Article

Systemic Blood Proteome Patterns Reflect Disease Phenotypes in Neovascular Age-Related Macular Degeneration

Steffen E. Künzel [1], Leonie T. M. Flesch [1], Dominik P. Frentzel [1], Vitus A. Knecht [1], Anne Rübsam [1], Felix Dreher [2], Moritz Schütte [2], Alexandre Dubrac [3], Bodo Lange [2], Marie-Laure Yaspo [4], Hans Lehrach [4], Antonia M. Joussen [1] and Oliver Zeitz [1,*]

[1] Department of Ophthalmology, Charité—Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Hindenburgdamm 30, 12203 Berlin, Germany; steffen-emil.kuenzel@charite.de (S.E.K.); leonie.flesch@charite.de (L.T.M.F.); dominik.frentzel@gmail.com (D.P.F.); vitus-andre.knecht@charite.de (V.A.K.); anne.ruebsam@charite.de (A.R.); antonia.joussen@charite.de (A.M.J.)
[2] Alacris Theranostics, Max-Planck-Straße 3, 12489 Berlin, Germany; f.dreher@alacris.de (F.D.); mo.schuette@gmail.com (M.S.); b.lange@alacris.de (B.L.)
[3] Département de Pathologie et Biologie Cellulaire, Université de Montréal, Montréal, QC H3C 3J7, Canada; alexandre.dubrac@umontreal.ca
[4] Max-Planck-Institute for Molecular Genetics, Ihnestrasse 63-73, 14195 Berlin, Germany; yaspo@molgen.mpg.de (M.-L.Y.); lehrach@molgen.mpg.de (H.L.)
* Correspondence: oliver.zeitz@charite.de; Tel.: +49-30-450-554302

Abstract: There is early evidence of extraocular systemic signals effecting function and morphology in neovascular age-related macular degeneration (nAMD). The prospective, cross-sectional BIOMAC study is an explorative investigation of peripheral blood proteome profiles and matched clinical features to uncover systemic determinacy in nAMD under anti-vascular endothelial growth factor intravitreal therapy (anti-VEGF IVT). It includes 46 nAMD patients stratified by the level of disease control under ongoing anti-VEGF treatment. Proteomic profiles in peripheral blood samples of every patient were detected with LC-MS/MS mass spectrometry. The patients underwent extensive clinical examination with a focus on macular function and morphology. In silico analysis includes unbiased dimensionality reduction and clustering, a subsequent annotation of clinical features, and non-linear models for recognition of underlying patterns. The model assessment was performed using leave-one-out cross validation. The findings provide an exploratory demonstration of the link between systemic proteomic signals and macular disease pattern using and validating non-linear classification models. Three main results were obtained: (1) Proteome-based clustering identifies two distinct patient subclusters with the smaller one ($n = 10$) exhibiting a strong signature for oxidative stress response. Matching the relevant meta-features on the individual patient's level identifies pulmonary dysfunction as an underlying health condition in these patients. (2) We identify biomarkers for nAMD disease features with Aldolase C as a putative factor associated with superior disease control under ongoing anti-VEGF treatment. (3) Apart from this, isolated protein markers are only weakly correlated with nAMD disease expression. In contrast, applying a non-linear classification model identifies complex molecular patterns hidden in a high number of proteomic dimensions determining macular disease expression. In conclusion, so far unconsidered systemic signals in the peripheral blood proteome contribute to the clinically observed phenotype of nAMD, which should be examined in future translational research on AMD.

Keywords: neovascular age-related macular degeneration; peripheral blood proteomics; intravitreal anti-VEGF injections; non-linear classification models; personalized medicine; systems biology; precision medicine

1. Introduction

Age-related macular degeneration (AMD) remains amongst the most prevalent sight-threatening conditions in the elderly population [1]. Intravitreal anti-vascular endothelial growth factor treatment (anti-VEGF, IVT) proves to be an effective measure in eyes affected by the exudative stage of the disease, neovascular AMD (nAMD), thereby halting or even reversing the loss of central vision in most subjects [1]. At the same time, the individual treatment response and course of disease under treatment is highly variable. At one end of the spectrum, there are subjects in which choroidal neovascularization (CNV) activity ceases after initial treatment with a small number of injections and the retina can be kept in this "dry" state with no or only a smaller number of injections—an 'effectively controlled CNV' (ECC). The other extreme are subjects in which the CNV remains active despite frequent injections every 4–6 weeks and a "dry" state of the retina is never achieved—'chronically active CNV' (CAC) [2,3]. The reasons for these heterogenous phenotypes of CNV activity are still incompletely understood today. The reasons may lay in the multifactorial pathogenesis of AMD.

In fact, it is generally accepted that neovascular AMD is caused by a complex interplay of genetic susceptibility, ageing, and environmental risk factors—with smoking and some nutritional aspects as consistent and modifiable conditions [1]. In contrast to the multifactorial nature of the disease, treatment of neovascular AMD has been largely mono-factorial over almost two decades. Substantial efforts were undertaken to identify novel targets, which close the clinical gaps that are left by VEGF-inhibition. Blocking additional pathways to VEGF such as Angiopoietin-2 resulted in some gradual clinical improvements in comparison to anti-VEGF only, mainly in terms of longer injection intervals, but have not eliminated the CAC phenotype as described above [4,5]. This indicates that consideration of other factors not previously considered may lead to improved therapeutic outcomes, but a fully satisfactory therapeutic option for all patients has not yet been established.

An alternative way to decipher the disease-specific processes of AMD is to view the eye as an integrated organ within the whole organism. It is well understood that a tight interlocking of inflammatory processes with neurodegeneration and angiogenesis are key processes in the onset and development of the disease [1,6–8]. None of these processes is specific to the eye, but all of them are highly preserved in terms of physiological or pathophysiological mechanisms occurring in all regions of the human body. Thereby, they are spanning a variety of mutually regulating pathways and molecular interrelationships including multiple organ systems—some of which are known to interact with the VEGF system or pathological growth of blood vessels [8–11]. Taking this into account, it is plausible that previously unappreciated systemic signals may influence nAMD disease expression, e.g., in terms of morphology and treatment response.

Proteins play a fundamental role in various biological processes, and alterations in their function and regulation can significantly impact disease onset and course [12,13]. Thus, one way to understand ocular phenotypes in a systems biology context is to explore entire sets of proteins present in the biological system—so called proteomics [14]. This methodology has been previously employed on diverse tissue specimens associated with AMD, encompassing the retinal pigment epithelium (RPE), Bruch's membrane, drusen, as well as on fluid samples such as vitreous humor, tear fluid, aqueous humor, urine, and peripheral blood [15–32]. Blood proteome profiles offer a comprehensive and unified approach to assess an individual's global molecular status, as they encompass information from multiple tissues and directly reflect disease-related molecular pathways and activities [12,13,24]. While the main focus of recent blood proteome approaches is on biomarker and target discovery with a distinct focus on single proteins instead of more holistic proteomic considerations, newer studies also employ complex modeling for disease course determination and subtype distinction and reveal altered proteomic profiles in nAMD patients [15,20–24]. However, a decryption of the two extreme manifestations regarding disease control under anti-VEGF IVT in nAMD has not yet been achieved on a proteomic level.

In this exploratory study, we hypothesize that there is a systemic molecular impact on the ocular nAMD phenotype. We conceptually examined proteomic whole-blood profiles of 46 nAMD patients under current anti-VEGF IVT, stratified with the effectiveness of CNV control (ECC vs. CAC), on the exploratory search for linkage between the individual proteomic blood profile with the retinal disease manifestation. Interestingly, we detect only a weak impact of single protein features on disease expression but rather high-dimensional similarity patterns of potential clinical relevance. This supports the idea of a multifactorial systemic nAMD understanding and should encourage more whole-organism research to better understand retinal diseases.

