the inflammation and to respond to antimicrobial occurrences not only in the oral cavity but also at systemic level. Our results showed that these abilities were reduced with the aging.

When adult controls were compared with elderly subjects with AD, not significant variations were observed in the abundances of the non-glandular components except for cystatin A, with lower level

in the patients, and the oxidized forms of S100A8, with higher levels in the patients. This last one appeared a disease-related difference and it has been highlighted also between the healthy elderly controls and the patients with AD together to other disease-related variations involving peptides and proteins included in the innate immune system pathway [40–48], which significantly increased in the patient group. In accordance with our previous study [29], they were statherin, histatin 1 among the secretory glandular peptides, α-defensins, thymosin β4, cystatins A and B, the S100A9 and S100A8 proteoforms, among the non-glandular components. The present investigation highlighted, in addition, that the oxidized forms of cystatins S2 and SN were found more abundant in AD patients than in elderly controls, suggesting a possible role of these proteins as oral ROS scavengers and the existence of oxidative state in oral cavity of the patients. It is noteworthy that S100A8 and S100A9 proteoforms, particularly the nitrosylated S100A8, as well as cystatins A and B, thymosin β4 and the α-defensins, mainly the α-defensin 3, resulted important for the classification of the groups based on the MDG scores, especially for the classification of subjects either in the AD group or in the control group of elderly. Interesting suggestions on the possible biological role in the pathology of these peptides and proteins were provided by the analysis of correlation and of functional network. The correlation analysis showed an association of proteins/peptides categories that was different in the three groups. Moreover, some specific categories of proteins/peptides were found to cluster together, as the cystatins A and B clustering in all the three groups. These two proteins share the same 3D structure, 80% sequence homology, and 52% identity [49], and as cathepsin inhibitors, both play an important role as regulatory factors of inflammation and in the innate immune response [48]. In particular cystatin B may exert additional functions in maintenance of cell homeostasis, reduction of oxidative stress [50], prevention of apoptosis [51], and neuronal protective role [52]. Interestingly, the high abundance of both cystatins A and B found in the patients and their constant co-regulation independently from age, could be in accordance with the hypothesis that they play a protective role in amyloid fibrilization, which probably is enhanced during AD occurrence. Indeed, it has been reported that both cystatins A and B co-localize in amyloid plaques of various origins [53, 54], bind Aβ and interrupt amyloid aggregation in cells [55]. The strong correlation between α-defensins 1–4 and thymosin β4 in saliva of the patients induced to reflect on the possible role that these antimicrobial and anti-inflammatory peptides could have in the brain, and on the hypothesis that the microbiota-induced neuronal inflammation may trigger Aβ deposition and AD development [56]. Indeed, the neuropathological alterations are often associated with abnormal expression and/or regulation of antimicrobial and anti-inflammatory peptides such as defensins [57]. This was in accordance with the antimicrobial functional pathway including the α-defensins 1–4 that was evidenced in the present study.

Interestingly, thymosin β4, which is a moonlighting peptide widely expressed in human tissues [58], exerts neuroprotective and neuro-regenerative effects [59], and it was found upregulated in reactive microglia of patients with AD, where it suppresses the pro-inflammatory signaling [60]. The strong correlation of the S100A8 proteoforms, and their clustering with S100A12 protein in both control groups, and, only in elderly subjects also with S100A9 proteoforms, appeared to be associated to two features: 1) they are all involved in the modulation of the inflammatory processes [47, 60]; 2) their co-participation in the inflammation appeared aging-related, being these proteins significantly less abundant in the elderly subjects. S100A12 is a potent chemoattractant for monocytic cells [61], and it activates various cell types inducing expression of adhesion molecules and pro-inflammatory cytokines [62]. S100A9 and S100A8 are constitutively expressed in immune cells and their expression and extracellular release are upregulated also in other cell types under inflammatory conditions [47]. The pathway analysis, performed on the most significant 20 components discriminating elderly controls from AD patients, indicated the participation of S100A9 and S100A8 in the regulation of the Toll-like receptors (TLRs) cascade. Indeed, extracellular S100A8 and S100A9 bind pattern recognition receptors including TLRs and receptor for advanced glycation end products (RAGE) to activate the innate immune system and mediate inflammation by inducing cytokine secretion, and by influencing monocyte and macrophage behavior [63]. Furthermore, they play an important role in protecting the body from pathogenic infection by triggering TLR4- or RAGE-mediated multiple inflammatory pathways [47]. It was demonstrated that S100A8 and A9 acts directly on cultured BV-2 microglial cells binding to TLR4 and RAGE and inducing the secretion of pro-inflamamtory cytokines through the activation of the nuclear factor κB

