Molecular Vision 2018; 24:875-889 <http://www.molvis.org/molvis/v24/875>
Received 9 July 2018 | Accepted 31 December 2018 | Published 31 December 2018

Comprehensive characterization of the adult ND4 Swiss Webster mouse retina: Using discovery-based mass spectrometry to decipher the total proteome and phosphoproteome

Jarrod C. Harman,[1,2,3] Jessie J. Guidry,[4,5] Jeffrey M. Gidday[1,2,3]

[1]Department of Ophthalmology, Louisiana State University Health Science Center (LSUHSC), New Orleans, LA; [2]Department of Physiology, LSUHSC, New Orleans, LA; [3]Neuroscience Center of Excellence, LSUHSC, New Orleans, LA; [4]Department of Biochemistry and Molecular Biology, LSUHSC, New Orleans, LA; [5]Proteomics Core Facility, LSUHSC, New Orleans, LA

Purpose: Diverse groups of proteins play integral roles in both the physiology and pathophysiology of the retina. However, thorough proteomic analyses of retinas of experimental species are currently unavailable. The purpose of the present paper is providing the field with a comprehensive proteomic characterization of the retina of a commonly used laboratory mouse using a discovery-based mass spectrometry (MS) approach.

Methods: Retinas from eight male and eight female 30-week-old outbred ND4 Swiss Webster mice were harvested and immediately processed for MS analysis on a Thermo Fisher (TF) Fusion Orbitrap MS. The retinal proteome and phosphoproteome were identified and subsequently analyzed using Proteome Discoverer 2.2 and Panther-GeneGo. SEQUEST-HT scoring was used for analysis, and the reference protein FASTA database was from *Mus musculus*. Specifically, three technical repeats were performed for each biological sample. For characterization, only high-scoring peptides were considered, with a false discovery rate (FDR) of <1%. Downstream bioinformatic analysis used Ingenuity Pathway Analysis (IPA; Qiagen).

Results: Using Proteome Discoverer 2.2, 4,767 different proteins were identified and segregated into 26 major protein classes, 9 functional molecular classes, and 12 categories of biological processes. The five largest protein classes included the following: nucleic acid binding (17%), hydrolases (13%), enzyme modulators (10%), transferases (9%), and oxidoreductases (6%). “Binding” and “catalytic” proteins contributed to 81% of the molecular function class at 37% and 42%, respectively. “Cellular processing” and “metabolic processes” contributed the most to biologic activity, at 31% and 26%, respectively. Phosphopeptide enrichment yielded the identification of 610 additional unique proteins that were not originally identified. The two datasets combined produced an adult mouse retinal proteome consisting of 5,377 unique proteins. Overall, 41% of the retinal proteome was phosphorylated. The overwhelming diversity of retinal protein functionality was reflected through further analyses revealing 2,086 unique pathway hits across 241 different pathways (TF). A core analysis summary report was performed in IPA (Qiagen) to analyze the top signaling networks, protein–protein interaction (PPI) enrichments, and canonical pathways.

Conclusions: Using this high-throughput technique, we have further deciphered and updated the diverse proteome of the mouse retina, including the phosphoproteome, thereby providing the most comprehensive proteomic profile for this tissue known to date. These findings, and the bioinformatic analyses we also provided, establish a platform for future studies, facilitating the elucidation of the relevance of these proteins to the molecular and cellular pathologies that underlie retinal function and disease.

To date, the retinal proteome of the normal adult mouse has yet to be comprehensively characterized. Given that the Human Eye Proteome Project (EyeOme) started in 2012 [1], we find it striking that the steady-state proteomic profile for the adult mouse retina, now so commonly used in ophthalmological research, has not been characterized and made publicly available. Ocular proteomic studies to date include analyses of tissues, or samples (e.g., tears, vitreous) [2], obtained from rodents, pigs, and even humans [3,4], as well as from cell culture [5]. Despite many animal models of ocular pathology, the proteomes associated with these diseases are still largely unstudied [6]. According to a recent review, only 23 proteomic papers and 27 human eye supplementary projects have been conducted, despite a global estimate of 253 million people worldwide living with some degree of visual impairment [1].

Advancements in both the technical and analytical capabilities of mass spectrometry (MS) have allowed for tremendous improvements in the detection and identification of low-abundance molecules. When coupled to liquid chromatography systems, in conjunction with fractionation and enrichment techniques, more information—and thus,

Correspondence to: Jeff M. Gidday, Department of Ophthalmology, Louisiana State University School of Medicine, 533 Bolivar St., CSRB Room 454, New Orleans, LA 70112; Phone: (504) 568-2360; email: jgidda@lsuhsc.edu

deeper analysis—can be obtained from samples dedicated to MS-based proteomics; this not only yields more proteins, but in turn, also allows for a deeper analyses of the proteome at the bioinformatics level. Advances in the field of proteomics forthcoming from high-throughput, low-abundance molecular enrichment techniques appear to be growing rapidly in other fields, but much less so in ophthalmology. Currently, UniProt estimates that there are over 20,000 expressed proteins in the mouse proteome [7]. In a 2011 MS-based proteomics study, a peptide-labeling technique was implemented, allowing researchers to detect and identify 2,173 proteins in the mouse cornea [8]. The same technique was applied 3 years later, and in conjunction with improvements in detection capabilities, 3,882 proteins were identified in the mouse retina during myopia progression [9]. Last year, the largest human RPE proteome to date, consisting of 5,309 identified proteins, was reported from cadaver tissue samples [3]. Thus, MS-based preclinical studies can provide powerful datasets that are difficult to match in terms of the insights they provide; not surprisingly, such studies are progressively growing in number.

The power of today's MS detection technology is manifested by the growing identification of tissue biomarkers for detecting and treating diseases across all disciplines of medicine [10,11]. However, improvements in the capabilities of detectors are not the only recent advancement in the proteomics realm. Enrichment techniques for specific post-translational modifications, such as sumoylation [12], acetylation, ubiquitination [13], and phosphorylation [14,15] events, have contributed significantly to identifying previously undetectable peptides, and in turn, they have facilitated the development of novel approaches targeting low abundance detection of these peptides.

The purpose of the present paper is to utilize the techniques described above to provide the field with a comprehensive proteomic characterization of the retina of a commonly used laboratory mouse. We share the belief advanced by other retinal "omic" scientists that, to understand pathophysiology, we must first comprehend normal baselines. Employing a discovery-based MS approach for qualitatively characterizing the resting proteome of the adult mouse retina, we report the largest number of mouse retinal proteins to date at 5,377, including over 2,000 phosphorylated proteins identified by an enrichment methodology we employed; of these, over 600 were exclusive to the enriched dataset (meaning that they would otherwise go undetected) and over 100 proteins were previously unreported in retinal tissue. In addition to meeting the demand for ongoing updates of tissue-specific databases and protein/peptide libraries that are necessary to advance algorithm-based bioinformatic platforms that give context and meaning to "big data"/omic experiments, our findings for provide a springboard for future ophthalmological research in fundamental retinal physiology and pathophysiology.

METHODS

Animals: Outbred ND4 Swiss Webster mice (Envigo, Indianapolis, IN) were shipped to our institution at 10 weeks of age and maintained in our institutional vivarium for 5 months in compliance with all Louisiana State University Health Sciences Center (LSUHSC) and Institutional Animal Care and Use Committee (IACUC) regulations. Both retinas from eight male and eight female 30-week-old mice were harvested and immediately processed for MS-based proteomic analysis. Retinal tissues was harvested in a manner such that Bruch's membrane, choriocapillaris, and the eyecup (RPE) were not dissected and included with the retina sample we analyzed. This was achieved by a single cut across the sclera with a #10-blade scalpel. The retina and lens were removed and separated from one another using sterile forceps.

MS sample collection and preparation: A comprehensive proteomic profile of the adult mouse retina was achieved by analyzing the combination of multiple animals' freshly harvested retinas, followed by peptide enrichment procedures, and finally, bioinformatic analysis. The overall workflow of the study is depicted in Figure 1. Freshly harvested retinas were placed in a 1% sodium dodecyl sulfate (SDS) solution and immediately homogenized. The samples were heated to 80 °C for 10 min, and the protein concentration was determined using a bicinchoninic acid (BCA) protein assay kit (Pierce, Thermo Fisher Scientific, Waltham, MA). Based on the protein concentration, 100 µg of protein sample was adjusted to a final volume of 100 µl using 100 mM tetraethylammonium bromide (TEAB), then reduced using 10 mM Tris (2-carboxyethyl) phosphine hydrochloride (TCEP) at 55 °C for 1 h. The reduced samples were subsequently alkylated using 5 µl of 375 mM iodoacetamide at room temperature for 30 min. Following this, the protein samples were precipitated by chloroform-methanol extraction, air dried, and digested with 2 µg of trypsin (Trypsin Protease, MS Grade, Thermo Fisher Scientific) at 37 °C overnight.