2. Results

2.1. Study Population and Baseline Clinical Characteristics

We recruited 46 nAMD patients stratified by the CNV activity under anti-VEGF IVT with 54% of participants ($n = 25$) in the CAC division versus 46% in the ECC division ($n = 21$, compare stratification strategy and criteria in methods). Supplementary Figure S3 presents the clinical and demographic characteristics of the study population which is well-representing the general nAMD population concerning multiple clinical features [1,33]. Expectedly, patients of the CAC cohort were more likely to have subretinal hyper-reflective material (SHRM, $p < 0.0001$), a higher central retinal thickness (CRT, $p = 0.0025$), more frequent anti-VEGF IVT ($p < 0.0001$, among stratification criteria), slightly more often subretinal fluid (SRF, ns, $p = 0.056$) and intraretinal cysts (IRC, $p = 0.29$), and gradually lower visual acuity (ns, $p = 0.2724$). No unexpected changes between any of the other epidemiological, functional, morphological, and medical-history-related meta-features were detected in our cohort, except for anticipated correlations between lower BCVA and intraretinal cysts (IRC, $p = 0.0013$), between lower BCVA and subretinal hyperreflective material (SHRM, $p = 0.0329$), between active smoking and respiratory dysfunction ($p = 0.0008$), and between diabetes mellitus and arterial hypertension ($p = 0.0371$), respectively.

2.2. Unbiased Dimensionality Reduction and Clustering Reveal a Distinct Patient Sub-Cohort with a Strong Proteomic Signature for Oxidative Stress Response and Respiratory Dysfunction as an Underlying Health Condition

Whole-blood samples of every patient were taken at the baseline. Samples were analyzed with mass spectrometry (LC-MS/MS)-based proteomics, thereby providing an individual dataset of 1182 protein values for each subject. To reveal local and global similarity patterns to identify patient subclusters ('similarity neighborhoods') of putative clinical relevance, we used the non-linear dimensionality reduction method UMAP (Uniform Manifold Approximation and Projection). Interestingly, UMAP separates the cohort into two distinct scatter clouds with the smaller one consisting of 10 patients (cluster 1; 21.7%) and the larger one of 36 patients (cluster 2; 78.3%). The clear distinction into two clouds in the UMAP visualization is reflected in 131 significantly different markers which represent more than 11% of the total number of examined proteins (FDR-adjusted p-value < 0.05, 88 of which with $p < 0.01$, Figure 1b,c). Notably, cluster 1 distinguishes itself with 115 significantly different proteins (85 with $p < 0.01$, all p-values are FDR-adjusted), while the larger cluster 2 yields 16 markers only (3 with $p < 0.01$; compare Figure 1b). The top ten detected proteins of cluster 1 include arachidonate 5-lipoxygenase-activating protein (AL5AP, log2 of protein ratio lr = 4.701, $p = 0.0006$), a putative marker for lung disease [34]; endoplasmic reticulum chaperone BiP (BIP, lr = 1.144, $p = 0.00025$), which is important for lung structure and function [35]; as well as multiple key components of the respiratory chain as responding elements to oxidative stress (cytochrome b-245 heavy chain, CY24B, lr = 1.113, $p = 0.00064$; mitochondrial phosphate carrier protein, MPCP, lr = 2.172, $p = 0.00038$; ATP synthase subunit alpha, ATPA, lr = 1.294, $p = 0.00062$, cytochrome c oxidase subunit II, COX2, lr = 5.365, $p = 0.00025$) and calnexin (CALX, lr = 2.796, $p = 0.00025$), which is another response molecule associated with oxidative stress [36]. Value ratios between the two clusters are extraordinarily high for the identified proteins (Figure 1d,e). The three top markers of cluster 2 include the nuclear receptor 2C2-associated protein (NR2CA, lr = 2.428,

p = 0.00538), tropomyosin alpha-4 chain (TPM4, lr = 0.246, p = 0.0057), and cartilage acidic protein 1 (CRAC1, lr = 1.254, p = 0.00645; Figure 1b–e).
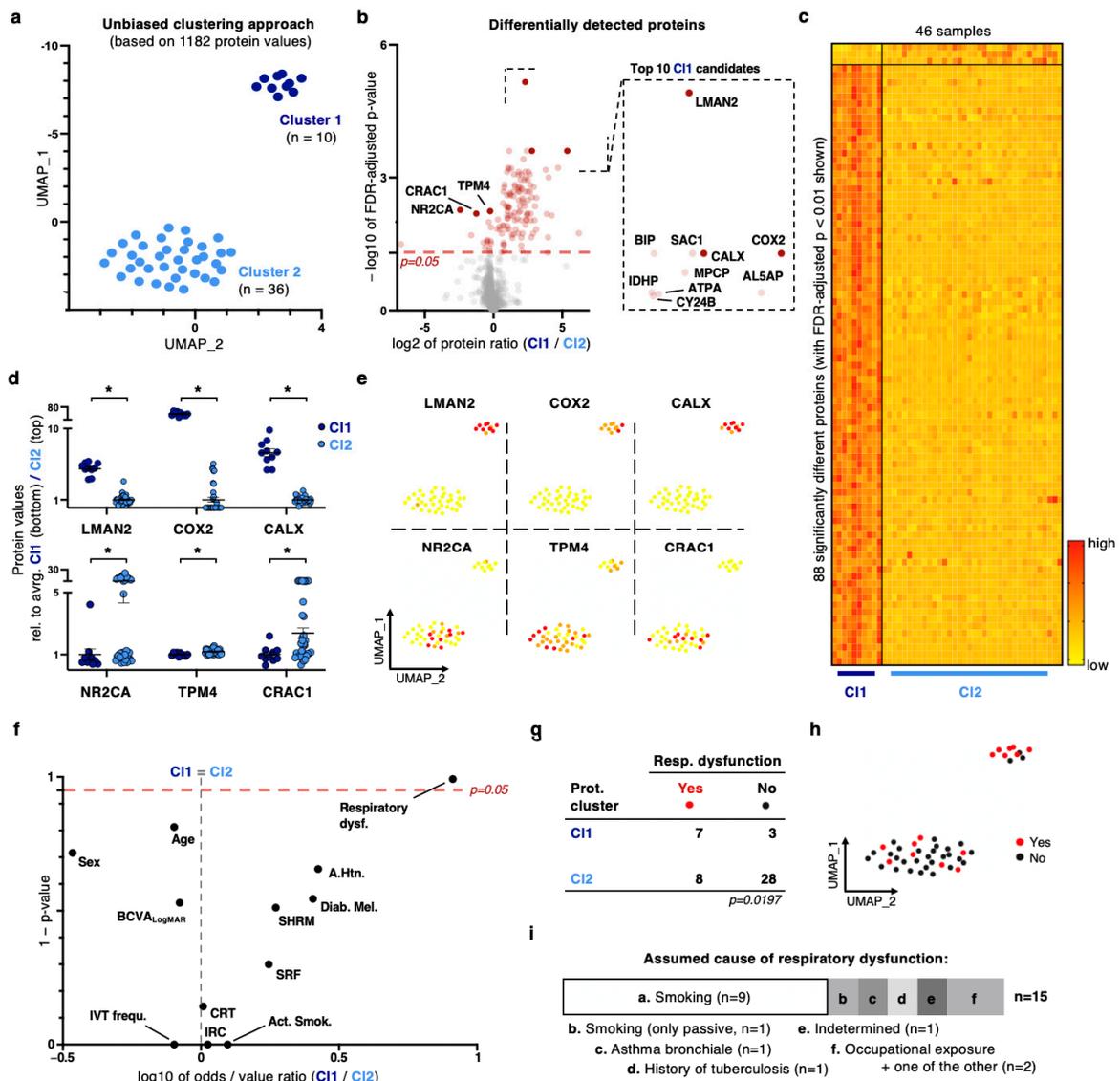

