Title: Data Independent Acquisition Mass Spectrometry of the Human Lens Enhances Spatiotemporal Measurement of Fiber Cell Aging

Lee S Cantrell^{1,2} and Kevin L Schey¹

¹Department of Biochemistry, Vanderbilt University, Nashville, TN 37232

²Chemical and Physical Biology Program, Vanderbilt University, Nashville, TN 37232

Current address:

Mass Spectrometry Research Center Vanderbilt University
465 21st Ave. So., Suite 9160 MRB III Nashville, TN 37232-8575

E-mail: kevin.schey@vanderbilt.edu Phone: 615.936.6861
Fax: 615.343.8372

Abstract

The ocular lens proteome undergoes post-translational and progressive degradation as fiber cells age. The oldest fiber cells and the proteins therein are present at birth and are retained through death. Transparency of the lens is maintained in part by the high abundance crystallin family proteins (up to 300 mg/mL), which establishes a high dynamic range of protein abundance. As a result, previous Data Dependent Analysis (DDA) measurements of the lens proteome are less equipped to identify the lowest abundance proteins. In an attempt to probe more deeply into the lens proteome, we measured the insoluble lens proteome of an 18-year-old human with DDA and newer Data Independent Analysis (DIA) methods. By applying library free DIA search methods, 4,564 protein groups, 48,474 peptides and 5,577 deamidation sites were detected: significantly outperforming the quantity of identifications in using DDA and Pan-Human DIA library searches. Finally, by segmenting the lens into multiple fiber cell-age related regions, we uncovered cell-age resolved changes in proteome composition and putative function.

Introduction

The ocular lens is a transparent tissue responsible for transmission and focusing of light to the retina. Unlike the majority of all tissue, lens fiber cells are not degraded after formation and organelles degrade in early cellular maturation, leading to cessation of protein synthesis. Since genetic material is degraded, proteomic measurement is exclusively capable of characterizing each stage of cellular maturation. Age-related measurements of lens fiber cell proteomes are facilitated by the inherent spatiotemporal gradient of the lens wherein fibers of the innermost nucleus of the lens are as old as the subject and surrounding concentric layers of fibers are progressively younger. Here, we divide the lens into three age-related regions: the inner nucleus that is composed of fiber cells formed *in utero* and infancy, the outer nucleus that contains fibers differentiated in early childhood, and the cortex that corresponds to fibers differentiated in early adulthood (Figure 1).

Throughout life, transparency of the lens must be maintained to prevent cataract.¹ In the absence of vasculature, metabolites are hypothesized to convect through the lens by a microcirculation system (MCS) where metabolite influx at the poles proceeds towards the inner nucleus and outflow proceeds through the equatorial region of the lens. In the proposed system, flux is established by a sodium gradient near the equatorial epithelium and water is transported through membranes by aquaporin and gap junction proteins. The MCS is critical for the delivery of glutathione to the inner nucleus of the lens, where oxidative stress may otherwise induce post-translational modifications (PTMs) that modify protein structure, function and induce light scattering by protein aggregation. In addition to glutathione conjugation to oxidized species, α -crystallin contributes to aggregation prevention by binding to partially unfolded substrates as a small heat shock protein.^{2,3} After the fifth decade of life, a barrier to extracellular metabolite transport is established, further challenging maintenance of lens transparency.^{4,5}

Historically, measurement of the lens proteome and its post-translational modifications have centered on the crystallin protein family due to their overwhelming abundance in the lens: estimated as high as 90% by mass and 3,000 mg/mL.^{6–9} The high concentration of lens crystallins reduces the technical ability to measure low abundance proteins in the lens, especially with intensity biased methods such as Data Dependent Acquisition (DDA) MS,¹⁰ making the lens an excellent test tissue for DIA methodology. Further, biological insights may be gained by solubility-fractionation to deplete soluble crystallin proteins, isolating insolubilized proteins, cytoskeletal components, or proteins localized to the plasma membrane – like those in the MCS.^{11,12} Using a 13-pulse MuDPIT separation coupled to a LTQ Velos linear ion trap mass spectrometer, solubility fractionation provided previous identification of 951 human lens protein groups.¹¹ More recently, 5,466 protein groups from young mouse lenses have been identified with a 9-step basic reverse phase (BRP) pre-separation and acquisition on a newer orbitrap instrument. For either mass analyzer, MuDPIT/BRP improves the quantity of protein and peptide identifications by approximately 400% over single-dimension analyses.^{13,14} While 2D separation is effective in increasing protein identifications, it is less quantitatively reproducible and acquisition times are significantly greater per samples.¹⁵ Further, if the lens is to be segmented into multiple fiber cell-age related regions for measurement, 2D separations become significantly more instrument intensive than single-dimension analyses.

Arguably the greatest limitation of DDA is its weakness for measuring low abundance peptides due to intensity biased selection of precursors.^{16,17} Alternatively, Data Independent Acquisition (DIA) methods including Sequential Window Acquisition of All Theoretical Mass Spectra (SWATH-MS), circumvent biased precursor selection by co-isolating all peptides within a small mass window along the LC gradient.¹⁰ The major limitation of DIA is data-processing where a spectral library is required to deconvolute individual product spectra which may represent multiple co-eluting precursors.¹⁸ Conventional spectral libraries are developed by combining a fraction of each sample in the cohort and sampling with

DDA. To improve spectral library size, 2D separations and gas-phase fractionation are frequently employed. Constructing a library is a laborious and instrument intensive process, so widely applicable libraries such as the pan-human SWATH-MS library may be used instead.¹⁹ Alternatively, recent software developments include *in silico* library preparation.^{20,21} Here, we examine the application of DIA to multiple age-related regions of the human lens and discuss the impact that DIA, pre-developed libraries, and *in silico* searches have on annotating age-related changes in lens fiber cells.

Materials and Methods

Materials

An 18-year-old frozen human lens was obtained from NDRI (Philadelphia, PA). Sequence-grade modified trypsin was obtained from Promega (Madison, WI). S-Trap columns were purchased from Protifi (Farmingdale, NY). All other chemicals were purchased from Fisher Scientific (Waltham, MA).

Urea Insoluble Protein Preparation

An 18-year-old, cataract free human lens (9.2 mm x 7.0 mm) was mounted with its equatorial axis parallel to the cryostat chuck before removal of the anterior and posterior poles of the lens yielding a 4.5 mm thick equatorial lens section. Concentric biopsy centerpunches were taken at 4.5- and 7-mm diameter to yield inner nucleus (0-4.5 mm), outer nucleus (4.5 – 7.0 mm) and cortex (7.0 – 9.2 mm) samples. Tissue was homogenized in buffer containing 25 mM Tris (pH8), 5 mM EDTA, 1 mM DTT, 150 mM NaCl, 1 mM PMSF. After homogenization, samples were centrifuged at 100,000g for 30 minutes and the supernatant discarded. Pellets were washed twice with the above homogenization buffer followed by washes with 3.5 M and 7 M urea added to the homogenization buffer. Centrifugation at 100,000g was performed to separate the supernatant and pellets for each wash. The remaining urea insoluble pellet was taken up in 50 mM TEAB with 5% SDS and protein concentration was measured with BCA assay.

Tryptic Digest

Membrane pellets of urea insoluble sample (200 µg total protein) were suspended in 5% SDS with 50 mM TEAB and DTT was added to 10 mM before incubation at 56°C for 1 hour to reduce disulfide bonds. Reduced cysteines were alkylated by adding IAA to 20 mM and incubating in the dark at room temperature for 30 minutes. Phosphoric acid was added to 2.5% to insolubilize proteins for efficient trapping. Proteins were taken up in 100 mM TEAB in 90% methanol before transferring 75 µg protein to the S-Trap micro. Trapping solution was centrifuged through the filter at 4,000g for 30 seconds. Trapped proteins were washed three times with 100 mM TEAB in 90% methanol with buffer removed by centrifugation at 4,000g. Proteins were digested by diluting in 50 mM TEAB with 5 µg (1:15) sequencing grade trypsin added and incubated for 1.5 hours at 47°C. To elute digested peptides, four sequential elution solutions of 50 mM TEAB, 0.2% aqueous formic acid, 50% ACN in water and 0.2% aqueous formic acid were added and centrifuged through the filter at 1,000g. Eluted peptides from each sample were pooled and dried with vacuum centrifugation. Peptides were taken up in 0.2% aqueous formic acid and stored at -80°C.

Liquid Chromatography

Peptides were analyzed using a Dionex Ultimate 3000 UHPLC coupled to an Exploris 480 tandem mass spectrometer (Thermo Scientific, San Jose, CA) with sample order randomized. An in-house pulled capillary column was created from 75 µm inner diameter fused silica capillary packed with 1.9 µm ReproSil-Pur C18 beads (Dr. Maisch, Ammerbuch, Germany) to a length of 250 mm. Solvent A was 0.1% aqueous formic acid and solvent B was 0.1% formic acid in acetonitrile. Peptides were loaded (cortex: 200 ng, outer nucleus: 300 ng, inner nucleus: 300 ng) and separated at a flow rate of 200 nL/min on a

95-minute gradient from 2 to 29% B, followed by a 14-minute washing gradient and 35-minute blank injection between runs. The exact gradient was determined by linearized separation of the top 50% most intense cortical peptide signals by Gradient Optimization Analysis Tool (GOAT 1.0.1).²² Full details of chromatographic gradient are provided in Supplemental Table 1.

Mass Spectrometry

For DDA injections, the Thermo Exploris 480 was set to acquire in top-20 configuration with auto dynamic exclusion. Precursor spectra were collected from 400 to 1600 m/z at 60,000 resolution (AGC target 3e6, max IIT of 50 ms). MS/MS spectra were collected on peptidic precursors between charge state +2 and +5 achieving a minimum AGC of 1e4. MS/MS scans were collected at 15,000 resolution (AGC target of 1e5, max IIT of 35 ms) with an isolation width of 1.6m/z. For DIA injections the Exploris 480 instrument was configured to acquire 35x20m/z (400-1100 m/z) precursor isolation window DIA spectra (30,000 resolution, AGC target 1e6, max IIT 55 ms, 27 NCE) using a staggered window pattern with window placements optimized by Skyline.²³ Precursor spectra (385-1115 m/z, 60,000 resolution, AGC target 3e6, max IIT 100 ms) were interspersed every 25 ms/ms spectra. In all experiments, default charge state was set to +3, S-Lens RF level set at 40%, NCE set at 27 and data collected in profile mode. See Supplemental Table 2 for complete windowing scheme.

Data Analysis

For analysis of DDA data, RAW files were searched in MaxQuant²⁴ (1.6.7.0) without variable modifications, match between runs enabled, and all other settings set as default. For analysis of DIA data, RAW files were converted to mzML files in MSConvert,²⁵ with staggered window deconvolution performed to improve precursor specificity. Processed DIA files were searched in DIA-NN²⁰ (1.7.16) with an Intel Core i7-7700 CPU at 3.60 GHz utilizing 8 threads. For all searches, up to one missed trypsin cleavage was allowed on peptides 7-30 residues in length with N-terminal M excision and cysteine carbamidomethylation enabled. All fragments between m/z 200 and 1800 were considered. In each search, the neural network classifier was run in double pass mode with likely interferences removed and retention time dependent normalization initiated between technical replicates from the same region of the lens with high accuracy quantitation enabled. An initial search of all files produced a library which was used to search the data a second time (termed match between runs). Precursors and protein groups were filtered at 1% FDR and contaminants in the MaxQuant contaminant list were removed. All library-free DIA and DDA analyses were searched against a UniProt SwissProt canonical human fasta database (UP000005640, downloaded 3/12/2021, 7,657 entries). For the library free search, deep learning-based spectra and retention time prediction was used for library generation. In a separate library free search, one variable deamidation of glutamine or asparagine was also allowed. For searches against the pan-human SWATH atlas, a consensus speclib file (SAL00035) from Rosenberger et al¹⁹ was downloaded and used for a library-based search in DIA-NN with identical search settings to library free analysis. Comparisons and analyses were initiated through custom R scripts with peptides having >1% q-value in any DIA-NN q calculation omitted.

