On the Origin of Proteins in Human Drusen: The Meet, Greet and Stick Hypothesis

Arthur A. Bergen\textsuperscript{a,b,c,*}, Swati Arya\textsuperscript{d}, Céline Koster\textsuperscript{a}, Matthew G. Pilgrim\textsuperscript{e}, Dagmara Wiatrek-Moumoulidis\textsuperscript{d}, Peter van der Spek\textsuperscript{f}, Stefanie M. Hauck\textsuperscript{g}, Camiel J. F. Boon\textsuperscript{b,h}, Eszter Emri\textsuperscript{i}, Alan J. Stewart\textsuperscript{d} and Imre Lengyel\textsuperscript{e,i}

Affiliations:
Departments of \textsuperscript{a} Clinical Genetics and \textsuperscript{b} Ophthalmology, Amsterdam UMC, University of Amsterdam, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands (NL).
\textsuperscript{c} Netherlands Institute for Neuroscience (NIN-KNAW), Amsterdam, NL.
\textsuperscript{d} School of Medicine, University of St Andrews, St Andrews, United Kingdom (UK).
\textsuperscript{e} UCL Institute of Ophthalmology, University College London, London, UK. Division of Biomaterials and Tissue Engineering, UCL Eastman Dental Institute, UCL, London, UK.
\textsuperscript{f} Dept. of Pathology, division Clinical Bioinformatics, Erasmus MC, Rotterdam, NL.
\textsuperscript{g} Research Unit Protein Science, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Neuherberg, Germany.
\textsuperscript{h} Department of Ophthalmology, Leiden University Medical Center, Leiden, NL.
\textsuperscript{i} Centre for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, Queen’s University, Belfast, Northern Ireland, UK

*Corresponding author:
Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands
Email address: aabergen@amc.uva.nl (A.A. Bergen).

Key words: Drusen proteins, Retinal Pigment Epithelium (RPE), Bruch’s membrane, Blood, Age-related Macular Degeneration (AMD), Alzheimer’s disease.

Acknowledgments:
This research was part-supported by de Algemene Nederlandse Vereniging ter Voorkoming van Blindheid (ANVVB), de Stichting Blinden-Penning, de Gelderse Blinden Stichting, de Landelijke Stichting voor Blinden en Slechtzienden (LSBS), Stichting Oogfonds Nederland, Stichting MD Fonds and Stichting Retina Nederland Fonds (represented by Uitzicht, grants 2011-6 and 2014-7 to A.A.B.), de Rotterdamse Stichting Blindenbelangen (RSB), de Haagse Stichting Blindenhulp, Stichting Lijf en Leven, Stichting Ooglijders (to A.A.B.); ZonMW grant nr 446001002 (to A.A.B. and C.K.); the Bill Brown Charitable Trust, Moorfields Eye Hospital Special Trustees, Mercer Fund from Fight for Sight, the Eye-Risk project funded by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 634479 (I.L. and E.E.), Fight for Sight project grant (I.L. and A.S.), the Bright Focus Foundation grant nr M2015370 (to SMH). The authors thank Dr J Booij for partly unpublished data and the reviewers for their invaluable commentsto improve the manuscript.

Conflict of interest statement:
The authors do not have any competing financial interest to declare.
Title: On the Origin of Proteins in Human Drusen: The Meet, Greet and Stick Hypothesis

Abstract: Retinal drusen formation is not only a clinical hallmark for the development of age-related macular degeneration (AMD) but also for other disorders, such as Alzheimer’s disease and renal diseases. The initiation and growth of drusen is poorly understood. Attention has focused on lipids and minerals, but relatively little is known about the origin of drusen-associated proteins and how they are retained in the space between the basal lamina of the retinal pigment epithelium and the inner collagenous layer space (sub-RPE-BL space). While some authors suggested that drusen proteins are mainly derived from cellular debris from processed photoreceptor outer segments and the RPE, others suggest a choroidal cell or blood origin.

Here, we reviewed and supplement the existing literature on the molecular composition of the retina/choroid complex, to gain a more complete understanding of the sources of proteins in drusen. These “drusenomics” studies showed that a considerable proportion of currently identified drusen proteins is uniquely originating from the blood. A smaller, but still large fraction of drusen proteins comes from both blood and/or RPE. Only a small proportion of drusen proteins is uniquely derived from the photoreceptors or choroid. We next evaluated how drusen components may “meet, greet and stick” to each other and/or to structures like hydroxyapatite spherules to form macroscopic deposits in the sub-RPE-BL space. Finally, we discuss implications of our findings with respect to the previously proposed homology between drusenogenesis in AMD and plaque formation in atherosclerosis.
Table of Contents

1. Drusen.

2. Functional annotation of drusen proteins.
   2.1. Biological or disease motifs and canonical pathways.
   2.2. Molecular networks.
   2.2.1. Network 1.1, 1.2, 1.3: Complement, collagens and crystallins.
   2.2.2. Network 2.4: development, genetics ophthalmic disorders.
   2.2.3. Network 3.5: Immunological response.
   2.2.4. Network 4.6: Cell-to-cell signaling and systemic involvement, lipid metabolism.

   3.1. The neural side of drusen.
   3.1.1. Histopathological, retinal imaging observations.
   3.1.2. Proteomic level observations.
   3.2. The systemic side of drusen.
   3.2.1. Bruch’s membrane.
   3.2.2. Choroidal capillaries.
   3.2.3. Contribution of blood proteins.

4. Selection of transcriptomic and proteomic datasets to determine the origin of drusen proteins.
   4.1. Selection criteria and considerations.
   4.2. Description of expression datasets used for drusenomics.
   4.3. Functional annotation of the photoreceptor (cPR-ET) and choroidal (cChor-ET) datasets.

5. Drusenomics, part II: Qualitative analysis.
   5.1. Comparative study design considerations.
   5.2. Where do proteins in drusen come from? A qualitative comparison.
   5.2.1. Network 1.1: The complement gene cluster.
   5.2.2. Network 1.2: The collagen cluster.
   5.2.3. Network 1.3: The crystallin cluster.
   5.2.4. Network 2.4: Genetic and developmental ophthalmic disorders.
   5.2.5. Network 3.5: Injury, inflammation and dermatological disease.
   5.2.6. Network 4.6: Cell to cell signaling; systemic involvement.

   6.1. Quantitative analysis and curation of datasets.
   6.1.1. Ten out of 89 drusen proteins originate uniquely from the PR/RPE.
   6.1.2. Twenty-three of 89 drusen proteins originate from both the "neural
and systemic side” of drusen.

6.2. Nineteen drusen proteins out of 89 were not assigned.

6.3. Blood proteins are an important source of drusen proteins.

7. Drusen and hydroxyapatite.

   8.2. Histological and pathobiological similarities.
   8.3. Genetics and molecular biology.

1. Drusen.

Drusen are extracellular deposits of bio-materials underneath the retinal pigment epithelium (RPE) in the eye (Farkas et al., 1971b; Sarks, 1976). They are considered clinical hallmarks for a number of diseases, including age-related macular degeneration (AMD) (Hogan, 1965; Sarks, 1976; Hageman et al., 2001; Khan et al., 2016), Alzheimer disease (AD) (Csincsik et al., 2018) and dense deposit disease (DDD) (Duvall-Young et al., 1989; Mullins et al., 2000; Boon et al., 2009). AMD is the leading cause of severe visual impairment, affecting 4% of the population over 60 years old (de Jong, 2006). AD is the biggest cause of dementia, affecting millions of people in the western world. DDD is a relatively rare juvenile disease characterized by kidney malfunction (Ito et al., 2017; Wang et al., 2017; Cunningham and Kotagiri, 2018). Despite the potential relevance for diseases, little is known about the composition of drusen and how and why biomaterials accumulate these deposits.

Drusen are heterogeneous in terms of size, shape, color on retinal imaging, retinal location and molecular content (Sarks, 1976; Sarks et al., 1980; Sarks et al., 1999; Crabb et al., 2002; Khan et al., 2016). In the clinic, drusen can be identified as yellow spots on funduscopy and color fundus images or dome shaped objects of different sizes under the RPE on Optical Coherence Tomography (OCT) (Marshall et al., 1992; Bird et al., 1995; Loeffler and Lee, 1998; Khan et al., 2016). Histopathological examination of drusen showed that are located between the basal lamina of the RPE cells and the Inner collagenous layer of the Bruch’s membrane, a space that had been termed recently as sub-RPE-BL space (Balaratnasingam et al., 2016; Li et al., 2018). Clinical definition of drusen depends on size, color, auto fluorescence, and retinal location (Sarks, 1976; Bird et al., 1995) (Figure 1). Drusen may appear in the macula, peri-macular area or in the mid-and/or far periphery (Lengyel et al., 2015; Domalpally et al., 2017; Csincsik et al., 2018). A particular druse can be termed as “hard”, when it’s appearance is small, round and well demarcated, with a size of <63 μm. “Intermediate” drusen have a size of approximately 63-125 μm, while “soft” drusen are >125 μm in size, and frequently have more ill-defined edges (Bird et al., 1995). A few (<5) small hard (sub-clinical) drusen in the macula does not raise alarm bells, but when numbers of hard drusen increase, or the size of drusen increases such that they become “intermediate” and/or- “soft” drusen, the likelihood to progression to AMD is increased significantly (Bird et al., 1995). Drusen
should be distinguished from reticular pseudodrusen (or subretinal drusenoid deposits) that occur between the RPE and photoreceptor (PR) in the subretinal space (Zweifel et al., 2010; Spaide et al., 2018). Relatively little is known about pseudo-drusen and as such, they are excluded from this review. Drusen are formed in the sub-RPE-BL space, between the basement membrane of the RPE and the inner collagenous layer of Bruch’s membrane (BrM).

The RPE is a multifunctional single neuro-epithelial cell-layer that act as a metabolic interface between the choroid and the neurosensory retina (Strauss, 2005). The RPE cells are connected by intercellular tight junctions, together forming the outer blood-retina barrier. On the apical side, the photoreceptor cells line the RPE. On the basal side the interposing BrM separates the basement membrane of the RPE from the choroidal micro-vasculature (choriocapillaris). The choroidal capillaries are fenestrated, and not surrounded by pericytes or smooth muscle cells. The BrM consists of three interleaved layers: the inner and outer collagenous layers with an elastic layer in between them (Booij et al., 2010a). Often, the basement membranes of the endothelium and the epithelium are classified as part of the BrM but we will refer here to the BrM structure as tri-laminar (rather than as penta-laminar). Embedded in the BrM are macromolecules such as proteins and proteoglycans to help remodeling the extra cellular matrix (with age)(Guo et al., 1999; Guymer et al., 1999; Del Priore et al., 2006; Beattie et al., 2010; Booij et al., 2010a; Hussain et al., 2011). The diffuse thickening of BrM is also a characteristic age-related feature (Hogan, 1965; Sarks et al., 1999). This is largely due to the entrapment of proteins and lipids within the ECM (Curcio et al., 2011; Curcio and Johnson, 2012). The diffuse build-up of extracellular bio-materials between the basement membrane of the RPE and the inner collagenous layer of the BrM is called basal linear deposits while the deposit formation between the basement membrane and the cell membrane of the RPE are called basal laminar deposits (Sarks, 1976; Sarks et al., 1980; van der Schaft et al., 1993; Abdelsalam et al., 1999; Curcio and Millican, 1999; Spraul et al., 1999). Due to the lack of information of the composition of these deposits, these specific classifications are excluded from our analysis. The deposits in BrM result in a decline in the conductivity of the membrane creating in a diffusion barrier that further enhances the accumulation of biomaterials (Green and Enger, 1993; Moore et al., 1995; Starita et al., 1997; Curcio and Millican, 1999; Curcio et al., 2011; Curcio, 2018b).
This phenomenon may be a general “passive” pathophysiological process that resembles plaque formation in disorders such as AD or atherosclerosis.

Even more detailed insights into sub-RPE-BL space deposits originated from molecular and histochemical studies on isolated drusen material. Recent investigations have shown that drusen contain lipids, trace elements, including zinc, iron and calcium, as well as a wide array of different proteins (Crabb et al., 2002; Curcio et al., 2011; Thompson et al., 2015; van Leeuwen et al., 2018). The distribution of these components is not uniform, neither within nor between drusen, further emphasizing the heterogeneous nature of the deposits (Thompson et al., 2015).

Oxidative modification of lipids and proteins may result in the cross-linking of these molecules and may contribute to deposit formation and drusenogenesis. Subsequently, local cellular damage at the very early onset of AMD, via the complement cascade attack on drusen compounds and the NLRP3 inflammasome (Edwards and Malek, 2007; Yuan et al., 2010; Doyle et al., 2012), can lead to retinal damage and more advanced AMD. Relatively few studies addressed the origin of proteins in the initiation and progression of drusen (Mullins et al., 2000; Nordgaard et al., 2006; Cryan and O’Brien, 2008; Wang et al., 2010; Crabb, 2014). A number of studies (Johnson et al., 2011; Kunchithapatham et al., 2014) have yielded conflicting data as to where drusen proteins originate from, and whether the accumulation of this apparent depositioning of biomaterials in BrM is a passive or an active process. Several questions remain, which include: to what extent do proteins in drusen originate from photoreceptors, RPE, choroidal endothelium or even the circulating blood? How do drusen form and how are drusen components recruited and deposited in the sub-RPE-BL space? What is the extent of the (molecular) heterogeneity that exists within and between drusen? Here, we will review and combine data from the existing literature, and supplement these with our own (new and recently published) data from subretinal transcriptomic, proteomic and immunohistochemical staining experiments. To enable this, we have functionally annotated a compiled list of drusen proteins and compared these proteins with those identified in specific transcriptomic and proteomic datasets derived from cells and tissues of the various relevant compartments. These include both subretinal and choroidal tissues, as well as the plasma proteome. Collectively, these analyses increase our understanding of drusenogenesis, which may provide clues for the prevention of drusen formation and, ultimately, for the prevention of drusen associated disorders (Khan et al., 2016).
2. Functional annotation of drusen proteins.

One of the key aims of this study is to identify the most likely original sources of drusen proteins. More specifically, do drusen proteins only come from the neural tissues (photoreceptors and RPE) or is there also a choroidal or systemic component? In the next chapters, we try to answer this question through a literature search and by using a variety of qualitative and quantitative transcriptomics and proteomics meta-analyses of the relevant genes and proteins involved.

We did not distinguish between various drusen types, sizes and/or drusen locations, since little –omics data are available for each drusen subtype. Essentially, we followed the (sub-clinical) drusen type description used by Crabb and coworkers (Crabb et al., 2002) who defined drusen to appear as opaque, 0 to 250 μm spherical to irregular deposits that remained attached to BrM after removing the RPE from human donor globes, both in the macular and the retinal periphery.

Based on relevant studies in the literature (Mullins et al., 2000; Crabb et al., 2002; Wang et al., 2010), we curated a list of 89 drusen proteins (Table 1). This was achieved by combining the published datasets and removing incomplete, duplicate or ambiguous entries. Several entries did not correspond to a single full-length cDNA annotated in the knowledge database Ingenuity (www.ingenuity.com) and were left out. Since the complement gene pathway is likely the best and most extensively studied pathway (compared to other pathways) we only added a few complement proteins to the list, to avoid bias toward one pathway and the “winner’s curse”. In addition, we also searched the literature for confirmatory immunohistochemistry (IHC) studies and manually added proteins from such smaller-scale studies. We realize this list may not be complete. For example, individual entries like the locally produced vitronectin (Hageman et al., 1999; Wasmuth et al., 2009) present in drusen is missing in Table 1 and an entry like elastin may be present as contamination of the BrM rather than a “specific” drusen protein. The problem with selecting these proteins lies with the heterogeneity of drusen (one protein may be present in one drusen but not in the other), lack of uniform criteria “what is drusen-specific (?)”, lack of uniformity in healthy or diseased stage of examined samples and overall, how much evidence is needed to assign proteins to drusen (see also discussion section). Nonetheless, we believe that, for the purposes of this study, our
selection of 89 proteins, largely based on the proteomic study of Crabb and colleagues (Crabb et al., 2002) provides us with a sufficient representative drusen protein dataset for the purpose of this study.

We used the 89 drusen protein data set first to investigate the molecular aggregation and the functional annotation of drusen proteins. A similar study was previously carried out by Crabb and coworkers (Crabb et al., 2002; Crabb, 2014). However, here we used a slightly different list of drusen proteins and subjected this to additional, advanced bioinformatics analysis. Consequently, we ran an Ingenuity knowledge database core analysis using our list of drusen components (Table 1) which yielded biological motifs, canonical pathways and molecular structural or functional networks. A summary of the results of this analysis is shown in Table 2.