Figure 1. Dimensionality reduction and unbiased clustering based on proteomic profiles reveal a strong signature for oxidative stress response with respiratory dysfunction as an underlying clinical condition. (**a**) UMAP plot based on 46 whole-blood proteome samples. Each dot represents one sample of one subject. Color-coding: unbiased clustering based on nearest-neighbor approach with two distinct subclusters. (**b**) Left: Volcano plot for visualization of differentially detected proteins between the two cohorts. Right: Highlight of the top 10 candidates. (**c**) Heatmap visualization. Each column represents one blood sample of one donor; each row represents one distinct protein. Only proteins with significant difference (FDR-adjusted p-value < 0.01) between clusters are shown. Color-coding reflects normalized protein detection level per individual sample: red represents highest, bright yellow lowest value. A color-coded scale is located at the bottom right. (**d**) Protein values of top three detected proteins per cluster compared to average of contrary cluster. (**e**) Feature plot of proteins shown in (**c**). Color-coding: compare (**c**). (**f**) Volcano plot of meta-data annotations between the two clusters. (**g**) Contingency table of the two dichotomic features proteome cluster and respiratory dysfunction. (**h**) Feature plot for respiratory dysfunction. (**i**) Frequency of assumed causes for respiratory dysfunction. For all graphs in this figure (if not specified otherwise): * p < 0.05, Mann–Whitney *U* test, FDR correction.

Next, we annotate meta-features on patient level to test whether there is an epidemiological or general health condition that might explain the strong proteomic patterning between the two clusters. Notably, we detect respiratory dysfunction as a significantly overrepresented feature in cluster 1 (log2-ratio = 0.912 and p = 0.0079, Figure 1f–h). None of the other conditions was significantly different in one of the sub-clusters and, in particular morphological nAMD disease features (e.g., SRF or stratification division), did not differ between clusters 1 and 2, so nAMD did not feature on this allocation (Figure 1f). When the 15 patients suffering from respiratory dysfunction (7 of 10 in cluster 1, 70% versus 8 out of 36 in cluster 2, 22.2%; p = 0.0197) were asked for the assumed cause of this condition, two-thirds of them declared active or passive smoking as the probable root condition (Figure 1i); however, there were more non-smoking patients suffering from respiratory dysfunction in cluster 1 (compare Figure 1h,i with Supplementary Figure S4a).

2.3. Exploiting Meta-Feature Annotations on a Patient Level Yields a Limited Number of Biomarker Discoveries

As a basic analytic screening similar to previous proteomic studies [16,18,32,37], we performed canonical, non-parametric testing to determine whether a particular protein is differentially detected between certain meta-conditions for the 36 patients of cluster 2 (Figures 2, S6 and S7). The focus on this cluster for further analytic steps is due to the strong proteomic distortion of cluster 1 samples (Figure 1), but we provide the same analysis for all 46 samples in Supplementary Figure S4.

After a rigorous FDR correction of p-values, we detected a small number of significant protein markers, inter alia the y-chromosomal eukaryotic translation initiation factor 1A (EIF1AY, lr = 1.61, p = 0.000069) for male patients, insulin-like growth factor binding protein 5 (IGFBP5, lr = −2.83, p = 0.0051) for non-smokers, and COP9 signalosome complex subunit 3 (COPS3, lr = 0.21, p = 0.006) for diabetics. The three identified proteins are known markers for the given conditions and thus serve as confirmatory anchors for our proteomic approach and data annotation [38,39]. We detect higher levels of prefoldin subunit 2 (PFDN2, lr = 0.42, p = 0.0014) in nAMD patients with SRF (Figure 2b), a molecule that has previously been weakly associated with permeability [40]. The significance of prefoldin subunit 2 for SRF features on cluster 2, as well as on all 46 patients (Figure 2c, Supplementary Figure S3c) was observed. Furthermore, we detect peroxiredoxin-6 (PRDX6) as a biomarker for SHRM in all 46 patients (lr = 0.29, p = 0.0493), although being non-significant for cluster 2 after FDR-correction (lr = 0.27, non-FDR-corrected p = 0.0017, $p \approx 1$). The protein is involved in redox regulation during protection against oxidative injury [41]. Files containing ratios and p-values for all proteins between any logged conditions are provided for patients in cluster 2 (Supplementary Tables S6 and S7).

As singular or less-frequent biomarkers do not enable us to reliably classify patients into the two extremes of the stratification divisions, we took an alternative approach in which we projected meta-annotations on the generated UMAP dimensionality reduction to visualize potential patterns based on proteomic similarity neighborhoods. For most meta-features, we detected a seemingly random allocation without apparent patterns (Figure 2a). However, concerning CNV activity (stratification), we yielded a bipolar division with patients from the CAC division on the one and from the ECC division on the other pole in cluster 2 (Figure 2a,d). We believe that this observation, as well as subsequent findings shown in Figures 2 and 3 might also be true for the smaller cluster 1. However, the small number of 10 patients prevents us from yielding significant results due to statistical underpower with many proteomic dimensions. In order to represent this bipolar division of cluster 2 with a classification boundary between the two divisions, we built an intuitive linear support vector machine (SVM) classifier without the specification of Kernel functions or other modifications [42]. We use this boundary line to screen for proteomic markers that might define this 'level of CNV control' axis (by regarding Euclidian distance between the UMAP coordinate and the SVM boundary). Interestingly, with Aldolase C—a key enzyme for glycolysis [43]—we find a biomarker that is significantly overrepresented on the ECC

pole (compare Figure 2d–f, Table S8). We rate the finding of Aldolase C supportive for this model, as it has previously been associated with AMD [44].
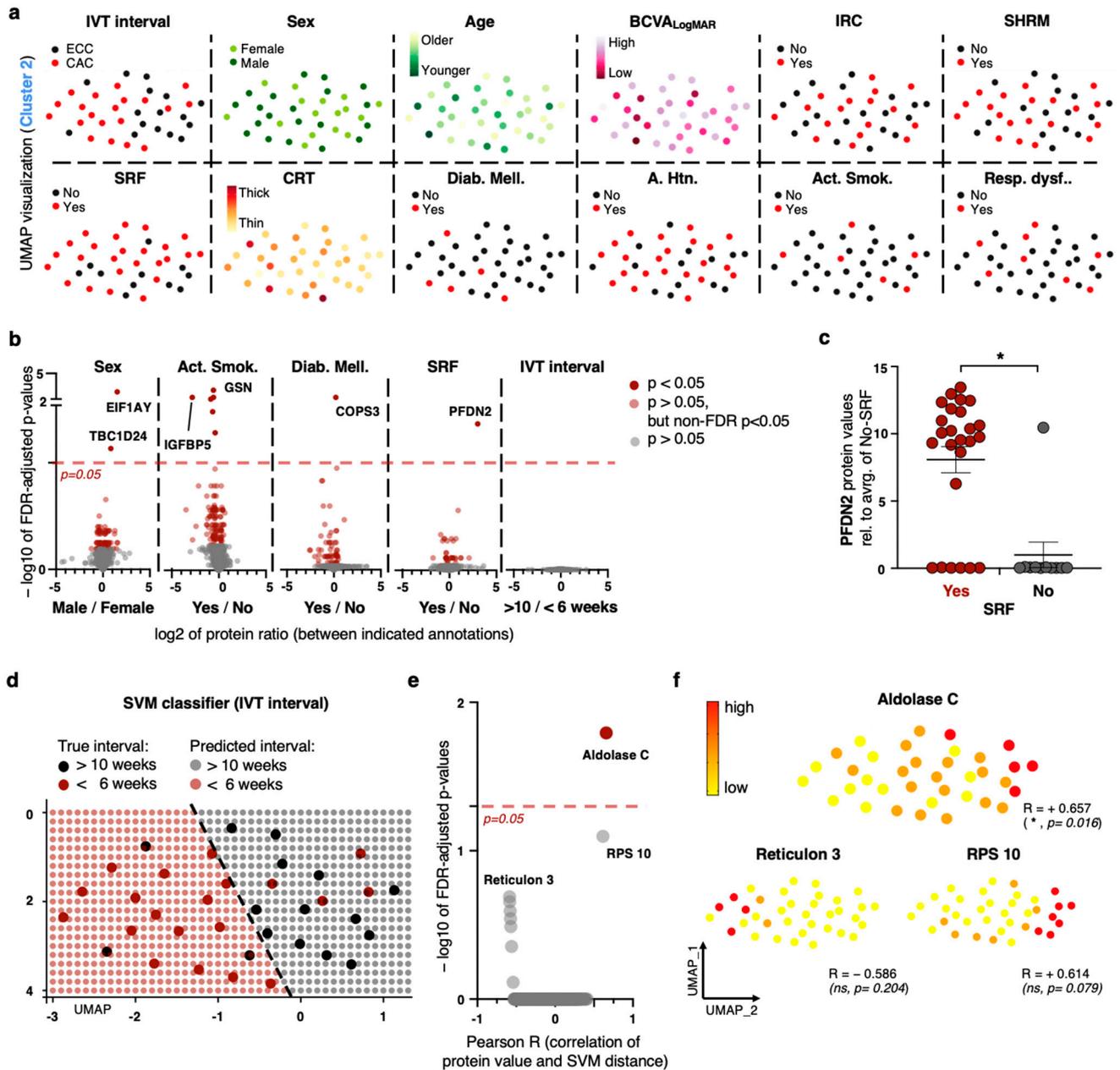