Intraregional peptide abundance measurements were independently normalized by DIA-NN abundance normalization based on the 40% of peptides with the lowest CV (<10% CV). To improve normalization, trimmed mean of M-values (TMM) normalization was performed within regions. Interregional peptide abundances were then median normalized. For comparisons of identification efficiency, imputation was not done. When ontology changes were considered in DIA data, the R package DMwR was used for knn-imputation with scaled data pre-processing. Peptides were considered for imputation if detected in 2 of 3 intraregional technical replicates. Protein and peptide abundance was calculated by the diann R package function `diann_maxlfq` available at <https://github.com/vdemichev/diann-rpackage>. Proteins were only assembled on

peptides considered proteotypic. For DDA comparison, raw intensity peptide abundance values were normalized as above and treated with the `diann_maxlfq` to retain similarity of protein group quantitation.

Gene association data was downloaded from the Uniprot database (04/06/2021). PSEA-quant²⁶ was initiated through a Java command line interface with R scripts for data preparation. Volcano plots were prepared with limma moderated p-values calculated in the `eb.fit` function.²⁷ Quantity of overlapping features were defined by unique razor protein identity or peptides with charge state not considered. For quantitative abundance comparison of DDA and DIA, a linear median normalization factor was determined for proteins identified in both modes and applied to all DDA protein groups. Searches with deamidation enabled were normalized as above, with an additional TMM normalization calculation based on unmodified peptides detected in each search scheme. For all statistical enrichment calculations, PANTHER²⁸ was used with the 02/01/2021 GO database applying a Fisher's exact test with FDR calculated. Significance was set at 0.01 for PANTHER searches. All proteins identified in the modification-free DIA-NN search were included as part of the statistical background.

Results and Discussion

Benchmarking against DDA

To evaluate the hypothesis that DIA outperforms DDA in lens protein quantitation, we injected three separate regions of the lens in triplicate to the Exploris 480 in both acquisition modes. DDA analysis was performed in MaxQuant and DIA analysis in DIA-NN. For this comparison, a library free DIA search with no variable modifications was used. In total, 2,161 protein groups and 14,222 peptides were identified in DDA experiments. Previously, as many as 1,251 protein groups had been identified in a cohort of lenses.¹⁴ The improvement in identifications is facilitated in part by the isolation of insoluble proteins and of the Exploris instrument compared to an older Q Exactive instrument. Prior measurements of the lens have shown significant decreases in protein identification quantity within the outer nucleus and inner nucleus regions of the lens due to organelle-associated protein degradation and targeted degradation of proteins that become misfolded with age. The trend of decreasing protein/peptide identifications in progressively older fibers is replicated in the optimized DDA dataset where there is an approximate 300% increase in proteins and peptides (Figure 2A,B).

Though lens proteomes are less complex than HeLa or other cell culture systems, BRP fractionation established identification of 5,466 proteins in a juvenile mouse lens in 10, 100-minute injections.²⁹ In this study, similar proteome coverage was achieved in a single 95-minute gradient where 4,652 protein groups and 33,516 peptides are accessible by single injection DIA experiments, the majority of which are detected in the lens cortex (Figure 2A,B). Juvenile lenses do not undergo significant aging and retain more proteins associated with early lens morphogenesis; therefore, it is not expected that an 18-year-old human lens be more complex than especially young mouse lenses.

Relative to DDA, DIA facilitated a 206% improvement in protein identifications and 235% improvement in total peptide identifications. As with DDA, there is a significant decrease in total identifiable proteins (Figure 2A) as cortical fibers mature to outer nuclear proteins. Unlike DDA results, DIA identification performance of the outer nucleus was nearly matched in the inner nucleus where fibers exceed a decade of age. It is suggested that while still present, a large subset of the nuclear proteome is not detectable with DDA due to biasing limitations. DIA also facilitated greater coverage of identified proteins, with an approximate 7 associated peptides to every protein in DIA, and 5 in DDA. With maturation of the nucleus, the nuclear peptidome, unlike the proteome becomes less populated by unique peptides (Figure 2B). Sequence alignment was not performed on the 1,610 peptides not measured in the inner nucleus; however, the near-

complete incorporation of common age-related modifications including deamidation, oxidation, N- and C-terminal truncation may contribute to this difference in searches, like this one, where PTMs are not considered.

To validate the specificity of DIA library free search relative to DDA search algorithms, overlap analysis of leading proteins and peptides was done (Figure 2C,D). Leading proteins were considered to reduce protein group assembly differences between MaxQuant and DIA-NN peptide assignment algorithms. From the 2,161 DDA proteins identified in total, 71 were not included in the DIA dataset, with 38 being identified in an assembled DIA protein group as non-leading or library FDR > 1%. The remaining 33 unmatched proteins (1.5% DDA IDs) show no ontological enrichment against the lens proteome, leading us to determine this discrepancy as insignificant towards protein-network insights. A landmark of success for DIA, 21,243 peptides were uniquely measured with DIA. This figure exceeds the total quantity of peptides identified with DDA, greatly improving the robustness of potential protein quantitation. For proteins measured in DIA and DDA experiments, there is strong regional agreement which further validates the specificity of DIA library free search for peptide assignment. Among peptides, 1,195 from the cortical DDA dataset were not matched to DIA experiments. While not tested, it is expected that the dynamic charge detector facilitated different peptide fragmentation by variable collision energy in DDA whereas DIA employs a static collision energy.

As a consequence of AGC target parameters and chromatographic co-elution of high abundance peptides, peptides of putatively low abundance are not identified in DDA experiments. To demonstrate the distribution of lens protein abundance, we compared the quantitative distribution of calculated protein abundances in each region of the lens (Supplemental Figure 1A,B,C). Search engine quantitation bias was reduced by protein assembly and normalization with the MaxLFQ algorithm in R. For proteins detected in paired studies, a constant necessary to median normalize DDA data to DIA was calculated and applied to all DDA peptides before protein group assembly. For proteins in both DDA and DIA experiments, a non-significant difference of abundance distribution was associated with each regional comparison (Wilcoxon Test). In all regions, the abundance distribution of proteins detected by DDA and DIA had higher means and medians than proteins exclusive to DIA or in the DIA dataset as a whole ($p < 2 \times 10^{-16}$). Therefore, we conclude that DIA enables enhanced accessibility to the low abundance lens proteome.

Evaluating the Use of Pan-Human or Library Free DIA search

Pan-human libraries provide a convenient, sample agnostic approach to DIA assay development. These libraries are highly effective in well-studied tissues or cell lines. Relative to library-free methods, the pan-human library has a significantly decreased search space, decreasing the time needed to process data. To evaluate the utility of pan-human libraries in lens proteomics, we independently searched each sample with and without the pan-human library provided by Rosenberger, Koh, et al.¹⁹ Since the pan-human library does not have variable modifications annotated, variable modification remained disabled in the library free search. As before, we compared regional feature identifications and performed overlap analysis (Figure 3). Consistent with library free DIA search, the pan-human cortex showed increase protein group diversity relative to the outer and inner nucleus samples. Though much more rapid for data analysis of all 9 replicates (162 minutes for pan-human, 1295 minutes for library free on a personal computer), it is clear that library free searching significantly improves identification quantity of protein groups and peptides. Although the pan-human library search identified more proteins and peptides (3,699/26,147 respective mean IDs in the cortex) than DDA, there is significant variation in nuclear peptide assignment (Figure 3B). In the outer nucleus, CV amounts to 8.6% of mean identifications, and 7.7% in the inner nucleus within technical replicates. Discrepancy is likely caused by a limited quantity of transitions used to define precursors but may be due to false matches.

In overlap analysis (Figure 3C,D), it was clear that both pan-human and library free methods effectively identify the majority of peptides and proteins expressed in all regions of the lens. However, a significant quantity of proteins (508) and peptides (5,467) were identified in the cortex of the library free search and not in pan-human search. A smaller quantity of cortical features (146 proteins, 2,827 peptides) was uniquely identified in the pan-human search. To determine the biological impact of these unique features, we performed statistical-enrichment analysis of unique library free proteins (859) and pan-human proteins (200) against the lens background (Table 1). No ontologies were statistically enriched from the pan-human search. However, the large quantity of lens-specific ontologies uniquely identified in the library free search demonstrate that the pan-human library does not adequately sample lens fiber cells for annotation of lens biology. Of great significance, Aquaporin-0, one of the most abundant lens proteins, is not in the pan-human library. For future lens studies and studies where PTMs are desirable to measure, it is then suggested that library free search be used in the absence of an exhaustive library.

Impact of fiber cell aging on proteome composition

A result of aging, deamidation of asparagine or glutamine is the most prevalent PTM in the lens.³⁰ While deamidation may be initiated by glutaminase-like enzymes, non-enzymatic deamidation is most common in the lens.^{31,32} To investigate changes in the lens proteome as a function of fiber cell age, we included one variable deamidation in the library free search parameters. Inclusion of a variable deamidation greatly increased the search space over a modification restricted search and may negatively impact FDR calculations. As a result, search time increased from 1,295 minutes with no modification enabled to 3,619 minutes with one variable deamidation allowed: calculations performed on a personal computer. Here, search time increased linearly with search space. In total, 4,564 protein groups and 48,474 peptides were detected. Among these, 1,681 proteins were partially deamidated at 5,577 unique deamidation sites. All of these figures exceed the capabilities of previous human lens proteomic measurements. In this study, we have not considered the peptide-centric effect that deamidation has on lens fiber aging, but rather included these peptides in protein assembly. While possible that some peptides are only identified as deamidated, the functional implication of site-specific modifications are beyond the scope of this discussion.

To evaluate the progressive changes in proteome composition and putative function as a function of fiber cell age, we first evaluated feature overlap between regions (Figure 4A,B). Regional segmentation of the lens allows clear comparisons of fiber cell aging whereas whole-lens measurements previously published are most representative of the high protein density cortical fibers. Exceptionally, 1,844 protein groups (41% of all identified) and 16,250 peptides (34%) were detected as common among lens regions. Here, DIA allows more comparisons in each region than proteins measured in any previous human lens DDA experiment. Thus, changes in whole-protein networks as a function of fiber cell age are best determined with DIA.

It is expected that few or no unique proteins are found in the inner nucleus relative to the outer nucleus, however 265 inner nucleus proteins were detected in the inner nucleus and not in the outer nucleus (Figure 4A). In reference to the lens proteome, enriched parent terms are generally translation related (nuclear-transcribed mRNA catabolic process, nonsense-mediated decay, SRP-dependent co-translational protein targeting to the membrane, viral transcription, translation initiation, rRNA processing, structural constituent of ribosome, cytosolic large ribosomal subunit). It is not expected that these modules, in the absence organelles, are active. It is possible that proteins unique to the inner nucleus are part of lens morphogenesis and are not degraded during early lens development. Nonetheless, in the absence of

organelles, it is unlikely that these proteins are metabolically active, and these changes should be considered non-functional. For the 238 proteins detected in the outer nucleus and not in the inner nucleus, no significant protein-network level enrichment was detected. For the 2,155 proteins measured in the cortex but not in either nuclear region, there is no significant functional enrichment in reference to the lens background, but cellular component ontologies including ER membrane, cytosol and integral component of the membrane are enriched in this protein subset.

After numeric comparisons of identification, we investigated protein abundance related differences in each region. With the 1,844 proteins identified in all regions, principal component analysis was performed (Figure 4C). The majority of variance in the dataset is explained by the transition of fiber cells from more classical cellular function in cortical fibers, to the less metabolically active fibers in nuclear regions of the lens, separating cortical and nuclear fibers along PC1. The 50 most descriptive cortical loadings are associated with extracellular matrix assembly, ATP:ADP antiporter activity and basement membrane formation while nuclear loadings are most correlated to cortical cytoskeleton organization and structural constituent of eye lens (data not shown). Secondly, proteins in the inner nucleus and outer nucleus separated along axis 2. The top 50 PC2 loadings correlated with the outer nucleus are weakly enriched for GDP/GTP binding and GTPase activity. There is weak enrichment of structural constituent of eye lens and phosphatase activity ontologies in loadings most associated with the inner nucleus. While these findings are consistent with previous lens biology studies, they do not take full advantage of the rich DIA dataset. Ontology searches based on a small subset of differentially expressed proteins may be too conservative to identify changes associated with low abundance proteins and analyses considering all identified proteins without considering associated abundance may increase type 1 error rate. Therefore, we set out to perform ontology analysis sensitive to each protein and its associated abundance.