2.1. Biological or disease motifs and canonical pathways.

The functional annotation of the 89 drusen proteins (Table 2), revealed that these proteins (motifs or aggregates) can be associated with a number of functional or disease entities, such as “hereditary disorders”, “ophthalmic disease”, “organismal injury or abnormalities” and “metabolic disease and developmental disorders”. Although these annotation categories are broad and not very specific, they do point to a wide range of potential sources of drusen components from both local and systemic origin.

Ingenuity analysis also yielded a number of canonical pathways. A canonical pathway is the simplest linear representation of an established chain of biochemically related molecules in a given system or cellular environment. The software recognizes enriched canonical pathways specific for “acute phase response signaling”, the “retinoid- and farnesoid X receptors (LXR/RXR and FXR/RXR) response”, “atherosclerosis signaling” and “IL-12 signaling in macrophages” in the drusen dataset.

“The acute phase response” is a fast-systemic inflammatory response triggered by infection, tissue injury and/or immunological disease (Serhan et al., 2015). The response is mediated by the hypothalamus and several acute phase plasma proteins. These proteins have a broad-working spectrum: they kill micro-organisms and modulate complement activation, enzyme activity and the immune response. How and why these proteins potentially end up in drusen is not clear (Johnson et al., 2000). Although not undisputed, Despriet and coworkers found independently, that AMD is associated with
acute phase plasma protein levels and with genetic variation in C-reactive protein (CRP),
one of the principal acute phase proteins (Despriet et al., 2006). Chirco and Potempa showed that CRP protein acts as a mediator of complement activation and inflammatory signaling in AMD (Chirco and Potempa, 2018). It is generally assumed that acute phase proteins are present in the blood; suggesting that some drusen proteins can originate from this pathway and have a systemic origin. While choroidal CRP apparently correlate to serum levels (Chirco et al., 2018), it cannot be said with certainty that these proteins are not (transiently and/or locally) produced by the choroid as well in cases of (nearby) low-grade inflammation.

“The retinoid X receptors (RXRs)” are nuclear retinoid receptors that regulate, via the ligand LXR, lipid and cholesterol metabolism as well as inflammation (Hiebl et al., 2018). Cholesterol metabolism is essential for many retinal functions (Pikuleva and Curcio, 2014), while ocular (para-) inflammation is crucial for maintaining retinal homeostasis (Xu et al., 2009). In the eye, retinoid X receptor activation contributes to retinal photoreceptor differentiation, survival, and disease (Forrest and Swaroop, 2012), and more specifically, for docosahexaenoic acid-mediated protection of photoreceptors (German et al., 2013). The presence of this protein signature in drusen points toward a local cellular origin of this protein. LXR can form heterodimers with “the farnesoid X receptor (FXR)” which is also a nuclear receptor, and is an important regulator of a variety of bile acid, glucose and lipid-related metabolic pathways, including the removal of cholesterol (Tu et al., 2000; Hiebl et al., 2018). FXR protein was detected in a variety of tissues, including heart, ovary, thymus and eye. Both LXR and FXR may be involved in cholesterol homeostasis in RPE and retina (Zheng et al., 2015). The presence of these receptor proteins in drusen points toward a local cellular origin.

“Atherosclerosis signaling”: Atherosclerosis is a low grade chronic inflammatory disorder characterized by local plaque deposition in the vessel wall, formed by a local accumulation of modified plasma lipoproteins and macrophage activation. The major cause of coronary events is rupture and thrombosis. Interestingly, clinical, epidemiological, pathobiological and molecular evidence suggest that an overlap exists between drusen in AMD and plaque formation in atherosclerosis. Indeed, like AMD, atherosclerosis is now considered as a low-grade chronic inflammatory process resulting from interaction (in) between plasma lipoproteins and the vascular wall (Mullins et al., 2000). In AMD, not only plasma lipoproteins, but also local lipoproteins
are involved. In section 8 of this manuscript, we describe the potential molecular and
pathobiological overlap between drusen/AMD and vascular plaques in detail. Taken
together, the homology between drusen and atherosclerotic plaques points toward a
systemic origin of some drusen proteins.

“IL-12 Signaling and Production in Macrophages”: The production of the cytokine IL-12
by activated (incoming) macrophages in damaged or diseased retinal tissue, is well
known (Zamiri et al., 2006; Chen et al., 2013). However, IL-12 exerts an autocrine effect
since macrophages and dendritic cells also respond to IL-12 by producing interferons
that stimulates T-helper cell differentiation. The RPE is apparently able to suppress
inflammation by modulating IL-12 production (Zamiri et al., 2006). Cao and coworkers
showed that cultured RPE cell \textit{in vitro} secrete several cytokines, including IL-12, under
conditions of oxidative stress and replicative senescence (Cao et al., 2013). Therefore,
the molecules identified in this category (IL-12 signaling) can originate from both the
circulation as well as from the local cellular environment.

2.2. Molecular networks.
Molecular networks in Ingenuity are built up from a myriad of relevant literature
connections and they are formed on the basis of most likely physical or functional
interactions between (input) genes and/or proteins. For example, see molecular
network 1 in Figure 2a. Based on millions of experimentally verified and curated data
points, these networks represent the most likely functional associations between
components of the “biological soup” in the context of the input molecules. Structural,
functional and mixed molecular networks exist. Structural networks contain primarily
networks of structurally and physically interacting entries. Functional networks are
dominated by functional relationship between participating molecules. A third, “mixed”
network, contains both structural and functional associations. The molecular network
analysis of drusen proteins yielded 4 significant networks, with 6 distinct functional
clusters. Note that the networks are not \textit{a priori} built through their possible relationship
with drusen or AMD \textit{per se}.

2.2.1. Network 1.1, 1.2 and 1.3: Complement, collagens and crystallins.
The most significant network formed in the data-driven Ingenuity drusen analysis, is
presented in Figure 2a. This network consists of three functionally more or less specific
molecular clusters: the complement protein cluster, the collagen protein cluster and the
crystallin heat shock protein cluster (Table 3). The presence of complement proteins in drusen (and the choriocapillaris) was previously shown in older and AMD affected eyes through immunohistochemistry, long before the genetic involvement of CFH and other complement factors in drusen formation and AMD became genetically apparent (Johnson et al., 2000; Hageman et al., 2001; Edwards et al., 2005; Hageman et al., 2005; Haines et al., 2005; Klein et al., 2005). Analyzing the drusen proteome, we confirmed the involvement of the terminal complement protein complex by identifying the complement factors C7, C8A, C8B, C8G, and the membrane attack complex (MAC) in drusen. Of note, the MAC was initially identified in drusen from unspecified retinal locations, but was later shown not to be present in macular drusen (Johnson et al., 2000; Mullins et al., 2014). The MAC is the final downstream event of the complement cascade. It results from the binding of C5b to blood plasma complement proteins C6, C7, C8, and C9, forming transmembrane pores that leads to cell lysis and death. In the same cluster, we found the Prolyl endopeptidase-like protein (PRELP), a small leucine-rich proteoglycan (SLRP) (Hultgardh-Nilsson et al., 2015), which among others, is involved in the inhibition of complement activation (Warwick et al., 2014). The main complement cascade regulator CFH is another member of the complement cascade that is present in the drusen proteome. Genetic variation in CFH may regulate complement activation on RPE cells (Radu et al., 2014). Of note, is that a certain degree of low-grade complement activation and para-inflammation is always present in healthy aging eyes, to maintain local health (Xu et al., 2009). In a recent review, Warwick et al. concluded that complement deposition in the retina could be of local and/or systemic origin (Warwick et al., 2014). The majority of complement genes are expressed in the liver, resulting in an abundance of complement proteins in the blood. However, the RPE expresses several key complement genes which may modulate the complement attack on RPE and drusen (Chen et al., 2007; Kim et al., 2009; Pao et al., 2018). Interestingly, locally produced CFH is, at least in cultured RPE cells, secreted apically, and not basally (Kim et al., 2009; Pao et al., 2018). Consequently, the potential regulating role of locally produced CFH in vivo, and other complement factors, potentially involved in the complement attack on drusen, needs to be further investigated. A diversity of collagen proteins, such as COLA1, A2, 6A1, 6A2 and 8A1 were previously consistently identified in basal laminar deposits and basal linear deposits, and, occasionally, in drusen (Newsome et al., 1987; Booij et al., 2010a; Curcio and Johnson, 2012). Newsome and coworkers (1987) noted that the
involvement of extracellular matrix components in drusen is variable but these findings have not been confirmed in other studies. These molecules may primarily present as a remnant from the (ab)normal turnover of BrM components (Newsome et al., 1987). Alternatively, they may be secreted by the RPE in response to challenges presented by drusen or by the conditions that lead to drusen formation. Interestingly, collagen IV is not present in our curated drusen dataset, despite the fact that collagen IV accumulations are found in autosomal dominant radiant drusen in Doyne Honeycomb Retinal Dystrophy caused by EFEMP1 (EGF-containing fibulin extracellular matrix protein 1) mutations (Sohn et al., 2015). In fact, in the absence of detailed electron microscopic examination it is not clear whether these are drusen or basal laminal deposits.

The presence of crystallin proteins in drusen had been shown by Crabb and co-workers (Crabb et al., 2002) and functionally studied by Nakata et al. (Nakata et al., 2005). These authors found that BrM, drusen and part of the choroidal connective tissue, when affected by AMD, showed higher immunoreactivity for α- and β-crystallins than healthy control tissues. Retinal crystallins are also up-regulated in a variety of other retinal pathologies, including diabetic retinopathy, ischemia, mechanical injury and uveitis. The α-crystallin family plays a crucial role in neuroprotection and inflammation (Fort and Lampi, 2011), while the β- and γ-crystallins are small proteins with a possible ganglion cell protective role in glaucoma (Anders et al., 2017) and a role in retinal tissue remodeling and repair (Thanos et al., 2014). Consequently, the presence of these proteins in drusen points to a local cellular origin.

2.2.2. Network 2.4 development, genetics of ophthalmic disorders.

The fourth cluster in our drusen protein analysis is actually similar to the entire network 2 which is functionally annotated as “network of genetic and developmental disorders”. The components of this cluster are functionally presented in detail in Figure 2b; Table 3. This network contains annexin A2 (ANXA2), a relatively small calcium and phospholipid-binding protein involved in multiple intra-cellular transport functions. The RPE secretion data set, called RPE-IVS (Table 3; STable 1) reveals that this protein is indeed secreted basally by the RPE (Pao et al., 2018). The protein was initially assigned to drusen (Crabb et al., 2002). However, the same authors showed, using IHC in a
number of human donor eyes, that ANXA2 is not associated with the interior of drusen, but with the basal lamina of the RPE close to the drusen surface.

2.2.3. Network 3.5: Immunological response.

The fifth functional cluster in our drusen protein analysis represents network three: “Injury and inflammatory response; dermatological disease” (Figure 2c). The components are given in Table 3. This network contains, among others, annexin A1 (ANXA1). ANXA1 antibodies intensely stained whole drusen, but also the BrM and choroid (Rayborn et al., 2006). Given its positive staining in entire drusen, we consider it here as a drusen protein. Apolipoprotein E (APOE) is also present in this group. APOE is classically thought of as a cholesterol carrier. Risk alleles of the APOE gene were associated with a variety of diseases including AMD, AD and atherosclerosis (Klaver et al., 1998; Ashford, 2004; Song et al., 2004; Tikellis et al., 2007). Its presence in laminar deposits and drusen was initially established by Klaver and colleagues (Klaver et al., 1998) and later confirmed by Anderson and Malek (Anderson et al., 2001; Malek et al., 2003). Interestingly, in an RPE cell culture model that mimics drusen formation, Pao (2018) and coworkers found that APOE is secreted basally by these cells. Subsequent exposure of these cultures to human serum led to heterogeneous sub-RPE-BL space deposits, some of which were rich in serum-derived proteins such as vimentin, clusterin, and amyloid P (Pao et al., 2018). In addition to ANXA1 and APOE, the serum amyloid proteins S100A7, S100A8 and S100A9 are part of this functional cluster and the drusen proteome. S100 proteins are a family of small calcium-binding proteins, produced in the nucleus and cytoplasm of a wide variety of cells (Gross et al., 2014; Narumi et al., 2015; Cunden et al., 2017).

2.2.4. Network 4.6 Cell-to-cell signaling and systemic involvement, lipid metabolism.

The sixth cluster “cell-to cell signaling and systemic involvements” (Figure 2d; Table 3) points to proteins which come from an extracellular environment. For example, The APOA1, APOA4 and SAA1 lipoproteins and S, and the protein-groups related to the LDL, HDL, VLDL metabolism (that have been added by Ingenuity to construct a meaningful network) are most likely derived from the blood, and not from the retina. However, cautious interpretation of these general data is warranted, since the RPE is also capable of secreting a number of lipoproteins, such as APOB (Li et al., 2005b). The mechanisms
of biogenesis of lipid-laden soft drusen has been recently reviewed elsewhere (Curcio, 2018a, b) as has the role of lipids in AMD (van Leeuwen et al., 2018).

In at least two blood proteomics datasets (Table 3) the ORM1 (acute phase plasma protein of unknown function; www.genecards.org) and the SERPINA1 (serine protease inhibitor; www.genecards.org) proteins occur, which point also at a systemic origin of these drusen proteins. Furthermore, in this cluster we see the drusen protein clusterin (CLU), which is expressed in many cell types, including photoreceptors or RPE, and is also present in blood (Garcia-Aranda et al., 2018). The presence of annexin 6 (ANXA6) in drusen (and BrM) was previously confirmed using immunohistochemistry (Rayborn et al., 2006). Finally, we observe also the presence of the (systemic) HRG protein, which is extensively discussed in section 7 of this manuscript.

Multiple epidemiological, genetic, biochemical and pathophysiological studies in the literature address the origin of drusen. While many studies address the origins of metal ions or lipids in drusen, here we focus on the likely source of proteins. Drusen proteins could originate from either the neural side of drusen (Photoreceptors, RPE), the systemic side (BrM, choroid complex, blood) BrM, or both (Penfold et al., 2001; Curcio and Johnson, 2012).

3.1. The neural side of drusen.

Theories on drusen accumulation from the neural side vary: proteins may either come from dying PR and RPE cells, or from (basal) secretion of proteins generated by the normal functions of the RPE (Crabb et al., 2002; Kinnunen et al., 2012). Respectively, cellular debris or secreted proteins may get trapped in BrM or drusen. Evidence for these origins was gathered from histopathological investigation, retinal imaging, and proteomics studies.

3.1.1. Histopathological and retinal imaging observations.

Drusen formation goes hand in hand with hypo- or hyperpigmentation (Curcio et al., 1998) of the RPE, especially in the early stages of AMD. Indeed, retinal cells overlying drusen exhibit numerous irregular structural and molecular abnormalities which are confined to areas directly internal to drusen (Farkas et al., 1971b; Hogan, 1972; Burns and Feeney-Burns, 1980; The Eye Disease Case-Control Study, 1992; Johnson et al., 2003). Deflection and shortening of rod inner and outer segments of rod photoreceptors have been postulated to contribute to sub-RPE deposit formation (Farkas et al., 1971a). Drusen have been also associated with more indirect changes, such as alterations in the synaptic terminals of photoreceptor cells and an increase in vimentin and glial fibrillary acidic (GFAP) protein within Müller cells (Johnson et al., 2003). Other retinal cells, such as bipolar, horizontal, amacrine and ganglion cells are most likely unaffected by drusenogenesis (Johnson et al., 2003).

Using immunohistochemical, molecular biological and biochemical approaches, Hageman and coworkers found that RPE cell loss is correlated with increasing drusen density (Hageman et al., 2001). More recent OCT studies, focusing on the integrity of the RPE layer directly internal to drusen showed that 41.3% of all drusen coincided with an
intact overlying RPE, and that in 28.1% of cases, the RPE was irregular but continuous (Schlanitz et al., 2018). In 30.6% of cases, the RPE layer adjacent to drusen was discontinuous. Larger drusen were associated with higher probability of RPE loss (Schlanitz et al., 2018). Taken together, these results suggest that RPE or PR cell death is associated with drusenogenesis. However, it is not clear whether the observed cellular damage is a cause or consequence of sub-RPE deposit formation.