Figure 2. Proteomic biomarker discoveries for nAMD meta-features. (**a**) Visualization of meta-features on UMAP dimensionality reduction of cluster 2. (**b**) Volcano plots for identification of proteomic biomarkers of indicated conditions. Meta-features without graphic representation yield no significant biomarkers. (**c**) Comparison of protein values of prefoldin subunit 2 for patients with and without SRF. (**d**) Linear SVM classifier categorizing samples into the two CNV control divisions (color-coded). (**e**) Volcano plot indicating proteomic biomarkers of SVM classifier from C based on Euclidian distance to SVM boundary. (**f**) Feature plot of top three significant proteins. Color-coding explained in Figure 1c. Color-coded scale provided at top left. For all graphs in this figure (if not specified otherwise): * $p < 0.05$, Mann–Whitney U test, FDR correction. Spearman correlation in (**f**).
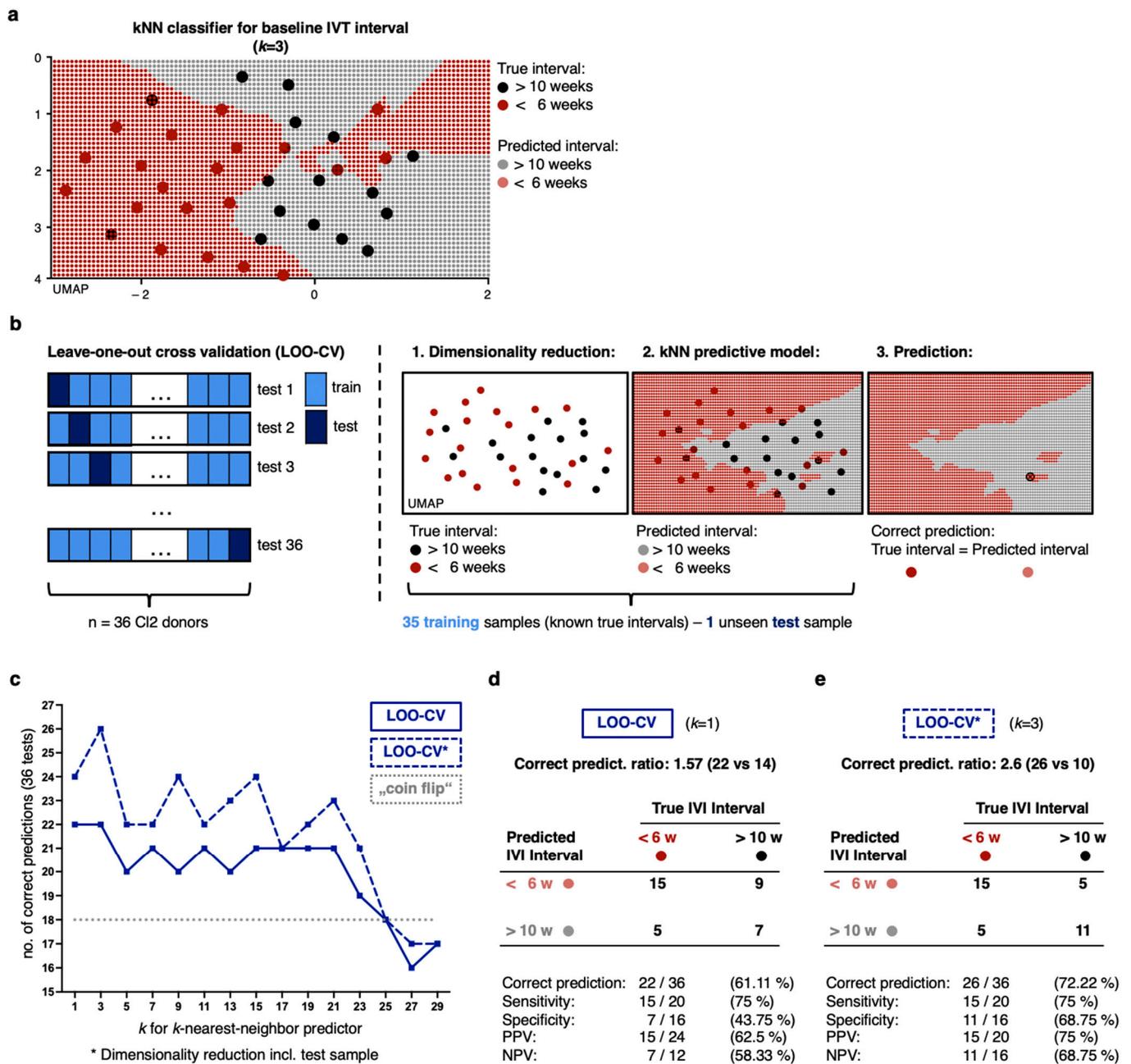
Figure 3. *k*-Nearest-Neighbor approach as a robust non-parametric tool to classify patients into divisions based on proteomic profiles. (**a**) UMAP dimensionality reduction and *k*NN approach (*k* = 3) for prediction of CNV control divisions. (**b**) Left: Leave-one-out cross validation as validation strategy. Center to right: Exemplary visualization for cross validation. (**c**) Number of correct predictions of 36 patients in cluster 2 as a function of the number of *k* (for *k*NN approach). (**d**,**e**) Contingency table for true and predicted division for LOO-CV$_{k=1}$ and LOO-CV*$_{k=3}$.

2.4. Non-Linear Classification Model Detects the Level of CNV Activity Based on Similarities in a High Number of Proteomic Dimensions

As a validation approach, we assessed how well our analysis will generalize to independent data using an adapted leave-one-out cross validation (LOO-CV) approach (compare methods and Figure 3) [45]. With this approach, we exclude one sample from cluster 2, perform the UMAP dimensionality reduction on the remaining 45 samples, and rebuild the *k*NN predictor (Figure 3b, compare methods). We then project the unseen test sample on the novel *k*NN map to classify this unseen test sample into the stratification

divisions. In the example for patient '54492928' as a test sample in Figure 3b, our classification coincides with the true division. We repeat this for all 36 patients of our cluster 2 cohort and receive for the full cross validation (with $k = 3$ nearest neighbors, compare Figures S5 and S6 for different k) 22 correct and 14 incorrect predictions (ratio: 1.57, rate: 61.11%). We also created the UMAP dimensionality reduction for all samples including the test-sample, thereby only testing the kNN classifier during the validation process and not the UMAP algorithm in a complete cross validatory setting (LOO-CV*). Expectedly, we yielded more correct predictions with this (compare below and in Figure 3c–e); however, there are limitations and risk of a potential overfit to our data. Furthermore, finding the right k for a kNN approach is a common challenge. Expectedly, due to the relatively small training set of 35 patients only, we yielded the best performance with a small k of 1 or 3 (compare Figures 3c–e and S6).