via ERK and JNK pathways [64]. For this reason, S100A8 and S100A9 were proposed as novel therapeutic targets for microglial-induced neuroinflammatory diseases. A recent review highlighted that TLRs are considered one of the major components of the AD pathogenesis, since signaling of TLRs affects synaptic plasticity, microglial activity, tau phosphorylation, and inflammatory responses, moreover, several genetic polymorphisms of TLRs have been also recognized as protective or risk factors for the AD [65].

Therefore, our results place the pro-inflammatory activity of S100A9 and S100A8 very well in the dynamics of AD pathogenesis. However, it is also worth reflecting on the fact that they may exert both pro- and anti-inflammatory effects by manipulating the cytokine profile of cells through TLRs- and RAGE-binding [66], and they can switch from pro-inflammatory to anti-inflammatory activity probably depending on the local microenvironment, the oxidative modifications, and the binding with metal ions [67]. The methionine oxidation of S100A9 can terminate the chemo-repulsive effect on peripheral neutrophils [68]. S100A8 exerts anti-inflammatory activity when modified by nitrosylation on its cysteine residue [69]. In addition, S100A9 and S100A8 can play a dual role also in oxidative stress conditions. Indeed, in activated granulocytes and macrophages, these two proteins are involved in the activation of NADPH oxidase 2, as suggested by the functional pathway analysis performed in this study, and thus contribute to the generation of reactive oxygen/nitrogen species (ROS/RNS) with a subsequent progression and exacerbation of the inflammatory status [70]. However, they can play protective roles against oxidative stress in several tissues as scavenger of ROS/RNS [71]. The very low levels of S100A12 and of all the proteoforms, unmodified and oxidized, of S100A8 and S100A9 in elderly controls, which we found strongly correlated, suggested an enhanced need of protection from oxidative damages, as well as a loss of capacity to regulate the inflammatory response. These results well agreed with the hypothesis that aging is a loss of homeostasis due to a chronic oxidative stress caused by an accumulation of ROS/RNS-induced damages, which is at the base of the "Oxidative stress theory of aging" [72]. Increased ROS/RNS levels lead to several age-related conditions affecting the regulatory systems, such as nervous, endocrine, and immune systems. The activation of the immune system induces an inflammatory state creating a vicious circle and, consequently, increases the age-related morbidity and mortality [72]. In this context, the highest levels of oxidized forms of S100A9 and S100A8, particularly the nitrosylated S100A8, in subjects with AD, reinforced the suggestion that these proteins, being particularly sensitive to oxidative cross-linking and massive oxidation [60], could act as ROS/RNS scavengers against oxidative damages in the patients, as well as part of an anti-inflammatory response.