The presence of cytoplasmic (Burns and Feeney-Burns, 1980), fibrous and membranous/lipoid material (Fine, 1981; Young, 1987; Green and Enger, 1993; Loeffler and Lee, 1998; Curcio and Millican, 1999) in drusen suggest that deposits are formed after cellular degeneration. According to Coats, small colloid bodies derived from degenerated RPE cells, develop into larger drusen due to uptake of biomolecules through a defective BrM (Coats, 1905) and clinical support was provided for the existence of these bodies (Pauleikhoff et al., 1990). Later, necrotic RPE cells were presumed to be incorporated into existing drusen (Young, 1987). However, these findings also did not distinguish between cause or consequence of deposit formation. To complicate matters further, there are a number of reports in the literature describing drusen regression; in an experimental study after laser photocoagulation and in clinical studies using fluorescein angiograms (FAs) fundus photography (Bressler et al., 1995) and OCT (Yehoshua et al., 2011). A similar observation were done in rhesus monkeys (Duvall and Tso, 1985) in APOE mice with thickened BrM as well as AMD patients (Jobling et al., 2015). This intriguing phenomenon may be linked to transiently increasing the RPE-mediated release of active MMP enzymes that alter the turnover of BrM (Zhang et al., 2012).

3.1.2 Proteomic level observations.

Proteomics studies into drusenogenesis can be divided into studies on (archived) human post-mortem eyes, in vitro RPE culture, and proteomic studies on retinas of animal models. A variety of techniques, such as 2D gels and LC-MS/MS analysis have been used. To date, up to over 500 healthy and AMD-affected post-mortem human eye tissue specimens (numerous contributions of Sarks, Hageman, Mullins, Lutty, Bergen, Lengyel, and Curcio) have been examined by light, confocal, or electron microscopy, in conjunction with proteomics and with antibodies to specific drusen-associated proteins (Curcio et al., 2017). These studies emphasize the heterogeneity of drusen, a concept
initially developed by Sarks and coworkers (Sarks et al., 1980; Sarks et al., 1994; Sarks et al., 1999) and strongly suggest that chronic local inflammation at the level of BrM is an important contributor to drusenogenesis. In vitro, the transcriptome and proteome of RPE cells, such as cultured primary retinal cells (fetal or from postmortem human donor eyes) (Alge et al., 2003; Oshikawa et al., 2011; Pao et al., 2018) has been determined. Stable isotope labeling of amino acids showed that these cells secrete a variety of extracellular matrix proteins, complement factors, and protease inhibitors, that have also been reported to be major constituents of drusen (An et al., 2006). In addition, abnormal protein secretion by human primary RPE cultures derived from AMD patients has been observed compared to age-matched controls (An et al., 2006). However, the fact that major components of drusen can be reproduced by RPE cells without the need for PR outer segments, supports a crucial role of RPE in drusen formation (Pilgrim et al., 2017). At the same time, it suggests that PRs may contribute but are not essential for drusenogenesis. Off note, it is important to emphasize that cells in culture were treated with heat-inactivated serum, and that the contribution of components from this material to drusenogenesis, as "dietary" contribution, is highly likely (Bretillon et al., 2008; Pikuleva and Curcio, 2014; Pilgrim et al., 2017).

Wang and coworkers found that, after simultaneous mass spectrometry analysis of both archived drusen and RPE material, similar protein profiles, but with higher intensities and greater variability in the drusen. Within the limits of unavoidable sample contamination, these data suggest that other than RPE alone, additional local cells or tissues contribute to formation of debris in the sub-RPE-BL space (Wang et al., 2010).

3.2. The systemic side of drusen.

Drusenogenesis theories have focused on the role of lipids and immune-mediated effects. Lipoproteins, neutral lipids (Curcio et al., 2011), complement-activating molecules and other immune mediators as well as monocyte-derived cellular processes have been identified within drusen (Hageman et al., 2001; Penfold et al., 2001; Anderson et al., 2010; Molins et al., 2018), which indicates the biogenesis or propagation of drusen from the systemic side.

3.2.1. Bruch’s membrane.
The main functions of BrM are structural, to support the RPE, and to regulate the transport of fluid, ions and biomolecules from the choroid to the RPE, and *vice versa* (Curcio and Johnson, 2012). BrM thickening and decline of hydraulic conductivity have been observed during aging (Hussain et al., 2010; Cankova et al., 2011). Studies suggest diffuse thickening of the inner aspect of BrM is associated with retinal pigment epithelial hypopigmentation, focal atrophy, and soft (large) drusen formation (Bressler et al., 1994). A variety of extracellular matrix components have been detected in diffuse thickenings of BrM (Fernandez-Godino et al., 2016). Immunohistochemical reactivity of BrM showed age-related accumulation of type I collagen and localized changes associated with some drusen (Newsome et al., 1987; Curcio and Johnson, 2012). The tissue inhibitor of metalloproteinases-3 (TIMP-3) protein, a major component of the drusen proteome, showed high immune-reactivity in human drusen and in BrM (Fariss et al., 1997). The continuous turnover of BrM during life could provide a continuous local supply of BrM proteins. Some of the remnants may be cleared to the blood but some of them might end up in drusen. Please note, that most studies on the aspects of BrM thickening have been performed by light microscopy on paraffin sections. In future studies, it will require TEM or high resolution light microscopy to confirm the majority of these findings, and to distinguish, for example, between "BrM thickening" and basal laminar deposits.

### 3.2.2. Choroidal capillaries.

The choriocapillaris is located directly underneath the RPE and BrM. It is composed of a unique vascular network which provides nutrients and fluid for the RPE and the retina (Bernstein and Hollenberg, 1965). The abundance of fenestrations on the RPE aspect of the choriocapillaris endothelium makes this vascular bed much leakier than non-fenestrated vessels (Bernstein and Hollenberg, 1965). A compromised interface can result in various abnormalities such as choroidal neovascularization (CNV) and AMD (Lutty et al., 2010).

With age and in AMD, the choroid thins. The choriocapillaris loses density and covers an increasingly smaller portion of BrM. At the same time, increased drusen deposition occurs, as witnessed by histopathological evidence (Ramrattan et al., 1994; Iida et al., 2004). OCT Angiography (OCTA) showed atrophy of choriocapillaris underneath and beyond the region of photoreceptors and RPE loss (Wakatsuki et al., 2015; Moreira-Neto...
et al., 2018), in agreement with previous and parallel histopathological studies (McLeod et al., 2009; Biesemeier et al., 2014). In human macular sections, histopathological evaluation of the sub-RPE-BL deposits together with potential vascular changes, showed that vascular density was inversely correlated with sub-RPE-BL deposit density (Biesemeier et al., 2014). Curcio and coworkers observed that modest endothelial cell loss in the choriocapillaris also occurred directly adjacent to basal linear deposits and subretinal drusenoid deposits (Curcio et al., 2013). Sub-RPE-BL deposits showed a positive correlation with the number of ghost vessels in the choroid, suggesting that vascular endothelial cell loss could contribute to deposit formation (Mullins et al., 2011). It has also been shown that the presence of complement components and specifically, MAC, in the choroid increases with aging, and increases even more in AMD-affected eyes (Mullins et al., 2014; Chirco et al., 2016). In fact, C5b-9 complement complexes are present in hard drusen, BrM, and extend to the choriocapillaris in some cases (Johnson et al., 2000; Anderson et al., 2002). C5b-9 complexes were not observed in soft drusen (Mullins et al., 2014).

On whole-mount hydrated preparations of the choroid and BrM, (hard) drusen were located to the intercapillary pillars of the choroid, suggesting a close relationship between drusen formation and the capillary bed (Lengyel et al., 2004). This was observed in earlier studies, but not systematically examined (Friedman et al., 1963). It was suggested that drusen are a manifestation of (a) disturbed transport mechanism(s) of substances across the capillary wall or BrM (Penfold et al., 2001). Whether this indicates that drusen deposition is the result of slower clearance at the intercapillary pillars or a manifestation of a disturbed transport mechanism of substances across the capillary wall, or both, needs additional investigation. Of note, further pathological compromise of the vascular bed and BrM leads eventually to the development of subretinal neovascularization and wet AMD.

3.2.3. Contribution of blood proteins.

Penfold and coworkers suggested that breakdown of the normal choroidal vascular function allows the movement of plasma proteins to the sub-RPE-BL space and this leakiness is one of the cause of initiating the progression to AMD (Penfold et al., 2001). Another study involved the analysis of age-related changes in various proteins and lipids in the BrM using multiplexed Raman spectroscopy and found age dependent change in
heme signals (Beattie et al., 2010). However, there are no detailed and definitive studies how these plasma molecules end up in the sub-RPE-BL space. Involvement of fenestrations, breakdown of tight junctions, active vesicle transport (caveola) and receptor-mediated endocytosis (for macromolecules) have been suggested. Fenestrations are found predominantly on the endothelial vessel wall closest to the RPE (Bernstein and Hollenberg, 1965; Pino, 1985; Mancini et al., 1986). Rodent studies suggested that the number of fenestrae initially increases with age; but in advanced age and in AMD the number of fenestrae decreases (Burns and Hartz, 1992; McLeod et al., 2009). Transport through fenestrae is likely to be tightly regulated but it is not yet fully characterized (Pino and Essner, 1981; Essner and Gordon, 1983). Tight junctions of the choroidal capillaries show a tendency to become leaky with age, and lack transport regulation which may facilitate movement of plasma proteins from the choroid towards to the RPE (Nakanishi et al., 2016) (Aiello et al., 1998). Finally, vesicle- or receptor-mediated transport of proteins also exist in the choroid. (Smith et al., 1989). Taken together, transport of proteins at the choroid/BrM interface is complex and warrants further investigation.

It has long been speculated that both blood plasma and incomplete digestion of photoreceptor outer segments contribute to the buildup of drusen material (Farkas et al., 1971a). It has also been suggested that drusen formation in the retina may be similar to plaque formation in arterial walls (Curcio et al., 2001), which, again, suggests that the contribution of blood proteins may be more important than previously thought (see section 8 on “drusen and plaques”). However, there is a paucity of information as to what extent proteins from the blood really contribute to drusen formation. It is thus plausible that some molecules exit the choroidal vessels into the extracellular space adjacent to the RPE, especially as the barriers in place to prevent such an event from happening, become compromised with age.
4. Selection of transcriptomic and proteomic datasets to determine the origin of drusen proteins

4.1. Exclusion criteria and considerations.

One of the main goals of this study was to compare subretinal cellular transcriptomics and proteomics as well as the blood proteome with proteins that are present in drusen. To achieve this, we made use of a subset of studies from the literature as well as our own data. Apart from the drusen protein studies, which date back to 2002, we only considered here mRNA and protein studies published over the last 8 years; we did not include retinal microRNA studies, non-coding RNA, metabolomics, imprinting studies and data from (differences in) single-cell expression studies, simply because there are relatively few confirmed and validated studies for the various types of retinal tissues available yet.

Multiple excellent transcriptomics and proteomics studies have been published on different layers of the retina/RPE/choroid complex, these are reviewed by a number of authors recently (Skeie and Mahajan, 2014; Tian et al., 2015; Zhang et al., 2015a). However, the studies currently available differ in many aspects, including study design, retinal area and retinal cell type examined, sample source selection, sample handling, sample numbers investigated, probe labeling methodology, microarray- or RNA sequencing- methodology as well as the platform, quality and type of bioinformatics programs used for analysis. It is not our goal here to describe and compare all the retinal transcriptomic or proteomic data in the literature. Nonetheless, if one wants to compare different sources (subretinal transcriptomics and proteomics, blood proteomics) and/or outcomes (drusen proteins), similarity of the components and parameters of the comparison(s) is obviously, highly desirable (Ahmad et al., 2018).

In the relevant transcriptomics literature, at least three phases can be observed: studies before and after the introduction of the MIAME (Minimum Information About a Microarray Experiment) quality guidelines studies (Brazma et al., 2001); studies before and after the introduction of whole genome microarrays (at least 22000 genes (22 K or more)) and studies before and after the introduction of RNA-Seq and GTex criteria. Over time, a similar technological development has taken place in the proteomics field: from 2-D gels to high pressure liquid chromatography columns coupled and high throughput...
mass-spectrometry-based studies (Geyer et al., 2016). In principle, the quality of large-scale transcriptomics and proteomics studies has continued to improve, and better and more complete datasets may become available in time that may change some of the interpretations described here.

There are several obvious differences between transcriptomics and proteomics studies. In principle, transcriptomics techniques are highly sensitive and highly quantitative, but as such, highly susceptible to RNA contamination or degradation. In addition, transcriptome changes may not equate with changes on coded proteins and as such are further away from biological function.

Proteomics studies, however, are usually less sensitive and quantitation can only be achieved under certain circumstances, but proteomes per se are closer to function. During disease progression, transcriptomics and proteomics profiles of a tissue can change rapidly depending on disease stage. Also, a single tissue under study can be affected by two or more consecutive disease stages at the same time. For example, in AMD, new hard drusen continue to appear in the sub-RPE-BL space, while other drusen in the same tissue already become confluent, and perhaps part of the same retina is already prone to neovascularization. Consequently, for a disease like AMD, where the RPE is subject to consecutive, insidious and overlapping disease stages, it is very difficult to sift out useful and consistent healthy and disease stage specific expression profiles for this cell layer.

Obviously, transcriptomics and proteomics studies cannot be translated one-to-one, due to, for example, differences in RNA and protein synthesis and turnover rates. The sound interpretation of both transcriptomics and proteomics is highly dependent on the use of advanced bioinformatics and knowledge databases, which combine millions of data-points from human, mouse, and rat studies. Nonetheless, it is the investigator, with knowledge of disease pathology, molecular biology and bioinformatics alike, who can make the difference.

There are two goals with most transcriptomics (or proteomics) studies: One type of study aims to find a complete molecular blueprint of the cells or tissues of interest; these studies usually yield an enriched expression data set for the cell of interest. This type of study usually includes both genes specifically expressed in the cell type of interest, but also genes expressed in similar cell types. For example, the RPE is probably defined by a few hundred RPE-specifically expressed genes, a few thousand neural cell-type
expressed genes, many expressed housekeeping genes for basic functions, as well as
many genes which are on “standby”. The genes that are on “standby” have a very low
(leaky) expression if the cell is in a state of homeostasis. However, if the environment
changes, these very low expressed genes can rapidly be expressed to adapt the cell to a
changing environment. For example, the RPE shares most likely the RNA expression of a
large portion of its transcriptome: neural cell type genes, the household genes, and low-
level expressed genes, with the other (neural) cell types in the retina (own
observations). Finally, there are many specific non-expressed genes in a certain cell-
type. An example of expression studies which aim to find a molecular blueprint of the
cell is the uncurated RPE expression dataset, RPE-ET (Table 3), which contains 10% of
the biologically highest expressed genes in the RPE (Booij et al., 2009).
The other type of study aims to find only a maximum of genes specifically expressed in
only the cells or tissue of interest. These few hundred genes, in the context of the more
generally expressed genes, give the cells of interest their specific cell type-associated
functionalities. An example is the dataset, RPE-ST (Bennis et al., 2015), which contains
170 RPE-specific expressed genes derived from previous RPE expression studies (Booij
et al., 2009; Booij et al., 2010b; Strunnikova et al., 2010) (STable 2).

4.2. Description of expression datasets used for drusenomics.
Apart from the 89 drusen protein data set, we used in this review 11 additional
subretinal and blood data-sets derived from previous transcriptomics and proteomics
studies; this is summarized in Table 3. We found that these transcriptomics and
proteomics databases complement each other and, together, give a more complete
overview of relevant expressed genes/proteins per tissue investigated. A common
feature of all high throughput studies is that they generate, by default, a small
percentage of misidentifications. This is due to cellular or molecular contaminations, or
mis-representation due to experimental sample handling. Therefore, individual gene
findings usually need to be confirmed by at least a second technique which focuses on
the analysis of single genes or proteins.

We used pure, enriched and curated cellular expression datasets. Pure datasets are
those without possible contaminations of other cell types while enriched datasets are
those datasets that have a certain degree of contamination of adjacent cell types. Finally,
curated datasets are those which are manually enriched either by bioinformatics or by
literature search to remove inevitable contaminations or irrelevant data as much as possible. The curation strategies employed are presented in Figure 3. Most of the (non-curated) data were used for qualitative studies, have been published and analyzed elsewhere, and are mentioned below for reference. For the quantitative studies, we used curated datasets. The photoreceptors and choroidal transcriptome datasets, cPR-ET and cChor/ET (Booij et al., 2010b) (GEO database accession number GSE20191) have not been fully published before and therefore, their description will receive a little more attention here.