Acidic reverse-phase conditions were employed for desalting the samples before drying them in a Speedvac. After drying, an offline basic pH reverse-phase fractionation step was employed to reduce the complexity. The samples were brought up in 260 ul of 10 mM Optima Grade ammonium hydroxide, pH 10, and separated using an Acquity Ultra Performance LC (UPLC) Peptide Ethylene Bridged Hybrid (BEH) column, 300A, 1.7 um, 2.1 × 50 mm (Waters,

Ireland) on a Dionex U3000 High Perfomance LC (HPLC) system (Thermo Fisher Dionex, Sunnyvale, CA). Following fractionation, samples were ultraviolet (UV) monitored at 215 nm for an injection of 130 µl at 0.1 ml/min with a gradient developed from 10 mM ammonium hydroxide, pH = 10, to 100% acetonitrile (ACN) over 96 min. A total of 48 fractions (200 µl each) were collected in a 96-well microplate and recombined in a checkerboard fashion to create 12 "super fractions" (original fractions 1, 13, 25, and 37 became new super fraction #1; original fractions 2, 14, 26, and 38 became new super fraction # 2; etc.) [16].

The 12 "super fractions" were then run on a Dionex U3000 nanoflow system coupled to a Thermo Fusion Orbitrap mass spectrometer. Each fraction was subjected to the following: 1) 120-min chromatographic method employing a gradient of 2–25% ACN in 0.1% formic acid (ACN/FA) over the course of 100 min, 2) a gradient to 50% ACN/FA for an additional 10 min, 3) a step to 90% ACN/FA for 5 min, and 4) a re-equilibration into 2% ACN/FA in a "trap-and-load" configuration. The trap column used an Acclaim C18 PepMap100, 5 um, 100A trap and the separation column used an Acclaim PepMap Rapid Separation Liquid Chromatography (RSLC) 75 um × 15 cm (Thermo Fisher Scientific). The entire run employed a 0.3 µl/min flow rate and sample ionization through a Thermo Nanospray Flex Ion Source. Data dependent scans were run on a Thermo Fisher Fusion Tribrid Mass Spectrometer. MS1 scans were performed in the Orbitrap using a resolution of 240,000. Data-dependent MS2 scans were performed in the Orbitrap using high-energy collision dissociation (HCD) of 30% using a resolution of 30,000. Automatic Gain Control was employed during MS2 acquisition with a setting of 5×10^4, and dynamic exclusion was set at 10 s. This was repeated for a total of three technical replicates per "super fraction."

Phosphopeptide enrichment: Protein phosphorylation patterns can be difficult to measure without enrichment techniques due to biochemical factors, including the following: low stoichiometry, poor ionization of peptides, incomplete fragmentation of peptides, and high hydrophilicity. Thus, we performed a sequential enrichment of metal oxide affinity

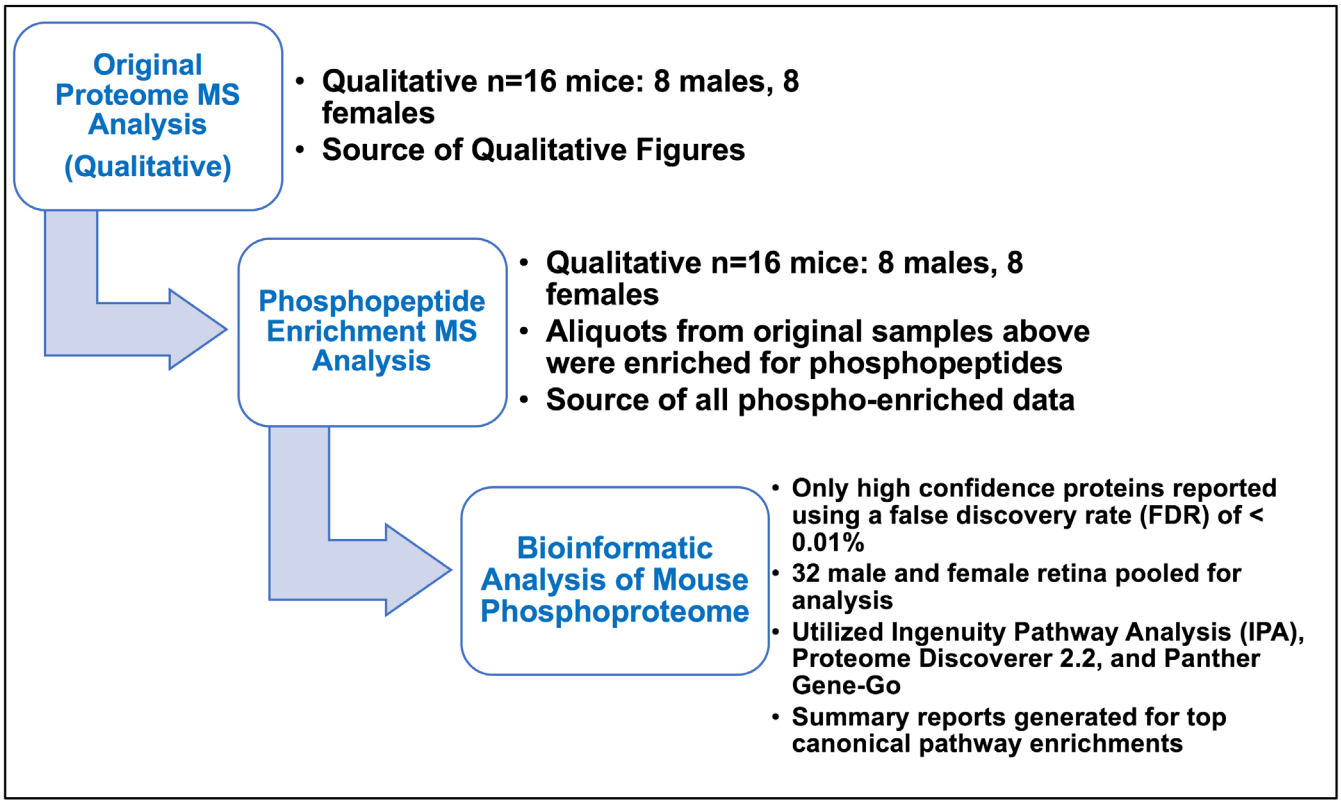

Figure 1. To characterize the retinal proteome/phosphoproteome in its entirety, we performed several qualitative mass spectrometry (MS) runs from eight male and eight female mice; pooling the retinas from each animal yielded a total of 32 retinas for each sample, with three technical repeats per sample. Following the initial MS analysis, 10-ul sample aliquots were pooled together to perform phosphopeptide-enrichment procedures. Upon acquisition of the MS data, Ingenuity Pathway Analysis (IPA) software was used to perform bioinformatic analysis.

chromatography (SMOAC) using both TiO_2 and Fe-NTA resins (Thermo Fisher Scientific). Following retina dissections and initial sample preparation, 10-µl sample aliquots from each animal were pooled together (1.6 mg of total protein). After reduction, alkylation, and trypsin digestion, the peptides were processed using a High-Select™ TiO_2 Phosphopeptide Enrichment Kit (Thermo Fisher Scientific, Cat. No. A32993) followed by a High-Select Fe-NTA phosphopeptide enrichment kit (Thermo Fisher Scientific, Cat. No. A32992) per the manufacturer's specifications. The enriched phosphopeptides were then analyzed via MS, as described previously. The overall phosphopeptide workflow (Figure 2) essentially consisted of lysis, reduction, alkylation, digestion, phosphopeptide enrichment, and reverse-phased peptide fractionation before loading into the MS.

Data analyses and bioinformatics: Initial data analysis was performed with Proteome Discoverer 2.2 using SEQUEST HT scoring and comparing PEP posterior error probability (PEP) scores and peptide spectral matches (PSMs). Only one unique high-scoring peptide was required for inclusion of protein identification in the results. High-scoring peptides have less than a 1% false discovery rate (FDR). Peptides on the lower scoring end were evaluated manually to confirm b and y ion presence in the MS2 spectra. Subsequent bioinformatic analyses was performed using GeneGo (Panther) before loading into Qiagen's Ingenuity Pathway Analysis (IPA). The Protein FASTA database is for *M. musculus* (TaxID = 10090) version 2017–07–05. Static modifications included carbamidomethyl on cysteines (= 57.021), and dynamic modification of oxidation of methionine (= 15.9949). Parent ion tolerance was set at 10 ppm and fragment mass tolerance to 0.6 Da, as recommended for the initial orbitrap tolerances. The maximum number of missed cleavages for trypsin was set to 2. Following completion of data acquisition and analysis in Thermo and GeneGo, our dataset was uploaded for bioinformatics using IPA. Multiple bioinformatic platforms were used to cross-reference proteins and resolve inconsistencies that may arise due to variable protein identifications (IDs) across different libraries, incorrect classifications, or scheduled update differences between software packages.