Finally, we can demonstrate a superiority of both LOO-CV cross validated models for any reasonably chosen k in comparison to a random ("coin flip") or even a selective classifier for expected value, e.g., CAC classification for all samples (Figure 3c). We yielded best outcome parameters of the fully cross validated model (LOO-CV) for $k = 1$ (the nearest neighbors) and for $k = 3$, respectively, with a ratio of 22 correct versus 14 false classifications and a Rand index (RI) of 0·51 for $k = 1$ (best outcome for LOO-CV* for $k = 3$ with 26 vs. 10, and a RI of 0.59). The provided classifier (LOO-CV$_{k=1}$) is superior for common test statistics, e.g., sensitivity of 75% (LOO-CV*$_{k=3}$: 75%), positive predictive value 62.5% (PPV; LOO-CV*$_{k=3}$: 75%), negative predictive value 58·33% (NPV; LOO-CV*$_{k=3}$: 68.75%), but yields a limited specificity of 43.75% (LOO-CV*$_{k=3}$: 68.75%). With an increasing k for the kNN classification algorithm, we detected even higher values for sensitivity at the expense of a lower specificity (Figure 3 and Figure S6).

3. Discussion

The main finding of this study is the detection of systemic signals in the peripheral blood of nAMD patients, which reflect the macular disease phenotype—in our example the level of CNV control. With this, we provide renewed conceptual support of a tangible influence of systemic factors on the nAMD manifestation and the course of disease. This observation may open new avenues to the mechanistic understanding of nAMD.

Epidemiologically, our study cohort is representative for a real-world nAMD population (compare Supplementary Figure S2 and patient meta-features with the relevant literature) [1,33]. For spectral library generation, a shotgun LC-MS/MS approach was applied, which is simple to perform, and a standard reproducible method of current mass spectrometry, thus feasible for mass diagnostics [46]. However, although the applied UMAP approach stands out as an efficient method with superior runtime performance [47], in silico analysis turned out to be more challenging: We demonstrate in Figure 1 that, for sensitive detection of hidden proteomic patterns of clinical relevance for nAMD, a simultaneous suppression of superior, confounding factors of paramount health conditions is necessary (e.g., systemic diseases, Figure 1). We believe this to be particularly important in an aged nAMD cohort suffering from a multitude of comorbidities with potentially strong systemic representation (compare Supplementary Table S5). In this context, we were able to identify oxidative stress response as molecular and respiratory dysfunctions as correlating clinical signals in a very distinct nAMD subcohort (10 patients in cluster 1, $p = 0.0197$ for respiratory dysfunction, and 131 significantly different proteins with $p < 0.05$, >11% of all proteins, Figure 1). With cytochrome c oxidase subunit II, we found a key component of the mitochondrial respiratory chain, and, with the membrane-bound chaperone calnexin, a protein that is known for its function for protein folding, quality control, glycoprotein maturation, and interaction with other proteins. Both of them are response molecules to oxidative stress, and early reports suggest retinal defects upon deficiency [48,49]. The molecular functions of cluster 1 markers (NR2CA, TPM4, CRAC1) are diverse and nonspecific: inhibition of cell proliferation, migration, epithelial–mesenchymal transition (EMT), cell–cell adhesion, and others [50,51]. Their role as clinical biomarkers remains unclear at this point. Notably,

oxidative stress and respiratory diseases are mutually dependant [52], with us providing the potential molecular link. With this, we identify the only two consistent modifiable risk conditions for nAMD in this subcohort: lung disease (smoking) and oxidative stress response (nutrition, Figure 1, and AREDS2 study results) [1,33,53]. Although not being able to provide a causal proof, our study suggests that, for some nAMD patients, the ophthalmological disease is substantially promoted by or even might represent a secondary complication of a stronger underlying molecular condition, e.g., oxidative stress or a state of systemic hypoxia due to lung disease. However, nAMD manifestation in these patients (cluster 1) did not differ from non-affected nAMD cases (cluster 2) in terms of morphology or functional or epidemiological features (compare Figure 1f). Furthermore, according to the literature, an epidemiological association between pulmonary (dys-)function and AMD is at least unclear; in our study as well, we cannot present a robust causal relationship [54,55]. Further research on this topic is needed to explore potential causalities and to understand how to integrate these findings in terms of clinical considerations.

In Figure 2, we can provide some significant marker correlations of morphological and functional nAMD features (Prefoldin Subunit 2 for SRF; Peroxiredoxin-6 for SHRM, both features determined on OCT scan) with Aldolase C being associated with effective control of CNV activity (Figure 2, compare CAC and ECC criteria in methods). The identified molecules have already been weakly associated with the given conditions but need to undergo further confirmation [40,41,43,44]. Interestingly, anti-retinal auto-antibodies against Aldolase C have been identified as potential disease drivers in nAMD [44]. Thus, higher blood levels of the enzyme as putative decoy molecules might have a protective potential, as we observe in our SVM approach (Figure 2d–f). However, although the provided singular markers have a certain value for target discovery research, they are insufficient to robustly assign patients into the clinically stratified divisions or to draw clinical conclusions regarding therapy or prognosis.

In contrast, in Figure 3 we provide a *k*NN classifier that works based on proteomic similarities (neighborhoods), thereby consulting complex high-dimensional data patterns concealed in numerous proteome features, not in singular biomarkers. With this, we support an nAMD understanding of multifactorial, personalized genesis—in which a high number of molecular factors define disease course and treatment response. At this point, the provided prediction model for the level of CNV controls is statistically inferior to clinical diagnostics. However, it proves that ulterior signals can be detected even in a 'small' dataset of only 46 patients and 1182 proteomic features measured at one baseline timepoint. This represents a cursory glance into the manifold molecular manifestation of nAMD pathogenesis.

Limitations of this study should be acknowledged. Firstly, the inclusion of real-world patients with coexisting diseases from a wide range of backgrounds certainly affects our results. This has upsides in terms of the translatability into clinical practice. On the other hand, the diverse comorbidities present in our patient cohort (Supplementary Table S5) introduce additional variables that may influence the observed proteomic profiles and disease phenotypes. Although our approach can partially but significantly factor out these comorbidities and detect hidden signals of the investigated disease (Figure 1), the presence of other concurrent medical conditions may also introduce some bias. The findings of the present study help to inform follow-up studies with additional strata of subjects including healthy controls and nAMD subjects before treatment. Furthermore, it will be important to expand the number of included molecular dimensions, such as incorporating more detected proteomic variants, genome, metabolome, and immunome parameters. Longitudinal test strategies, coupled with an increase in the number of donors, are crucial to ensure sufficient data density for robust inference. The relatively small sample size of 46 blood samples, divided by CNV activity, is a further limitation. However, BIOMAC should be seen as a pilot investigation showing that the blood proteome in larger AMD cohorts deserves attention. Moreover, the presented classifier model is expected to perform even better when trained on a larger cohort and with additional dimensions, as machine

learning models generally improve with more data. Additionally, the applied machine learning techniques have inherent complexity, making them challenging to understand for patients, clinicians, and scientists alike [56–58]. The specific proteome features that contribute to the classification and how they exert their effects remain largely elusive, limiting conventional target discovery research. However, the outcome of our study sheds light on the heterogeneity of individual courses of nAMD. Considering this individual heterogeneity is crucial in the pursuit of identifying new therapeutic targets beyond VEGF and closely related pathways.