First of all, we used (1) a combined data set for drusen proteins, curated by hand as described above (Table 1). Furthermore, we used (2) a photoreceptor outer segment proteomics dataset published by Kiel and coworkers (Kiel et al., 2011), which contains proteins reflecting a multiscale signaling network associated with rhodopsin, the major protein component of rod photoreceptor outer segments. It was constructed by combining relevant proteomics datasets, structural and functional literature mining and bioinformatics approaches (Table 3; STable3). Most likely, this database listing contains some contamination from adjacent cell types, the RPE and choroid. Therefore, a curated list was used for the quantitative studies: we subtracted the most highly expressed sequences of the choroid (top 10% chor Booij; Chor-ET; and the uniquely expressed sequences of the RPE (RPE-ST, Bennis)) from this database listing. The acronym used for this dataset in this manuscript is PRos-EP (Photoreceptor outer segment-enriched proteomics). The annotation of the curated version (c) of this dataset is cPRos-EP.

(3) The RPE-specific database with 170 entries was constructed by bioinformatic curating and combining other (highly) enriched RPE gene expression databases (Booij et al., 2010b; Strunnikova et al., 2010; Bennis et al., 2015). This database listing should be viewed as a minimal number of RPE-specific expressed genes based on previous –omics studies; the acronym used here is RPE-ST (Specific Transcriptomics) (Table 3; STable 2).

(4) The RPE secretome data from (Pao et al., 2018) that was published recently. RPE cells were grown in vitro to confluency while adding various amounts of zinc to the culture medium. Both the apically and basally secreted RPE proteomes were determined. Here, we use the basal secretome proteomics listing which contains 276 entries. (Table 3; STable 1). Due to its nature, this dataset does not contain contamination from other cell types but may contain contaminants from the culture
medium. In addition, its in vitro basis may not be fully representative of the in vivo situation, particularly in the disease state. The acronym for this database in this study is RPE-IVS (in vitro secreted) (Table 3).

(5) The RPE/choroid proteomics dataset from Zhang and coworkers that contain proteins extracted from RPE/choroid tissues of eyes from five individuals, fractionated and separated using SDS-PAGE and analyzed using mass spectrometry (Zhang et al., 2016). In the RPE/choroid the authors identified 2755 non-redundant proteins. This dataset is rather large in components and is likely to contain entries from multiple cell-types (RPE, choroid, blood and possibly PR), and not only (RPE/choroid), given the inevitable contaminations of the PR sample with RPE and vice versa, and the contamination of the choroid with blood. The authors deposited their data to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers PXD001424 and PXD002194. The acronym for this database in this manuscript is RPE/chor-EP (RPE/choroid-enriched proteomics) Table 3.

(6) The blood proteome listing by Geyer and coworkers was produced by a new efficient plasma proteome profiling pipeline (Geyer et al., 2016). Using a modified mass spectrometry-based workflow they were able to identify and quantify at least 1000 plasma proteins. Given the nature of the samples, it is unlikely to contain other retinal cells or proteins as contamination. The acronym for this database in this study is BL-SP1 (Blood plasma-specific proteomics; no 1) (Table 3; STable 4) (Geyer et al., 2016).

(7) The blood proteome dataset by Farrah and coworkers contains a non-redundant set of 1929 protein sequences from human plasma detected by tandem MS (Farrah et al., 2011). The full data are available via PeptideAtlas, a large, international database of publicly accessible peptides identified in tandem MS experiments in a multitude of organisms. This is also a “pure” database listing. The original dataset contains endogenous chemicals, which we removed for our analyses. The acronym for this database in this study is BL-SP2 (Blood plasma-specific proteomics, no 2) (Table 3).

(8) The BL-PHP blood proteome dataset consists of 262 HAP binding proteins from AMD patients and controls, as recently described (Arya et al., 2018). Plasma samples were taken from 23 individuals aged 65-90 with late stage AMD, each displaying drusen and choroidal neovascularization in clinical images and attending the anti-VEGF injection clinic at Moorfields Eye Hospital, London (STable 5).
(9) The atherosclerosis plaque proteomics dataset contains 3196 entries based on a comprehensive review of the literature in this field (Bleijerveld et al., 2013). The acronym used in this study is AS-EP (Atherosclerosis-enriched proteomics) (Table 3). The (large) dataset is available as supplementary file to the authors’ publication.

(10-12) Transcriptomics datasets of the photoreceptor (acronym: PR-ET: Photoreceptor; enriched transcriptomics), the choroid (acronym: Chor-ET: Choroid-enriched transcriptomics.) (Table 3), and RPE (acronym: RPE-ET: RPE-enriched transcriptomics) were produced using the same Agilent methodology and platform. For functional annotation and quantitative analyses, curated versions of these databases were constructed, named, respectively, cPR-ET (STable 6) and cChor-ET (STable 7). The (c)RPE(-ET) database has been extensively published elsewhere (Booij et al., 2009; Booij et al., 2010b).

4.3. Functional annotation photoreceptor (cPR-ET) and choroidal (cChor-ET) datasets. The PR-ET and Chor-ET datasets contain, respectively, the averaged top 10% highest expressed genes in the photoreceptor and choroid. The isolation methods, study design and methodological issues for these datasets have been extensively discussed elsewhere (Booij et al., 2009; Booij et al., 2010b). These raw datasets were used for the qualitative studies in this manuscript. The experimental studies were performed in agreement with the declaration of Helsinki concerning the use of human material for research and followed both MIAME and GTex criteria (Brazma et al., 2001; Consortium, 2013). We curated both datasets PR-ET and Chor-ET according to scheme C in Figure 3. In order to obtain cell-specific datasets for photoreceptor and choroid, which are useful for both cell-specific functional annotation and for quantitative studies described elsewhere in this manuscript. Consequently, we removed from the PR-ET and Chor-ET datasets all expressed genes that overlap between them (either contaminations or truly shared gene expression). This resulted in two smaller curated datasets. Subsequently, we also removed all potentially present RPE-expressed unique sequences (RPE-ST dataset) to generate the cPR-ET and cChor-ET datasets. Thus, the resulting cPR-ET and cChor-ET datasets contain less, but highly cell-specific entries compared to PR-ET and Chor-ET. Hence, we ended up with a highly photoreceptor-enriched gene expression dataset consisting of 745 genes (STable 6) and a highly enriched expression dataset for the choroid of 848 entries (STable 7).
We ran an Ingenuity core analysis (www.ingenuity.com) on both cPR-ET and cChor-ET datasets. This type of analysis typically yields data-driven functional annotations (i.e. it produces biological motifs, canonical pathways and molecular networks enriched in the dataset). The results of the cPR-ET analysis are presented in Table 4; PDF summary. We found some very basic and very specific functional features related to established photoreceptor function. The basic annotations included “cancer”, “cellular function and maintenance” as well as “tissue morphology”. One could speculate that these relate to the unique shape of the photoreceptor cell, and its unique ability to renew its photoreceptor outer segments. More specific (highly ranked) annotations included “photo transduction cascade”, “visual system development and function” and “neurological disease”. These data-driven results clearly fit with reported specific photoreceptor functionalities from the literature (Diamond, 2017; Musser and Arendt, 2017; Fain and Sampath, 2018).

The choroidal transcriptomics dataset cChor-ET was generated in a similar way to the photoreceptor cPR-ET dataset described above. Obviously, the choroid is not a single tissue, but consists of multiple cell types, including endothelial cells, fibroblast cells, melanocytes, macrophages, and resident lymphocytes. The choroid is unavoidably contaminated with blood cells and proteins. Nevertheless, after curation, we obtained 848 genes with a highly enriched choroidal expression in the cChor-ET dataset (STable 7). Following Ingenuity core analysis, the resulting functional picture of the choroid is, as expected, completely different from that of the photoreceptors (Table 5; PDF summary). We found that two of the top five biological motifs (“inflammatory response” and “inflammatory disease”) and three canonical pathways (“antigen presentation pathway”, “acute phase response signaling” and “complement system”) are all involved with the immune system. This confirms the crucial role of the choroid and blood in external immune surveillance of the eye (Dick, 2017). The second highlight of this analysis was the canonical pathway “atherosclerosis signaling” which again points to an important resemblance between healthy or disease processes going on at the BrM (choroidal-RPE interface) and the vessel walls (see also section 8).

Finally, both the canonical pathways and the highest ranked networks identified in this cChor-ET analysis indicate tissue damage and injury. One possible explanation is that this damage refers to early molecular complement attack already present or setting in, which may be well before any morphological changes or damage may be visible.
Alternatively, although we used data from healthy post-mortem eyes, performed the studies according to MIAME and GTex guidelines, 3’ primer design which avoids potential problems due to 5’ directed degradation, as well as very stringent RNA quality controls, the tissue damage and injury might still be due to post-mortem damage. Detailed data relating to both cPR-ET and cChor-ET analyses are available on request.
5. Drusenomics, part II: Qualitative analysis.

5.1. Comparative study design considerations.

From the previous sections, it has become clear that a systematic investigation into the origin of proteins in drusen is lacking in the literature. Authors suggest a variety of protein sources, frequently on the basis of single observations. Systematic investigation of this phenomenon is hampered by the heterogeneity of source samples, methodology, and analysis. Large scale transcriptomics or proteomics studies frequently end up with a rather abstract annotation analysis, allowing a certain error rate and lack of detail; small scale studies frequently lack sufficient technical, methodological or biological replicates.

In a first attempt to investigate the origin of drusen-proteins systematically, we used the presence of functional protein clusters identified in drusen (described in section 2). Subsequently, we investigated whether expression of entities in these clusters also (partly) occurs in the various non-curated expression database sets selected for this study. An overview of this comparison is presented in Table 3. This comparison serves two purposes: Firstly, the presence or absence of clusters in subcellular databases or blood may give a qualitative indication of the origin of (the proteins in) the cluster and secondly, it gives an indication of the completeness, quality, and contamination in each of the databases listed.

We argued above (section 4.2.) in detail that all individual transcriptomic and proteomics database lists used here (and those in the literature) are incomplete and, as a rule, have a degree of RNA/protein/cellular contamination due to original mixed cell sampling. How do we compare incomplete, contaminated datasets?

First of all, one should have some knowledge of the study-design and character of the dataset under study, to understand why certain entries do, or do not, appear. The characteristics of the databases used are described in section 4.2 above. As an example, in the study of RPE/Chor-EP expression dataset (Table 3) we systematically identified that a large number of functional cluster queries/entries are indeed present. However, this is most likely due to the fact that the RPE/Chor-EP study contains proteins from photoreceptor, RPE, choroid and blood. Thus, we decided to use this dataset as a positive control (i.e., a dataset where almost all genes relevant to the study are expressed/present). Similarly, we used either the unique sequences from the RPE-ST
dataset, or the consequent absence of an entry in a set of similar dataset listings, as a negative control.

In our qualitative comparison, the incompleteness and potential contaminations of the various non-curated datasets (Table 3) may be largely overcome by considering similar data from different studies at the same time. For example, a specific query may be present in all photoreceptor studies, and, at the same time be absent from all blood proteomics studies.

5.2. Where do proteins in drusen come from? A qualitative comparison.

We will now turn to the interpretation of the highlights of the comparative study results presented in Table 3. In the first column (top to bottom) the functional gene clusters from the molecular networks of drusen (Figure 2a-2d) are presented together with their individual gene content (column 2) and functional annotation (column 3). On the top rows (fourth column onward), subretinal transcriptomics and proteomics as well as blood proteomics data(sets) from the literature are given. For full length names of the abbreviated gene/proteins, see Table 1.

5.2.1. Network 1.1: The complement gene cluster.

The first functional cluster (Table 3; Figure 2a), in our analysis consists of the complement end proteins: C7, C8 isoforms as part of the Membrane Attack (MAC) protein group in general as well as the multifunctional PRELP protein. Mutations in PRELP cause myasthenic syndrome (Engel, 2018). Among other functions, PRELP is involved in regulation of the complement cascade (Engel, 2018). As expected, we observed that the (alternative pathway) complement gene transcripts/proteins are absent from all photoreceptor and/or RPE transcription and proteomics datasets (PRosEP, PR-ET, RPE-ET, RPE-ST, RPE-IVS). The only possible exception is the presence of these proteins in the RPE/chor-EP dataset (positive control), where they most likely originate from the choroid/blood component of the sample (please note again that the choroid sample is inevitably contaminated with blood). In contrast, in 2 out of 3 blood-plasma datasets (BLP-SP1 and BLPHP) this cluster is present, except for PRELP. The latter entry is apparently uniquely present in the Chor-ET listing (and in the positive control RPE/Chor-EP) and is probably produced in the choroid. Interestingly, this leads us to suggest that the systemic driven complement attack from the blood is locally regulated by PRELP produced by choroidal cells (Happonen et al., 2012).
5.2.2. Network 1.2: The collagen cluster.

The second functional cluster is that of the collagens and related molecules (Table 3; Figure 2a). Here, the picture directly becomes, perhaps understandably, more complicated. Collagen proteins are likely to be produced (RNA, protein) by the basal secretion of the RPE and apical secretion of choroidal cells, and become part of the a-cellular BrM (Booij et al., 2010a) and, at least theoretically, by the apical RPE and by the PR for the interphotoreceptor matrix (IPM). Indeed, proteomics of apical RPE secretion \textit{in vitro} (Fort and Lampi, 2011) and functional annotation of the \textit{in vivo} human RPE and photoreceptor predicted secretomes (based on transcriptomics of the RPE and photoreceptor cells; Bergen, unpublished), suggest that several specific collagen proteins may (transiently) be present in the IPM, although their presence was never detected by immunohistochemistry (yet).

During life, there is a constant turnover of ECM's, resulting in a mix of newly synthesized and (partly) digested collagen fragments shuttling around the subretinal area. Interestingly, the unique RPE-ST database (170 entries) does not contain any collagen related entries, thereby confirming that, if the RPE produces collagen, none of these collagens are made exclusively by the RPE (but also by adjacent tissues like the choroid). Further to this, it is of interest to note that collagen (type 8A1) is produced or present in the PR-ET, RPE-ET, Chor-ET and the RPE/CHOR-ET datasets, which supports the hypothesis that the (two collagenous layers of the) BrM, at least in part, are built from both the RPE and choroid sides side (Booij et al., 2010a). The data in Table 3 further suggest that the BrM proteins COL1A2, COLA1 and COLA2 are produced or are present exclusively (or at least mainly) in the Chor/blood, and not in the PR and RPE datasets.

COL6A1 is secreted basally by the RPE \textit{in vitro} (Pao et al., 2018) (Table 3), and is part of the BrM (Booij et al., 2010a). Consequently, the protein may end up in the drusen dataset either as a contamination, or as a remnant of the turnover of BrM components. The proteins THBS4 and TNC occur only in one of the blood proteomics datasets (BL-SP1). The presence of other entities (RBP3, EFEMP1, PLG, GPNMB, SEMA3B, TBHS4) from this cluster in multiple PR/RPE and Chor/blood listings suggest that these genes/proteins can be derived from different sources.

5.2.3. Network 1.3: The crystallin cluster.
The third cluster of drusen proteins to be discussed are the crystallins (Crabb et al., 2002; Nakata et al., 2005) (Figure 2a), frequently referred to as heat-shock proteins, which act as chaperones to prevent or reduce protein degradation in stressed or aging cells. Although they may have a more structural role, it is possible that the expression of crystallins is increased only in those studies in which cells or tissues have been exposed to a relatively large amount of stress. This would mean that further systematic and methodological analysis of consistent (differences in) expression does not make sense. It is remarkable however, that the CRYBB2 protein is present in the PRos-PT, RPE-IVS, and the RPE/Chor-EP studies and in drusen. Consequently, this protein may originate from the PR outer segments, processed, transported and secreted by the RPE and then accumulates in drusen. To our knowledge, this is the only photoreceptor protein known to possibly make it through phagocytosis and lysosomal processing in the RPE and end up in drusen (Feeney-Burns et al., 1988; Hoppe et al., 2001).

5.2.4. Network 2.4: Genetic and developmental ophthalmic disorders.

The fourth functional gene cluster (Network 2, Figure 2b) can be considered as a pathobiological cluster of developmental and ocular disease. The comparative analysis (Table 3) shows that ATP5F1B, ACTB and annexin2 (ANXA2), are present in a number of PR/RPE and Chor/blood expression datasets. These entries may thus be expressed in multiple cell types or blood. ANXA2 is included here “within brackets”, since it was initially assigned to drusen using proteomics, but later the same authors stained for ANXA2 in human donor eyes and concluded it was not present in drusen (Crabb et al., 2002; Nakata et al., 2005). The CRYAB, ENO2 and SPTAN1 proteins cannot be clearly assigned, but appear to be of a local cellular origin (PR, RPE, or Chor) and not from the blood. The subcellular/systemic assignment of the FN1 and MYH9 entries are not clear. The BFSBP1 and BFSP2 proteins neither occur in the subretinal datasets, the blood proteomics lists, nor the positive control (RPE/Chor-EP; Table 3). According to the literature, both are structural proteins that specifically form filaments in the cytoskeleton of lens-cells (www.ingenuity.com). We therefore conclude that these are very weakly expressed genes which express proteins that build up slowly and/or with a long half-life. The only other explanation that could be offered is that they are contaminations within the drusen dataset.