It is clear from pairwise comparisons of protein expression (Figure 5A,B) that the density of most proteins in the cortex exceeds that of either the outer or inner nucleus. As a validation of quantitative comparisons, there is a measured significant increase in γ -crystallin density in each nuclear region respective to the cortex. Unlike the majority of other lens proteins, γ -crystallin is sustained at higher quantities in the lens nucleus due to its role in lens morphogenesis and establishment of the lens gradient of refractive index.³³ In contrast to previous discussion of insignificant change between the outer nucleus and inner nucleus of the lens, an insignificant, but appreciable trend emerges in the outer nucleus wherein outer nucleus proteins have higher abundance than inner nuclear proteins, which is expected with gradual protein degradation (Figure 5C). In the young lens, the maturation differences between the inner and outer nucleus are proportionally more significant than in an older lens. We suspect that with age, the moderate change in protein abundance between nuclear regions will approach homogeneity except for those proteins differentially expressed by biological function or spatially targeted cellular stress.

To evaluate changes in biology at a protein-network level, abundance-linked ontology analysis was performed. Crystallin proteins make up a disproportionate part of the proteome, so comparisons based on binary expression/non-expression do not represent the dynamic range of the lens proteome. Instead, a rank-based comparison was performed in PSEA-Quant – a modification of the GSEA algorithm.²⁶ As may be expected, each region was independently enriched for lens ontologies such as visual perception (GO:0007601) and lens fiber cell development (GO:0070307). Several ontologies were independently enriched in each region but are better delineated in between region calculations.

Between region calculations suffer from the same limitations as previously mentioned; the dramatic dynamic range of the lens proteome abundances does not warrant binary enriched/non-enriched ontology analysis. To quantify changes while

between region abundance changes, we again employed PSEA-quant. Briefly, the ratio of protein abundances was computed between each combination of technical replicates of two regions. Only leading proteins that were detected or imputed in all samples were used ($n = 1,954$). The rank-based enrichment method then identifies changes that are linked to changes between samples, with additional weighting given to proteins with significant changes. For our discussion here, we limit ontological enrichment comparisons to those between the cortex and inner nucleus.

In total, 133 ontologies were enriched in cortex in relation to the inner nucleus, and 73 of the inverse relationship (Table 2, Supplemental Table 4). While it is to be expected that organelle specific proteins are more abundant in the outer cortex relative to the inner nucleus, several inferences related to MCS function can be made. First, there is enrichment of cell-to-cell junctions in young fibers, inferring gradual age-related degradation of extracellular transport networks, even in a young lens, consistent with the formation of a barrier between the cortex and outer nucleus.^{4,34} Next, in the absence of a complex proteome, proteins in the inner nucleus are best suited for oxidoreductase activity by glutathione conjugation. In this model, oxidized glutathione is reduced by glutathione reductase, yielding NADP⁺ which can be reduced and used by glucose 6-phosphate dehydrogenase (G6PDH) to convert glucose-6-phosphate to 6-phosphogluconate as part of the pentose phosphate pathway. Previous results in an aged rat lens demonstrated that G6PDH is catalytically inactive in the inner nucleus³⁵ leaving fibers devoid of glycolysis function, even though the protein machinery is present.³⁶

A benefit of greater proteome coverage is that a greater proportion of networks may be considered for ontological significance. In PSEA-quant, networks that meet p-value and FDR significance fail to be deemed significant when <3 proteins from the studied ontology are present. For ontology networks enriched in the inner nucleus (Supplemental Table 4, Table 2), 22 of the significant ontologies are represented by less than 10 proteins in the dataset. Indeed, when DDA data was processed through the PSEA-quant pipeline with identical treatment, only 61 ontologies were enriched in either the cortex or inner nucleus (data not shown) – all with redundancy to the 206 ontologies from DIA results. Among cortically enriched DIA ontologies, few are sparsely populated as a result of the significant improvement in identifications. Thus, the enhanced proteome coverage improves preliminary identification of protein networks that are influential in lens fiber aging. Lastly, the enhanced peptidome coverage afforded by DIA provides a more robust quantitative assembly of protein groups, leading to more reproducible and biologically representative protein abundance measurement.

Finally, we evaluated the impact of fiber cell age-related deamidation on total age-related protein abundance in the lens. Peptide level search results with deamidation enabled and disabled were TMM normalized before protein assembly and significance testing (Figure 5D,E,F). In both the outer cortex and outer nucleus, the vast majority of the proteome is not highly deamidated, while the inner nucleus shows a trend towards deamidation. In the relatively young 18-year-old lens, it is not expected that the majority of the lens would be deamidated, but the gradual accumulation of deamidation demonstrated here establishes specificity of library-free search towards deamidated peptides. The young lens is less age-affected than an older lens but sustains deamidation on 1,681 proteins. The majority of these proteins are believed to be deamidated non-enzymatically, and prior DDA evidence (Schey, unpublished data) suggests that there will be even more peptides deamidated as a function of age. Given the assumption of non-enzymatic deamidation, ontological assignment of differences is not relevant until a greater quantity of biological replicates are available.

Conclusion

Here, we have shown that the use of DIA methodology for lens proteomics greatly improves identification of low abundance proteins and have demonstrated an effective way to search the data with lens biology in mind. The gradual

shift towards protein deamidation and proteome specialization for oxidoreductase activity clearly demonstrates that this method is sensitive to age-related fiber maturation measurements – and the enhanced ability of DIA to measure low abundance peptides improves protein-network level characterization of lens aging. Additionally, the application of PSEA-quant here demonstrates that an ontological search that is sensitive to protein abundance is well suited towards assessing lens biology. It is reasonable to suggest that the improved protein network coverage afforded by DIA paired to sensitive analysis pipelines may allow improved insights into human lens aging and cataract formation.

Data Availability

Reports from DIA-NN, MaxQuant evidence file and normalized protein group assemblies are available as Supplemental Files 1-8.

Acknowledgements

We would like to acknowledge the Vanderbilt Mass Spectrometry Research Center, particularly Kristie L. Rose and W. Hayes McDonald for assistance in developing DIA methods for this experiment. We also thank Zhen Wang and all other members of the Schey lab for thoughtful conversations in the development of this work. Finally, we acknowledge Vadim Demichev for his ongoing support of DIA data analysis. Financial support is acknowledged from NIH grants EY013462 (KLS), EY008126, and T32 GM065086 (LSC).

References

- (1) Schmid, P. W. N.; Lim, N. C. H.; Peters, C.; Back, K. C.; Bourgeois, B.; Pirolto, F.; Richter, B.; Peschek, J.; Puk, O.; Amarie, O. V.; Dalke, C.; Haslbeck, M.; Weinkauff, S.; Madl, T.; Graw, J.; Buchner, J. Imbalances in the Eye Lens Proteome Are Linked to Cataract Formation. *Nat Struct Mol Biol* **2021**, 28 (2), 143–151. <https://doi.org/10.1038/s41594-020-00543-9>.
- (2) Horwitz, J. Alpha-Crystallin Can Function as a Molecular Chaperone. *Proceedings of the National Academy of Sciences* **1992**, 89 (21), 10449–10453. <https://doi.org/10.1073/pnas.89.21.10449>.
- (3) Wu, S.-Y.; Zou, P.; Fuller, A. W.; Mishra, S.; Wang, Z.; Schey, K. L.; Mchaourab, H. S. Expression of Cataract-Linked γ -Crystallin Variants in Zebrafish Reveals a Proteostasis Network That Senses Protein Stability. *J Biol Chem* **2016**, 291 (49), 25387–25397. <https://doi.org/10.1074/jbc.M116.749606>.
- (4) Sweeney, M. H.; Truscott, R. J. An Impediment to Glutathione Diffusion in Older Normal Human Lenses: A Possible Precondition for Nuclear Cataract. *Exp Eye Res* **1998**, 67 (5), 587–595. <https://doi.org/10.1006/exer.1998.0549>.
- (5) Moffat, B. A.; Landman, K. A.; Truscott, R. J. W.; Sweeney, M. H. J.; Pope, J. M. Age-Related Changes in the Kinetics of Water Transport in Normal Human Lenses. *Experimental Eye Research* **1999**, 69 (6), 663–669. <https://doi.org/10.1006/exer.1999.0747>.
- (6) Wilmarth, P. A.; Tanner, S.; Dasari, S.; Nagalla, S. R.; Riviere, M. A.; Bafna, V.; Pevzner, P. A.; David, L. L. Age-Related Changes in Human Crystallins Determined from Comparative Analysis of Post-Translational Modifications in Young and Aged Lens: Does Deamidation Contribute to Crystallin Insolubility? *Journal of Proteome Research* **2006**, 5 (10), 2554–2566. <https://doi.org/10.1021/pr050473a>.
- (7) Aquilina, J. A.; Benesch, J. L. P.; Ding, L. L.; Yaron, O.; Horwitz, J.; Robinson, C. V. Phosphorylation of AB-Crystallin Alters Chaperone Function through Loss of Dimeric Substructure. *Journal of Biological Chemistry* **2004**, 279 (27), 28675–28680. <https://doi.org/10.1074/jbc.M403348200>.
- (8) Schaefer, H.; Marcus, K.; Sickmann, A.; Herrmann, M.; Klose, J.; Meyer, Helmut E. Identification of Phosphorylation and Acetylation Sites in Alpha-A Crystallin of the Eye Lens (Mus Musculus) after Two-Dimensional Gel Electrophoresis. *Analytical and Bioanalytical Chemistry* **2003**, 376 (7), 966–972. <https://doi.org/10.1007/s00216-003-1983-1>.
- (9) Ueda, Y.; Duncan, M. K.; David, L. L. Lens Proteomics: The Accumulation of Crystallin Modifications in the Mouse Lens with Age. *Invest. Ophthalmol. Vis. Sci.* **2002**, 43 (1), 205–215.
- (10) Gillet, L. C.; Navarro, P.; Tate, S.; Röst, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R. Targeted Data Extraction of the MS/MS Spectra Generated by Data-Independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Molecular & Cellular Proteomics* **2012**, 11 (6), O111.016717. <https://doi.org/10.1074/mcp.O111.016717>.
- (11) Wang, Z.; Han, J.; David, L. L.; Schey, K. L. Proteomics and Phosphoproteomics Analysis of Human Lens Fiber Cell Membranes. *Investigative Ophthalmology & Visual Science* **2013**, 54 (2), 1135. <https://doi.org/10.1167/iovs.12-11168>.
- (12) Cantrell, L. S.; Schey, K. L. Proteomic Characterization of the Human Lens and Cataractogenesis. *Expert Review of Proteomics* **2021**, 1–17. <https://doi.org/10.1080/14789450.2021.1913062>.
- (13) MacCoss, M. J.; McDonald, W. H.; Saraf, A.; Sadygov, R.; Clark, J. M.; Tasto, J. J.; Gould, K. L.; Wolters, D.; Washburn, M.; Weiss, A.; Clark, J. I.; Yates, J. R. Shotgun Identification of Protein Modifications from Protein Complexes and Lens Tissue. *Proc Natl Acad Sci U S A* **2002**, 99 (12), 7900–7905. <https://doi.org/10.1073/pnas.122231399>.
- (14) Wu, X.; Liu, Z.; Zhang, X.; Wang, D.; Long, E.; Wang, J.; Li, W.; Lai, W.; Cao, Q.; Hu, K.; Chen, W.; Lin, H.; Liu, Y. Proteomics Analysis and Proteogenomic Characterization of Different Physiopathological Human Lenses. *BMC Ophthalmol* **2017**, 17 (1). <https://doi.org/10.1186/s12886-017-0642-9>.
- (15) Lee, H.-J.; Kim, H.-J.; Liebler, D. C. Efficient Microscale Basic Reverse Phase Peptide Fractionation for Global and Targeted Proteomics. *J. Proteome Res.* **2016**, 15 (7), 2346–2354. <https://doi.org/10.1021/acs.jproteome.6b00102>.
- (16) Doerr, A. DIA Mass Spectrometry. *Nat Methods* **2015**, 12 (1), 35–35. <https://doi.org/10.1038/nmeth.3234>.
- (17) Michalski, A.; Cox, J.; Mann, M. More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority Is Inaccessible to Data-Dependent LC-MS/MS. *J. Proteome Res.* **2011**, 10 (4), 1785–1793. <https://doi.org/10.1021/pr101060v>.
- (18) Searle, B. C.; Pino, L. K.; Egertson, J. D.; Ting, Y. S.; Lawrence, R. T.; MacLean, B. X.; Villén, J.; MacCoss, M. J. Chromatogram Libraries Improve Peptide Detection and Quantification by Data Independent Acquisition Mass Spectrometry. *Nat Commun* **2018**, 9 (1). <https://doi.org/10.1038/s41467-018-07454-w>.
- (19) Rosenberger, G.; Koh, C. C.; Guo, T.; Röst, H. L.; Kouvonen, P.; Collins, B. C.; Heusel, M.; Liu, Y.; Caron, E.; Vichalkovski, A.; Faini, M.; Schubert, O. T.; Faridi, P.; Ebhardt, H. A.; Matondo, M.; Lam, H.; Bader, S. L.; Campbell, D. S.; Deutsch, E. W.; Moritz, R. L.; Tate, S.; Aebersold, R. A Repository of Assays to Quantify 10,000 Human Proteins by SWATH-MS. *Scientific Data* **2014**, 1 (1). <https://doi.org/10.1038/sdata.2014.31>.
- (20) Demichev, V.; Messner, C. B.; Vernardis, S. I.; Lilley, K. S.; Ralser, M. DIA-NN: Neural Networks and Interference Correction Enable Deep Proteome Coverage in High Throughput. *Nature Methods* **2020**, 17 (1), 41–44. <https://doi.org/10.1038/s41592-019-0638-x>.