In conclusion, we show that there are so far not considered signals in the blood proteome of patients that correlate with the ocular phenotype of nAMD. The study also indicates that singular biomarkers are insufficient to explain clinical response when considered in isolation. Rather, we assume that decisive molecular processes are controlled by a multitude of proteomic factors in a complex interplay. The presented observations constitute an interesting starting point for future translational research on systemic factors in pathogenesis and the course of AMD.

4. Materials and Methods

4.1. Study Design

This study was part of the BIOMAC study, a cross-sectional observational study on nAMD biomarkers at Charité University Hospital, Berlin, Germany. The research protocol was conducted in accordance with the valid versions of the study protocol, ICH Good Clinical Practice Guidelines (ICH-GCP), and the tenets of the Declaration of Helsinki and was approved by the competent ethics committee of the Charité University Hospital, Berlin. All included participants provided written informed consent and were recruited prospectively.

4.2. Study Protocol and Subject Recruitment

From November 2018 through June 2020, eligible participants meeting all inclusion criteria and none of the exclusion criteria were enrolled at the time of their regular appointments at the department of ophthalmology at Campus Benjamin Franklin (CBF) of Charité University Hospital, Berlin. Charts including imaging results of patients that recently (within preceding 6 month of this study) received anti-VEGF IVT were reviewed retrospectively. Inclusion criteria included both genders $\geq$ 51 years of age, active subfoveal choroidal neovascularization (CNV) secondary to nAMD (all lesion types) in the study eye, $BCVA_{LogMAR} \geq +0.1$ and $\leq +1.3$ in the study eye (in the case that both eyes of an individual patient met the inclusion criteria, the eye with the lower visual acuity was included; in the case of both eyes having equal VA, the eye with the clearest lens and ocular media and the least amount of subfoveal scar or geographic atrophy was selected), and informed written consent. Exclusion criteria included any causes of CNV other than neovascular AMD in the study eye; subretinal hemorrhage in the study eye, which warrants surgical intervention except for intravitreal therapy with anti-VEGF IVT; any contraindication for continuous intravitreal therapy; and any kind of dependency on the investigator or employment by the sponsor or investigator.

As a stratification strategy during the recruiting process, we assigned patients to two distinct divisions based on their CNV activity under anti-VEGF IVT: chronically active CNV (CAC) versus effectively controlled CNV (ECC). Criteria for assignment to the CAC divisions included the following: IVT intervals between the current and the last as well as between the last and penultimate intravitreal injection in the study eye were $\leq$42 days (6 weeks), and CNV was regarded as active in the study eye as evidenced by residual fluid present on OCT at current and the last two visits before the injections. Assignment criteria for the ECC divisions were as follows: The intervals between the current and the last as well as between the last and penultimate intravitreal injection in the study eye were $\geq$70 days (10 weeks), and CNV activity was regarded as controlled in the study eye as evidenced by absent or stable fluid on OCT at current and the last two visits before injections.

4.3. Clinical Examination and Meta-Feature Logging

Visual function of the study eye and the fellow eye were assessed using the ETDRS protocol (The Early Treatment Diabetic Retinopathy Study Group, 1985). All participants received complete bilateral ophthalmologic examination, including a dilated fundus exam. Recruited subjects were bilaterally imaged with Fundus Autofluorescence Imaging (FAF), Optical Coherence Tomography (OCT), Fluorescein Angiography (FA, all: Spectralis, Heidelberg Engineering, Heidelberg, Germany), and Optical Coherence Tomography Angiography (OCT-Angiography, ZEISS Angioplex). Imaging was performed by highly experienced technicians following standard procedures to ensure consistency and high quality in image acquisition. For meta-feature annotation in terms of epidemiological (age, sex) and general health features (smoking status, pulmonary dysfunction, history of smoking, diabetes mellitus, arterial hypertension, profession, medication plans), information was extracted from the electronic patient record. All data relevant to the study were documented soon after measurement by the investigatory team in the clinical software database. Match of meta and proteomic features on individual-patient level occurred at later analysis steps (compare below).

4.4. Sample Collection and Mass Spectrometry Analysis

Due to the length of the paragraph describing this methodological approach in detail, we have summarized it in separate files (File S1, Figures S1 and S2, and Tables S1–S5).

4.5. Statistical Approach and Data Analysis

Due to the length of the paragraph describing this methodological approach in detail, we have summarized it in separate files (File S1, Figures S1 and S2, and Tables S1–S5).

Abbreviations

anti-VEGF	anti-vascular endothelial growth factor
BCVA	best-corrected visual acuity
CAC	chronically active CNV
CNV	choroidal neovascularization
CRT	central retinal thickness
ECC	effectively controlled CNV
ETDRS	The Early Treatment Diabetic Retinopathy Study Group
FAF	fundus autofluorescence
FA	fluorescein angiography
IRC	intraretinal cysts
IVT	intravitreal treatment
*k*NN	*k*-nearest neighbors
LC-MS	liquid chromatography with tandem mass spectrometry
LOO-CV	leave-one-out cross validation
ML	machine learning
nAMD	neovascular age-related macular degeneration
OCT	optical coherence tomography
SHRM	subretinal hyperreflective material
SRF	subretinal fluid
UMAP	Uniform Manifold Approximation and Projection

References

1. Fleckenstein, M.; Keenan, T.D.; Guymer, R.H.; Chakravarthy, U.; Schmitz-Valckenberg, S.; Klaver, C.C.; Wong, W.T.; Chew, E.Y. Age-related macular degeneration. *Nat. Rev. Dis. Prim.* **2021**, *7*, 31. [CrossRef]
2. Chong, V. Ranibizumab for the treatment of wet AMD: A summary of real-world studies. *Eye* **2016**, *30*, 270–286. [CrossRef] [PubMed]
3. Rosenfeld, P.J.; Brown, D.M.; Heier, J.S.; Boyer, D.S.; Kaiser, P.K.; Chung, C.Y.; Kim, R.Y. Ranibizumab for Neovascular Age-Related Macular Degeneration. *N. Engl. J. Med.* **2006**, *355*, 1419–1431. [CrossRef] [PubMed]
4. Mullard, A. FDA approves bispecific antibody for two eye diseases. *Nat. Rev. Drug Discov.* **2022**, *21*, 170. [CrossRef]
5. Heier, J.S.; Khanani, A.M.; Ruiz, C.Q.; Basu, K.; Ferrone, P.J.; Brittain, C.; Figueroa, M.S.; Lin, H.; Holz, F.G.; Patel, V.; et al. Efficacy, durability, and safety of intravitreal faricimab up to every 16 weeks for neovascular age-related macular degeneration (TENAYA and LUCERNE): Two randomised, double-masked, phase 3, non-inferiority trials. *Lancet* **2022**, *399*, 729–740. [CrossRef]
6. Li, M.; Dolz-Marco, R.; Messinger, J.D.; Ferrara, D.; Freund, K.B.; Curcio, C.A. Neurodegeneration, gliosis, and resolution of haemor-rhage in neovascular age-related macular degeneration, a clinicopathologic correlation. *Eye* **2021**, *35*, 548–558. [CrossRef]
7. Kauppinen, A.; Paterno, J.J.; Blasiak, J.; Salminen, A.; Kaarniranta, K. Inflammation and its role in age-related macular degeneration. *Cell. Mol. Life Sci.* **2016**, *73*, 1765–1786. [CrossRef]
8. Ferrara, N. Molecular and biological properties of vascular endothelial growth factor. *J. Mol. Med.* **1999**, *77*, 527–543. [CrossRef]
9. Apte, R.S.; Chen, D.S.; Ferrara, N. VEGF in Signaling and Disease: Beyond Discovery and Development. *Cell* **2019**, *176*, 1248–1264. [CrossRef] [PubMed]
10. Miller, J.W.; Le Couter, J.; Strauss, E.C.; Ferrara, N. Vascular endothelial growth factor A in intraocular vascular disease. *Ophthalmology* **2013**, *120*, 106–114. [CrossRef] [PubMed]
11. Ferrara, N.; Gerber, H.-P.; LeCouter, J. The biology of VEGF and its receptors. *Nat. Med.* **2003**, *9*, 669–676. [CrossRef] [PubMed]
12. Lamb, J.R.; Jennings, L.L.; Gudmundsdottir, V.; Gudnason, V.; Emilsson, V. It's in Our Blood: A Glimpse of Personalized Medicine. *Trends Mol. Med.* **2021**, *27*, 20–30. [CrossRef] [PubMed]
13. Emilsson, V.; Ilkov, M.; Lamb, J.R.; Finkel, N.; Gudmundsson, E.F.; Pitts, R.; Gudnason, V. Co-regulatory networks of human serum proteins link genetics to disease. *Science* **2018**, *36*, 769–773. [CrossRef] [PubMed]
14. Suhre, K.; McCarthy, M.I.; Schwenk, J.M. Genetics meets proteomics: Perspectives for large population-based studies. *Nat. Rev. Genet.* **2020**, *22*, 19–37. [CrossRef]
15. García-Quintanilla, L.; Rodríguez-Martínez, L.; Bandín-Vilar, E.; Gil-Martínez, M.; González-Barcia, M.; Mondelo-García, C.; Fernández-Ferreiro, A.; Mateos, J. Recent Advances in Proteomics-Based Approaches to Studying Age-Related Macular Degeneration: A Systematic Review. *Int. J. Mol. Sci.* **2022**, *23*, 14759. [CrossRef]
16. Yuan, X.; Gu, X.; Crabb, J.S.; Yue, X.; Shadrach, K.; Hollyfield, J.G.; Crabb, J.W. Quantitative proteomics: Comparison of the macular bruch membrane/choroid complex from age-related macular degeneration and normal eyes. *Mol. Cell. Proteom.* **2010**, *9*, 1031–1046. [CrossRef]
17. Joo, J.-H.; Kim, H.; Shin, J.-H.; Moon, S.W. Aqueous humor cytokine levels through microarray analysis and a sub-analysis based on optical coherence tomography in wet age-related macular degeneration patients. *BMC Ophthalmol.* **2021**, *21*, 399. [CrossRef]