The 79.8% accuracy in the classification of elderly subjects with and without AD was relatively higher than that reported for clinical diagnostic methods [73], but lower than that obtained using ultrasensitive assays of biomarkers, namely Aβ and tau proteins, in plasma [74], and in salivary samples [75]. Nevertheless, present data may be of interest because they were obtained from a panel of proteins/peptides present in saliva, not associated with classical markers of AD, and that were recently related to the disease [29]. For the classification, RF was preferred over other methods because of several reasons: (a) it is not conditioned by the data distribution; (b) it has a low risk of overfitting; (c) it does not require supplementary samples for the validation of results, as each tree is built up omitting nearly one-third of the samples that are subsequently used to test the misclassification rate; (d) it provides a measure of the relative importance of each feature in the classification of samples (MDG); (e) it provides an estimation of the proximity (i.e., similarity) between each two samples that opens the possibility of applying multidimensional scaling and hierarchical cluster analyses to obtain a visual representation of the classification. In addition, the use of this approach provides the possibility to correlate the position of misclassified subjects with their specific clinical profiles (i.e., severity of the disease, therapy, presence of comorbidities, etc.). This will be the subject of future investigations. The results of the RF classification were in accordance with those obtained by the analyses of the differential abundances and of correlation. Surprisingly, some proteins identified as important for RF classification were not found to be significantly changed by Mann-Whitney tests (e.g., statherin desF_{43} and PRP1, comparing eHC and AD subjects), and *vice versa* some proteins significantly changed were not found important for RF classification (e.g., statherin des1-13 and cystatin C, comparing aHC and eHC subjects). This apparent paradox is due to the *modus operandi* of RF, and more in general of methods based on decision trees, that are able to detect changes in variables using multiple split points, although the averages or medi-

ans of the groups being compared are approximately equivalent. Because of this fact, proteins important for RF classification could not be directly regarded as candidate markers for AD. For diagnostic purposes, the panel of differentially expressed proteins identified by Mann-Whitney tests appears to provide very reliable indications.

Conclusions

The results obtained in the present study, confirmed those of our previous study [29] that is the individuation of potential biomarkers candidates of AD in peptides and proteins principally involved in defense mechanisms of the innate immune system, in inflammation regulation and in protection against oxidative stress. Moreover, interesting new finding demonstrated that salivary protein profile strongly changes its composition, mainly at quantitative point of view, in relation to the aging, and that variation involved almost all the analyzed peptides and proteins and their proteoforms. The mainly fascinating results were produced by RF analysis, which demonstrated the feasibility of the salivary proteome to discriminate groups of subjects who are different in their age and health status, and thus patients with AD in comparison with healthy controls. This outcome drives to the potential use of the salivary protein profile for diagnostics and classification purposes.

ACKNOWLEDGMENTS

This study was supported by funds of Cagliari University (FIR-2019 and FIR 2020 Cabras Tiziana, FIR 2019 Alessandra Olianas).

Authors' disclosures available online (https://www.j-alz.com/manuscript-disclosures/22-0246r2).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: https://dx.doi.org/10.3233/JAD-220246.

REFERENCES

[1] Campisi J, Kapahi P, Lithgow GJ, Melov S, Newman JC, Verdin E (2019) From discoveries in ageing research to therapeutics for healthy ageing. *Nature* **571**, 183-192.

[2] Hoffman JM, Lyu Y, Pletcher SD, Promislow DEL (2017) Proteomics and metabolomics in ageing research: From biomarkers to systems biology. *Essays Biochem* **61**, 379-388.

[3] Rivero-Segura, NA, Bello-Chavolla OY, Barrera-Vázquez OS, Gutierrez-Robledo LM, Gomez-Verjan JC (2020) Promising biomarkers of human aging: In search of a multi-omics panel to understand the aging process from a multidimensional perspective. *Ageing Res Rev* **64**, 101164.

[4] Johnson AA, Shokhirevb MN, Wyss-Corayc T, Lehallier B (2020) Systematic review and analysis of human proteomics aging studies unveils a novel proteomic aging clock and identifies key processes that change with age. *Ageing Res Rev* **60**, 101070.

[5] Santos AL, Lindner AB (2017) Protein posttranslational modifications: Roles in aging and age-related disease. *Oxid Med Cell Longev* **2017**, 5716409.

[6] Tabak LA (2001) A revolution in biomedical assessment: The development of salivary diagnostics. *J Dental Educ* **65**, 1335-1339.

[7] Cabras T, Iavarone F, Manconi B, Olianas A, Sanna MT, Castagnola M, Messana I (2014) Top-down analytical platforms for the characterization of the human salivary proteome. *Bioanalysis* **6**, 563-581.

[8] Boroumand M, Olianas A, Cabras T, Manconi B, Fanni D, Faa G, Desiderio C, Messana I, Castagnola M (2021) Saliva, a bodily fluid with recognized and potential diagnostic applications. *J Sep Sci* **44**, 3677-3690.