All confidently identified peptides (those with SEQUEST-HT scores > 0 or PEP scores < 10 that were individually validated by spectral viewing), including proteins identified by one unique peptide, were exported to Microsoft

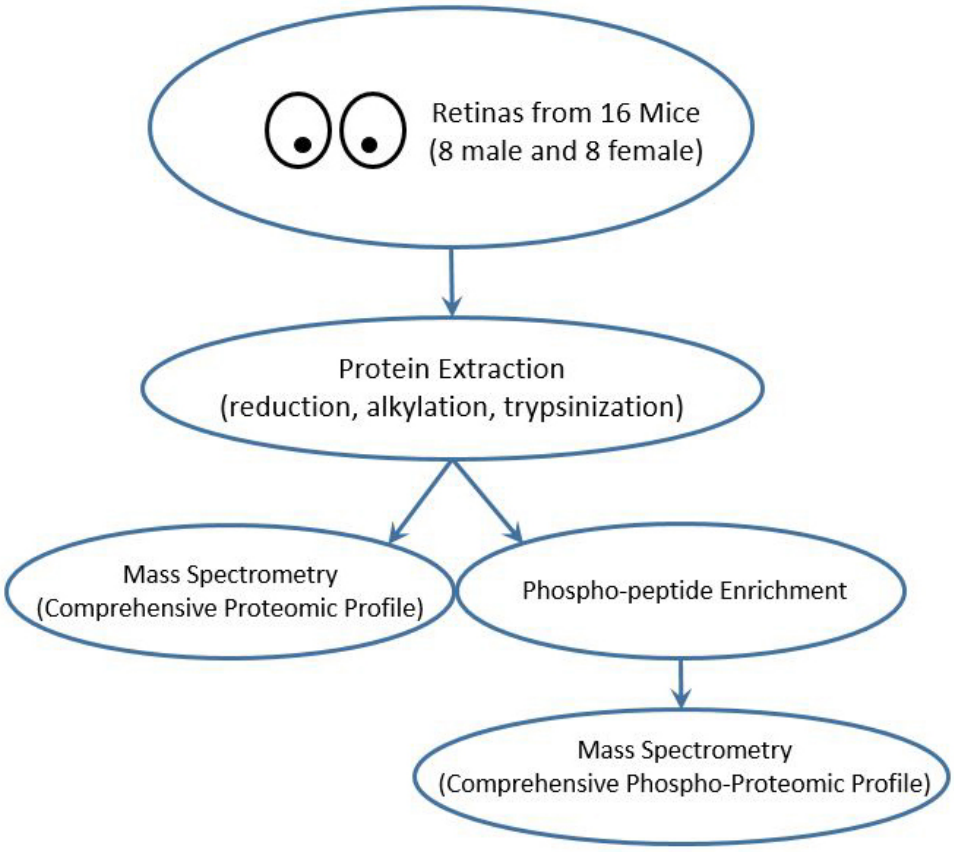

Figure 2. Simplified comprehensive proteome and phosphoproteome workflow: Retinas from 16 ND4 Swiss Webster mice (eight males and eight females) were dissected, pooled together, proteins extracted, reduced, alkylated, and trypsinized for proteomic identification.

Excel. When annotating and applying filters for the dataset uploaded to IPA, the following criteria were applied: Only < 1% FDR SEQUEST-HT scoring peptides were analyzed, only high experimentally observed molecules were factored, the IPA organism settings were mouse (in vivo), and high confidence only. These filters were applied in a "core-analysis" of the entire dataset. The content version release date of the IPA used was June 14th, 2018. The bioinformatics reported herein are descriptive of the entire pooled dataset and are not individual comparisons. Herein, we report a core analysis summary consisting of the following: top canonical pathways; signaling networks; and the disease, molecular, and biological functions of the MS-derived retinal proteome.

RESULTS

Pre-enriched retinal proteome: There were 4,767 different proteins identified in the normal, steady-state, adult ND4 Swiss Webster retina. As shown in Figure 3, the identified proteins could be segregated into 26 major protein classes, 9 functional molecular classes, and 13 categories of biological processes (Panther Gene Ontology, GeneGo). The five largest major protein classes included the following: nucleic acid binding (17%), hydrolases (13%), enzyme modulators (10%), transferases (9%), and oxidoreductases (6%). Binding and catalytic proteins contributed to 79% of the molecular function class, at 37% and 42%, respectively. Cellular processing and metabolic processes contributed the most to biological activity, at 31% and 26%, respectively. The overwhelming diversity of retinal protein functionality was reflected through subsequent bioinformatic analyses, revealing 2,169 unique pathway hits across 307 different pathways (Proteome Discoverer 2.2 software, Thermo Fisher Scientific). Of the 4,767 proteins identified, only 269 phosphopeptides (5.6%) were detected (Figure 4A). Table 1 lists the top 25 scoring (abundance ranked by SEQUEST-HT) proteins identified in the MS run.

Retinal phosphoproteome: The phosphoproteome dataset contained a total of 2,043 proteins (Figure 4B). We identified 1,864 proteins with at least one phosphorylation site. Of that number, 1,540 (82.6%) contained a phosphorylated serine, 904 (48.5%) contained a phosphorylated threonine, and 208 (11.2%) contained a phosphorylated tyrosine. This phosphoenriched dataset identified 610 unique proteins that were not found in the initial MS run. For the phosphopeptides in this dataset, PSMs were used for confirming phosphorylation on serine, threonine, or tyrosine residues. Comparing the enriched dataset to the original proteome, there was an overall reduction of approximately 2.5-fold in the total proteins identified (purification) and a sixfold increase in phosphorylated proteins detected (phosphopeptide enrichment). It is noteworthy that 179 (9%) of the phosphoenriched peptides were not phosphorylated. There are two viable explanations for this observation, as follows: (A) the proteins had liable phosphorylation (histidine and other noncanonical amino acids can be modified but do not hold together during MS analysis), or (B) the peptides had chemical properties with affinity to the phosphorylation enrichment chromatography media (those containing a high number of residues, such as aspartic or glutamic acids) and copurified with phosphorylated peptides. Table 2 lists the top 25 scoring (abundance ranked by SEQUEST-HT) proteins identified in this MS run.

Total retinal proteome: Comparing the nonenriched retinal proteome with the phosphoproteome indicated that, overall, 41% of the mouse retina contains peptides that have been posttranslationally modified by phosphorylation events. Combining the two datasets and removing duplicates produced a master list of 5,377 different retinal proteins identified in the mouse retina (Figure 4C). It is noteworthy that 1,698 out of the 5,377 (32%) proteins were identified by one unique peptide. Of these 1,698, 466 were phosphorylated proteins from our enrichment that were also characterized by the "one unique peptide" criterion, likely as a result of this phosphorylation status. Of the remaining 1,222 proteins, over 100 of these (118, or 9.7% of the total) have yet to be identified in the retina, based on the dataBase for gene expression evolution (Bgee) analysis platform linked to UniProt).

Bioinformatic analysis: Due to the qualitative and descriptive nature of this study, we provide a core analysis summary report that was performed in IPA (Qiagen) to analyze the top signaling networks, protein–protein interaction (PPI) enrichments, and canonical pathways—identified and defined by IPA—that represent protein enrichments of the normal retina of the adult Swiss Webster mouse. Figure 5 illustrates the top 22 canonical pathways enriched in mouse retinal tissue. These were selected based on the coverage and overlap of the number of proteins known to be associated with the described pathways, as well as the confidence of the peptide FDR (<1%) we used. Each signaling pathway represented shares a minimum of 50 proteins with one or more neighboring signaling pathways. Table 3 includes the top five canonical pathways with coverage and confidence scores, as well as the top five associated network functions.

Figure 6 expands on the prior core analysis and canonical pathway report, illustrating protein enrichments in our dataset among a host of other protein networks that IPA organizes by three distinct categories, as follows: Diseases and Disorders (Figure 6A), Cellular and Molecular Functions (Figure 6B), and Physiologic System Development and Functions (Figure

6C). We have also assembled cellular geographic distribution figures for protein networks that are only derived from IPA's core analysis, as shown in Figure 7. These proteins were chosen from an IPA-defined network of ophthalmic disease pathways specific to the retina and known retinal proteins, including the following: abnormal morphology of the retina (Figure 7A), containing 99 proteins; abnormal morphology of photoreceptors (Figure 6B), containing 50 proteins; retinal degeneration (Figure 6C), containing 71 proteins; degeneration of photoreceptors (Figure 6D), containing 42 proteins; and abnormal electrophysiology of the eye (Figure 6E), containing 46 proteins. This also allowed us to further curate a table, Appendix 1, consisting of categorical function annotations of proteins identified in this dataset that are classified under "Ophthalmic Diseases" by Qiagen.