18. Rinsky, B.; Beykin, G.; Grunin, M.; Amer, R.; Khateb, S.; Tiosano, L.; Almeida, D.; Hagbi-Levi, S.; Elbaz-Hayoun, S.; Chowers, I. Analysis of the aqueous humor proteome in patients with age-related macular degeneration. *Investig. Opthalmol. Vis. Sci.* **2021**, *62*, 18. [CrossRef]

19. Erie, J.C.; Good, J.A.; Butz, J.A.; Hodge, D.O.; Pulido, J.S. Urinary Cadmium and Age-related Macular Degeneration. *Am. J. Ophthalmol.* **2007**, *144*, 414–418.e1. [CrossRef]

20. Sivaprasad, S.; Chong, N.V.; Bailey, T.A. Serum elastin-derived peptides in age-related macular degeneration. *Investig. Opthalmology Vis. Sci.* **2005**, *46*, 3046–3051. [CrossRef]

21. Kim, H.J.; Ahn, S.J.; Woo, S.J.; Hong, H.K.; Suh, E.J.; Ahn, J.; Park, J.H.; Ryoo, N.K.; Lee, E.J.; Kim, K.W.; et al. Proteomics-based identification and validation of novel plasma biomarkers phospholipid transfer protein and mannan-binding lectin serine protease-1 in age-related macular degeneration. *Sci. Rep.* **2016**, *6*, 32548. [CrossRef] [PubMed]

22. Zhang, P.; Zhu, M.; Geng-Spyropoulos, M.; Shardell, M.; Gonzalez-Freire, M.; Gudnason, V.; Eiriksdottir, E.G.; Schaumberg, D.; Van Eyk, J.E.; Ferrucci, L.; et al. A novel, multiplexed targeted mass spectrometry assay for quantification of complement factor H (CFH) variants and CFH-related proteins 1–5 in human plasma. *Proteomics* **2017**, *17*, 1600237. [CrossRef] [PubMed]

23. Krishnakumar, S.; Sivagurunathan, S.; Selvan, L.N.; Khan, A.; Parameswaran, S.; Bhattacharjee, H.; Gogoi, K.; Gowda, H.; Prasad, T.K.; Pandey, A.; et al. Proteomics-based approach for differentiation of age-related macular degeneration sub-types. *Indian J. Ophthalmol.* **2021**, *69*, 647–654. [CrossRef] [PubMed]

24. Emilsson, V.; Gudmundsson, E.F.; Jonmundsson, T.; Jonsson, B.G.; Twarog, M.; Gudmundsdottir, V.; Li, Z.; Finkel, N.; Poor, S.; Liu, X.; et al. A proteogenomic signature of age-related macular degeneration in blood. *Nat. Commun.* **2022**, *13*, 3401. [CrossRef]

25. Senabouth, A.; Daniszewski, M.; Lidgerwood, G.E.; Liang, H.H.; Hernández, D.; Mirzaei, M.; Keenan, S.N.; Zhang, R.; Han, X.; Neavin, D.; et al. Transcriptomic and proteomic retinal pigment epithelium signatures of age-related macular degeneration. *Nat. Commun.* **2022**, *13*, 4233. [CrossRef]

26. Flores-Bellver, M.; Mighty, J.; Aparicio-Domingo, S.; Li, K.V.; Shi, C.; Zhou, J.; Cobb, H.; McGrath, P.; Michelis, G.; Lenhart, P.; et al. Extracellular vesicles released by human retinal pigment epithelium mediate increased polarised secretion of drusen proteins in response to AMD stressors. *J. Extracell. Vesicles* **2021**, *10*, e12165. [CrossRef]

27. Cai, B.; Zhang, Z.; Sun, S.; Lin, T.; Ke, Y.; Li, Z.; Yang, J.; Li, X. A Pilot Application of an iTRAQ-Based Proteomics Screen Estimates the Effects of Cigarette Smokers' Serum on RPE Cells with AMD High-Risk Alleles. *Transl. Vis. Sci. Technol.* **2022**, *11*, 15. [CrossRef]

28. Koss, M.J.; Hoffmann, J.; Nguyen, N.; Pfister, M.; Mischak, H.; Mullen, W.; Husi, H.; Rejdak, R.; Koch, F.; Jankowski, J.; et al. Proteomics of Vitreous Humor of Patients with Exudative Age-Related Macular Degeneration. *PLoS ONE* **2014**, *9*, e96895. [CrossRef]

29. Nobl, M.; Reich, M.; Dacheva, I.; Siwy, J.; Mullen, W.; Schanstra, J.P.; Choi, C.Y.; Kopitz, J.; Kretz, F.T.; Auffarth, G.U.; et al. Proteomics of vitreous in neovascular age-related macular degeneration. *Exp. Eye Res.* **2016**, *146*, 107–117. [CrossRef]

30. Harrell, C.R.; Markovic, B.S.; Fellabaum, C.; Arsenijevic, A.; Djonov, V.; Volarevic, V. Molecular mechanisms underlying therapeutic potential of pericytes. *J. Biomed. Sci.* **2018**, *25*, 21. [CrossRef]

31. Schori, C.; Trachsel, C.; Grossmann, J.; Zygoula, I.; Barthelmes, D.; Grimm, C. the proteomic landscape in the vitreous of patients with age-related and diabetic retinal disease. *Investig. Opthalmol. Vis. Sci.* **2018**, *59*, AMD31–AMD40. [CrossRef]