[9] François M, Karpe A, Liu JW, Beale D, Hor M, Hecker J, Faunt J, Maddison J, Johns S, Doecke J, Rose S, Leifert WR (2021) Salivaomics as a potential tool for predicting Alzheimer's disease during the early stages of neurodegeneration. *J Alzheimers Dis* **82**, 1301-1313.

[10] Uchida H, Ovitt CE (2022) Novel impacts of saliva with regard to oral health. *J Prosthet Dent* **127**, 383-391.

[11] Cabras T, Pisano E, Boi R, Olianas A, Manconi B, Inzitari R, Fanali C, Giardina B, Castagnola M, Messana I (2009) Age-dependent modifications of the human salivary secretory protein complex. *J Proteome Res* **8**, 4126-4134.

[12] Castagnola M, Inzitari R, Fanali C, Iavarone F, Vitali A, Desiderio C, Vento G, Tirone C, Romagnoli C, Cabras T, Manconi B, Sanna MT, Boi R, Pisano E, Olianas A, Pellegrini M, Nemolato S, Heizmann CW, Faa G, Messana I (2011) The surprising composition of the salivary proteome of preterm human newborn. *Mol Cell Proteomics* **10**, M110.003467.

[13] Manconi B, Cabras T, Pisano E, Sanna MT, Olianas A, Fanos V, Faa G, Nemolato S, Iavarone F, Castagnola M, Messana I (2013) Modifications of the acidic soluble salivary proteome in human children from birth to the age of 48months investigated by a top-down HPLC-ESI-MS platform. *J Proteomics* **91**, 536-543.

[14] Messana I, Cabras T, Iavarone F, Manconi B, Huang L, Martelli C, Olianas A, Sanna MT, Pisano E, Sanna M, Arba M, D'Alessandro A, Desiderio C, Vitali A, Pirolli D, Tirone C, Lio A, Vento G, Romagnoli C, Cordaro M, Manni A, Gallenzi P, Fiorita A, Scarano E, Calò L, Passali GC, Picciotti PM, Paludetti G, Fanos V, Faa G, Castagnola M (2015) Chrono-proteomics of human saliva: Variations of the salivary proteome during human development. *J Proteome Res* **14**, 1666-1677.

[15] Morzel M, Palicki O, Chabanet C, Lucchi G, Ducoroy P, Chambon C, Nicklaus S (2011) Saliva electrophoretic protein profiles in infants: Changes with age and impact of teeth eruption and diet transition. *Arch Oral Biol* **56**, 634-642.

[16] Johnson DA, Yeh CK, Dodds MWJ (2000) Effect of donor age on the concentrations of histatins in human parotid and submandibular/sublingual saliva. *Arch Oral Biol* **45**, 731-740.