5.2.5. Network 3.5: Injury, inflammation and dermatological disease.
The fifth functional drusen cluster (Network 3; Figure 2c) is related to interacting genes and proteins involved in injury, inflammation and dermatological disease. A substantial number of entries of this group seem to have both a cellular as well as a systemic presence or origin since they are present in at least two database listings from PR/RPE and Chor/blood category. These include ANXA1, ANXA5, CKB, GAPDH, PRDX1 and S100A8.

The SERPINA3 and ALDH1A1 proteins appear only in at least two of the RPE-Chor-EP, Chor-ET, BL-SP1, BL-SP2 and the BL-PHP datasets, but not in the PR/RPE dataset. Thus, both proteins appear to come from the systemic side of drusen. FRZB (SFRP3) (www.genecard.org) is present in 3 PR/RPE listings (Table 3), including the RPE-specific listing (RPE-ST), and in only one choroid-enriched list (Chor-ET). We tentatively assign this drusen protein primarily to the RPE, and as a contamination in the (PR and/or) Chor database listings. S100A7 is only once present in the PRos-EP proteomics dataset. Assignment of S100A9, TYRP1, LAMB2, APOE, and FrzB or LUM to a single source cannot be done on the basis of this comparison.

5.2.6. Network 4.6: Cell to cell signaling; systemic involvement

As can be expected from the functional annotation “cell to cell signaling; systemic involvement”, almost all of the entries of this category of drusen proteins, appear in the Chor/blood datasets (Table 3). The exceptions are clusterin (CLU), ANXA6 and HRG.

From the literature, we know that CLU is a ubiquitously expressed gene that is expressed in all cell types (Wilson and Zoubeidi, 2017). It is therefore not surprising that it features in both the PR/RPE as well as the Chor/blood listings. The final assignment of ANXA6 and HRG, on the basis of this comparison is not clear. The role of the systemically derived HRG drusen protein is discussed in detail below (section 7). Of particular interest is the expression of CFH, given its central regulatory role in the complement attack on (chemically modified) drusen components. There is compelling evidence in the literature that CFH is present in the blood, the neural retina and that it is also expressed by the RPE (Li et al., 2014; Mullins et al., 2014; Whitmore et al., 2014; Chirco et al., 2016; Chirco and Potempa, 2018; Toomey et al., 2018). The presence of CFH protein in the blood corresponds with the data and proteomics listings of blood in Table 3. What is not entirely clear is why CFH does not pop up in the RPE listings. This can perhaps be explained as follows: The enriched RPE-ET transcriptomics list contains only the highest
10% expressed genes in the RPE. Apparently, CFH is somewhat lower expressed and so does not belong to this group (Warwick et al., 2014). Also, the RPE-ST specific listing only contains 170 entries uniquely expressed by the RPE; whilst CFH is produced in other cells or blood as well. Finally, CFH does not occur in the RPE-IVS basal secretion proteomics listing, which is not entirely unexpected as recent evidence suggests that CFH is secreted apically, not basically by the RPE (Kim et al., 2009; Pao et al., 2018).

5.2.7: Conclusion.

On the basis of our qualitative comparison, we suggest that a number of (functional clusters of) drusen proteins come from the blood, while others come from a subretinal cellular compartment. These results are in line with the findings in the literature. However, it is not clear yet how many of the drusen proteins come from each particular compartment. The latter may be estimated by a more quantitative analysis, which is the subject of the next section.
6. Drusenomics, part III: A quantitative approach

6.1. Quantitative analysis and curation of datasets.

In this chapter, we quantitatively compare drusen proteins with transcripts and proteins from adjacent retinal compartments (photoreceptor, RPE, choroid) and blood. As described in section 2 above, the drusen protein list was compiled manually according to the curation strategy presented in Figure 3, scheme A. Similar to the qualitative studies, the quantitative analysis of the origin of drusen proteins is also hampered by two problems: (a) most large scale cellular transcriptomics and proteomics datasets contain some contamination (both RNA and/or protein) from adjacent cells or tissues, and (b) most of the datasets are incomplete due to differences in the study design and methodology used in contributing studies. In other words, we need to use a quantitative comparison strategy that maximizes the signal (number of entries to be compared) and minimizes the noise (number of contaminations in datasets). We overcame the incompleteness of various datasets by pooling the entries from various similar (cell-type specific) studies, to get a more complete numerical picture (Figure 3, scheme 3A).

With regard to possible contaminations, we used two types of datasets. The first category includes datasets that, by definition or by previous curation in the literature, contain cell-specific expressed entries only, such as the RPE-ST, RPE-IVS, BL-SP1 and the BL-SP2 datasets (section 4.2 and/or Table 3). The curation of datasets PR-ET and Chor-ET into cPR-ET (Table 6) and cChor-ET (Table 7) was already described above (section 4.2). The other category datasets used (PRos-EP, PR-ET, Chor-ET) were newly curated, as presented in Figure 3 and 4, in such a way that they, after curation, also only contained cell-type specific entries.

The PRos-EP dataset was curated according to the curation strategy presented in Figure 3, scheme 3C: We removed from the PRos-EP dataset (in principle containing photoreceptor outer segment expressed genes only) all choroidal highly expressed genes (from Chor-ET) as well as potentially present uniquely RPE expressed entries (from RPE-ST) resulting in the curated cPRos-EP dataset (Table 6). The removed choroidal and RPE entries were (potentially) present in the PRos-EP dataset due to truly overlapping gene expression between these different cell types and/or due to contaminations in the original cell sample. Of note, cPRos-EP does still contain entries from both photoreceptor and RPE since contamination between these two is inevitable.
Together, the database listings cPRos-EP, cPR-ET, RPE-ST and RPE-IVS form the neural side of drusen database listings (Figure 4). Similarly, the cChor-ET, the BL-SP1 and the BL-SP2 constitute the systemic side of proteins found in the drusen dataset. In summary, we ended up with three large datasets suitable for further robust analysis: The drusen protein dataset, the “neural source of drusen” database, and the ‘systemic source” of drusen dataset (Figure 4).

Next, we compared the entries present in the “neural source listing” and in “the systemic source listing” with the proteins present in drusen. The result of this analysis is summarized in the Venn-diagram in Figure 5. The comparison revealed that 10 proteins appeared to be uniquely derived from the neural side (Table 6) and 37 proteins are derived from the systemic drusen side (Table 8). In addition, there were 23 proteins that come (potentially) from both the neural and systemic side (Table 9). For 19 drusen proteins (out of the 89), the origin remained unclear as they were neither present in the “neural source” nor in the “systemic source” expression datasets (Table 10).

6.1.1. Ten out of 89 drusen proteins originate uniquely from the PR/RPE.

Our analysis yielded 10 drusen proteins that originate from the PR or RPE (Figure 5; Table 6). They are both uniquely present in the drusen proteomics dataset and the neural source of drusen database listing. We traced these proteins back to their original source(s), and we observed that three of them (FRZB, RDH5 and RGR) originally came from the unique entries in the RPE-ST dataset, five came from the RPE-IVS dataset (CRYBA1, CRYBA4, CRYBB2, ENO2 and TUBB3), and the remainder from the cPRos; cPR-ET datasets. Taken together, 8 out of the 89 drusen proteins originated uniquely from the RPE, while 2 came from the curated PR/RPE database listings (cPROS; CPR-ET) (Table 3). Finally, we also reviewed the psychochemical properties and molecular weight (Mw) of these 10 proteins (Table 6). We do not know the HAP binding properties of these proteins, but they do not occur in the BH-PLP HAP-binders’ dataset (Table 5).

In conclusion, we did not observe any common signatures of these proteins that would explain why they in particular are trapped in the sub-RPE-BL space (whilst other proteins are not).

6.1.2. Twenty-three of 89 drusen proteins originate from both neural and systemic sources.
From our analysis, 23 drusen proteins were present in both the “neural source” as the “systemic source” datasets (Figure 5; STable 9). From these, only 1 protein (S100A9) falls in the curated PR/RPE category (Figure 4). Additional groups of two and twenty proteins come from the unique RPE RPE-ST and the RPE-IVS datasets, respectively. At the same time, all 23 of these proteins are also present in the blood. Remarkably, in this shared category the majority of proteins are either secreted basally by the RPE or present in a soluble form in the blood plasma. We hypothesize that the proteins in this category enter the sub-RPE-BL from both sides, where they “meet, greet and stick”, i.e. form aggregates that cannot be cleared and therefore contribute to drusenogenesis. Functional and pathobiological annotation of (combinations of) these proteins can be found in STable 9a.

6.1.3. Thirty-seven out of 89 drusen proteins originate from the choroid/blood.

We found that 37 out of 89 drusen proteins uniquely originate from choroid or blood datasets (STable 8). From these, 31 proteins came from the plasma-proteomics datasets (BL1-SP1 and BL-SP2). The remaining six entries (ANXA6, FLBN5, HLA-DRA, MFAP4, PPRELP, SEMA3B) are present in cChor-ET database listing and thus originate from either the choroid or blood. Functional and pathobiological annotation of (combinations of) these proteins can be found in STable 8a.

In summary, we again observed a large proportion of drusen proteins are most likely originating from the blood. If we take this unique category (31 proteins from plasma) and the shared contribution of plasma (23 proteins) from the previous paragraph into account, we can conclude that as many as 54 out of 89 drusen proteins (>60 %) are (co-) derived from blood plasma.

6.2. Nineteen drusen proteins out of 89 were not assigned.

We can, in the end, still not determine the possible origin of 19 out of 89 drusen proteins (STable 10). Why is this not possible? Do these proteins have a number of characteristics in common that prevents us to determine their origin? To attempt to answer these questions we need to take a closer look at these remaining drusen proteins. The functional annotation of these proteins is presented in STable 10a and they can be divided in five groups: (1) a gamma-crystallin group; (2) a histone cluster group; (3) remnants from) BrM turnover group (4) a beaded filament group; (5) a rest group containing a variety of proteins that do not belong to a specific functional group.
In group one, we observed several gamma-crystallin-isoforms in drusen, which have not been assigned to a specific source (as yet). Crystallins are commonly found in the lens but are also present in soluble form in the retina (Jones et al., 1999) and probably act as chaperone proteins after (oxidative) stress. Indeed, in the mouse retina, crystallin expression has a binary nature in which either they are highly upregulated, or their expression is extremely low (Templeton et al., 2013). Gamma-crystallins may have a neuroprotective role (Thanos et al., 2014). At least one specific type of crystalline (alphaB type) is known to be secreted by the RPE through micro-vesicle release (Kannan et al., 2016). In conclusion, the source assignment of gamma-crystallin isoforms in drusen, in our comparison, may be hampered by this binary expression. More specifically, it will simply be absent from a number of subretinal expression datasets and, as such, too little evidence exists to make a definite assignment.

Next, we found a group consisting of HIST1H1E, HIST1H2BJ, HIST1H2BL and HIST2H2BE. Histones are highly basic proteins that have an essential role in the maintenance of nuclear DNA structure and gene transcription. HIST1H1E is a 219-amino acid protein that binds to the linker DNA stretch between nucleosomes, while HIST2, together with HIST3 and HIST 4, are part of the nucleosome core (Tessarz and Kouzarides, 2014). Damaged or dying cells (potentially RPE or endothelial cells of the choroid) can release cellular as well as nuclear fragments that may contain histones. Alternatively, high concentrations of serum histones have been detected in several human diseases (Yang et al., 2015). These extracellular histones may get trapped in BrM and drusen. Interestingly, extracellular histones trigger activation of multiple signaling pathways related to cell death, growth and inflammation and may play a role in autoimmunity, aging and disease (Allam et al., 2014; Kalbitz et al., 2015; Zhang et al., 2015b). Why these specific histone proteins (and not others) are trapped in drusen and cannot be assigned to a source remains to be elucidated.

The third drusen protein group with an as yet unassigned source contains elastin (ELN), collagen 8A1 (COL8A1), biglycan (BGN) and tissue inhibitor of metalloproteinase 3 (TIMP-3) proteins. These proteins may come from an as yet little considered drusen protein source: the BrM and its (turnover) components (Booij et al., 2010a; Curcio and Johnson, 2012). TIMP-3 is expressed in the RPE (Ruiz et al., 1996) and is crucial for the maintenance of BrM. Mutations in TIMPO-3 caused Sorsby Fundus dystrophy, a
monogenic disease that resembles the phenotype of AMD (Weber et al., 1994). Indeed, as described previously, BrM is dynamic in nature, not only in a physiological sense, but its composition and properties vary with age. Proteins involved in BrM and its turnover may be absent from (some of) our subretinal transcriptomics and proteomics datasets, if the relative expression levels of such entries are low. The middle layer of the BrM consists of elastin so it is conceivable that the RPE and/or choroidal cells make this protein. Within the BrM, elastin turnover might be relatively low, thus little “new” elastin is needed. While elastin protein fragments (tropo-elastin) might be present in some drusen as a remnant from BrM-turnover, there is little evidence in the literature that they accumulate in drusen.

Fourth, two members of the beaded filament structural protein family, BFSP1 and BFSP2, remain unassigned. Similar to crystallins, these proteins were initially discovered as lens fiber proteins. To our knowledge, it is not clear whether they also play a role in retina/RPE or maybe even the BrM. How these proteins end up in drusen and their origin remains unclear.

The fifth, yet unassigned group contains a number of, apparently, unrelated proteins, including retinol binding protein 3 (RBP3), tyrosinase protein-like 1 (TYRP1), spectrin alpha, non-erythrocytic 1 (SPTAN1), disco interacting protein homologue (DIP2C), forkhead-associated phosphor peptide (FHAD1), and scavenger receptor class B member 2 (SCARB2).

The RBP3 gene is transcribed in the PR and its protein is located in the interphotoreceptor matrix (IPM). It binds to retinoids which are shuttled from the PR to the RPE, and vice versa (Gonzalez-Fernandez et al., 1993). The TYRP1 gene belongs to the tyrosinase family transcribed in the RPE and encodes an enzyme in the melanin biosynthetic pathway (Lai et al., 2018). Mutations in this gene are one of the causes of albinism (Kamaraj and Purohit, 2014; Kruijt et al., 2018). Both RBP3 and TYRP1 genes may be absent from our datasets given their relatively low, transient or binary expression in the relevant tissues.

Finally, FHAD1 is a small protein that recognizes phosphorylated epitopes on a wide range of proteins as part of an evolutionarily ancient mechanism enabling assembly of protein complexes (Durocher and Jackson, 2002). The expression of FHAD1 is very low in many tissues including the retina but is high in the testis and lungs. Since the
expression is very low, the transcript and protein production or presence may go undetected in the subretinal and blood transcriptomics and proteomics studies we used in this study. However, once the FHAD1 protein has accumulated, as apparently in drusen, it may be (more) detectable there. DIPC2 is a ubiquitously expressed protein that shares homology with a *Drosophila* protein that interacts with the transcription factor disco (www.ncbi.nlm.nih.gov). It is possible that the expression of this type of protein is transient or binary; and it may go undetected in our retinal compartment and blood transcriptomics and proteomics studies for that reason.

Finally, the last two proteins, SPTAN1 and SCARB2 may have related functionalities. The SPTAN1 protein is a part of the cytoskeletal spectrin protein family that is involved in stabilizing membranes of both cell and organelles (Tohyama et al., 2015). It is highly expressed in the brain, and still expressed to a significant level in multiple other tissues. *SCARB2* is a ubiquitously expressed gene that encodes a lysosomal type III plasma membrane glycoprotein (Gonzalez et al., 2014). Given the involvement of this type of proteins in the lysosomal digestion of cellular material, it is tempting to speculate that these proteins come from (transiently present in high numbers) lysosomal membrane fragments basally secreted by the RPE.

6.3. **Blood proteins are an important source of drusen proteins.**

If we summarize the combined data from our literature search and our qualitative and quantitative analyses, we conclude that blood proteins are an important protein source for drusen development. Further studies are needed to confirm and enhance our data, especially on the single-protein level. Given the apparent contribution of blood to the formation of drusen, the next chapters will discuss the role of hydroxyapatite as a retainer of blood proteins during drusen formation, and the similarities that exist between drusen and atherosclerotic plaques, which occur exclusively in the vasculature.
Our analyses provide strong evidence that proteins in drusen come from multiple sources. The next logical step was to consider how proteins arrive and how they are retained in the sub-RPE-BL space. One possibility is that proteins may bind to constituents of the BrM (Tabas et al., 2007). Another is the formation of large oligomers in the sub-RPE-BL space in the presence of the high concentration of trace metals (Lengyel et al., 2007; Nan et al., 2013; Flinn et al., 2014). In addition, it was recently hypothesized that proteins might be retained in the sub-RPE-BL space due to their binding to hydroxyapatite spherules recently identified in human drusen (Thompson et al., 2015) (Figure 6). Since this hypothesis is relatively new, it is described in more detail below.