DISCUSSION

The detailed proteome of the normal, steady-state retina of the adult ND4 Swiss Webster mouse, including the phosphoproteome, that we provide herein (see Appendix 2) is of paramount importance for advancing our understanding of the normal biochemistry and physiology of this tissue, as well as providing a baseline for the modification of proteins and pathways/networks induced by disease and lifestyle interventions (diet, exercise). Elucidating drug mechanisms of action, and discovering and identifying new therapeutic targets, are also facilitated by such analyses. Our investigation yielded

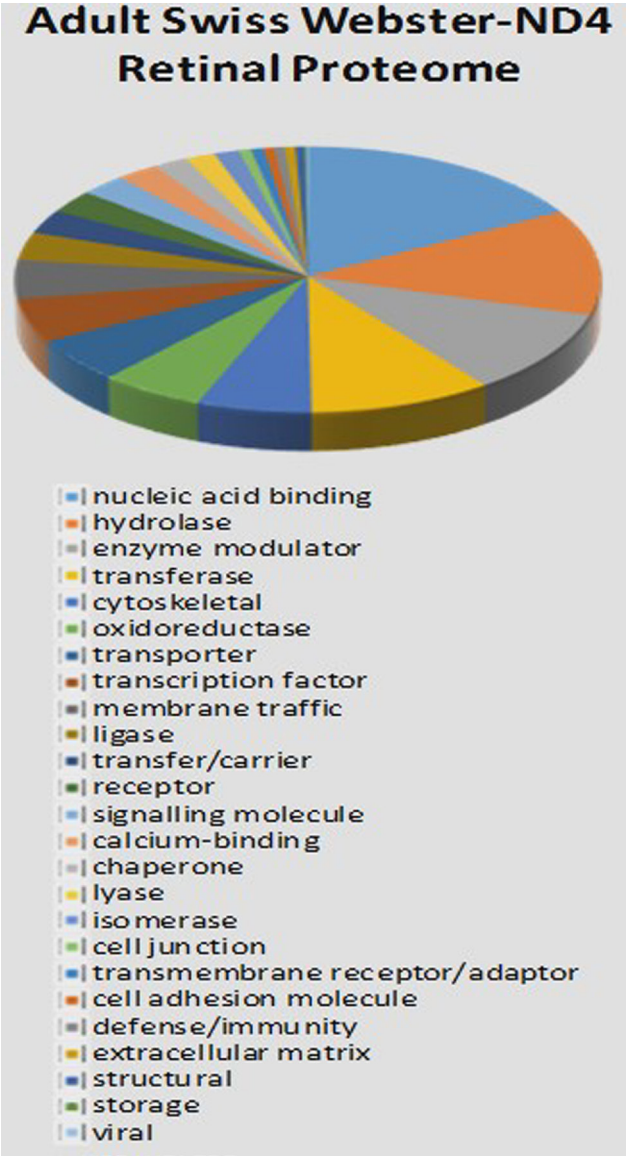

Figure 3. Simple visualization of the retinal proteome by biological function: This pie chart illustrates the top 26 retinal protein functional classifications of our nonenriched retinal proteome (4,767 proteins).

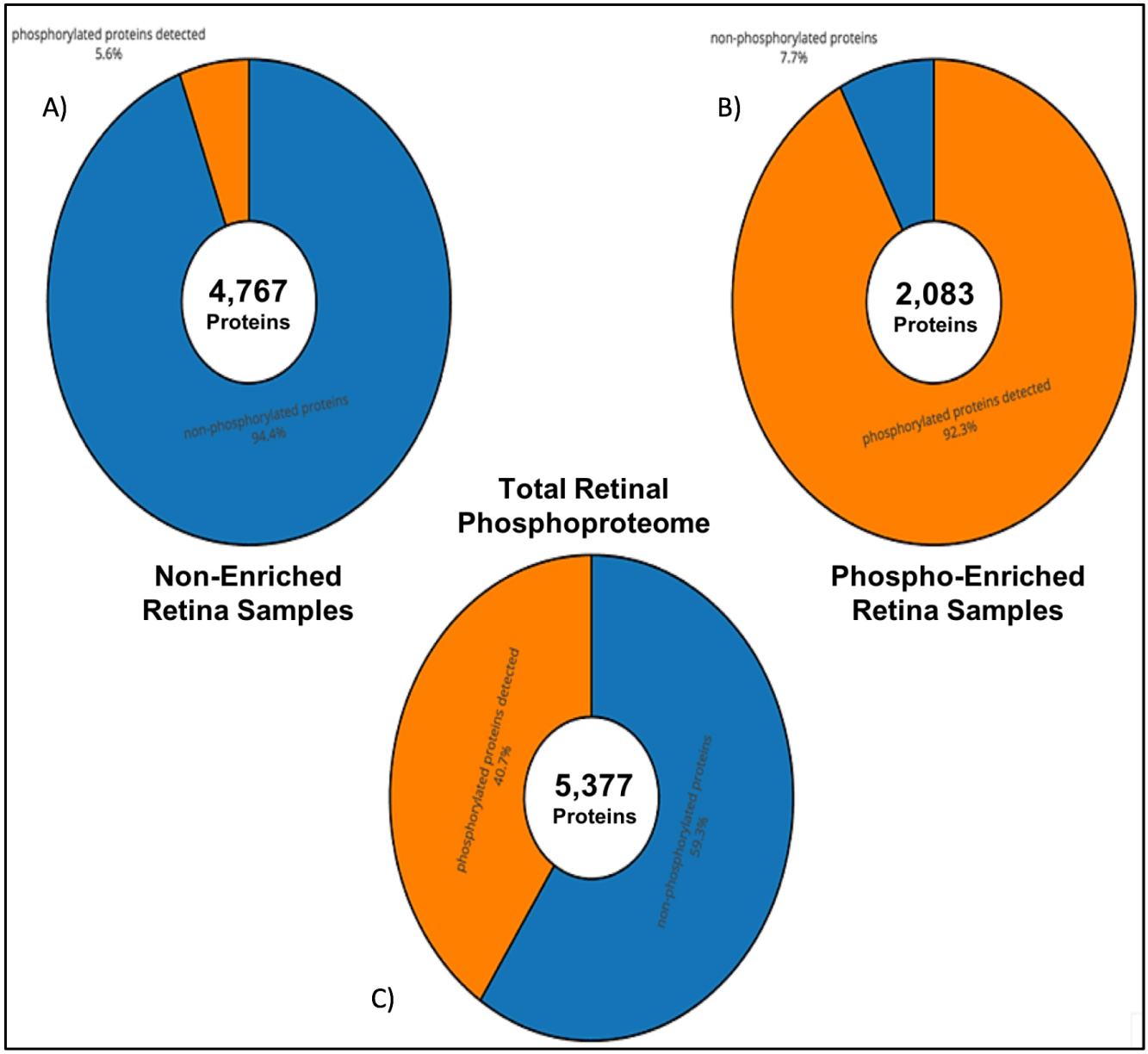

Figure 4. Comparison of phosphorylated/non-phosphorylated proteins across datasets. **A**: Original qualitative proteome composition; **B**: phosphopeptide enrichment performed on (A). **C**: Total retinal proteome composition from all three datasets. (Figure produced using Bio Vinci data visualization software, BioTuring, San Diego, CA).

the largest proteome for this tissue to date, and as such, can serve as a foundation for future quantitative comparisons secondary to experimental intervention, disease, or treatment.

It is important to note at the outset that several methodological, technological, and analytical differences can affect proteome quality and utility, in addition to the experimental rigor needed on the "front end" in terms of reporting the factors that can affect protein expression in a given cell or tissue in a given animal at a given time. These biological variables often go unreported or are assumed to be of minor or no significance, but in most instances, this is unlikely to be true. Especially, animal vendor, strain, environmental factors (light, temperature, etc.), housing conditions (singly or group housed), circadian rhythms, diet, age, gender, and estrous cycle [17] can influence the proteome of interest. For example, Lee et al. [18] reported light-induced differential expression of crystalline protein phosphorylation events in

Table 1. Top 25 highest scoring proteins from the original non-enriched data set. SEQUEST-HT scoring was used to rank proteins identified.