- (21) Bekker-Jensen, D. B.; Bernhardt, O. M.; Hoglebe, A.; Martinez-Val, A.; Verbeke, L.; Gandhi, T.; Kelstrup, C. D.; Reiter, L.; Olsen, J. V. Rapid and Site-Specific Deep Phosphoproteome Profiling by Data-Independent Acquisition without the Need for Spectral Libraries. *Nat Commun* **2020**, *11* (1). <https://doi.org/10.1038/s41467-020-14609-1>.
- (22) Trudgian, D. C.; Fischer, R.; Guo, X.; Kessler, B. M.; Mirzaei, H. GOAT - A Simple LC-MS/MS Gradient Optimization Tool. *PROTEOMICS* **2014**, *14* (12), 1467–1471. <https://doi.org/10.1002/pmic.201300524>.
- (23) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. *Bioinformatics* **2010**, *26* (7), 966–968. <https://doi.org/10.1093/bioinformatics/btq054>.
- (24) Cox, J.; Mann, M. MaxQuant Enables High Peptide Identification Rates, Individualized p.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nature Biotechnology* **2008**, *26* (12), 1367–1372. <https://doi.org/10.1038/nbt.1511>.
- (25) Adusumilli, R.; Mallick, P. Data Conversion with ProteoWizard MsConvert. *Methods Mol Biol* **2017**, *1550*, 339–368. https://doi.org/10.1007/978-1-4939-6747-6_23.
- (26) Lavallée-Adam, M.; Rauniyar, N.; McClatchy, D. B.; Yates, J. R. PSEA-Quant: A Protein Set Enrichment Analysis on Label-Free and Label-Based Protein Quantification Data. *J. Proteome Res.* **2014**, *13* (12), 5496–5509. <https://doi.org/10.1021/pr500473n>.
- (27) Ritchie, M. E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C. W.; Shi, W.; Smyth, G. K. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Research* **2015**, *43* (7), e47–e47. <https://doi.org/10.1093/nar/gkv007>.
- (28) Thomas, P. D. PANTHER: A Library of Protein Families and Subfamilies Indexed by Function. *Genome Research* **2003**, *13* (9), 2129–2141. <https://doi.org/10.1101/gr.772403>.
- (29) Zhao, Y.; Wilmarth, P. A.; Cheng, C.; Limi, S.; Fowler, V. M.; Zheng, D.; David, L. L.; Cvekl, A. Proteome-Transcriptome Analysis and Proteome Remodeling in Mouse Lens Epithelium and Fibers. *Experimental Eye Research* **2019**, *179*, 32–46. <https://doi.org/10.1016/j.exer.2018.10.011>.
- (30) Lampi, K. J.; Ma, Z.; Hanson, S. R. A.; Azuma, M.; Shih, M.; Shearer, T. R.; Smith, D. L.; Smith, J. B.; David, L. L. Age-Related Changes in Human Lens Crystallins Identified by Two-Dimensional Electrophoresis and Mass Spectrometry. *Experimental Eye Research* **1998**, *67* (1), 31–43. <https://doi.org/10.1006/exer.1998.0481>.
- (31) Hains, P. G.; Truscott, R. J. W. Age-Dependent Deamidation of Lifelong Proteins in the Human Lens. *Invest. Ophthalmol. Vis. Sci.* **2010**, *51* (6), 3107–3114. <https://doi.org/10.1167/iovs.09-4308>.
- (32) Forsythe, H. M.; Vetter, C. J.; Jara, K. A.; Reardon, P. N.; David, L. L.; Barbar, E. J.; Lampi, K. J. Altered Protein Dynamics and Increased Aggregation of Human Γ S-Crystallin Due to Cataract-Associated Deamidations. *Biochemistry* **2019**, *58* (40), 4112–4124. <https://doi.org/10.1021/acs.biochem.9b00593>.
- (33) Slingsby, C.; Wistow, G. J. Functions of Crystallins in and out of Lens: Roles in Elongated and Post-Mitotic Cells. *Prog Biophys Mol Biol* **2014**, *115* (1), 52–67. <https://doi.org/10.1016/j.pbiomolbio.2014.02.006>.
- (34) Grey, A. C.; Jacobs, M. D.; Gonen, T.; Kistler, J.; Donaldson, P. J. Insertion of MP20 into Lens Fibre Cell Plasma Membranes Correlates with the Formation of an Extracellular Diffusion Barrier. *Exp Eye Res* **2003**, *77* (5), 567–574. [https://doi.org/10.1016/s0014-4835\(03\)00192-1](https://doi.org/10.1016/s0014-4835(03)00192-1).
- (35) Dovrat, A.; Scharf, J.; Eisenbach, L.; Gershon, D. G6PD Molecules Devoid of Catalytic Activity Are Present in the Nucleus of the Rat Lens. *Exp Eye Res* **1986**, *42* (5), 489–496. [https://doi.org/10.1016/0014-4835\(86\)90008-4](https://doi.org/10.1016/0014-4835(86)90008-4).
- (36) Beebe, D. C.; Truscott, R. J. W. Counterpoint: The Lens Fluid Circulation Model—A Critical Appraisal. *Invest. Ophthalmol. Vis. Sci.* **2010**, *51* (5), 2306. <https://doi.org/10.1167/iovs.10-5350a>.

Figures and Tables

Figure 1 - Cartoon of lens with regions annotated. Fiber cells are differentiated from epithelial cells (purple) into the cortex (red) where organelles (teal circles) degrade. After the fifth decade of life, an extracellular diffusion barrier forms between the outer nucleus (green) and cortex, and the oldest fibers, formed in utero may be found in the inner nucleus (orange). The entire lens is encompassed by a collagenous capsule (blue).

Figure 2 - Comparison of Data Dependent (DDA) and Data Independent (DIA) acquisition scan modes for 3 technical replicates of 3 lens regions, PTMs disabled. Only unique leading proteins considered. In total, 4,535 proteins and 35,465 peptides were detected across all experiments. DIA clearly outperforms DDA for identifications and search provides low CV values in protein (A) and peptide (B) comparisons. Upset plots of proteins(C, 82% represented) and peptides (D, 80% represented) reveals near complete coverage of DIA proteome/peptidomes.

Figure 3 - Comparison of Pan-Human (PH) library search with Library Free (LF) DIA-NN Searching, PTMs disabled. Only unique leading proteins considered for identification. A) Protein identifications in each lens region, B) Peptide identifications in each lens region; no modifications were considered. Library-free search produces a greater quantity of IDs filtered at 1% FDR. Upset plots of protein groups (C, 85% represented) and peptides (D, 83% represented) identified in at least one technical replicate of each search method.

Table 1 – Ontological enrichment of proteins detected in library free DIA search and not in Pan-Human DIA search (n = 859). No ontologies were enriched among proteins identified in Pan-Human DIA search and not in library free DIA search (n = 200). Ontology enrichment and statistical analysis performed by PANTHER²⁸, using all identified lens proteins as the statistical background.

GO Term	Annotation	# in Set	Fold Enrichment	FDR
Unique in Library Free Search				
GOBP				
GO:0002088	Lens development in camera-type eye	26	3.56	0.0003
GO:0050953	Sensory perception of light stimulus	43	3.11	<1x10 ⁻⁴
GO:0007601	Visual perception	43	3.11	<1x10 ⁻⁴
GO:0043010	Camera-type eye development	51	2.66	0.0006
GO:0007600	Sensory perception	52	2.30	<1x10 ⁻⁴
GO:0001654	Eye development	55	2.56	0.0003
GO:0048880	Sensory system development	55	2.51	<1x10 ⁻⁴
GO:0150063	Visual system development	55	2.51	<1x10 ⁻⁴
GO:0007423	Sensory organ development	66	2.37	<1x10 ⁻⁴
GO:0050877	Nervous system process	69	1.83	0.0024
GOMF				
GO:0005212	Structural constituent of eye lens	15	4.11	0.0001
GO:0004888	Transmembrane signaling receptor activity	34	2.13	0.0003
GO:0022857	Transmembrane transporter activity	80	1.65	0.0001
GO:0005215	Transporter activity	90	1.57	0.0002
GOCC				
GO:0005877	Integral component of plasma membrane	93	1.80	<1x10 ⁻⁴
GO:0031226	Intrinsic component of plasma membrane	101	1.78	<1x10 ⁻⁴
GO:0016021	Integral component of membrane	334	1.41	<1x10 ⁻⁴
GO:0031224	Intrinsic component of membrane	345	1.41	<1x10 ⁻⁴
GO:0005886	Plasma membrane	310	1.22	0.0060
GO:0071944	Cell periphery	351	1.21	0.0024

Figure 4 - The spatial distribution and abundance of peptides and proteins across the lens displays significant differences as a function of fiber cell age. Data with one variable deamidation used for analysis, intact protein groups used for comparisons. A) Lens protein group overlap with one technical replicate required for consideration, B) Lens peptide overlap, charge states combined, one technical replicate required for consideration. C) PCA demonstrated that cortical fiber and nuclear fiber compositions show significant differences on PC axis 1, and abundance of nuclear fiber proteins contributes to separation on PC axis 2.