[17] Denny PC, Denny PA, Klauser DK, Hong SH, Navazesh M, Tabak LA (1991) Age-related changes in mucins from human whole saliva. *J Dent Res* **70**, 1320-1327.
[18] Chang WI, Chang JY, Kim YY, Lee G, Kho HS (2011) MUC1 expression in the oral mucosal epithelial cells of the elderly. *Arch Oral Biol* **56**, 885-890.
[19] Tanida T, Ueta E, Tobiume A, Hamada T, Rao F, Osaki T (2001) Influence of aging on candidal growth and adhesion regulatory agents in saliva. *J Oral Pathol Med* **30**, 328-335.
[20] Salvolini E, Martarelli D, Di Giorgio R, Mazzanti L, Procaccini M, Curatola G (2000) Age-related modifications in human unstimulated whole saliva: A biochemical study. *Aging Clin Exp Res* **12**, 445-448.
[21] Nagler RM, Hershkovich O (2005) Relationships between age, drugs, oral sensorial complaints and salivary profile. *Arch Oral Biol* **50**, 7-16.
[22] Manconi B, Liori B, Cabras T, Iavarone F, Manni A, Messana I, Castagnola M, Olianas A (2017) Top-down HPLC-ESI–MS proteomic analysis of saliva of edentulous subjects evidenced high levels of cystatin A, cystatin B and SPRR3. *Arch Oral Biol* **77**, 68-74.
[23] Fanali C, Inzitari R, Cabras T, Pisano E, Castagnola M, Celletti R, Manni A, Messana I (2008) alpha-Defensin levels in whole saliva of totally edentulous subjects. *Int J Immunopathol Pharmacol* **21**, 845-849.
[24] François M, Bull CF, Fenech MF, Leifert WR (2019) Current state of saliva biomarkers for aging and Alzheimer's disease. *Curr Alzheimer Res* **16**, 56-66.
[25] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack Jr CR, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillot MC, Thiest B, Weintraub S, Phelps CH (2011) The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 263-269.
[26] Aisen PS, Cummings J, Jack Jr CR, Morris JC, Sperling R, Frölich L, Jones RW, Dowsett SA, Matthews BR, Raskin J, Scheltens P, Dubois B (2017) On the path to 2025: Understanding the Alzheimer's disease continuum. *Alzheimers Res Ther* **9**, 60.
[27] Gleerup HS, Hasselbalch SG, Simonsen AH (2019) Biomarkers for Alzheimer's disease in saliva: A systematic review. *Dis Markers* **2019**, 4761054.
[28] Huan T, Tran T, Zheng J, Sapkota S, MacDonald SW, Camicioli R, Dixon RA, Li L (2018) Metabolomics analyses of saliva detect novel biomarkers of Alzheimer's disease. *J Alzheimers Dis* **65**, 1401-1416.
[29] Contini C, Olianas A, Serrao S, Deriu C, Iavarone F, Boroumand M, Bizzarro A, Lauria A, Faa G, Castagnola M, Messana I, Manconi B, Masullo C, Cabras T (2021) Top-down proteomics of human saliva highlights anti-inflammatory, antioxidant, and antimicrobial defense responses in Alzheimer disease. *Front Neurosci* **15**, 668852. Corrigendum in *Front Neurosci* **15**, 743596.
[30] Serrao S, Firinu D, Olianas A, Deidda M, Contini C, Iavarone F, Sanna MT, Boroumand M, Amado F, Castagnola M, Messana I, Del Giacco S, Manconi B, Cabras T (2020) Top-down proteomics of human saliva discloses significant variations of the protein profile in patients with mastocytosis. *J Proteome Res* **19**, 3238-3253.
[31] Ong SE, Mann M (2005) Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* **1**, 252-262.
[32] Messana I, Inzitari R, Fanali C, Cabras T, Castagnola M (2008) Facts and artifacts in proteomics of body fluids. What proteomics of saliva is telling us? *J Sep Sci* **31**, 1948-1963.
[33] Zaiontz C, Exact Mann-Whitney tests using the Real Statistics Resource Pack software (Release 7.6), https://www.real-statistics.com, 2021.
[34] Kendall M (1938) A new measure of rank correlation. *Biometrika* **30**, 81-89.
[35] Breiman L (2001) Random forests. *Mach Learn* **45**, 5-32.
[36] Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B Stat Methodol* **57**, 289-290.
[37] Kursa MB, Jankowski A, Rudnicki WR (2010) Boruta - A system for feature selection. *Fundam Inform* **101**, 271-285.
[38] Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagès F, Trajanoski Z, Galon J (2009) ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091-1093.
[39] Messana I, Cabras T, Pisano E, Sanna MT, Olianas A, Manconi B, Pellegrini M, Paludetti G, Scarano E, Fiorita A, Agostino S, Contucci AM, Calò L, Picciotti PM, Manni A, Bennick A, Vitali A, Fanali C, Inzitari R, Castagnola M (2008) Trafficking and postsecretory events responsible for the formation of secreted human salivary peptides. *Mol Cell Proteom* **7**, 911-926.
[40] Fábián TK, Hermann P, Beck A, Fejérdy P, Fábián G (2012) Salivary defense proteins: Their network and role in innate and acquired oral immunity. *Int J Mol Sci* **13**, 4295-4320.
[41] Scott J, Flower EA, Burns J (1987) A quantitative study of histological changes in the human parotid gland occurring with adult age. *J Oral Pathol* **16**, 505-510.
[42] Vissink A, Spijkervet FKL, Van Nieuw Amerongen A (1996) Aging and saliva: A review of the literature parotid gland submandibular gland. *Spec Care Dentist* **16**, 95-103.
[43] Xu F, Laguna L, Sarkar A (2019) Aging-related changes in quantity and quality of saliva: Where do we stand in our understanding? *J Texture Stud* **50**, 27-35.
[44] Sun S, Zhao F, Wang Q, Zhong Y, Cai T, Wu P, Yang F, Li Z (2014) Analysis of age and gender associated n-glycoproteome in human whole saliva. *Clin Proteomics* **11**, 1-10.
[45] de la Rubia Ortí JE, Prado-Gascó V, Castillo SS, Julián-Rochina M, Romero Gómez FJ, García-Pardo MP (2019) Cortisol and IgA are involved in the progression of Alzheimer's disease. A pilot study. *Cell Mol Neurobiol* **39**, 1061-1065.
[46] Yeaman MR, Bayer AS (2007) Antimicrobial host defense. In *Platelets*, Michelson AD, 2nd ed. Academic Press, New York, pp. 727-755.
[47] Wang S, Song R, Wang Z, Jing Z, Wang S, Ma J (2018) S100A8/A9 in inflammation. *Front Immunol* **9**, 1298.
[48] Magister S, Kos J (2013) Cystatins in immune system. *J Cancer* **4**, 45-56.
[49] Žerovnik E, Staniforth RA, Turk D (2010) Amyloid fibril formation by human stefins: Structure, mechanism & putative functions. *Biochimie* **92**, 1597-1607.
[50] Butinar M, Prebanda MT, Rajkovic J, Jeric B, Stoka V, Peters C, Reinheckel T, Kruger A, Rurk V, Turk B, Vasiljeva O (2014) Stefin B deficiency reduces tumor growth via sensitization of tumor cells to oxidative stress in a breast cancer model. *Oncogene* **33**, 3392-3400.
[51] Sun T, Turk V, Turk B, Kopitar-Jerala N (2012) Increased expression of stefin B in the nucleus of T98G astrocytoma Cells Delays caspase activation. *Front Mol Neurosci* **5**, 93.