Using confocal microscopy and hydroxyapatite (HAP)-specific fluorescent dyes, small hollow spherical structures ranging from 0.5 µm to 20 µm in diameter were identified within sub-RPE-BL deposits in retinal tissue sections of human cadaveric eyes (Thompson et al., 2015). The HAP spherules were present in all deposits examined (Thompson et al., 2015). Protein constituents of drusen, such as amyloid-beta, vitronectin and complement factor H, were localized to the surface of the HAP spherules, either individually or in combination (Thompson et al., 2015). Although not all investigated drusen proteins appeared to bind to the surface of HAP (Thompson et al., 2015), this finding proved that the retention of proteins can, at least partly, occur through this protein-HAP interaction. These results also suggested that the binding of proteins to HAP spherules is a wide ranging, though selective, process and thus understanding which proteins can bind to HAP might be important. The plasma protein-binding capacity and selectivity of HAP was recently examined using a quantitative proteomic approach called Sequential Window Acquisition of all theoretical fragment-ion spectra-Mass Spectrometry (SWATH-MS) (Arya et al., 2018). Using this approach, 242 proteins with the propensity to binding HAP were identified and quantified (Table 3; STable 5) (Arya et al., 2018). Taking advantage of the quantitative nature of the analysis the binding of samples from participants with wild type and the AMD associated high risk CFH variant, T1277C were compared. Quantitative differences in the abundance of at least 34 proteins were identified, suggesting that the genetic background is likely to affect the protein composition of drusen “simply” due to the
availability of proteins in the blood. This approach also highlighted that there are proteins, whose presence and potential role in sub-RPE deposit formation and in AMD had not previously been explored. One such example is the pregnancy zone protein (Arya et al., 2018), a plasma protein whose levels are known to increase in pregnancy and some disease states such as AD (Nijholt et al., 2015).

It appears therefore that while drusen deposition is a hallmark of AMD, HAP deposition is a hallmark of drusen formation. The study by Thompson and coworkers was not the first to identify calcified components of drusen (Thompson et al., 2015). Spherical particles of similar size were previously identified within drusen and BLinD (Green and Key, 1977), and electron microscopy (Ulshafer et al., 1987; van der Schaft et al., 1992; van der Schaft et al., 1993; Thompson et al., 2015). Particles size observed in these studies ranged from 0.5 µm to 10 µm in diameter and contained calcium and phosphate as determined by elemental analysis (Ulshafer et al., 1987). More recent studies using von Kossa staining, a silver enhancement technique that identifies phosphates salts also indicated that calcium phosphates were present within deposits in the sub-RPE-BL space (Suzuki et al., 2015).

The precipitation of calcium phosphate from an aqueous solution is a complex process (Kani et al., 1983; Tas, 2000; Jang et al., 2014). At neutral pH, HAP is considered the most thermodynamically stable form of calcium phosphate. In fact, it is possible that HAP is stable enough that once the lipid or protein components of the drusen regress (Sallo et al., 2009; Toy et al., 2013; Novais et al., 2015), HAP still remains and continues to interact with its environment. This may suggest that HAP interactions, not only with the BrM and the RPE but also with the remanence of photoreceptor cells or other parts of the neurosensory retina (Bird et al., 2014) may require further investigation.

Figure 7 summarizes the model of HAP associated deposit initiation. Under normal circumstances, there is a physiological exchange of material between the RPE and the choroidal circulation (Fig.7A and A'), and this includes the exchange of lipid particles (Curcio et al., 2011). With age and disease lipid particles start accumulating in the sub-RPE space including the BrM (Fig.7B and B') (Curcio et al., 2011). In the presence of lipid droplets and homeostatic changes in calcium and phosphate availability in the sub-RPE-BL space, HAP can precipitate on the surface of the lipid droplets (Fig.7C and C'). Then, on the surface of the HAP spherules drusen proteins can accumulate (Fig.7D and D') via...
directly interacting with HAP (Arya et al., 2018). Based on fluorescence labeling of HAP in human eyes, it is appeared that HAP spherules can exist without drusen (Fig.7C'), but drusen have not been seen without HAP spherules (Fig.7D and D') (when specifically looked for) thus far (Thompson et al., 2015). Based on these observations it was proposed that HAP deposition is a seeding point for drusen formation (Thompson et al., 2015).

The next obvious question to ask is where the HAP spherules are originating from? Could they be blood or RPE derived? Do they exist as spherules only in the sub-RPE-BL space or is the material present in the surrounding tissues? Spherules have not been detected in any of the cellular or intercellular spaces although the calcium phosphate crystals had been showed in mitochondria (Carafoli, 2010). Therefore, it appears that HAP spherules are deposited in the retina exclusively in the sub-RPE-BL space. In fact, it had been shown that HAP deposition can occur in primary RPE cell models which showed that HAP deposition can be initiated by the RPE alone, although contribution of the culture medium cannot be ruled out (Pilgrim et al., 2017). Whether spherical structures can develop in a cell culture system that are co-cultured with endothelial cells and/or fed with photoreceptor outer segments will need to be investigated.

HAP mineralization in the retina clearly differs from classical mineralization in bone, but it may, or may not, share some key features with general soft tissue/elastin calcification (Figure 8). Obviously, the retina lacks extracellular matrix forming osteoblasts. Also, no relationship has been found between spherule mineralization and general HAP deposition on elastin and/or collagen. However, systemic driven HAP deposition can take place in the BrM, as reported before (Gorgels et al., 2012). Indeed, a systemic lack of inorganic PPi in the blood (Jansen et al., 2013) may be involved in local HAP deposition in BrM, facilitated by local conditions, such as oxidative stress (Mungrue et al., 2011). Interestingly, investigation of the ultrastructure and composition of vascular micro-calculcations associated with uremia showed the presence of spherical particles in the media of the kidney, with internal structures comparable to those observed in the human eye (Schlieper et al., 2010). Similarly, the loose stroma of the choroid plexus of the aging or Alzheimer's disease brain contain psammoma bodies, which are entities with distinct HAP cores and multiple concentric rings or swirls of collagen wrapped around it (Alcolado et al., 1986). More recently, similarly structured spherules were also
identified within patients with osteoporosis and in cardiovascular disease (Bertazzo et al., 2013; Shah et al., 2017). Thus, comparable mineralization mechanisms in a variety of non-osseous tissues appear to be associated with a number of different disease conditions.

Alternatively, transcriptomic data suggests that part of the elements of the physiological mineralization process are (also) present in the RPE cells (Booij et al., 2009), or at the RPE/choroid interface (Whitmore et al., 2014). This evidence suggests that the molecular machinery required for general physiological mineralization (depicted in Figure 8) could (also in part) be assembled in the outer retina. Given that the bulk of the calcium is extracellular, while phosphate is mainly localized intracellularly, the conditions that allow mineralization to happen could be present locally, in the sub-RPE-BL space. This concept is novel and has not been investigated previously but may lead to HAP-based treatment strategies and/or new early detection mechanisms.
Drusen and plaques: age-related macular degeneration and atherosclerosis

The finding that a substantial number of drusen proteins are blood-borne prompted us to re-summarize a possible relationship between initiation and propagation of drusen and atherosclerotic plaques. A possible link between these two diseases was previously suggested based on (controversial) epidemiological evidence, the involvement of similar lipoproteins in the formation of extracellular deposits in AMD and atherosclerosis and structural commonalities between the vessel wall and Bruch's membrane (Curcio et al., 2001; Sivaprasad et al., 2005).

AMD is a disease starting with (multiple macular) drusen formation. Drusen consist of (oxidatively-modified) lipids and proteins as well as minerals (Sarks et al., 1988; Green and Enger, 1993; Curcio and Millican, 1999; Crabb, 2014; Flinn et al., 2014; Handa et al., 2017; Pilgrim et al., 2017; Spaide et al., 2018). Drusen constituents most likely invoke a complement attack and sustain a continuous low-grade inflammation, which leads to serious events such as RPE cell loss, neovascularization and ultimately, central vision loss (Bird et al., 1995; de Jong, 2006).

Atherosclerosis is a disease associated with the build-up of plaques also composed of (oxidatively-modified) lipids, proteins as well as minerals in the vessel wall of arteries that, via complement attack and low-grade inflammation, can lead to serious events including heart attack, stroke or aneurysm (Simmons et al., 2016).

There are a number of clear differences between AMD and atherosclerosis, such as location of the deposition, the local metabolic physiology (Stefansson et al., 2011), involvement of other (different) genes or molecules and obviously, aspects of the pathological consequences of deposition build up (Hageman et al., 2001; Hopkins, 2013).

For example, other than in plaques in atherosclerosis, deposit (drusen) formation in AMD most likely blocks the exchange of biomolecules between the retina and the choroid. It is thought that this interferes with the "nourishment" of the sensory cells in the retina causing them to die. Over time the cells cannot be replaced leading to a loss of vision, typically within the macula, which progressively deteriorates over time contributing to the AMD pathology (Bhutto and Lutty, 2012).

Nonetheless, a relatively large number of commonalities have been found between both drusen and atherosclerotic plaque formation and their associated diseases. Available
evidence comes from clinical, epidemiological, genetic, histological and pathobiological investigations (see below).

Verhoeff and Grosmann were the first to suggest a relationship between vascular disease and AMD (Verhoeff and Grossman, 1937). Except for a few reports (Gass, 1967; Kornzweig, 1977), this observation was largely ignored for over forty years when Maltzman and Hyman (Maltzman et al., 1979; Hyman et al., 1983) pioneered a plethora of subsequent epidemiological studies on the subject (Vidaurri et al., 1984; Vingerling et al., 1995; Snow and Seddon, 1999). Some studies between atherosclerosis (and similar diseases) and AMD showed positive associations (The Eye Disease Case-Control Study, 1992; Klein et al., 1993), while others did not (Hyman et al., 1983). These controversial results, especially in the early investigations, were partly due to differences in description of the clinical phenotype, use of different end phenotypes, study design, population size, lack of suitable replication populations, and insufficient knowledge of possible confounders. Nonetheless, this issue has not been resolved up until today. A wide range of epidemiological studies have also suggested that there are certain risk factors which are common to both AMD and atherosclerosis (associated cardiovascular disease). These most consistently include environmental factors, such as age and tobacco smoking (Woodell and Rohrer, 2014). These studies indicate that, mechanistically, oxidative stress and potentially lipid metabolism may play an important role in both disorders (Serban and Dragan, 2014; Gehlbach et al., 2016; George et al., 2018; van Leeuwen et al., 2018; Wilson et al., 2018).

8.2. Histological and pathobiological similarities.
Histological and pathobiological similarities between drusen and atherosclerotic plaques and their associated diseases include similarities of lipid and mineral content and structural similarities between the BrM and the vascular wall, endothelial cell dysfunction, and proteoglycan turnover. Curcio and coworkers proposed, for the first time, a relationship between drusen and atherosclerotic plaques since both contain similar neutral lipids and both accumulate cholesterol esters (Curcio et al., 2001). This finding was confirmed by others (Chung et al., 2005; Wang et al., 2010). These and subsequent studies made clear that diseases related to this type of accumulations may be mediated by genetic variation, oxidative stress and inflammation. The accumulation
of lipids in drusen and atherosclerosis has recently been reviewed elsewhere in detail (Pikuleva and Curcio, 2014; van Leeuwen et al., 2018; Xu et al., 2018).

Sivaprasad and coworkers observed that the BrM and the vascular intima share a number of common structural modalities, and age-related changes (Sivaprasad et al., 2005). Indeed, similar to the vessel wall, and given the presence of local fenestrated choroidal capillaries, BrM acts a collagen and elastin rich physical barrier for the blood. Both the BrM and vascular intima thicken through accumulation of extracellular lipids and other debris and become less flexible with age (Chung et al., 2005; Curcio and Johnson, 2012).

Another important feature of the ECM of both the vessel wall and BrM are the presence and turnover of a variety of proteoglycans. In BrM, the ratio between several proteoglycan types, most notably heparan sulfate and chondroitin sulfate, changes dramatically during aging and the development of AMD (Barzegar-befroei et al., 2012). In atherosclerosis and AMD, (oxidatively) modified proteoglycans may bind and retain specific apolipoproteins from the circulation in, respectively the artery wall (Williams and Tabas, 1995; Tabas et al., 2007) and the BrM (Curcio et al., 2009; Al Gwairi et al., 2016) BrM. Indeed, proteoglycans may play an, as so far underestimated, role in regulating the complement response and the development of both AMD and atherosclerosis pathology (Tate et al., 1993; Toomey et al., 2018). Happonen and coworkers (2012) recently showed that small proteoglycans, such as PRELP, are regulators of the complement cascade (Happonen et al., 2012). Please note that choroidal cells produce PRELP (as suggested in the current study) and that this protein apparently accumulates in drusen. There are a few reports which have established the different patterns of distribution of large (Clark et al., 2011) and small proteoglycans (Keenan et al., 2012); the latter including biglycan, decorin, fibromodulin, lumican, mimecan, opticin, and prolargin in post-mortem eye or vascular tissue.

Both atherosclerosis and AMD patients may suffer from endothelial cell dysfunction. Accumulating evidence suggests that endothelial cell dysfunction may be the initiating step in atherosclerosis (Miteva et al., 2018). In their AMD studies, Schaumberg and coworkers provided epidemiological evidence that at least one marker for endothelial dysfunction and inflammation, sICAM-1 is linked to drusen formation and neovascularization (Schaumberg et al., 2007). Interestingly, higher levels of circulating endothelial cells (CECs), a biomarker for a diversity of systemic complications, including
vascular disorders, were found in AMD patients compared to controls (Machalinska et al., 2011). While these studies focused on common risk factors and parallel development of drusen and plaques, the possibility that atherosclerosis plays a direct role in the development of AMD cannot be ruled out. Using FA, a slow filling of the choroidal capillaries over time has been observed in AMD patients (Pauleikhoff et al., 1990). This may be due to (a combination of) thickening of the BrM, a declining function of the RPE or by decreased atherosclerosis-driven perfusion of these capillaries. Reduced capillary blood flow could directly enhance the initiation of drusen or development of AMD.

As with drusen, deposition of calcified mineral, that includes hydroxyapatite (Lee et al., 2012) is associated with the formation of atherosclerotic plaques (Doherty et al., 2003). Such mineral is readily quantifiable using radiography and even serves as a marker for atherosclerosis. It has been reported that the presence of mineral in cardiovascular soft tissue can be used to predict mortality (Okuno et al., 2007; Kestenbaum et al., 2009), and morbidity of cardiovascular disease in various forms (Arad et al., 2000; Keelan et al., 2001). The specific molecular mechanisms underlying mineral formation in such tissues remains to be fully elucidated. However, both AMD and atherosclerosis are associated with low grade inflammation in the respective affected tissues (Hansson et al., 2006; Kauppinen et al., 2016), and it has been proposed that soft tissue mineralization may be best conceptualized as a convergence of bone biology with inflammatory pathobiology (Doherty et al., 2003).

8.3. Genetics and molecular biology.

Early candidate gene association studies found an association between genetic variation in APOE in both AMD (Klaver et al., 1998; Toops et al., 2016) and atherosclerosis (Zhang et al., 2018), thereby implicating lipid metabolism and transport in both disorders. Genetic variations in apolipoproteins and complement factors showed strong associations with AMD and CVD conditions. For example, polymorphisms in the CFH and a number of other complement factor genes confer at-risk genotypes for AMD (Klein et al., 2005), whilst similar associations between complement C5 and the complement receptor 1 genes confer an increased risk of atherosclerosis (Hoke et al., 2012; de Vries et al., 2017). Of note, although the same genes may be frequently associated with both (or other) diseases, different alleles are frequently implicated in the associations found
between these disorders. A well-known example is the APOE4 allele, that increases the risk of Alzheimer’s disease, and perhaps atherosclerosis (Mahley, 2016), but is protective in age-related macular degeneration (Klaver et al., 1998). Indeed, these observations were confirmed and extended by large GWAS studies that implicated regulation of lipid metabolism, extracellular matrix remodeling and the immune system low-grade inflammation in both AMD and atherosclerosis (Fritsche et al., 2016; Schunkert et al., 2018). A recent study by the International AMD consortium explored the overlap between 34 AMD-associated loci with other complex diseases (Grassmann et al., 2017). Surprisingly, the authors found that an increased risk of AMD correlates with a reduced risk for cardiovascular disease.