Table 2 - Reduced table of ontologies specifically enriched in the cortex or inner nucleus over respective partner. Less specific, redundant and terms less relevant to lens biology were filtered. Full list of significant annotations found in Supplemental Table 4. Calculations performed with PSEA-quant.²⁶

GO Term	Annotation	PES	p-value	Proteins in Set
Enriched in Cortex over Inner Nucleus				
GO:0000184	Nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	42.48	<1.0x10 ⁻⁴	54
GO:0022625	Cytosolic large ribosomal subunit	22.62	<1.0x10 ⁻⁴	27
GO:0031982	Vesicle	115.25	<1.0x10 ⁻⁴	171
GO:0005886	Plasma membrane	233.86	<1.0x10 ⁻⁴	363
GO:0006414	Translational elongation	40.01	<1.0x10 ⁻⁴	49
GO:0005912	Adherens junction	34.49	<1.0x10 ⁻⁴	47
GO:0043624	Cellular protein complex disassembly	40.06	<1.0x10 ⁻⁴	50
GO:0006886	Intracellular protein transport	118.86	<1.0x10 ⁻⁴	176
GO:0070161	Anchoring junction	38.60	<1.0x10 ⁻⁴	53
GO:0006415	Translational termination	38.29	<1.0x10 ⁻⁴	46
GO:0015075	Ion transmembrane transporter activity	39.36	<1.0x10 ⁻⁴	53
GO:0016071	mRNA metabolic process	95.93	<1.0x10 ⁻⁴	144
GO:0031090	Organelle membrane	222.46	<1.0x10 ⁻⁴	347
GO:0032984	Protein-containing complex disassembly	41.32	<1.0x10 ⁻⁴	52
GO:0005544	Calcium-dependent phospholipid binding	9.07	0.0002	11
GO:0008324	Cation transmembrane transporter activity	26.73	0.0002	37
GO:0045296	Cadherin binding	8.91	0.0008	10
GO:0015077	Monovalent inorganic cation transmembrane transporter activity	19.90	0.0008	27
KEGG	ECM receptor interaction	15.31	<1.0x10 ⁻⁴	17
KEGG	Ribosome	38.05	<1.0x10 ⁻⁴	46
KEGG	Focal adhesion	35.02	0.0001	49
Reactome	Translation	75.19	0.0003	112
Reactome	Metabolism of amino acids and derivatives	86.99	0.0004	132
Reactome	L1CAM interactions	28.91	0.001	41
Enriched in Inner Nucleus over Cortex				
GO:0007601	Visual perception	19.74	<1.0x10 ⁻⁴	33
GO:0005212	Structural constituent of eye lens	12.84	<1.0x10 ⁻⁴	17
GO:0016765	Transferase activity, transferring alkyl or aryl (other than methyl) groups	9.38	<1.0x10 ⁻⁴	15
GO:0019318	Hexose metabolic process	32.67	<1.0x10 ⁻⁴	62
GO:0007600	Sensory perception	28.03	<1.0x10 ⁻⁴	51
GO:0002088	Lens development in camera-type eye	8.34	<1.0x10 ⁻⁴	12
GO:0016491	Oxidoreductase activity	60.55	0.0001	120
GO:0046983	Protein dimerization activity	71.11	0.0001	141
GO:0097354	Prenylation	3.53	0.0001	5
GO:0008318	Protein prenyltransferase activity	3.53	0.0001	5
GO:0045762	Positive regulation of adenylate cyclase activity	3.51	0.0002	5
GO:0016056	Rhodopsin mediated signaling pathway	4.54	0.0002	7
GO:0016614	Oxidoreductase activity, acting on CH-OH group of donors	17.62	0.0003	33
GO:0006749	Glutathione metabolic process	9.19	0.0003	16
GO:0016616	Oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	17.62	0.0005	33
GO:0008238	Exopeptidase activity	14.65	0.0006	27
GO:0006096	Glycolytic process	11.22	0.0006	20
GO:0006733	Oxidoreduction coenzyme metabolic process	12.75	0.0008	24
GO:0046872	Metal ion binding	172.67	0.0009	355
GO:0006656	Phosphatidylcholine biosynthetic process	2.17	0.001	3
KEGG	Glutathione metabolism	11.72	0.0006	21
KEGG	Glycolysis gluconeogenesis	14.76	0.0008	27
Reactome	Metabolism of carbohydrates	37.91	<1.0x10 ⁻⁴	73
Reactome	Beta catenin independent WNT signaling	27.14	0.0005	53
Reactome	PCP CE pathway	23.13	0.0007	45
Reactome	Fructose metabolism	2.65	0.0008	4
Reactome	Biological oxidations	16.75	0.0009	32

Supplemental Figures

Supplemental Figure 1 - Histogram comparison of Log2 abundances between DDA and DIA, each calculated with the diann R package. For each region of the lens A) Cortex, B) Outer Nucleus, C) Inner Nucleus, quantitative median normalization factors were calculated for matched protein groups between DDA and DIA and applied to all unmatched protein groups. DDA abundances (grey) visually resemble the distribution of matched ID DIA proteins (red). The majority of proteins uniquely identified in DIA (orange) were of lower mean abundance compared to the DDA proteome. The resulting complete DIA experiment (beige) has a moderately decreased median and wider quantitative distribution than DDA.

Supplemental Table 1 - Full chromatography details with gradient calculated at 3-minute intervals by GOAT²² with an assumed 11-minute retention time delay.

Time (min)	%B
0.0	2.0
1.0	2.0
2.0	2.0
3.0	2.0
6.0	7.1
9.0	8.4
12.0	9.4
15.0	10.1
18.0	11.1
21.0	11.8
24.0	12.3
27.0	12.6
30.0	13.7
33.0	14.4
36.0	14.7
39.0	15.0
42.0	15.5
45.0	16.2
48.0	16.5
51.0	17.0
54.0	17.7
57.0	18.4
60.0	18.7
63.0	19.2
66.0	19.9
69.0	20.6
72.0	21.5
75.0	22.4
78.0	24.1
81.0	90.0
83.0	90.0
84.0	2.0
85.0-95.0	2.0

Supplemental Table 2 - Calculated mass windows for DIA experiments. Optimization done in Skyline²³ from 400-1100 m/z with staggered window deconvolution selected. Windows exceeding the specified range were discarded.

Calculated m/z Window	
400.4319-420.441	410.4364-430.4455
420.441-440.4501	430.4455-450.4546
440.4501-460.4592	450.4546-470.4637
460.4592-480.4683	470.4637-490.4728
480.4683-500.4774	490.4728-510.4819
500.4774-520.4865	510.4819-530.491
520.4865-540.4956	530.491-550.5001
540.4956-560.5047	550.5001-570.5092
560.5047-580.5138	570.5092-590.5183
580.5138-600.5229	590.5183-610.5274
600.5229-620.5319	610.5274-630.5365
620.5319-640.541	630.5365-650.5456
640.541-660.5501	650.5456-670.5547
660.5501-680.5592	670.5547-690.5638
680.5592-700.5683	690.5638-710.5729
700.5683-720.5774	710.5729-730.582
720.5774-740.5865	730.582-750.5911
740.5865-760.5956	750.5911-770.6002
760.5956-780.6047	770.6002-790.6093
780.6047-800.6138	790.6093-810.6183
800.6138-820.6229	810.6183-830.6274
820.6229-840.632	830.6274-850.6365
840.632-860.6411	850.6365-870.6456
860.6411-880.6502	870.6456-890.6547
880.6502-900.6593	890.6547-910.6638
900.6593-920.6684	910.6638-930.6729
920.6684-940.6775	930.6729-950.682
940.6775-960.6866	950.682-970.6911
960.6866-980.6957	970.6911-990.7002
980.6957-1000.7048	990.7002-1010.7093
1000.7048-1020.7138	1010.7093-1030.7184
1020.7138-1040.7229	1030.7184-1050.7275
1040.7229-1060.732	1050.7275-1070.7366
1060.732-1080.7411	1070.7366-1090.7457
1080.7411-1100.7502	

Supplemental Table 3 - Complete protein group and gene group ID of each protein identified as differentially enriched in supplemental figure 2 (n=190). Only proteins with limma moderated p-values below 1E-6 and Log2 FC in excess of 4 are shown.

Identifier		Cortex – Inner Nucleus		Cortex – Outer Nucleus		Outer Nucleus – Inner Nucleus	
Protein Group	Gene Name	Log2FC	p-value	Log2FC	p-value	Log2FC	p-value
P08670	VIM	6.66892192	2.86E-13	3.39293944	9.32E-13	3.27598247	5.99E-12
P98160	HSPG2	4.95910813	6.04E-13	7.79252315	7.31E-14	-2.833415	8.40E-11
P07315	CRYGC	-6.5520202	6.80E-13	-6.4574087	5.57E-13	-0.0946115	0.00715305
P02458	COL2A1	9.20151563	8.27E-13	6.70124025	1.13E-12	2.50027538	2.07E-08
P35555	FBN1	4.58564351	2.95E-12	6.62156761	3.27E-13	-2.0359241	8.63E-10
P07197; P08670; P17661	NEFM	6.76827947	4.19E-12	3.72234638	2.86E-10	3.04593309	2.05E-10
P39060	COL18A1	5.17064277	6.44E-12	7.67359385	4.32E-11	-2.5029511	5.55E-08
P62805	H4C1	8.16936591	9.25E-12	7.68137705	1.81E-12	0.48798886	0.00035163
P11844	CRYGA	-4.5561994	1.21E-11	-1.1113656	4.81E-06	-3.4448338	4.01E-10
P22352	GPX3	4.58345127	1.55E-11	5.21791167	8.15E-11	-0.6344604	0.00026116
O15230	LAMA5	4.63967449	1.84E-11	6.59226578	5.17E-12	-1.9525913	4.38E-07
P68371	TUBB4B	6.45206734	1.88E-11	5.68257606	1.69E-09	0.76949127	0.00388688
Q14767	LTBP2	5.39539249	2.98E-11	9.52965947	8.97E-11	-4.134267	1.59E-08
Q9BXM0	PRX	4.33953939	2.98E-11	2.39384614	1.17E-10	1.94569324	1.94E-09
P08133	ANXA6	5.58346148	3.69E-11	5.06885014	1.75E-11	0.51461134	0.00184285
Q14031	COL4A6	4.67805838	3.72E-11	7.58270217	1.75E-12	-2.9046438	3.29E-09
O95715	CXCL14	4.19268339	8.42E-11	4.02206237	4.88E-08	0.17062102	0.3069832
P05141; P12235; P12236;	SLC25A5	6.31984217	9.25E-11	6.02403811	3.02E-11	0.29580406	0.03960826
P35555; P35556	FBN1	5.30032314	9.98E-11	6.9658835	6.14E-11	-1.6655604	1.36E-06
P18206	VCL	4.39095105	1.08E-10	2.5013772	2.72E-10	1.88957385	3.14E-08
P07437	TUBB	4.02354901	1.14E-10	3.917119	2.46E-10	0.10643001	0.20177378
P07355	ANXA2	4.68813924	1.20E-10	4.26528574	3.05E-10	0.4228535	0.00312475
P55001	MFAP2	5.00986917	1.27E-10	6.69832329	3.97E-10	-1.6884541	4.70E-06
P14923	JUP	5.0866772	1.28E-10	4.19834959	9.72E-11	0.88832761	0.00024065
P07942	LAMB1	4.75554961	1.30E-10	NA	NA	NA	NA
P07437; P68371; Q13885; Q9BVA1	TUBB	5.46275397	1.70E-10	4.84495964	1.47E-10	0.61779433	0.00164571
P07315; P07316	CRYGC	-4.14956	1.76E-10	-3.9748889	4.60E-11	-0.1746711	0.00574235
P26232	CTNNA2	4.68398297	1.80E-10	2.58988669	5.06E-11	2.09409628	1.32E-08
Q9UHG3	PCYOX1	5.47404263	1.90E-10	5.92480186	2.11E-10	-0.4507592	0.02151154
P24539	ATP5PB	4.64893694	1.92E-10	6.44102653	1.19E-08	-1.7920896	5.25E-05
P14625	HSP90B1	5.49856524	2.06E-10	5.9147676	3.88E-09	-0.4162024	0.04956131
P17931	LGALS3	6.68427944	2.13E-10	5.2428097	1.70E-10	1.44146975	1.10E-05
P35579	MYH9	4.12594807	2.17E-10	3.47823561	2.00E-09	0.64771245	0.00017733
P14543	NID1	4.65034162	2.25E-10	NA	NA	NA	NA
P62249	RPS16	4.42709456	2.33E-10	NA	NA	NA	NA
Q00325	SLC25A3	5.65800838	2.56E-10	6.33367123	9.22E-13	-0.6756629	0.0008649
P04083	ANXA1	4.30677202	4.07E-10	4.83509457	2.15E-11	-0.5283226	0.0013305
P23284	PPIB	5.09808908	4.10E-10	4.5714633	2.03E-11	0.52662578	0.00114678
P55268	LAMB2	4.69969161	5.79E-10	6.18931871	7.53E-12	-1.4896271	7.45E-06
Q9UDW1	UQCR10	4.46314175	7.54E-10	NA	NA	NA	NA
Q14011	CIRBP	4.03638458	7.97E-10	6.15225661	1.69E-05	-2.115872	0.0061563
Q86UX2	ITIH5	5.97660612	8.03E-10	NA	NA	NA	NA
P54709	ATP1B3	4.24814109	9.76E-10	3.03777921	2.58E-09	1.21036188	1.48E-05
P07339	CTSD	4.10941342	1.00E-09	4.34191593	3.38E-08	-0.2325025	0.21194613
Q02978	SLC25A11	4.75375269	1.23E-09	5.60417155	6.91E-10	-0.8504189	0.00192228
P25705	ATP5F1A	7.56814332	1.23E-09	7.00909659	6.74E-12	0.55904673	0.03626309
P12236	SLC25A6	8.18681714	1.26E-09	6.88851149	7.02E-08	1.29830565	0.00515544
P35268	RPL22	5.58254275	1.27E-09	5.40407634	1.97E-09	0.17846641	0.42182063
Q9H9B4	SFXN1	4.27695478	1.29E-09	5.69038357	1.18E-07	-1.4134288	0.00106241
P12235	SLC25A4	5.18738214	1.60E-09	5.727416	9.88E-10	-0.5400339	0.01917183
Q9NZU5	LMCD1	4.18156219	2.30E-09	3.72538195	9.24E-09	0.45618024	0.01405188