Accession	Description	Score Sequest HT	Sum PEP Score	# PSMs	MW [kDa]	calc. pI
P17182	alpha-enolase	17,064.94	486.77	6359	47.1	6.80
P20443	S-arrestin	15,901.53	558.30	5005	44.9	5.80
P16858	glyceraldehyde-3-phosphate dehydrogenase	12,287.32	462.88	4346	35.8	8.25
P62806	histone H4	9289.99	144.90	3966	11.4	11.36
Q8R1M2	Histone H2A.J	11,790.53	93.21	3664	14.0	11.05
P52480	Pyruvate kinase PKM	8683.77	639.67	3663	57.8	7.47
P24622	Isoform 2 of Alpha-crystallin A chain	9122.78	218.74	3301	19.8	6.20
P52480	Isoform M1 of Pyruvate kinase PKM	7578.32	601.67	3297	57.9	7.14
P24622	Alpha-crystallin A chain	7796.96	231.30	2853	22.5	6.86
P99024	tubulin beta-5 chain	6137.55	511.38	2707	49.6	4.89
P27661	Histone H2AX	5508.91	70.28	2570	15.1	10.74
P68372	Tubulin beta-4B chain	5916.98	532.34	2554	49.8	4.89
Q8CGP2	Isoform 2 of Histone H2B type 1-P	5841.76	104.21	2503	15.6	10.24
Q7TMM9	Tubulin beta-2A chain	5662.93	522.47	2448	49.9	4.89
Q9CWF2	Tubulin beta-2B chain	5564.75	526.13	2400	49.9	4.89
P68369	tubulin alpha-1A chain	5739.92	350.80	2364	50.1	5.06
Q04447	Creatine kinase B-type	6342.71	447.14	2261	42.7	5.67
P20152	Vimentin	4970.02	432.00	2188	53.7	5.12
P17183	Gamma-enolase	6245.82	390.49	2147	47.3	5.11
P0C0S6	Histone H2A.Z	3881.54	32.31	1969	13.5	10.58
P62696	Beta-crystallin B2	4394.09	306.20	1921	23.4	7.01
Q9D6F9	Tubulin beta-4A chain	4210.18	486.37	1887	49.6	4.88
Q64524	Histone H2B type 2-E	4786.94	91.90	1875	14.0	10.32
P20612	Guanine nucleotide-binding protein G(t) subunit alpha-1	3889.71	344.74	1833	39.9	5.62
P68368	Tubulin alpha-4A chain	4309.22	325.60	1785	49.9	5.06

the RPE of over 60 proteins associated predominantly with chaperone activity after exposures to various light intensities.

In terms of methodological considerations, although the workflow of an MS experiment will be relatively similar, sample collection and preparation are key components that drive the quality of the output. More specifically, we refer here to the critical role played by the extraction and separation protocols used to prepare the samples for MS. For example, Karn and Laukatis [2] performed an MS analysis on the tear proteome of the C57Bl/6 mice and reported 139 proteins using an Orbitrap. A clinical study reported 677 proteins in human tear samples [19], despite both groups using Orbitrap instrumentation, as we employed herein. The notable difference between the two studies lies in the sample preparation, and more significantly, the samples themselves. Karn and Laukatis used two-dimensional (2D) gels, whereas Schori [19] used a method referred to as filter-assisted sample preparation (FASP), originally adapted from Wiśniewski [20]. Generally, the number of proteins identified in the proteome dataset has increased in parallel to the steady improvements and advances in both instrumentation and analytical capabilities that define MS technology. Improvements in high-resolution accurate mass (HRAM) MS have driven the increase in detection and corresponding identification of larger protein datasets. We estimate that, with continued instrumentation improvement, the entire mouse retina proteome will be identified, including low-abundance proteins, and eventually, single-cell proteomic analyses will be possible, providing deeper insights into “subretinal” phenotypes.

Table 2. Top 25 highest scoring proteins from the phosphopeptide-enriched data set. SEQUEST-HT scoring was used to rank proteins identified.

Accession	Description	Score Sequest HT	Sum PEP Score	# PSMs	MW [kDa]	calc. pI
P14873	microtubule-associated protein 1B	12,723.65	958.59	8536	270.1	4.83
P17095-P2	High mobility group protein HMG-I/HMG-Y	7846.16	103.95	2633	11.6	10.32
P07901	Heat shock protein HSP 90-alpha	6057.90	172.91	1818	84.7	5.01
P15409	Rhodopsin	5794.45	62.49	5315	39.0	6.65
P51859	hepatoma-derived growth factor	4958.06	158.42	2047	26.3	4.83
Q99JF8	PC4 and SFRS1-interacting protein	4043.27	218.72	2325	59.7	9.13
P20357	microtubule-associated protein 2	4000.63	231.90	2248	199.0	4.91
P47955	60S acidic ribosomal protein P1	3970.19	52.05	938	11.5	4.32
Q8CJ40	Rootletin	3479.92	196.60	2171	226.8	5.55
P24622-P2	Isoform 2 of Alpha-crystallin A chain	3355.14	51.93	1617	19.8	6.20
P99027	60S acidic ribosomal protein P2	3329.04	44.32	790	11.6	4.54
Q62188	Dihydropyrimidinase-related protein 3	3159.62	91.63	1700	61.9	6.49
O08553	Dihydropyrimidinase-related protein 2	3074.13	61.41	1106	62.2	6.38
P54227	Stathmin	3042.13	98.24	1314	17.3	5.97
Q9JIX8	Apoptotic chromatin condensation inducer in the nucleus	2842.98	139.20	1645	150.6	5.91
Q52KI8–1	Serine/arginine repetitive matrix protein 1	2757.02	138.94	1241	106.8	11.87
P27546-P1	Microtubule-associated protein 4	2656.20	311.12	1885	117.4	4.98
Q5XG69	Soluble lamin-associated protein of 75 kDa	2483.44	217.05	1975	73.2	4.68
O55022	Membrane-associated progesterone receptor component 1	2414.01	57.20	1374	21.7	4.70
P10637-P1	Microtubule-associated protein tau	2390.67	151.41	1070	76.2	6.79
P61264	Syntaxin-1B	2372.25	63.03	741	33.2	5.38
P10637-P5	Isoform Tau-D of Microtubule-associated protein tau	2359.53	138.35	1027	38.9	9.50
Q8BTI8	serine/arginine repetitive matrix protein 2	2316.97	269.79	1692	294.7	12.03
O70318	band 4.1-like protein 2	2290.07	197.13	865	109.9	5.43
Q80X80	Phospholipid transfer protein C2CD2L	2131.65	166.44	1372	76.3	7.21

Given all the aforementioned biological and methodological variables that can affect the results of a proteomics study, studies focused on interrogating the proteome of the same tissue will likely yield distinct results. This, in turn, makes fair and objective comparisons between our findings and others difficult. As examples, in 2009, Gao [21] identified 1,792 proteins in the C57Bl/6 retina using a streptozotocin (STZ)-induced diabetic mouse strain. Notably, the mice used in our study were of a different strain and significantly older than the early adolescent (7- to 8-week-old) mice used in the previous study were. Four years later, Skeie and Mahajan [22] reported 1,680 proteins using Jax's C57Bl/6J mouse and acquired approximately the same number of proteins. However, advancements in downstream bioinformatic analysis revealed 675 proteins unique to vitreous in their dataset versus those reported earlier. The following year, Barathi [9] nearly doubled the number of proteins identified, with 3,882 in the same strain. Ly [23] reported ~3,000 unique proteins in an rd10 model of retinitis pigmentosa at pre-, peak-, and post-degenerative timepoints; while impressive in terms of interrogating the proteome during the course of retinal degeneration, the label-free quantitative approach used remains concerning, given reported inaccuracies ranging up to 40% and potential issues with the FDR rates of the spectra [24]. In contrast, our current study identified 4,767 proteins before any enrichment procedure and 5,377 proteins after identifying 610 additional proteins by phosphoenrichment.

Our current study stands alone with respect to proteomics analyses that include both enrichment techniques

and bioinformatic analyses specific to normal, adult mice. Regarding the former, the proteome dataset quality and utility can be enhanced by enriching for specific protein modifications, revealing proteins that would not be identified even by state-of-the-art methods and instrumentation. The purification and isolation involved in enrichment procedures have the advantage of increasing the ability to detect low abundance molecules that would normally be missed, especially when coupled with improvements in instrument detection sensitivity. Our study produced an additional 610 unique phosphorylated proteins that were not previously identified in our original dataset. Time and cost are two obvious disadvantages for running enrichment studies, as well as the requirement for the rather large amounts of "raw tissue" needed to perform the enrichment steps. Our nonenriched data were acquired with 100 µg of tissue, while 1.6 g was used during the phosphoenrichment procedure. The latter limited us to one enrichment procedure (i.e., phosphorylation). How ours and other proteomic profiles change by enriching for sumoylated, acetylated, and so on proteins remains of considerable interest, given the importance of posttranslational modifications in regulating cellular function and as sites of pathological signaling. For example, an enrichment approach provided investigators the ability to determine kinase activity in photoreceptor outer segments and how these kinases regulate phagocytosis in the retinal pigment epithelium [25].