A key similarity between atherosclerotic plaque and drusen formation are the molecular components involved. Both types of deposit have a significant lipid component (including cholesterol and neutral fats) and mineral content, as described above. It has also been reported that drusen contains a number of proteins that are also common to atherosclerotic deposits (Mullins et al., 2000; Klein et al., 2005; Booij et al., 2010a). To gain information as to the degree of this overlap in proteins contributing to these pathologies we compared a dataset of 3196 proteins known to be present in atherosclerotic plaques from Bleijerveld and coworkers (Table 3) with our drusen data set, as shown in Figure 9a (Bleijerveld et al., 2013). The resultant Venn-diagram revealed that out the 89 drusen-associated proteins, 64 of these (72%) were also present in atherosclerotic plaques. Indeed, 50 out of 60 drusen proteins derived from blood are also present in atherosclerotic plaques (Figure 9b). Details of proteins found to be common to both plaques and drusen can be found in STable 11 and STable 11a.

Closer inspection of proteins common to both atherosclerotic plaques and drusen as defined in this manuscript revealed a number of functional classes of protein in this group including apolipoproteins (APOA1, A2 and E), complement factors (C7, C8A, C8B, C8G and complement factor H), as well as lipid- and Ca\(^{2+}\)-binding annexins (annexins-1, 2, 5 and 6). Obviously, our analysis may not be fully comprehensive, since it is limited to the entries which are present in both database listings. Another limitation is that similar proteins still may originate from different sources. For example, the previously suggested presence of APOB as principal protein of LDL in both sub-RPE-BL deposits and cardiovascular plaques (Curcio et al., 2001) is missing from the current overlap,
since detailed investigation of the (presence and origin) of this lipoprotein (Li et al., 2005a) suggested that APOB isolated from BrM thickenings is (also) present in a distinct, non-LDL lipid profile. Consequently, it was suggested that APOB in BrM thickenings is made locally, while APOB in plaques is probably from systemic origin. Cytoskeletal proteins (actinin α1, tubulin α1c and tubulin β3) as well as extracellular matrix proteins such as collagens (type 1 α2, type 6 α1, type 6 α2 and type 8 α1), tenascin C, microfibril-associated protein 4 and vimentin were found to be present in both BlamD deposits and atherosclerotic plaques (Fernandez-Godino et al., 2016; Pelisek et al., 2016). Analysis of proteins common to both plaques and drusen in biological processes revealed significant contribution of this group of proteins to other diseases and processes including various cancers, development of the vasculature, cell movement and AD (tauopathy and amyloidosis; see STable 11a). The involvement of these proteins in AD is particularly interesting as it is another disorder of which extracellular deposits are a feature (Figure 9c). Furthermore, drusen reside on the interface between the neural and cardiovascular system, so it may share properties of both types of atherosclerosis and Alzheimer’s plaques (Booij et al., 2010a).

Our review of the literature and the qualitative and quantitative meta-analysis of retinal and blood transcriptomic and proteomic data all point in the same direction: proteins in drusen originate from multiple sources. Based on the data we have available, the largest number of protein contribution from a single source appears to be the blood. The second-most prominent source of number of specific proteins in drusen is from the RPE, while the contribution from the choroid and the photoreceptors appears to be relatively modest. However, the varying number of proteins cannot be directly translated to concentration. There is the possibility that a relatively small number of proteins contribute the bulk of proteins in drusen.

How proteins get recruited to and retained in the sub-RPE-BL space is still not fully understood. In vivo and in vitro BrM conductance studies suggest that human proteins of average size, such as proteins of 53 kDa (source: NCBI), can readily diffuse through healthy BrM, while macromolecular migration through BrM is slower and/or limited (Curcio and Johnson, 2012). Moore and Clover found that proteins of 200 kDa could readily cross young BrM (Moore and Clover, 2001). More recent work suggested that the transport exclusion size limit in healthy young BrM can be as high as 180-500 kDa, well over the size of macro molecules like HDL (Hussain et al., 2010; Cankova et al., 2011). In our current study, we found that 9 out of 10 drusen proteins that are uniquely derived from the RPE had a Mw of less than 50 kDa (Table 6). Moreover, we took a random sample of 30 proteins from the RPE-IVS basal secretion dataset (Table 3), containing proteins which are likely to encounter BrM in vivo, and determined their average Mw: 95 kDa. After taking three extremely large proteins out (APOB, AHNAK, and C4B; proteins that we did not identify in drusen in this study) that average dropped to 60 kDa (data not shown). Six of these 30 proteins are present in drusen (ALB, ANXA1, ANXA2, APOA4, APOE, ATP5F1B) and have a MW <66.5 kDa. Although the overall transport capability BrM decreases substantially in the AMD-affected and aging retina (Hussain et al., 2010; Cankova et al., 2011; Curcio and Johnson, 2012; Lee et al., 2015), older BrM was found to be still permeable to proteins in excess of 100 kDa (Moore and Clover, 2001). Taken together, entrapment of proteins in the sub-RPE-BL space is unlikely to be due to size if single molecules. They might become entrapped by forming aggregates that are no longer capable of leaving through BrM, as was suggested for CFH (Nan et al.,...
drusen proteins, especially the ones that come from multiple sources, "meet, greet and stick" to form sub-RPE-BL space deposits.

There are several ways proteins can interact in BrM to form larger aggregates: they can interact among themselves, with other lipids, proteins and/or mineral deposits, or stick to the ECM of BrM itself. These interactions may be enhanced by chemical modification, including oxidative damage and glycosylation of lipids, proteins and carbohydrates (Blaum et al., 2010; Crabb et al., 2002; Hollyfield et al., 2010) and they may be further facilitated by the structure and dynamic nature of BrM (Booij et al., 2010a). Over time, several changes in BrM occur, that may hinder protein clearance from the sub-RPE-BL space. Remodeling of BrM ECM takes place, including proteoglycan changes and turnover, elastin changes and eventually mineralization takes place. BrM becomes laden with lipids to form a hydrophobic barrier ("lipid wall") and accumulates other debris (Curcio and Johnson, 2012). Consequently, the role of the structure and function of BrM and the chemical state of the sub-RPE-BL space may be even more important in sub-RPE-BL space deposit formation than its exact protein composition.

An important source of entrapment of proteins in BrM may be the formation of HAP surfaces in the sub-RPE-BL space (Thompson et al., 2015). In our current study, at least 30% of the 89 drusen proteins can bind to HAP. This percentage increases towards at least 50% if only the blood borne proteins are counted (data not shown). Thus, HAP readily binds a substantial number, but not all drusen proteins (Arya et al., 2018). The finding that blood proteins are seemingly the most important contributors to drusen formation provides a new target to prevent the initiation and propagation of sub-RPE-BL space deposits. Reducing the concentration of blood proteins that interact with HAP may lead to a reduction of the source of drusen components and ultimately postpone, or potentially even stop, the progression to AMD.

Finally, it is also important to mention that the non-specific interaction of proteins with HAP will also affect their ability to carry out their physiological function. For example, once CFH binds to the HAP surface it may not be able to regulate the alternative complement pathway. Therefore, this interaction with HAP could be a double whammy: it increases the bulk of sub-RPE-BL space deposits and stops the local protein function. It will be important to understand the role of the blood-derived proteins in the sub-RPE-
The study and potential modification of these interactions is now possible and could lead to intervention strategies through modified diet, supplementation or through manipulation of retinal molecular or cellular processes.

An important specific question in the context of this study that needs still to be resolved is how plasma proteins find their way into the sub-RPE-BL space. Apart from the mechanisms already described above (chemical modification of interacting biomolecules, dynamic structure and functional changes BrM, and HAP-binding) it is tempting to speculate that not only blood composition but also blood pressure plays a role. Why blood pressure? It was previously shown that a relationship exists between drusen location and choriocapillary pillars. Indeed, by investigating retina whole mounts, initially Friedman, and subsequently, Lengyel and coworkers concluded that drusen deposition is the result of a lower clearance at the choroidal intercapillary pillars (Friedman et al., 1963; Lengyel et al., 2004). Thus, in other words, higher clearance of sub-RPE-BL space debris corresponds with the vascular lumen, through which the blood flows and directly encounters BrM. Much in line with the reflections of Penfold and others (Penfold et al., 2001), we hypothesize that the pulsating blood pushes debris through endothelial fenestrations into BrM, through relatively open BrM pores; and at the same time, clears debris which was already present in BrM. One could compare that, by analogy, with the sea bringing and taking, wave after wave, debris to and from the beach. Changes in blood composition, choroidal endothelial cell compromise and rising blood pressure with age (Pinto, 2007) may negatively change the dynamics of this proposed “debris-exchange”.

Our review further underlines the importance of comparative studies between drusen deposition and atherosclerosis plaque formation. Clinical, (genetic) epidemiological pathobiological and molecular similarities between these two disorders have been highlighted previously (see section 8). Such similarities include that both are extracellular lipid/protein/mineral-based depositions that invoke a low-grade immune response leading to further disease. Several molecular similarities between drusen and plaques have also been described. We currently add the observation that most drusen (and plaque) components are blood-borne. Therefore, the genesis of drusen and plaques may be similar, and should be subject of further multidisciplinary studies.
While studying the literature for this review, we have made a number of additional observations that may guide future research directions: First, while the number of retinal (cell-type-specific) transcriptomics studies are large and proteomics information is emerging, there are very few proteomics studies on different types of (human) sub-RPE-BL space deposits (types). For example, additional proteomics studies of hard versus soft drusen or macular versus peripheral drusen might improve our understanding of deposition formation in the sub-RPE-BL space and their association with different disorders or disease stages. Next, transcriptomics, proteomics, and immunohistochemical studies have their own conceptual and technical advantages and limitations. However, in the literature, the description of these strengths and weaknesses are not always clear and standardization is lacking. International agreements such as MIAME and MISFISHIE (minimum information specification for in situ hybridization and immunohistochemistry experiments) guidelines (Deutsch et al., 2008) are a step in the right direction, but must be seen as initial steps for further standardization. A few examples for illustration: How many confirmatory transcriptomics or proteomics studies should be performed before a definite subcellular assignment can be made? How do we define cellular specificity and cellular enrichment? How many drusen types should be screened and how many different antibodies should be used before proteins are clearly assigned as drusen proteins (or as sub-types). When should we designate labeling drusen specific? Do we consider staining of the border of hydroxyapatite or drusen important; or is only the staining of the whole inner mass of drusen relevant? Given the heterogeneity of drusen: what is the exact location of the drusen under study and its appearance? Indeed, in line with recent similar calls by Curcio and co-workers (Curcio, 2018a, b) we call here for better considerations, agreements and definitions of these issues.

Last but not least, it will be interesting to understand whether drusen heterogeneity is a direct feature of a disease or a reflection of the change in the (micro-) environment that results in initiation and growth of the deposits. While drusen deposition clearly is a hallmark of AMD and is associated with a number of other diseases (Khan et al., 2016), its actual composition might reflect the disease state at the RPE/choroid interface more than (cause) the disease. The identification of why and not necessarily what proteins
and lipids are deposited in the sub-RPE-BL space might therefore be an important question to consider for future studies.
References


465-481.

60. Curcio, C.A., Johnson, M., Huang, J.D., Rudolf, M., 2009. Aging, age-related macular degeneration,
and the response-to-retention of apolipoprotein B-containing lipoproteins. Prog Retin Eye Res 28,
393-422.

The British journal of ophthalmology 95, 1638-1645.


Subretinal drusenoid deposits in non-neovascular age-related macular degeneration: morphology,
prevalence, topography, and biogenesis model. Retina 33, 265-276.

64. Curcio, C.A., Millican, C.L., 1999. Basal linear deposit and large drusen are specific for early age-

human Bruch's membrane. Investigative ophthalmology & visual science 42, 265-274.

Pigment Epithelium, an Optical Coherence Tomography Biomarker for Progression in Age-Related


pigment epithelium cell changes and expression of alphaB-crystallin: a biomarker for retinal pigment

69. de Vries, M.A., Trompet, S., Mooijaart, S.P., Smit, R.A., Bohringer, S., Castro Cabezas, M., Jukema,
Atherosclerosis 257, 16-21.

degeneration: reengineering Bruch's membrane and the human macula. Prog Retin Eye Res 25, 539-
562.

S.S., Vingerling, J.R., Hofman, A., Oostra, B.A., Uitterlinden, A.G., Stijnen, T., van Duijn, C.M., de Jong,
P.T., 2006. Complement factor H polymorphism, complement activators, and risk of age-related
macular degeneration. Jama 296, 301-309.

72. Deutsch, E.W., Ball, C.A., Berman, J.J., Bova, G.S., Brazma, A., Bumgarner, R.E., Campbell, D.,
Causton, H.C., Christiansen, J.H., Daian, F., Dauga, D., Davidson, D.R., Gimenez, G., Goo, Y.A.,
Grimmmond, S., Henrich, T., Herrmann, B.G., Johnson, M.H., Korb, M., Mills, J.C., Oudes, A.J.,
Parkinson, H.E., Pascal, L.E., Pollet, N., Quackenbush, J., Ramialison, M., Ringwald, M., Salgado, D.,
in situ hybridization and immunohistochemistry experiments (MISFISHIE). Nat Biotechnol 26, 305-
312.


(Lond) 31, 87-96.

inflammation at the arterial crossroads. Proc Natl Acad Sci U S A 100, 11201-11206.

E.Y., 2017. Peripheral Retinal Changes Associated with Age-Related Macular Degeneration in the Age-
Related Eye Disease Study 2: Age-Related Eye Disease Study 2 Report Number 12 by the Age-Related


Lake, S., Mackey, D.A., Constable, I.J., Craig, J.E., Kitchner, T.E., Yang, Z., Su, Z., Luo, H., Chen, D.,
Rudolph, G., Olden, M., Morrison, M.A., Morgan, D.J., Schu, M., Ahn, J., Silvestri, G., Tsironi, E.E.,
P., Katsanis, N., Holz, F.G., Blond, F., Blanche, H., Deleuze, J.F., Igo, R.P., Jr., Truitt, B., Peachey, N.S.,
Meuer, S.M., Myers, C.E., Moore, E.L., Klein, R., Hauser, M.A., Postel, E.A., Courtenay, M.D.,
Schwartz, S.G., Kovach, J.L., Scott, W.K., Liew, G., Tan, A.G., Opitz, C., Smith, R.T.,
Khan, J.C., Shahid, H., Moore, A.T., McGrath, J.A., Mahtani, K., Agarwal, A., Ersoy, L.,
Caramoy, A., Langmann, T., Sachsens, N.T., de Jong, E.K., Hoyng, C.B., Cain, M.S., Richardson, A.J.,
Martin, T.M., Blangero, J., Weeks, D.E., Dhillon, B., van Duijn, C.M., Doheny, K.F., Romm, J., Klaver,
C.C., Hayward, C., Gorin, M.B., Klein, M.L., Baird, P.N., den Hollander, A.I., Fauser, S., Yates, J.R.,
Leveillard, T., Zhang, K., Brilliant, M.H., Hewitt, A.W., Swaroop, A., Chew, E.Y., Pericak-Vance, M.A.,
A large genome-wide association study of age-related macular degeneration highlights contributions
of rare and common variants. Nat Genet 48, 134-143.
Protein Pept Sci 19, 612-622.
63, Suppl:1-139.
Database Syst Rev, Cd006927.
hypothesis for treating inflammation and oxidative stress with hydrogen sulfide during age-related
activation is essential for docosahexaenoic acid protection of retina photoreceptors. J Lipid Res 54,
2236-2246.
Grainger, R.M., 1993. Interphotoreceptor retinoid-binding protein (IRBP), a major 124 kDa
glycoprotein in the interphotoreceptor matrix of Xenopus laevis. Characterization, molecular cloning
Gorgels, T.G., Teeling, P., Meeldijk, J.D., Nillesen, S.T., van der Wal, A.C., van Kuppevelt, T.H.,
Bergen, A.A., 2012. Abcc6 deficiency in the mouse leads to calcification of collagen fibers in Bruch's
Grassmann, F., Kiel, C., Zimmermann, M.E., Gorski, M., Grassmann, V., Stark, K., Heid, I.M.,
diseases and traits. Genome Med 9, 29.
Green, W.R., Enger, C., 1993. Age-related macular degeneration histopathologic studies. The
25, 180-250; discussion 250-184.
for better or for worse, in sickness and in health. Cell Mol Life Sci 71, 1551-1579.
metalloproteinase activity of isolated human Bruch's membrane and choroid. Investigative
ophthalmology & visual science 40, 2676-2682.


differential distribution of proteoglycan core proteins in the adult human retina, choroid, and sclera.
Investigative ophthalmology & visual science 53, 7528-7538.
Incidence and progression of coronary calcification in chronic kidney disease: the Multi-Ethnic Study
2015 150. Khan, K.N., Mahroo, O.A., Khan, R.S., Mohamed, M.D., McKibbin, M., Bird, A., Michaelides, M.,
with ageing, age-related macular degeneration, inherited eye disease and other pathological
2016 151. Kiel, C., Vogt, A., Campagna, A., Chatr-aryamontri, A., Swiatek-de Lange, M., Beer, M., Bolz, S.,
complement factor H by RPE and its role in RPE migration. Graefes Arch Clin Exp Ophthalmol 247,
651-659.
of retinal pigment epithelium damage and development of age-related macular degeneration. Acta
Ophthalmol 90, 299-309.
C., de Jong, P.T., 1998. Genetic association of apolipoprotein E with age-related macular
cynomolgus monkey (Macaca fascicularis) pedigrees with early onset drusen formation. Exp Anim 63,
305-310.
2024 159. Krujit, C.C., de Wit, G.C., Bergen, A.A., Florijn, R.J., Schalij-Delfos, N.E., van Genderen, M.M.,
reticulum stress and lipid accumulation in retinal pigment epithelium through oxidative stress and
Tyrosinase and Tyrosinase-Related Proteins. Chemistry 24, 47-55.
tissues. Atherosclerosis 224, 340-347.
Characteristics of Bruch’s Membrane With Steroidal Glycosides and its Relevance to Age-Related
Macular Degeneration (AMD). Investigative ophthalmology & visual science 56, 8403-8418.
Based Ultra-Widefield Digital Image Grading Study for Age-Related Macular Degeneration-Like
2030 165. Lengyel, I., Flinn, J.M., Peto, T., Linkous, D.H., Cano, K., Bird, A.C., Lanzione, A., Frederickson,
Eye Res 84, 772-780.