Identifier		Cortex – Inner Nucleus		Cortex – Outer Nucleus		Outer Nucleus – Inner Nucleus	
<i>Protein Group</i>	<i>Gene Name</i>	<i>Log2FC</i>	<i>p-value</i>	<i>Log2FC</i>	<i>p-value</i>	<i>Log2FC</i>	<i>p-value</i>
P09543	CNP	4.89091723	2.31E-09	2.18373171	1.12E-08	2.70718552	2.57E-07
P62269	RPS18	4.92251802	2.66E-09	6.42948495	2.47E-09	-1.5069669	0.00015161
P05141; P12236	SLC25A5	4.3982441	3.08E-09	4.50908265	7.17E-08	-0.1108385	0.59114278
P08574	CYC1	4.54820943	5.21E-09	5.05520506	3.94E-09	-0.5069956	0.03981098
P02768	ALB	4.42283915	5.95E-09	4.14540882	1.16E-09	0.27743032	0.08862824
P02792	FTL	5.73106316	6.03E-09	3.44856498	4.02E-11	2.28249818	1.47E-06
Q9Y6C2	EMILIN1	7.40671534	6.83E-09	NA	NA	NA	NA
O00571	DDX3X	4.01030195	7.21E-09	3.60472786	6.05E-09	0.40557409	0.04747256
P11142; P54652	HSPA8	4.03350703	8.62E-09	2.20698612	1.91E-08	1.82652091	1.26E-06
P30101	PDIA3	6.49876066	9.65E-09	6.44912587	6.83E-09	0.04963479	0.86702537
P84095	RHOG	4.04872968	9.68E-09	3.36683069	1.32E-09	0.68189899	0.00228158
P14543; Q14112	NID1	5.61162594	1.00E-08	NA	NA	NA	NA
P26232; P35221	CTNNA2	4.63967841	1.01E-08	2.65536953	6.90E-10	1.98430888	1.34E-06
Q09666	AHNAK	4.15792839	1.12E-08	4.99306058	1.66E-09	-0.8351322	0.00090172
O60762	DPM1	4.69258982	1.13E-08	5.67932685	3.50E-08	-0.986737	0.00898065
O95837; P29992; P50148	GNA14	4.19636458	1.35E-08	2.7809385	2.84E-08	1.41542608	2.05E-05
P62829	RPL23	5.72309474	1.48E-08	6.89451421	3.67E-10	-1.1714195	0.00201671
P26885	FKBP2	4.69428022	1.77E-08	4.65526496	1.61E-08	0.03901525	0.85265304
P04350; P07437; P68371; Q13885; Q9BUF5; Q9BVA1	TUBB4A	4.3801227	2.01E-08	4.30810166	6.25E-10	0.07202104	0.63422784
O00468	AGRN	4.95652906	2.04E-08	NA	NA	NA	NA
P12235; P12236	SLC25A4	4.69024973	2.07E-08	4.82185047	1.26E-09	-0.1316007	0.50493794
P53007	SLC25A1	4.4487355	2.19E-08	4.88862886	4.53E-11	-0.4398934	0.02747621
Q02318	CYP27A1	4.65684008	2.19E-08	NA	NA	NA	NA
P29400	COL4A5	4.80333051	2.40E-08	NA	NA	NA	NA
P26373	RPL13	4.46852755	2.45E-08	NA	NA	NA	NA
O15397; O95373	IPO8	-4.1288738	2.52E-08	-3.1468444	3.67E-07	-0.9820294	4.42E-05
P62701	RPS4X	6.15415181	2.83E-08	7.18539784	2.95E-08	-1.031246	0.02001256
P62424	RPL7A	4.3625053	3.20E-08	NA	NA	NA	NA
Q6ZMP0	THSD4	4.55883866	3.56E-08	NA	NA	NA	NA
O96000	NDUFB10	4.3475031	3.71E-08	5.06760997	1.79E-08	-0.7201069	0.01594434
P60866	RPS20	5.73454418	3.72E-08	NA	NA	NA	NA
P08670; P17661	VIM	8.32382654	3.90E-08	3.53490929	1.15E-11	4.78891724	5.67E-07
P09038	FGF2	4.65917747	4.07E-08	4.45860064	2.11E-09	0.20057683	0.4063334
O75923	DYSF	4.60192761	4.25E-08	0.89878401	3.10E-06	3.70314359	9.77E-08
P45880	VDAC2	9.02016459	5.00E-08	5.59167903	3.16E-12	3.42848556	1.92E-05
Q14112	NID2	4.71592388	5.33E-08	NA	NA	NA	NA
P35613	BSG	4.67458128	5.44E-08	4.345681	1.48E-10	0.32890028	0.12606528
P16035	TIMP2	-4.4159174	5.61E-08	-3.695219	2.62E-07	-0.7206985	2.90E-05
P50914	RPL14	5.63576557	5.79E-08	NA	NA	NA	NA
P46821; P78559	MAP1B	5.14941112	7.57E-08	3.28706816	4.76E-09	1.86234296	9.74E-05
Q07960	ARHGAP1	4.88553221	8.11E-08	4.03352219	3.31E-10	0.85201002	0.00326824
P30050	RPL12	5.0826088	1.00E-07	NA	NA	NA	NA
P46821	MAP1B	4.37042052	1.06E-07	3.34071983	6.61E-08	1.02970069	0.00093862
P18124	RPL7	5.28816994	1.18E-07	NA	NA	NA	NA
Q14697	GANAB	4.42825477	1.35E-07	4.97849761	1.99E-07	-0.5502428	0.10755977
P46779	RPL28	5.22619463	1.76E-07	NA	NA	NA	NA
P62266	RPS23	4.85707181	2.13E-07	NA	NA	NA	NA
O95837; P19086; P29992; P50148;	GNA14	4.85940675	2.61E-07	5.01694152	1.12E-06	-0.1575348	0.6892738

Identifier		Cortex – Inner Nucleus		Cortex – Outer Nucleus		Outer Nucleus – Inner Nucleus	
<i>Protein Group</i>	<i>Gene Name</i>	<i>Log2FC</i>	<i>p-value</i>	<i>Log2FC</i>	<i>p-value</i>	<i>Log2FC</i>	<i>p-value</i>
Q14344							
P05556	ITGB1	5.57202156	4.55E-07	5.362389	7.22E-09	0.20963256	0.53514291
P08670;	VIM	6.53286156	4.58E-07	2.92949443	1.28E-08	3.60336714	9.49E-06
P41219							
P04350;	TUBB4A	4.65175265	4.60E-07	3.3457707	7.29E-08	1.30598195	0.00065271
P68371							
P62330	ARF6	4.9411772	4.66E-07	4.66494806	8.67E-10	0.27622914	0.3606126
P30044	PRDX5	4.33777052	5.58E-07	5.14310766	0.00011356	-0.8053371	0.26126416
Q9NZ01	TECR	5.06019438	5.81E-07	5.43432971	3.33E-09	-0.3741353	0.28075157
Q96IX5	ATP5MK	5.49219814	6.92E-07	NA	NA	NA	NA
Q96BM9;	ARL8A	6.09178967	8.14E-07	3.03726579	1.30E-07	3.05452388	5.99E-05
Q9NVJ2							
O00560	SDCBP	4.12743525	9.59E-07	3.9772657	6.62E-09	0.15016955	0.60681353
P08572	COL4A2	3.97453512	8.69E-12	4.28057939	6.42E-12	-0.3060443	0.00020156
P07305	H1-0	NA	NA	7.13489232	7.98E-12	NA	NA
P02794	FTH1	NA	NA	6.1656295	1.26E-11	NA	NA
O15173	PGRMC2	NA	NA	5.63883423	2.58E-11	NA	NA
O76024	WFS1	NA	NA	4.43323186	2.93E-11	NA	NA
Q99623	PHB2	NA	NA	6.98114024	3.04E-11	NA	NA
P40939	HADHA	NA	NA	6.00072438	6.12E-11	NA	NA
P16435	POR	NA	NA	5.20238238	1.83E-10	NA	NA
P11021	HSPA5	NA	NA	6.80995877	2.25E-10	NA	NA
Q6UW68	TMEM205	NA	NA	5.62722074	2.30E-10	NA	NA
P27824	CANX	NA	NA	7.06318521	2.55E-10	NA	NA
P49755	TMED10	NA	NA	5.16593136	2.85E-10	NA	NA
P02462;	COL4A1	3.69388751	7.52E-11	4.68922754	3.36E-10	-0.99534	1.61E-06
P29400							
P21796	VDAC1	NA	NA	6.29781512	3.90E-10	NA	NA
P23396	RPS3	1.01421853	2.33E-07	4.00880311	5.63E-10	-2.9945846	1.25E-09
P36542	ATP5F1C	NA	NA	7.17525037	6.35E-10	NA	NA
Q9P035	HACD3	NA	NA	5.84821076	7.86E-10	NA	NA
P11047	LAMC1	3.71933978	1.15E-09	5.9108017	9.05E-10	-2.1914619	8.96E-07
P07237	P4HB	NA	NA	6.23541458	9.65E-10	NA	NA
O95747	OXSRI	2.87849609	1.39E-07	4.20259913	1.09E-09	-1.324103	9.16E-05
P00367	GLUD1	NA	NA	4.79945714	1.25E-09	NA	NA
Q9UBM4	OPTC	NA	NA	6.01053977	1.34E-09	NA	NA
P62917	RPL8	7.51695601	3.63E-05	6.02655735	1.50E-09	1.49039866	0.09380128
P11117	ACP2	NA	NA	4.34861181	1.55E-09	NA	NA
Q08495	DMTN	NA	NA	4.07940598	1.59E-09	NA	NA
P62906	RPL10A	1.43688052	2.65E-06	4.11977132	1.87E-09	-2.6828908	5.77E-08
P35221	CTNNA1	3.19492745	1.89E-09	4.16878797	2.35E-09	-0.9738605	8.52E-05
P30084	ECHS1	NA	NA	4.09318026	2.37E-09	NA	NA
P07437;	TUBB	NA	NA	4.75094195	2.56E-09	NA	NA
P68371;							
Q13509;							
Q13885;							
Q9BVA1							
P48047	ATP5PO	NA	NA	6.96140685	3.55E-09	NA	NA
P38646	HSPA9	NA	NA	5.57905461	5.01E-09	NA	NA
P06576	ATP5F1B	NA	NA	6.49352282	6.02E-09	NA	NA
O94874	UFL1	3.41673933	2.38E-07	4.17094076	6.52E-09	-0.7542014	0.00525994
Q9UBI6	GNG12	NA	NA	4.10944918	6.68E-09	NA	NA
P07384	CAPN1	NA	NA	4.16959463	6.96E-09	NA	NA
P62851	RPS25	NA	NA	7.61811559	7.52E-09	NA	NA
Q9BVK6	TMED9	NA	NA	6.36130713	8.61E-09	NA	NA
Q12905	ILF2	3.52441715	3.10E-07	4.62714075	1.43E-08	-1.1027236	0.0002999
P15880	RPS2	1.23289059	2.22E-05	5.14744204	1.47E-08	-3.9145514	9.94E-08
P35232	PHB	NA	NA	7.69798711	1.60E-08	NA	NA
Q9H8H3	METTL7A	NA	NA	4.91374406	1.60E-08	NA	NA
P20849	COL9A1	NA	NA	7.48589185	1.76E-08	NA	NA
Q12797	ASPH	NA	NA	5.45140951	1.90E-08	NA	NA
P53420	COL4A4	4.12173371	3.96E-06	4.73116402	2.05E-08	-0.6094303	0.07122235
P55084	HADHB	NA	NA	7.02897407	2.23E-08	NA	NA
O95865	DDAH2	NA	NA	4.18881313	2.78E-08	NA	NA
P51884	LUM	2.33875944	2.43E-08	4.26582997	3.65E-08	-1.9270705	4.95E-06
P46781	RPS9	1.93593971	2.61E-07	5.76440429	3.73E-08	-3.8284646	1.70E-07