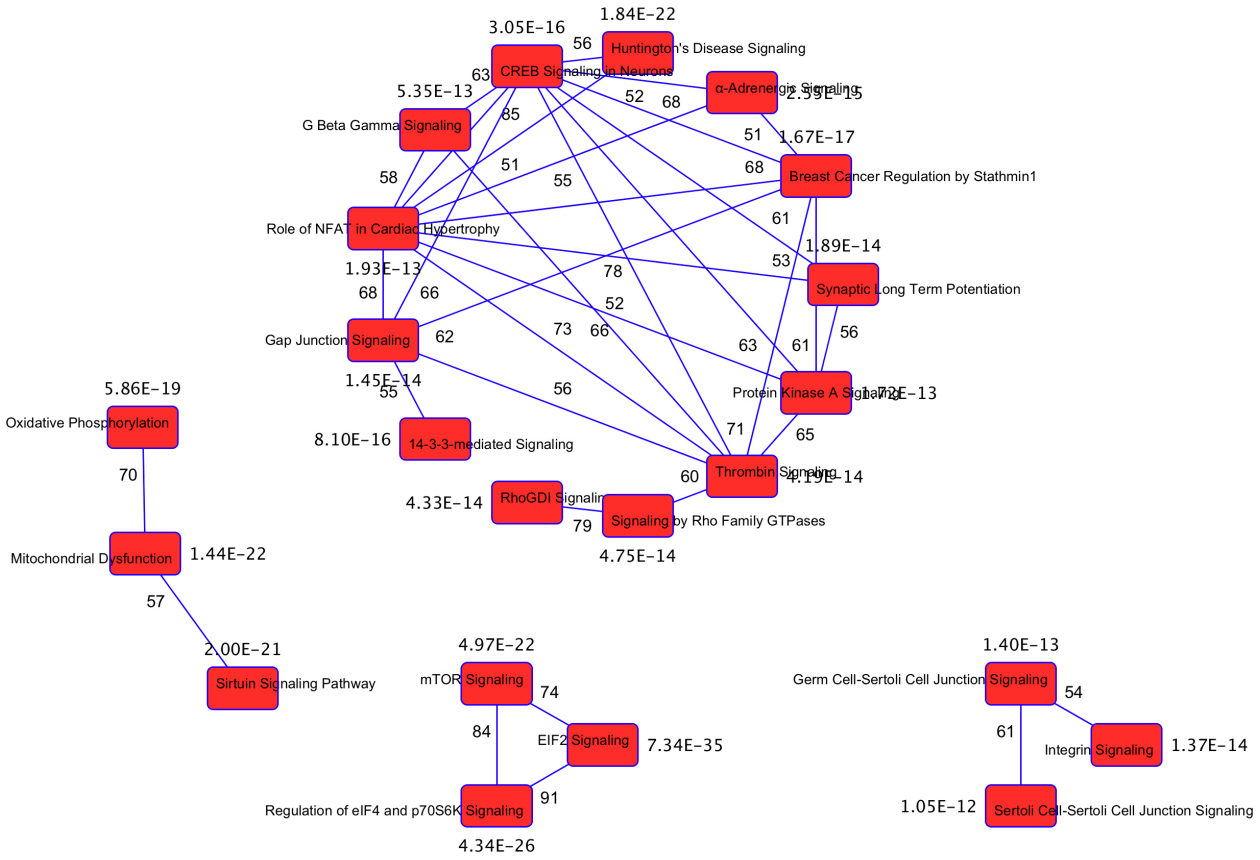

Figure 5. Ingenuity pathway analysis (IPA) core analysis of the top 22 canonical pathways, most enriched, in our mouse retinal total phosphoproteome data set, reflecting the diverse functional classifications of identified retinal proteins. Each canonical listed has a minimum of 50 interacting proteins. The actual number or interacting proteins is indicated by the number above the lines connecting the different pathways. "Top" canonicals are determined based on the number of molecules identified in a dataset that are also found in IPA's predefined canonicals. P values listed adjacent to each canonical describe the "p-value of overlap" of the proteins in our dataset that are contained in that canonical relative to the proteins that define the entire canonical.

Table 3. Summary report generated using IPA.

Name	p-value	Overlap
EIF2 Signaling	3.63E-33	67.6% (140/207)
Regulation of eIF4 and P70S6K Signaling	2.06E-24	67.8% (101/149)
Mitochondrial Dysfunction	1.58E-22	64.8% (103/159)
Huntington's Disease Signaling	2.07E-22	57.5% (138/240)
Sirtuin Signaling Pathway	2.26E-21	55.0% (148/269)

Top 5 Associated Network Functions

1. **Metabolic Disease, Neurologic Disease, Organismal Injury and Abnormalities**
2. **Cancer, Cell Death and Survival, Organismal Injury and Abnormalities**
3. **Neurologic Disease, Ophthalmic Disease, Developmental Disorder**
4. **Protein Synthesis, Gene Expression, Hereditary Disorder**
5. **Cell-to-Cell Signaling and Interaction, Nervous System Development and Function, Neurologic Disease**

Table depicts the top 5 canonical pathways enriched in the data set with confidence (p value of overlap) and coverage (overlap of proteins) of total pathway. The top 5 associated network functions is listed additionally below.

MS-based proteomics studies yield "big data" sets that require subsequent bioinformatic analyses. The latter may initially appear to be a rather simple process involving acquiring a dataset, loading data into software, and reporting the findings in the output files. However, proprietary software may be necessary to provide accurate—and thus, useful—outcomes, as freeware can limit the size of the uploaded dataset, and consequently, produce unreliable results. For example, we initially used the freeware "Panther-Gene Ontology," which grouped our data into the 26 functional classes shown in Figure 3. Interestingly, the contents of this retinal proteome output file was a close match to that resulting from the evaluation of the effects of 2D versus three-dimensional (3D) cell culture models in the well-studied colon cancer cell line HT29 [11]—a highly unlikely possibility. The validation and crossexamination of bioinformatics results that are required may not be immediately obvious to many investigators, potentially compromising the purpose of performing such analyses. In contrast, bioinformatics software like IPA (Qiagen) can accept large datasets; together with their broad and deep, frequently updated, software, the program allows users to identify and investigate large canonical pathways and signaling networks that may otherwise be missed using inferior software. That said, all bioinformatics analyses are only as powerful as the company-defined proprietary algorithms on which they are based, which in turn, are based on the literature. Thus, in true iterative fashion, the findings in our present dataset will ultimately improve the utility of future bioinformatic analyses.

The power of such comprehensive bioinformatics analyses is evident in our study relative to investigations using PCR or immunoblotting to focus on a single protein. Specifically, our top 22 canonical pathways in Figure 5 illustrate the central nervous system (CNS) origin of retinal tissue and depict the high dependence of the retina on diverse mechanisms for resolving oxidative stress. These results highlight the relative enrichment of proteins in the retina that are better known for their involvement in other established pathways, for example, Huntington's disease or breast cancer, thereby providing insights into the diverse roles many of these retinal proteins may serve. To clarify this point, one should not infer that the top canonical pathways of Figure 5 indicate that retinal proteins cause Huntington's or breast cancer pathology, but rather, that there is a select group of retinal proteins that may be highly correlated with a select group of proteins that are highly studied with respect to their roles in glutamine metabolism or a select group of well-known proteins involved in cytoskeleton microtubule dynamics, respectively. Similarly, in terms of "cancer" and "tumor morphology" being among the top hits for our proteome for the regulatory pathways associated with disease (Figure 6), consider the tumor-suppressing protein p53: Although aberrant p53 signaling is central to many cancers, healthy cells also express p53, which is one protein among a larger complement of proto-oncogenes, tumor-suppressing proteins, chaperone proteins, and "guardian-like" regulatory proteins that are highly redundantly expressed in nearly every tissue.

For the limitations of our study, because there are no protein amplification methods available for MS-based proteomics projects to perform single-cell proteomics, it should be recognized that our analysis reveals the "collective" proteome of a heterogeneous tissue represented by many distinct cell types. In addition, our proteome dataset was derived from only one strain of mouse, and we did this using

a label-free analysis. Concerning the tolerance thresholds used to ensure appropriate protein identification, we applied a parent ion tolerance of 10 ppm and a fragment mass tolerance of 0.6 Da; we recognize that some investigators may feel that more stringent tolerances than those we applied should routinely be used. MS data acquisition in orbitraps can use the linear ion, or the orbitrap, for protein identification in label-free approaches, which should include more stringent tolerances. When using labeled methods, such as tandem-mass tags, the linear ion trap is used in conjunction with the orbitrap to identify and quantify specific reporter ions. What thresholds should be applied is currently considered subjective; despite the attention given to transparent reporting, it is not uncommon for such criteria to go underreported or be listed as "default parameters." At any rate, this is another reason direct database comparisons between retinal proteomics studies can be challenging.