[52] Soond SM, Kozhevnikova MV, Zamyatnin AA, Townsend, PA (2019) Cysteine cathepsin protease inhibition: An update on its diagnostic, prognostic and therapeutic potential in cancer. *Pharmaceuticals* **12**, 87.

[53] Bernstein HG, Rinne R, Kirschke H, Järvinen M, Knöfelh B, Rinne A (1994) Cystatin A-like immunoreactivity is widely distributed in human brain and accumulates in neuritic plaques of Alzheimer disease subjects. *Brain Res Bull* **33**, 477-481.

[54] Ii K, Ito H, Kominami E, Hirano A (1993) Abnormal distribution of cathepsin proteinases and endogenous inhibitors (cystatins) in the hippocampus of patients with Alzheimer's disease, parkinsonism-dementia complex on Guam, and senile dementia and in the aged. *Virchows Arch A Pathol Anat Histopathol* **423**, 185-194.

[55] Skerget K, Taler-Vercic A, Bavdek A, Hodnik V, Ceru S, Tusek-Znidaric M, Kumm T, Pitsi D, Pompe-Novak M, Palumaa P, Soriano S, Kopitar-Jerala N, Turk V, Anderluh G, Zerovnik E (2010) Interaction between oligomers of stefin B and amyloid-β *in vitro* and in cells. *J Biol Chem* **285**, 3201-3210.

[56] Mawanda F, Wallace R (2013) Can infections cause Alzheimer's disease? *Epidemiol Rev* **35**, 161-180.

[57] Williams WM, Castellani RJ, Weinberg A, Perry G, Smith MA (2012) Do β-defensins and other antimicrobial peptides play a role in neuroimmune function and neurodegeneration? *ScientificWorldJournal* **2012**, 905785.