Identifier		Cortex – Inner Nucleus		Cortex – Outer Nucleus		Outer Nucleus – Inner Nucleus	
<i>Protein Group</i>	<i>Gene Name</i>	<i>Log2FC</i>	<i>p-value</i>	<i>Log2FC</i>	<i>p-value</i>	<i>Log2FC</i>	<i>p-value</i>
Q9NSE4	IARS2	NA	NA	4.68880516	5.46E-08	NA	NA
Q15691	MAPRE1	2.99216406	3.57E-05	4.20060159	5.56E-08	-1.2084375	0.00911053
P05141	SLC25A5	6.49287205	4.29E-05	5.48536542	5.58E-08	1.00750664	0.19295539
O75964	ATP5MG	3.10312588	3.10E-10	4.63128393	8.07E-08	-1.528158	5.23E-05
Q14055	COL9A2	NA	NA	7.21218152	8.44E-08	NA	NA
P62888	RPL30	1.85710834	4.00E-07	5.15329817	8.70E-08	-3.2961898	1.02E-06
P04181	OAT	NA	NA	4.13717889	9.19E-08	NA	NA
P17301	ITGA2	1.21610701	5.80E-05	4.16435211	9.43E-08	-2.9482451	2.66E-06
O00767	SCD	NA	NA	5.27758418	1.03E-07	NA	NA
O95197	RTN3	NA	NA	5.4242471	1.06E-07	NA	NA
P39656	DDOST	NA	NA	7.08565904	1.07E-07	NA	NA
Q9NZJ7	MTCH1	2.65900242	2.82E-07	4.55700313	1.08E-07	-1.8980007	5.71E-05
O14949	UQCRCQ	1.96224264	2.80E-06	5.28040577	1.08E-07	-3.3181631	2.06E-06
P00403	MT-CO2	NA	NA	5.65293292	1.09E-07	NA	NA
Q15084	PDIA6	NA	NA	4.90151614	1.23E-07	NA	NA
P61978	HNRNPK	3.69765913	7.09E-07	5.19935643	1.40E-07	-1.5016973	0.00199832
O75396	SEC22B	NA	NA	5.89614431	1.52E-07	NA	NA
Q9Y3B3	TMED7	NA	NA	4.02205229	2.44E-07	NA	NA
Q15363	TMED2	NA	NA	4.95960976	2.65E-07	NA	NA
Q07157	TJP1	2.90110173	2.34E-07	4.17794088	2.85E-07	-1.2768391	0.00117743
Q6PCB7	SLC27A1	NA	NA	-4.9552436	2.91E-07	NA	NA
P55285	CDH6	0.57908995	0.00208561	4.41846074	2.94E-07	-3.8393708	6.34E-07
Q6DD88	ATL3	NA	NA	5.92820064	3.13E-07	NA	NA
Q14165	MLEC	2.65173892	6.13E-09	4.29760099	3.25E-07	-1.6458621	0.00013788
P51991	HNRNPA3	2.61790945	1.72E-08	4.3274265	4.81E-07	-1.7095171	0.00015481
P32969	RPL9	3.37732933	5.50E-09	7.42479569	4.86E-07	-4.0474664	1.81E-05
Q9BUF5	TUBB6	NA	NA	4.22862282	5.51E-07	NA	NA
P26196	DDX6	3.94961338	1.55E-06	4.09595554	5.77E-07	-0.1463422	0.67656516
P31930	UQCRC1	NA	NA	5.07515167	6.95E-07	NA	NA
O75251	NDUFS7	2.33117088	6.66E-08	4.54749073	7.33E-07	-2.2163199	5.09E-05
O75369	FLNB	NA	NA	4.63322342	9.70E-07	NA	NA
Q93088	BHMT	NA	NA	NA	NA	-6.2832154	1.02E-13
P32320	CDA	1.27635782	8.66E-05	-2.9752652	3.02E-09	4.25162302	1.68E-09
Q93088; Q9H2M3	BHMT	NA	NA	NA	NA	-7.6049309	1.21E-08

Supplemental Table 4 – Complete list of ontologies (n=206) enriched in either the cortex or inner nucleus over the other region. Ontology calculations performed by PSEA-quant²⁶ with CV tolerance 0.5, 10,000 iterations used for p-value calculation.

GO Term	Annotation	PES	p-value	Proteins in Set
Enriched in Cortex over Inner Nucleus				
GO:0031974	Membrane-enclosed lumen	70.5365	<1.0x10 ⁻⁴	95
GO:0005604	Basement membrane	15.54	<1.0x10 ⁻⁴	19
GO:0006612	Protein targeting to membrane	51.185	<1.0x10 ⁻⁴	62
GO:0000184	Nuclear-transcribed mrna catabolic process, nonsense-mediated decay	42.485	<1.0x10 ⁻⁴	54
GO:0022625	Cytosolic large ribosomal subunit	22.6225	<1.0x10 ⁻⁴	27
GO:0034655	Nucleobase-containing compound catabolic process	110.8015	<1.0x10 ⁻⁴	161
GO:0006605	Protein targeting	65.0145	<1.0x10 ⁻⁴	86
GO:0071702	Organic substance transport	225.7785	<1.0x10 ⁻⁴	351
GO:0031982	Vesicle	115.25	<1.0x10 ⁻⁴	171
GO:0051234	Establishment of localization	326.6375	<1.0x10 ⁻⁴	509
GO:0005886	Plasma membrane	233.866	<1.0x10 ⁻⁴	363
GO:0006414	Translational elongation	40.015	<1.0x10 ⁻⁴	49
GO:0043062	Extracellular structure organization	43.002	<1.0x10 ⁻⁴	55
GO:0005759	Mitochondrial matrix	27.1465	<1.0x10 ⁻⁴	36
GO:0000956	Nuclear-transcribed mRNA catabolic process	47.017	<1.0x10 ⁻⁴	63
GO:0043933	Protein-containing complex subunit organization	160.473	<1.0x10 ⁻⁴	246
GO:0005912	Adherens junction	34.4925	<1.0x10 ⁻⁴	47
GO:0043624	Cellular protein complex disassembly	40.069	<1.0x10 ⁻⁴	50
GO:0015935	Small ribosomal subunit	17.7745	<1.0x10 ⁻⁴	22
GO:0016043	Cellular component organization	360.0055	<1.0x10 ⁻⁴	569
GO:0019866	Organelle inner membrane	46.838	<1.0x10 ⁻⁴	63
GO:0044391	Ribosomal subunit	41.923	<1.0x10 ⁻⁴	52
GO:0031012	Extracellular matrix	33.3745	<1.0x10 ⁻⁴	45
GO:0046700	Heterocycle catabolic process	115.055	<1.0x10 ⁻⁴	169
GO:0005789	Endoplasmic reticulum membrane	39.407	<1.0x10 ⁻⁴	52
GO:0019083	Viral transcription	37.7415	<1.0x10 ⁻⁴	45
GO:0005200	Structural constituent of cytoskeleton	22.6465	<1.0x10 ⁻⁴	30
GO:0046907	Intracellular transport	177.6925	<1.0x10 ⁻⁴	269
GO:0044270	Cellular nitrogen compound catabolic process	114.819	<1.0x10 ⁻⁴	169
GO:0006886	Intracellular protein transport	118.8685	<1.0x10 ⁻⁴	176
GO:0070161	Anchoring junction	38.602	<1.0x10 ⁻⁴	53
GO:0071840	Cellular component organization or biogenesis	361.837	<1.0x10 ⁻⁴	573
GO:0006415	Translational termination	38.2925	<1.0x10 ⁻⁴	46
GO:0072594	Establishment of protein localization to organelle	63.31	<1.0x10 ⁻⁴	87
GO:1901361	Organic cyclic compound catabolic process	120.529	<1.0x10 ⁻⁴	179
GO:0030574	Collagen catabolic process	10.9265	<1.0x10 ⁻⁴	13
GO:0005198	Structural molecule activity	116.5215	<1.0x10 ⁻⁴	160
GO:0006613	Cotranslational protein targeting to membrane	41.004	<1.0x10 ⁻⁴	49
GO:0045184	Establishment of protein localization	188.9685	<1.0x10 ⁻⁴	291
GO:0042470	Melanosome	38.7975	<1.0x10 ⁻⁴	50
GO:0006810	Transport	322.4605	<1.0x10 ⁻⁴	503
GO:0022411	Cellular component disassembly	68.8565	<1.0x10 ⁻⁴	92
GO:0070062	Extracellular exosome	32.9365	<1.0x10 ⁻⁴	44
GO:0022627	Cytosolic small ribosomal subunit	16.119	<1.0x10 ⁻⁴	19
GO:0006402	mRNA catabolic process	48.1525	<1.0x10 ⁻⁴	65
GO:0015075	Ion transmembrane transporter activity	39.3665	<1.0x10 ⁻⁴	53
GO:0015031	Protein transport	185.297	<1.0x10 ⁻⁴	286
GO:0003735	Structural constituent of ribosome	41.793	<1.0x10 ⁻⁴	54
GO:0072599	Establishment of protein localization to endoplasmic reticulum	41.5735	<1.0x10 ⁻⁴	50
GO:0019439	Aromatic compound catabolic process	116.7855	<1.0x10 ⁻⁴	173
GO:0032963	Collagen metabolic process	10.9265	<1.0x10 ⁻⁴	13
GO:0016020	Membrane	421.909	<1.0x10 ⁻⁴	674
GO:0070013	Intracellular organelle lumen	56.1445	<1.0x10 ⁻⁴	72
GO:0005201	Extracellular matrix structural constituent	13.0005	<1.0x10 ⁻⁴	14
GO:0065010	Extracellular membrane-bounded organelle	32.9365	<1.0x10 ⁻⁴	44
GO:0016482	Cytosolic transport	92.9645	<1.0x10 ⁻⁴	135
GO:0016071	mRNA metabolic process	95.9325	<1.0x10 ⁻⁴	144
GO:0031090	Organelle membrane	222.463	<1.0x10 ⁻⁴	347
GO:0032984	Protein-containing complex disassembly	41.322	<1.0x10 ⁻⁴	52