Although defining the steady-state proteome and phosphoproteome in the nondiseased retina, and the networks and pathways they associate with, as we provided here, represents an important starting point for understanding retinal phenotypic conservation and variation, it is ultimately the

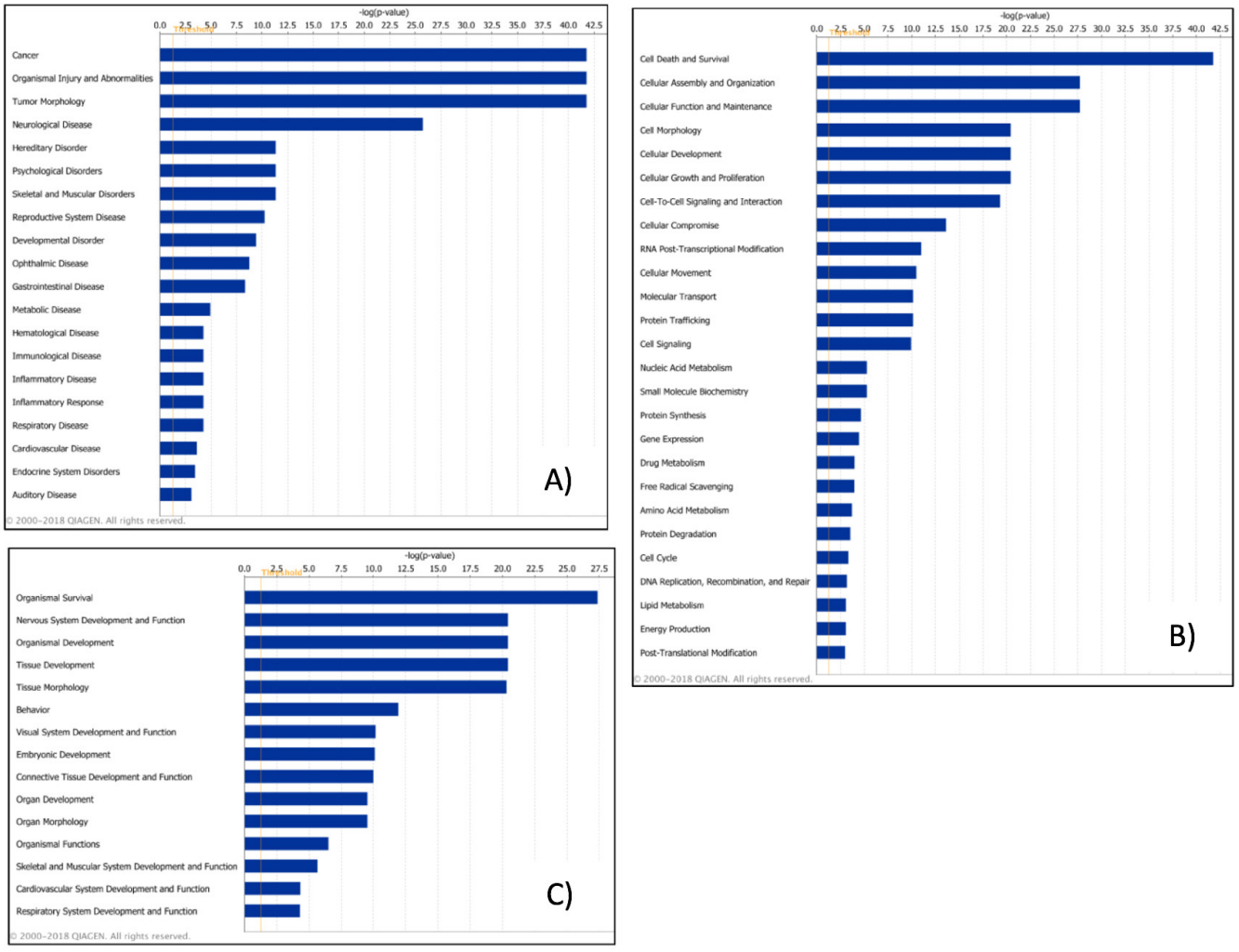

Figure 6. Ingenuity pathway analysis (IPA) core analysis downstream effects reveal diseases and functions network enrichments. This figure is based on proteins identified in the dataset and their known (published) regulatory associations across the following categories: (**A**) Diseases and Disorders Enrichment; (**B**) Cellular and Molecular Functions Enrichment; and (**C**) Physiological System Development and Functions, related specifically to Ophthalmic Diseases. Bar graphs illustrate the "p-value of overlap" of the proteins in our dataset relative to the IPA's predefined categories. Each histogram bar can be expanded into larger canonical signaling networks within the software for further details including protein-protein interactions (both upstream and downstream from a selected node), as well as their respective geographic distributions within the cell. A detailed table of this information has been provided as Appendix 1.

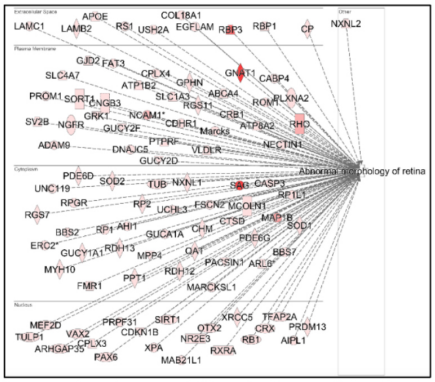

Figure 7A: Abnormal Morphology of the Retina

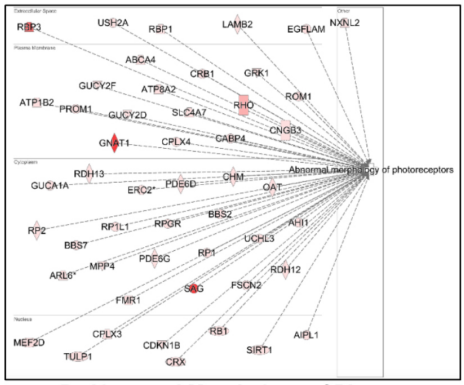

Figure 7B: Abnormal Morphology of Photoreceptors

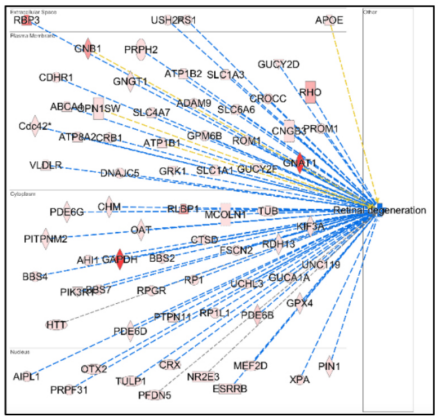

Figure 7C: Retinal Degeneration

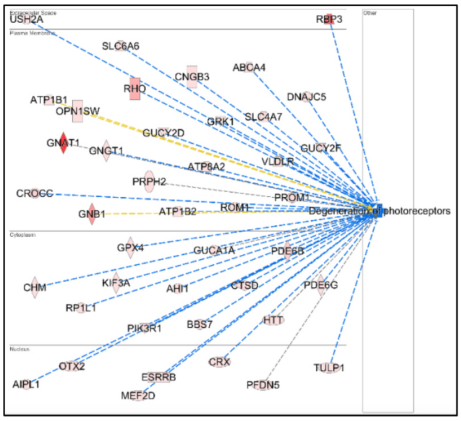

Figure 7D: Degeneration of Photoreceptors

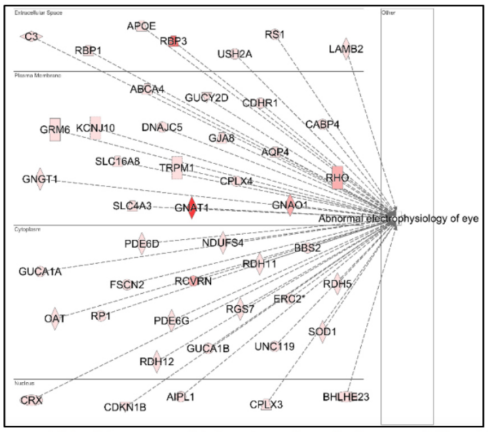

Figure 7E: Abnormal Electrophysiology of the Eye

Figure 7. Ingenuity pathway analysis (IPA) Core Analysis downstream effects depicts protein enrichments within our total retinal phosphoproteome across the top five retina-specific components of a larger network of proteins categorized within IPA under "Ophthalmic Disease". These proteins can be found in Appendix 1. Each protein network is displayed by individual function (the single item listed on the right side) and geographic distribution within the cell. These graphics are intended to aid researchers to identify proteins of interest as they relate to ophthalmic pathologies. The five retinal networks are: (**A**) Abnormal morphology of the retina; # of molecules: 99, P < 4.28E-08. **B**: Abnormal morphology of photoreceptors; # of molecules: 50, P < 8.36E-09. **C**: *Retinal Degeneration; # of molecules: 71, P < 4.09E-06. **D**: *Degeneration of photoreceptors; # of molecules: 42, P < 2.27E-06. **E**: Abnormal electrophysiology of the eye; # of molecules: 46, P < 6.52E-09. *graphic contains proteins with z-scores high enough for IPA to predict signaling effects (blue lines, which also represent an inhibitory effect), whereas gray lines reflect no predicted effect, and yellow lines indicate findings are inconsistent with the state of downstream molecule. Color intensity of protein node reflects expression ratio magnitude (here, Sequest-HT score).

interrogation of changes in the total proteome, including posttranslationally modified proteins, in response to injury or disease [6] or an efficacious therapeutic intervention, that will provide robust insights into disease mechanisms and treatment. In turn, this will require enhanced experimental rigor for incorporating and reporting detailed biological information, providing methodological and instrumentation specifics, using experimental designs that include technical and biological replicates, and ultimately, involving improved data resource sharing to maximize authenticity and reproducibility. The latter remains a goal that must be met more consistently to maximize the translational utility of proteomics-based investigations.