[58] Hannappel E (2007) beta-Thymosins. *Ann N Y Acad Sci* **1112**, 21-37.

[59] Zhang G, Murthy KD, Pare RB, Qian Y (2020) Protective effect of Tβ4 on central nervous system tissues and its developmental prospects. *Eur J Inflamm* **18**, 1-11.

[60] Goyette J, Geczy CL (2011) Inflammation-associated S100 proteins: New mechanisms that regulate function. *Amino Acids* **41**, 821-842.

[61] Miranda LP, Tao T, Jones A, Chernushevich I, Standing KG, Geczy CL, Alewood PF (2001) Total chemical synthesis and chemotactic activity of human S100A12 (EN-RAGE). *FEBS Lett* **488**, 85-90.

[62] Hofmann MA, Drury S, Fu C, Qu W, Taguchi A, Lu Y, Avila C, Kambham N, Bierhaus A, Nawroth P, Neurath MF, Slattery T, Beach D, McClary J, Nagashima M, Morser J, Stern D, Schmidt AM (1999) RAGE mediates a novel proinflammatory axis. *Cell* **97**, 889-901.

[63] Foell D, Wittkowski H, Roth J (2007) Mechanisms of disease: A 'DAMP' view of inflammatory arthritis. *Nat Clin Pract Rheumatol* **3**, 382-390.

[64] Li M, Peng S, Zhang JC, Zhang Q, Shang-Long Y (2017) Proinflammatory effects of S100A8/A9 via TLR4 and RAGE signaling pathways in BV-2 microglial cells. *Int J Mol Med* **40**, 31-38.

[65] Momtazmanesh S, Perry G, Rezaei N (2020) Toll-like receptors in Alzheimer's disease. *J Neuroimmunol* **348**, 577362.

[66] Donato R, Cannon BR, Sorci G, Riuzzi F, Hsy K, Weber DJ, Geczy CL (2013) Functions of S100 proteins. *Curr Mol Med* **13**, 24-57.

[67] Gazzar M (2015) Immunobiology of S100A8 and S100A9 proteins and their role in acute inflammation and sepsis. *Int J Immunol Immunother* **2**, 2.

[68] Sroussi HY, Berline J, Palefsky JM (2007) Oxidation of methionine 63 and 83 regulates the effect of S100A9 on the migration of neutrophils *in vitro*. *J Leukoc Biol* **81**, 818-824.

[69] Lim SY, Raftery M, Cai H, Hsu K, Yan WX, Hseih HL, Watts RN, Richardson D, Thomas S, Perry M, Geczy CL (2008) S-Nitrosylated S100A8: Novel anti-inflammatory properties. *J Immunol* **181**, 5627-5636.

[70] Schenten V, Melchior C, Steinckwich N, Tschirhart EJ, Bréchard S (2011) Sphingosine kinases regulate NOX2 activity via p38 MAPK-dependent translocation of S100A8/A9. *J Leukoc Biol* **89**, 587-596.

[71] Gomes LH, Raftery MJ, Yan WX, Goyette JD, Thomas PS, Geczy CL (2013) S100A8 and S100A9 - Oxidant scavengers in inflammation. *Free Radic Biol Med* **58**, 170-186.

[72] Liguori I, Russo G, Curcio F, Bulli G, Aran L, Della-Morte D, Gargiulo G, Testa G, Cacciatore F, Bonaduce D, Abete P (2018) Oxidative stress, aging, and diseases. *Clin Interv Aging* **13**, 757-772.

[73] Beach TG, Monsell SE, Phillips LE, Kukull W (2012) Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005-2010. *J Neuropathol Exp Neurol* **71**, 266-273.

[74] Kim K, Kim MJ, Kim DW, Kim SY, Park S, Park CB (2020) Clinically accurate diagnosis of Alzheimer's disease via multiplexed sensing of core biomarkers in human plasma. *Nat Commun* **11**, 119.

[75] Liang Q, Liu H, Zhang T, Jiang Y, Xing H, Zhang A (2015) Metabolomics-based screening of salivary biomarkers for early diagnosis of Alzheimer's disease. *RSC Adv* **5**, 96074.