32. Winiarczyk, M.; Kaarniranta, K.; Winiarczyk, S.; Adaszek, S.; Winiarczyk, D.; Mackiewicz, J. Tear film proteome in age-related macular degeneration. *Graefe's Arch. Clin. Exp. Ophthalmol.* **2018**, *256*, 1127–1139. [CrossRef] [PubMed]

33. Mitchell, P.; Liew, G.; Gopinath, B.; Wong, T.Y. Age-related macular degeneration. *Lancet* **2018**, *392*, 1147–1159. [CrossRef] [PubMed]

34. Kowal-Bielecka, O.; Chwieśko-Minarowska, S.; Bernatowicz, P.L.; Allanore, Y.; Radstake, T.; Matucci-Cerinic, M.; Broen, J.; Hesselstrand, R.; Krasowska, D.; Riemekasten, G.; et al. The arachidonate 5-lipoxygenase activating protein gene polymorphism is associated with the risk of scleroderma-related interstitial lung disease: A multicentre European Scleroderma Trials and Research group (EUSTAR) study. *Rheumatology* **2017**, *56*, 844–852. [CrossRef]

35. Nakada, E.M.; Sun, R.; Fujii, U.; Martin, J.G. The Impact of Endoplasmic Reticulum-Associated Protein Modifications, Folding and Degradation on Lung Structure and Function. *Front. Physiol.* **2021**, *12*, 665622. [CrossRef]

36. Choi, B.-H.; Kim, J.-S. Age-related decline in expression of calnexin. *Exp. Mol. Med.* **2004**, *36*, 499–503. [CrossRef] [PubMed]

37. Lynch, A.M.; Wagner, B.D.; Weiss, S.J.; Wall, K.M.; Palestine, A.G.; Mathias, M.T.; Siringo, F.S.; Cathcart, J.N.; Patnaik, J.L.; Drolet, D.W.; et al. Proteomic profiles in advanced age-related macular degeneration using an aptamer-based proteomic technology. *Transl. Vis. Sci. Technol.* **2019**, *8*, 14. [CrossRef]

38. Garcia, I.P.L.; Alfaro-Arnedo, E.; Canalejo, M.; Urtubia, A.; Roncero, A.; Ruiz-Martínez, C.; Hoeflich, A.; Höflich, C.; Gotera, C.; Barba, G.P.; et al. Insulin-Like Growth Factors and IGF-Binding Proteins levels in serum from smokers and patients with different grades of COPD, COPD and lung cancer and exacerbated COPD. *Eur. Respir. J.* **2021**, *58*, PA1952. [CrossRef]

39. Zhang, Y.; Kent, J.; Olivier, M.; Ali, O.; Broeckel, U.; Abdou, R.; Dyer, T.; Comuzzie, A.; Curran, J.; Carless, M.; et al. QTL-based association analyses reveal novel genes influencing pleiotropy of metabolic syndrome (MetS). *Obesity* **2013**, *21*, 2099–2111. [CrossRef]

40. Liang, J.; Xia, L.; Oyang, L.; Lin, J.; Tan, S.; Yi, P.; Han, Y.; Luo, X.; Wang, H.; Tang, L.; et al. The functions and mechanisms of prefoldin complex and prefoldin-subunits. *Cell Biosci.* **2020**, *10*, 87. [CrossRef]

41. Huu, T.N.; Park, J.; Zhang, Y.; Park, I.; Yoon, H.J.; Woo, H.A.; Lee, S.-R. Redox Regulation of PTEN by Peroxiredoxins. *Antioxidants* **2021**, *10*, 302. [CrossRef]

42. Noble, W.S. What is a support vector machine? *Nat. Biotechnol.* **2006**, *24*, 1565–1567. [CrossRef] [PubMed]

43. Adamus, G. Impact of autoantibodies against glycolytic enzymes on pathogenicity of autoimmune retinopathy and other au-toimmune disorders. *Front. Immunol.* **2017**, *8*, 505. [CrossRef] [PubMed]

44. Morohoshi, K.; Ohbayashi, M.; Patel, N.; Chong, V.; Bird, A.C.; Ono, S.J. Identification of anti-retinal antibodies in patients with age-related macular degeneration. *Exp. Mol. Pathol.* **2012**, *93*, 193–199. [CrossRef] [PubMed]

45. Lever, J.; Krzywinski, M.; Altman, N. Model selection and overfitting. *Nat. Methods* **2016**, *13*, 703–704. [CrossRef]

46. Bian, Y.; Zheng, R.; Bayer, F.P.; Wong, C.; Chang, Y.-C.; Meng, C.; Zolg, D.P.; Reinecke, M.; Zecha, J.; Wiechmann, S.; et al. Robust, reproducible and quantitative analysis of thousands of proteomes by micro-flow LC–MS/MS. *Nat. Commun.* **2020**, *11*, 157. [CrossRef]

47. McInnes, L.; Healy, J.; Saul, N.; Großberger, L. UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **2018**, *3*, 861. [CrossRef]

48. Brischigliaro, M.; Zeviani, M. Cytochrome c oxidase deficiency. *Biochim. Biophys. Acta Bioenerg.* **2021**, *1862*, 148335. [CrossRef]

49. Kraus, A.; Groenendyk, J.; Bedard, K.; Baldwin, T.A.; Krause, K.H.; Dubois-Dauphin, M.; Dyck, J.; Rosenbaum, E.E.; Korngut, L.; Colley, N.J.; et al. Calnexin deficiency leads to dysmye-lination. *J. Biol. Chem.* **2010**, *285*, 18928–18938. [CrossRef]

50. Li, Y.; Zhang, Y.; Wu, Z.; Sun, P. Overexpression of TPM4 is associated with worse prognosis and immune infiltration in patients with glioma. *BMC Neurol.* **2023**, *23*, 17. [CrossRef]

51. Jaeger, E.; members of the CORGI Consortium; Webb, E.; Howarth, K.; Carvajal-Carmona, L.; Rowan, A.; Broderick, P.; Walther, A.; Spain, S.; Pittman, A.; et al. Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. *Nat. Genet.* **2007**, *40*, 26–28. [CrossRef] [PubMed]

52. Forman, H.J.; Zhang, H. Targeting oxidative stress in disease: Promise and limitations of antioxidant therapy. *Nat. Rev. Drug Discov.* **2021**, *20*, 689–709. [CrossRef] [PubMed]

53. Chew, E.Y.; Clemons, T.E.; SanGiovanni, J.P.; Danis, R.; Ferris, F.L.; Elman, M. Lutein + zeaxanthin and omega-3 fatty acids for age-related macular degeneration: The Age-Related Eye Disease Study 2 (AREDS2) randomized clinical trial. *JAMA* **2013**, *309*, 2005–2015.

54. Klein, R.; Klein, B.E.; Linton, K.L. Prevalence of age-related maculopathy. The Beaver Dam Eye Study. *Ophthalmology* **1992**, *99*, 933–943. [CrossRef] [PubMed]

55. Moorthy, S.; Cheung, N.; Klein, R.; Shahar, E.; Wong, T.Y. Are lung disease and function related to age-related macular degeneration? *Am. J. Ophthalmol.* **2011**, *151*, 375–379. [CrossRef] [PubMed]

56. De Fauw, J.; Ledsam, J.R.; Romera-Paredes, B.; Nikolov, S.; Tomasev, N.; Blackwell, S.; Ronneberger, O. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat. Med.* **2018**, *24*, 1342–1350. [CrossRef]

57. Stuart, T.; Butler, A.; Hoffman, P.; Hafemeister, C.; Papalexi, E.; Mauck, W.M.; Hao, Y.; Stoeckius, M.; Smibert, P.; Satija, R. Comprehensive Integration of Single-Cell Data. *Cell* **2019**, *177*, 1888–1902. [CrossRef]

58. Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.S.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* **2018**, *172*, 1122–1131.e9. [CrossRef]