APPENDIX 1.

A supplementary table for the "ophthalmic disease network" procured has additionally been included in this appendix and consisting of 1,213 categorized function annotations. To access the data, click or select the words "Appendix 1."

APPENDIX 2.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository Vizcaino [26] with the data set identifier PXD009909 (comprehensive) and PXD009981 (phosphoproteome). PRoteomics IDEntifications (PRIDE), a *Nature*-preferred repository. See table here for the complete proteome/phosphoproteome. To access the data, click or select the words "Appendix 2."

ACKNOWLEDGMENTS

Supported by NIH R01EY018607 (JMG); BrightFocus (JMG); NIH GM103514–11 (JJG); and Sigma Xi GIAR [Grant-In-Aid-of Research] (JCH). The authors thank Melodie Baddoo of Tulane University's School of Medicine for bioinformatic resources and technical support. JCH would like to thank Dr. Anshuman Singh and Dr. Filipe Muhale for unwavering mentoring and guidance with writing this manuscript. Authors have no commercial or financial interests to disclose at this time.

REFERENCES

1. Ahmad M, Semba RD, Enghild JJ, Venkatraman V, Dyrlund TF, Van Eyk JE. The Human Eye Proteome Project: Perspectives on an emerging proteome. Proteomics 2018; 13:2500-11. [PMID: 23749747].
2. Karn R, Laukaitis C. Comparative Proteomics of Mouse Tears and Saliva: Evidence from Large Protein Families for Functional Adaptation. Proteomes 2015; 3:283-97. [PMID: 28248272].
3. Dammalli M, Murthy KR, Pinto SM, Murthy KB, Nirujogi RS, Madugundu AK, Keshava Prasad TS. Toward Postgenomics Ophthalmology: A Proteomic Map of the Human Choroid–Retinal Pigment Epithelium Tissue. OMICS 2017; 21:114-22. [PMID: 28186866].
4. Winiarczyk M, Kaarniranta K, Adaszek Ł, Winiarczyk D. Tear film proteome in age-related macular degeneration. Graefes Arch Clin Exp Ophthalmol 2018; 256:1127-39. [PMID: 29696386].
5. Grosche A, Hauser A, Lepper MF, Mayo R, von Toerne C, Merl-Pham J, Hauck SM. The Proteome of Native Adult Müller Glial Cells From Murine Retina. Mol Cell Proteomics 2016; 15:462-80. [PMID: 26324419].
6. Cehofski LJ, Honoré B, Vorum H. A review: Proteomics in retinal artery occlusion, retinal vein occlusion, diabetic retinopathy and acquired macular disorders. Int J Mol Sci 2017; 18:e907[PMID: 28452939].
7. Church D.M., Goodstadt L., Hillier L.W., Zody M.C., Goldstein S., Xinwe S., Bult C.J., Agarwala R., Cherry JL, DiCuccio M., Hlavina W., Kapustin Y., Meric P., Maglott D.. Zoe¨ Birtle, Marques A.C., Graves T., Zhou S., Teague B., Potamousis K., Churas C., Place M., Herschleb J., Runnheim R., Forrest D., Amos-Landgraf J., Schwartz D.C., Cheng Z., LindbladToh K., Eichler E.E., Ponting C.P., The Mouse Genome Sequencing ConsortiumLineage-Specific Biology Revealed by a Finished Genome Assembly of the Mouse. PLoS Biol 2009; 7:e1000112-[PMID: 19468303].
8. Shao HJ, Chaerkady R, Chen S, Pinto SM, Sharma R, Delanghe B, Chakravarti S. Proteome profiling of wild type and lumican-deficient mouse corneas. J Proteomics 2011; 74:1895-905. [PMID: 21616181].
9. Barathi VA, Chaurasia SS, Poidinger M, Koh SK, Tian D, Ho C, Zhou L. Involvement of GABA transporters in atropine-treated myopic retina as revealed by iTRAQ quantitative proteomics. J Proteome Res 2014; 7:4647-58. [PMID: 25211393].
10. Zhang L, Masetti G, Colucci G, Salvi M, Covelli D, Eckstein A, Biscarini F. Combining micro-RNA and protein sequencing to detect robust biomarkers for Graves' disease and orbitopathy. Sci Rep 2018; 8:8386-[PMID: 29849043].
11. Koss MJ, Hoffmann J, Nguyen N, Pfister M, Mischak H, Mullen W, Siwy J. Proteomics of vitreous humor of patients with exudative age-related macular degeneration. PLoS One 2014; 9:e96895-[PMID: 24828575].
12. Chen D, Gomes F, Abeykoon D, Lemma B, Wang Y, Fushman D, Fenselau C. Top-Down Analysis of Branched Proteins Using Mass Spectrometry. Anal Chem 2018; 90:4032-8. [PMID: 29513006].
13. Zolg DP, Wilhelm M, Schnatbaum K, Zerweck J, Knaute T, Delanghe B, Kuster B. Building ProteomeTools based on a complete synthetic human proteome. Nat Methods 2017; 14:259-[PMID: 28135259].

14. Yue X, Lukowski JK, Weaver EM, Skube SB, Hummon AB. Quantitative Proteomic and Phosphoproteomic Comparison of 2D and 3D Colon Cancer Cell Culture Models HHS Public Access. J Proteome Res 2016; 15:4265-76. [PMID: 27696853].

15. Greenwood EJ, Matheson NJ, Wals K, van den Boomen DJ, Antrobus R, Williamson JC, Lehner PJ. Temporal proteomic analysis of HIV infection reveals remodelling of the host phosphoproteome by lentiviral Vif variants. eLife 2016; 5:18296-[PMID: 27690223].

16. Yang F, Shen Y, Camp DG, Smith RD. High pH reversed-phase chromatography with fraction concatenation as an alternative to strong-cation exchange chromatography for two-dimensional proteomic analysis. Expert Rev Proteomics 2012; 9:129-34. [PMID: 22462785].

17. Chaychi S, Polosa A, Lachapelle P. Differences in Retinal Structure and Function between Aging Male and Female Sprague-Dawley Rats Are Strongly Influenced by the Estrus Cycle. PLoS One 2015; 10:e0136056-[PMID: 4552560].

18. Lee H, Chung H, Lee SH, Jahng WJ. Light-induced phosphorylation of crystallins in the retinal pigment epithelium. Int J Biol Macromol 2011; 48:194-201. [PMID: 21094180].

19. 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. Invest Ophthalmol Vis Sci 2018; 59:AMD31-40. [PMID: 30025106].

20. Wiśniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Methods 2009; 6:359-62. [PMID: 19377485].

21. Gao BB, Phipps JA, Bursell D, Clermont AC, Feener EP. Angiotensin AT1 receptor antagonism ameliorates murine retinal proteome changes induced by diabetes. J Proteome Res 2009; 8:5541-9. [PMID: 19845401].

22. Skeie JM, Mahajan VB. Proteomic interactions in the mouse vitreous-retina complex. PLoS One 2013; 8:e82140-[PMID: 24312404].

23. Ly A, Merl-Pham J, Priller M, Gruhn F, Senninger N, Ueffing M, Hauck SM. Proteomic Profiling Suggests Central Role of STAT Signaling during Retinal Degeneration in the rd10 Mouse Model. J Proteome Res 2016; 15:1350-9. [PMID: 26939627].

24. Käll L, Storey JD, MacCoss MJ, Noble WS. Posterior Error Probabilities and False Discovery Rates: Two Sides of the Same Coin. J Proteome Res 2008; 7:40-4. [PMID: 18052118].

25. Chiang CK, Tworak A, Kevany BM, Xu B, Mayne J, Ning Z, Palczewski K. Quantitative phosphoproteomics reveals involvement of multiple signaling pathways in early phagocytosis by the retinal pigmented epithelium. J Biol Chem 2017; 292:19826-39. [PMID: 28978645].

26. Vizcaino JA, Csordas A, del-Toro N, Dianes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternet T, Xu QW, Wang R, Hermjakob H. 2016 update of the PRIDE database and related tools. Nucleic Acids Res 2016; 44:D1D447-56. [PMID: 26527722].

Articles are provided courtesy of Emory University and the Zhongshan Ophthalmic Center, Sun Yat-sen University, P.R. China. The print version of this article was created on 31 December 2018. This reflects all typographical corrections and errata to the article through that date. Details of any changes may be found in the online version of the article.