GO Term	Annotation	PES	p-value	Proteins in Set
GO:0022617	Extracellular matrix disassembly	11.734	<1.0x10 ⁻⁴	14
GO:0045047	Protein targeting to ER	41.5735	<1.0x10 ⁻⁴	50
GO:0006413	Translational initiation	52.02	<1.0x10 ⁻⁴	72
GO:0048770	Pigment granule	38.7975	<1.0x10 ⁻⁴	50
GO:0005788	Endoplasmic reticulum lumen	17.189	<1.0x10 ⁻⁴	19
GO:0043230	Extracellular organelle	32.9365	<1.0x10 ⁻⁴	44
GO:0042383	Sarcolemma	11.763	0.0001	15
GO:0005581	Collagen trimer	9.606	0.0001	10
GO:0006928	Movement of cell or subcellular component	122.56	0.0001	187
GO:0006412	Translation	66.6785	0.0001	98
GO:0031966	Mitochondrial membrane	60.38	0.0001	88
GO:0032774	RNA biosynthetic process	89.8865	0.0001	135
GO:0022857	Transmembrane transporter activity	49.8445	0.0001	71
GO:0005544	Calcium-dependent phospholipid binding	9.0785	0.0002	11
GO:0050839	Cell adhesion molecule binding	14.1955	0.0002	18
GO:0019898	Extrinsic component of membrane	20.6575	0.0002	28
GO:0005911	Cell-cell junction	39.452	0.0002	56
GO:0034329	Cell junction assembly	31.771	0.0002	45
GO:0008324	Cation transmembrane transporter activity	26.736	0.0002	37
GO:0044291	Cell-cell contact zone	12.7445	0.0003	16
GO:0051271	Negative regulation of cellular component movement	13.782	0.0003	18
GO:0031410	Cytoplasmic vesicle	86.503	0.0003	131
GO:0044764	Multi-organism cellular process	119.078	0.0003	183
GO:0014704	Intercalated disc	11.04	0.0004	14
GO:0051235	Maintenance of location	15.787	0.0004	21
GO:0016323	Basolateral plasma membrane	21.3535	0.0006	29
GO:0030054	Cell junction	78.201	0.0007	118
GO:0016032	Viral process	119.078	0.0007	183
GO:0044419	Interspecies interaction between organisms	126.0995	0.0007	194
GO:0044403	Symbiotic process	119.078	0.0007	183
GO:0017111	Nucleoside-triphosphatase activity	106.7195	0.0007	164
GO:0045296	Cadherin binding	8.917	0.0008	10
GO:0015077	Monovalent inorganic cation transmembrane transporter activity	19.902	0.0008	27
GO:0051651	Maintenance of location in cell	12.4125	0.0009	16
KEGG	ECM receptor interaction	15.3155	<1.0x10 ⁻⁴	17
KEGG	Ribosome	38.0505	<1.0x10 ⁻⁴	46
KEGG	Focal adhesion	35.026	0.0001	49
KEGG	Parkinsons disease	22.586	0.0001	30
KEGG	Huntingtons disease	28.949	0.0002	40
KEGG	Small cell lung cancer	13.32	0.0004	17
Reactome	Laminin interactions	15.4535	<1.0x10 ⁻⁴	16
Reactome	Nonsense mediated decay nmd	42.6985	<1.0x10 ⁻⁴	54
Reactome	Extracellular matrix organization	43.111	<1.0x10 ⁻⁴	55
Reactome	Nervous system development	127.377	<1.0x10 ⁻⁴	190
Reactome	Influenza infection	48.068	<1.0x10 ⁻⁴	63
Reactome	SRP dependent cotranslational protein targeting to membrane	41.7885	<1.0x10 ⁻⁴	50
Reactome	Collagen degradation	9.8835	<1.0x10 ⁻⁴	11
Reactome	Degradation of the extracellular matrix	20.3545	<1.0x10 ⁻⁴	25
Reactome	Response of EIF2AK4 GCN2 to amino acid deficiency	40.775	<1.0x10 ⁻⁴	50
Reactome	Regulation of expression of slits and ROBOs	62.3505	<1.0x10 ⁻⁴	91
Reactome	ECM proteoglycans	18.5665	<1.0x10 ⁻⁴	21
Reactome	Non integrin membrane ECM interactions	19.3275	<1.0x10 ⁻⁴	22
Reactome	Eukaryotic translation elongation	40.2355	<1.0x10 ⁻⁴	49
Reactome	Integrin cell surface interactions	18.6175	<1.0x10 ⁻⁴	21
Reactome	Developmental biology	141.9425	<1.0x10 ⁻⁴	212
Reactome	Met activates PTK2 signaling	7.052	<1.0x10 ⁻⁴	8
Reactome	Collagen biosynthesis and modifying enzymes	10.1835	<1.0x10 ⁻⁴	11
Reactome	Collagen formation	11.9575	<1.0x10 ⁻⁴	14
Reactome	Assembly of collagen fibrils and other multimeric structures	9.6625	<1.0x10 ⁻⁴	11
Reactome	Eukaryotic translation initiation	51.365	<1.0x10 ⁻⁴	70
Reactome	rRNA processing	43.077	<1.0x10 ⁻⁴	56
Reactome	Selenoamino acid metabolism	45.43	<1.0x10 ⁻⁴	58
Reactome	Binding and uptake of ligands by scavenger receptors	10.6175	0.0001	13
Reactome	Molecules associated with elastic fibres	7.0035	0.0001	8
Reactome	NCAM1 interactions	7.92	0.0001	9
Reactome	Disease	188.331	0.0001	291
Reactome	Infectious disease	130.6035	0.0001	197

GO Term	Annotation	PES	p-value	Proteins in Set
Reactome	Elastic fibre formation	7.0035	0.0002	8
Reactome	NCAM signaling for neurite out growth	14.678	0.0002	19
Reactome	Smooth muscle contraction	8.1835	0.0003	10
Reactome	Translation	75.192	0.0003	112
Reactome	Metabolism of amino acids and derivatives	86.999	0.0004	132
Reactome	Signaling by met	14.9195	0.0007	20
Reactome	L1CAM interactions	28.918	0.001	41
Enriched in Inner Nucleus over Cortex				
GO:0007601	Visual perception	19.743	<1.0x10 ⁻⁴	33
GO:0003008	System process	56.885	<1.0x10 ⁻⁴	112
GO:0003824	Catalytic activity	412.0705	<1.0x10 ⁻⁴	851
GO:0050877	Nervous system process	37.025	<1.0x10 ⁻⁴	70
GO:0042803	Protein homodimerization activity	57.578	<1.0x10 ⁻⁴	112
GO:0005212	Structural constituent of eye lens	12.841	<1.0x10 ⁻⁴	17
GO:0016765	Transferase activity, transferring alkyl or aryl (other than methyl) groups	9.383	<1.0x10 ⁻⁴	15
GO:0050953	Sensory perception of light stimulus	19.743	<1.0x10 ⁻⁴	33
GO:0019318	Hexose metabolic process	32.67	<1.0x10 ⁻⁴	62
GO:0007600	Sensory perception	28.0335	<1.0x10 ⁻⁴	51
GO:0044281	Small molecule metabolic process	238.9625	<1.0x10 ⁻⁴	491
GO:0002088	Lens development in camera-type eye	8.3425	<1.0x10 ⁻⁴	12
GO:0005996	Monosaccharide metabolic process	35.7905	<1.0x10 ⁻⁴	68
GO:0006006	Glucose metabolic process	28.6985	<1.0x10 ⁻⁴	54
GO:0016052	Carbohydrate catabolic process	21.9255	<1.0x10 ⁻⁴	40
GO:0046365	Monosaccharide catabolic process	17.72	0.0001	32
GO:0019320	Hexose catabolic process	16.126	0.0001	29
GO:1900544	Positive regulation of purine nucleotide metabolic process	5.173	0.0001	8
GO:0006575	Cellular modified amino acid metabolic process	21.9645	0.0001	41
GO:0042398	Cellular modified amino acid biosynthetic process	7.661	0.0001	13
GO:0016491	Oxidoreductase activity	60.552	0.0001	120
GO:0046983	Protein dimerization activity	71.1105	0.0001	141
GO:0043436	Oxoacid metabolic process	87.658	0.0001	176
GO:0019362	Pyridine nucleotide metabolic process	11.969	0.0001	22
GO:0097354	Prenylation	3.5375	0.0001	5
GO:0008318	Protein prenyltransferase activity	3.5375	0.0001	5
GO:0031281	Positive regulation of cyclase activity	3.5115	0.0002	5
GO:0045762	Positive regulation of adenylate cyclase activity	3.5115	0.0002	5
GO:0051349	Positive regulation of lyase activity	3.5115	0.0002	5
GO:0030810	Positive regulation of nucleotide biosynthetic process	4.496	0.0002	7
GO:0045981	Positive regulation of nucleotide metabolic process	5.173	0.0002	8
GO:1900373	Positive regulation of purine nucleotide biosynthetic process	4.496	0.0002	7
GO:0022400	Regulation of rhodopsin mediated signaling pathway	4.5435	0.0002	7
GO:0019752	Carboxylic acid metabolic process	82.262	0.0002	164
GO:0016056	Rhodopsin mediated signaling pathway	4.5435	0.0002	7
GO:0046496	Nicotinamide nucleotide metabolic process	11.969	0.0002	22
GO:0004659	Prenyltransferase activity	4.638	0.0002	7
GO:0016614	Oxidoreductase activity, acting on CH-OH group of donors	17.6295	0.0003	33
GO:0006749	Glutathione metabolic process	9.1955	0.0003	16
GO:0042802	Identical protein binding	89.5835	0.0003	179
GO:0006520	Cellular amino acid metabolic process	52.6775	0.0003	103
GO:0006007	Glucose catabolic process	13.9	0.0004	25
GO:0008652	Cellular amino acid biosynthetic process	17.2205	0.0004	32
GO:0016616	Oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	17.6295	0.0005	33
GO:0006082	Organic acid metabolic process	89.6605	0.0005	180
GO:0018342	Protein prenylation	3.5375	0.0005	5
GO:0008238	Exopeptidase activity	14.657	0.0006	27
GO:0006790	Sulfur compound metabolic process	26.116	0.0006	50
GO:0006096	Glycolytic process	11.2235	0.0006	20
GO:0005615	Extracellular space	43.2595	0.0006	85
GO:0044238	Primary metabolic process	467.2315	0.0006	976
GO:0043603	Cellular amide metabolic process	19.016	0.0008	36
GO:0006733	Oxidoreduction coenzyme metabolic process	12.754	0.0008	24
GO:0072524	Pyridine-containing compound metabolic process	13.153	0.0008	25
GO:0071482	Cellular response to light stimulus	6.559	0.0008	11
GO:0031279	Regulation of cyclase activity	4.904	0.0009	8

GO Term	Annotation	PES	p-value	Proteins in Set
GO:0090330	Regulation of platelet aggregation	2.164	0.0009	3
GO:0006081	Cellular aldehyde metabolic process	8.882	0.0009	16
GO:0046872	Metal ion binding	172.6725	0.0009	355
GO:0006656	Phosphatidylcholine biosynthetic process	2.1765	0.001	3
KEGG	Glutathione metabolism	11.721	0.0006	21
KEGG	Glycolysis gluconeogenesis	14.7665	0.0008	27
Reactome	Metabolism of carbohydrates	37.9195	$<1.0 \times 10^{-4}$	73
Reactome	TP53 regulates transcription of cell death genes	5.4785	0.0001	9
Reactome	The phototransduction cascade	4.5435	0.0001	7
Reactome	Glycolysis	14.465	0.0003	27
Reactome	ROBO receptors bind AKAP5	2.771	0.0003	4
Reactome	Signaling by WNT	35.3235	0.0004	70
Reactome	Beta catenin independent WNT signaling	27.1485	0.0005	53
Reactome	PCP CE pathway	23.138	0.0007	45
Reactome	Fructose metabolism	2.654	0.0008	4
Reactome	Biological oxidations	16.75	0.0009	32
Reactome	Transcriptional regulation of TP53	26.5975	0.0009	52