Legends to Figures and (S)Tables.

Figure legends:

Figure 1. Heterogeneity of drusen. Imaging of drusen and drusen content with various clinical and laboratory methods. On color fundus images, the yellow spots identify drusen (A). On OCT image the elongated RPE is reflective. Drusen appear as homogeneous and hyper-reflective sub-RPE-BL space focal entities. (B); scale bar: 500 µm. On hematoxylin-eosin staining, drusen appears between the brown pigments of the RPE and the Bruch’s membrane (C); scale bar: 10 µm. Note the inclusions without staining. Drusen contain numerous von Kossa positive spherule structures identifying hydroxyapatite spherules (D); scale bar: 10 µm. Auto-fluorescence of the more or less circular drusen is indicative of the protein and lipid accumulation. Auto-fluorescence of the RPE is more intense and yellowish. Auto-fluorescence of the BrM adjacent to the RPE is greenish (E); scale bar: 10 µm.

Figure 2 a-d. Functional molecular network analysis (4 networks). For all four networks: Molecular network analysis of physical or functional interactions between 89 drusen genes/proteins using Ingenuity. The knowledge database generated four most likely functional networks from the given input (drusen proteins). The functionalities in the figures are generated based on a combination of available molecular and cellular experimental data in human, mouse, rat, and in vitro data. In each network the circles, squares and other symbols represent proteins from a homologue from either human, rat or mouse. The systematic name of the gene/protein is printed on each symbol. Double circles in a single symbol denotes a group or /family of entries with a specific function and are sometimes introduced by the knowledge database to make networks possible or to simplify them. Solid lines represent strong physical or functional interaction between the entries, taken from published peer reviewed literature and/or transcriptomics and proteomics databases. Dotted lines represent weaker, associated relationships between the genes/proteins based on published peer reviewed experimental data (for example co-upregulation of expression in an in vitro experiment). The lines represent thus functionalities found in either, both, or all experiments on human, rat or mice (tissues) and in vitro findings. For example, the functional relationship between molecules A, B, C, could possibly be defined as follows A-functional mouse finding-B-functional human
finding-C. The underlying hypothesis is that the functionalities in human, mouse and rat are very similar. One can also generate networks of human (or rat or mice) functional data separately, but they are frequently quite similar, but less extensive. In the first network (Figure 2a), clearly three functional clusters of closely related entries can be recognized: The complement cluster, the collagen cluster and the crystallin cluster. The second network (Figure 2b) is much more complicated and heterogeneous and is a network of genes and proteins related to development and genetic or ophthalmic disorders. The common theme of the third network (Figure 2c) is the immune response. Finally, the fourth and last network (Figure 2d) presents functional and structural relationships between entries involved in cell-cell interactions and systemic involvement.

Figure 3. Schematic overview of various strategies used for dataset curation. This figure shows several ways to curate (pre-existing) transcriptomics or proteomics datasets to form an improved, thorough or more specific dataset. For example, in (A) several datasets are merged into a cumulative new one by simply combining the datasets. The possible overlap is counted only once in the new merged database. One dataset was deleted, because it did not adhere to quality standard or had a different signature as the other ones. Strategy (B) has been published before (Booij et al., 2010b). In this case the original enriched RPE database contains 10% of the highest expressed genes in the cell. Some of the expressed genes in the RPE10% dataset overlap with the genes expressed in the adjacent tissues (photoreceptors and/or choroid). These “overlapping expressed genes” are therefore not specific of the RPE. Thus, to obtain a more specific (smaller) dataset, we discard of all the “overlapping expressed genes” in the RPE dataset, to obtain a highly enriched RPE dataset. Curation strategy C shows the breakdown of two datasets into desired subfamilies: The overlap between datasets X and Z is Y. Dataset Y can be used if overlap between X and Z is desired. Dataset X minus Y can be used to obtain unique entries from X (compared to set Z).

Figure 4. Scheme of the relationships between the respective transcriptomic and proteomic datasets used for quantitative studies. The curated drusen protein dataset represents 89 proteins known to be present in drusen/sub-RPE-BL space deposits (black box in the middle). These were compared with the entries present or produced
uniquely from the neural side of drusen (photoreceptor/RPE neural source), and with
entries uniquely from the systemic side of drusen (blood /choroid basal source).
The “neural source” and “systemic source” merged data-sets each consist of non-curated
datasets and curated datasets. The non-curated (pure) dataset contain, by virtue of their
nature or previous curation in the literature) only entries from, respectively, the neural
(RPE-ST; RPE-IVS) and systemic side (BL-SP1; BL-SP2) of drusen. The curated datasets
(cPRos; cPR-ET and cChor-ET)) contained, before curation, a number of entries
expressed/present in both the neural and systemic side (see PRos; PR-ET and Chor-ET).
Thus, we removed all overlapping expressed genes between the PRos; PR-ET and Chor-
ET datasets, to obtain unique datasets from both sides of drusen.

Figure 5. Potential contribution of neural and systemically expressed/present proteins
to drusen formation. Venn diagram showing overlap between (A) neural RPE and
photoreceptor-derived proteins, (B) systemically derived choroid and blood proteins
and (C) drusen-associated proteins.

Figure 6. Hydroxyapatite spherules can retain proteins originating from blood in the
sub-RPE-BL space. (A) Immunocytochemical labelling of histidine-rich glycoprotein
(HRG) using a specific anti-HRG primary antibody (green) on the surface of a HAP
spherule labelled by LiCor680 (magenta); scale bar: 10 µm. (B) Binding of purified
human HRG to HAP-coated magnetic beads. Binding assays were performed using 0.3
mg beads per sample. HAP-beads were washed with 50 mM Tris, 140 mM NaCl, pH 7.4
and incubated with 400 µl of 0-1µM human HRG for an hour at room temperature. The
protein-bound beads were washed with the same buffer twice followed by blocking with
1% BSA for an hour. Rabbit anti-human HRG antibody (1:1000 dilution) and HRP-
conjugated anti-rabbit antibody (1:10000 dilution) were respectively used as primary
and secondary antibodies. Detection was done at 492 nm using o-phenylenediamine
dihydrochloride (OPD, Sigma Aldrich) substrate.

Figure 7. A model for drusen formation. Top row (A-E) is adopted from the schematic
diagram proposed for sub-RPE-BL space deposit formation by Thompson and colleagues
(Thompson et al., 2015). (A) Healthy eyes show no sub-RPE-BL space deposit formation.
(B) At Stage 1 lipid droplets are retained in the sub-RPE-BL space (black dot). (C) At
Stage 2 mineralization occurs surrounding the lipid droplets (magenta ring). (D) At Stage 3 proteins bind to the HAP surfaces (blue ring). (E) At Stage 4 proteins and lipids start accumulating around the “seed” (yellow material). The bottom row (A’-E’) shows morphological evidence for the prediction in the top row. (A’) Retinal pigment epithelium forms a monolayer along the inner collagenous layer of the Bruch’s membrane in healthy eyes (scanning electron microscopic image); scale bar: 10 µm. (B’) Transmission electron micrograph of lipid droplets that accumulate in the sub-RPE-BL space; reproduced with permission from Curcio and Millican (Curcio and Millican, 1999); scale bar: 2 µm. (C’) Scanning electron microscopic identification of a single spherule located between the RPE basement membrane and the inner collagenous layer of Bruch’s membrane; scale bar: 2 µm. (D’) Immunofluorescent labelling of HRG (green) on the surface of a HAP spherule (magenta); scale bar: 2 µm. (E’) An immunofluorescent labelling of complement factor H on a spherule surrounded by the autofluorescence of drusen (green) and RPE cells (yellow) (blue is DAPI staining the cell nuclei); scale bar: 10 µm.

Figure 8. Schematic showing factors that are identified to contribute to mineralization of soft tissues and may contribute to HAP deposition in the sub-RPE-BL space.

Figure 9. Proteins present in atherosclerotic plaques and drusen. A. Venn diagram showing 64 out of 89 drusen proteins overlap with the atherosclerotic plaque proteome, while 25 entries are unique to drusen in this comparison. B. Venn diagram showing 50 out of 60 proteins (from the 89 drusen proteins) that come from blood (as unique source or shared with the PR/RPE) are actually present in atherosclerotic plaques. C.
Venn diagram displaying the uniqueness and overlap of proteins between drusen (C.A), Alzheimer plaque proteins (C.B.) and atherosclerotic plaque proteins (C.C). The corresponding Table 1 and Table 1a present the corresponding entries in detail.

Table Legends:

Table 1. List of proteins present in the curated drusen dataset. We assembled a list of 89 drusen proteins, mostly derived from the macular area, from the literature. For each entry the Gene symbol, Entrez gene name, location and type, human immunohistochemistry source and literature references are provided based on information found via the Ingenuity knowledge database (Qiagen, all rights reserved), relevant literature (PubMed searches) and other public databases, such as Genecard (www.genecard.org) and DAVID (https://david.ncifcrf.gov/).

Crabb 2002 (Crabb et al., 2002); Wang 2002 (Wang et al., 2010); Entries with *: although assigned to drusen by proteomics, IHC studies suggest a more likely protein location around or directly external from drusen. Further detailed investigation is warranted for these entries. **First detected in cynomolgus monkeys, afterwards in human drusen.

Table 2. Summary of Ingenuity knowledge database core analysis of 89 proteins present in the curated drusen protein dataset. Summary of enriched motifs present in the dataset presented as top disease and biological functions, canonical pathways and discrete molecular networks. Note that these functional annotations types relate to either cellular (LRX/RXR/FXR activation; macrophages) or systemic (acute phase, atherosclerosis) entities. In the top disease and biological functions, we see that the dataset is enriched for hereditary disorders, ophthalmic disease, injury, metabolic diseases and developmental disorders. Finally, in the top functional or structural molecular networks, we find combinations of very basic functions (cancer and cellular) to more specific pathobiological ones (ophthalmic and neurological disease etc.

Table 3. Summary of datasets used in this study and their respective functional clusters. Table displaying various datasets used in this study, along with their characteristics. In the first column, the result of the Ingenuity network analysis of drusen proteins is given in 4 significant molecular networks (N=1-4) corresponding to the networks shown in
Figure 2a-2d. Within these networks, six functional molecule clusters can be observed. For example, Network 1 (N1) contains 3 functional clusters: the complement (Network 1. cluster 1), the collagen (1.2) and the crystallines (heatshock) (1.3). Network 2 consist of 1 large cluster (2.4) being genetic and developmental ophthalmic disorders. Network 3 can be viewed as a cluster (# 5) of injury and inflammatory response and dermatological disease. Network 4 (N4) contains a cluster (4.6) of cell-to cell-signaling and systemic involvement. Column B gives the actual gene/protein names in these clusters. Column C states the overall functional annotation of these clusters. The first row of the Table from column D onward states the compartment of the datasets to be compared with drusen proteins in the functional clusters (within brackets, the number of entire in each dataset are given). In row 2 (acronym) from column D onward, the short and systematic acronym of each dataset is given. Row 3 (reference) contains from column D onward, the reference where the dataset can be found. Row 4 (methodology), from column D onward, contains the method by which the data were generated (transcriptomics, proteomics). Row five (source), from column D onward, contains the primary author who submitted the data or who can be contacted to obtain further information. The remaining boxes contains information which entries of the functional cluster are present both in drusen as well as in the transcriptomics or proteomics dataset(s). Combined analysis of the clusters in different datasets gives a qualitative idea from which cell type(s) drusen protein are derived.

Table 4. Summary of Ingenuity knowledge database core analysis of the curated photoreceptor gene expression (cPR-ET) dataset. Functional annotation of the curated and highly enriched photoreceptor cPR-ET database using the ingenuity knowledge database. The data driven top canonical pathways are highly relevant for photoreceptor function: Phototransduction pathway, glutamate receptor signaling, cholesterol biosynthesis and Wnt/Ca2+ signaling. The only surprise in our data-driven analysis could be the Huntinton disease signaling pathway. However, it has recently become clear that in Huntington's disease (HD), an inherited neurodegenerative disorder resulting in motor disturbances, cognitive and behavioral changes, deficits in retinal and visual processing function are significantly present (Coppen et al., 2018). Although we curated the PR database quite extensively, and thus selected for specific photoreceptor molecular signature and function, it is interesting to see that these motifs occur also in a
number of other (top) diseases and functions, such as cancer, organismal injury, gastrointestinal disease, Hepatic disease and reproductive system disease. This may reflect the accumulating evidence that a substantial number of genetic or metabolic disease are also affect photoreceptor function. Similar to the canonical pathways and the biological motifs, the functional annotation of the photoreceptor selected molecular machinery apparently reflects a broad spectrum of biological and disease processes.

Table 5. Summary of Ingenuity knowledge database core analysis of the curated choroid, cChor-ET datasets. In this Table, we present the summary of the functional annotation of the choroid. Of course, the choroid is not a single tissue, but contains multiple cell types (endothelial cells, fibroblasts, macrophages, etc.) and the sample is inevitably contaminated with the blood. Within these limitations, data driven analysis of this specifically curated data set yielded a number of interesting enriched motifs, which indeed can be contributed to the choroid or blood: The canonical pathways indicate enriched immunological themes, such as the complement system, acute phase response signaling, and antigen-presenting cells, which is confirmed by several biological motifs (inflammatory disease and response, injury). Further, the canonical pathways generated, suggest an overlap between the molecular machinery of the choroid and atherosclerosis signaling. Indeed, in this manuscript, we devoted a whole section (8) to the pathobiological and molecular similarities between drusen and atherosclerotic plaques, and their –in time-associated diseases: AMD and atherosclerosis. Finally, a homology between hepatic function and choroid was observed. Indeed, there are a number of reports in the literature of cross-talk between liver and choroidal function, but that potential relationship remains to be elucidated. The final biological motifs are cancer and connective tissue disorders. Cancer, is of course very broad and frequently relates to blood vessel metabolism or (abnormal) cell division, while the connective tissue motif may relate to the action of local fibroblasts. The choroidal networks, show, again a very broad spectrum of molecular interactions, but this spectrum is quite distinct from the functional annotation of the photoreceptor networks presented in Table 5.

Table 6. Drusen proteins expressed or present in the PR/RPE and their characteristics. Overview and characteristics of ten drusen proteins, which most likely originate from
the neural side of drusen (namely PR and Chor). In the first column (A), general used
abbreviations (according Gen bank) for gene/protein names are given. In column B, C, D
respectively systematic Entrez number, cellular location and protein type corresponding
to these proteins are presented. Column E and F contain the amino acid (aa) size and
Molecular weight (Mw) of the proteins. Further the isoelectric point (pl; column G), the
number of negative and positive charged aa residues (column H), the protein instability
Index number (column I); the Alipathic index for solubility (J), and the GRAVY
(hydrophobicity and hydrophilicity index). These are all standard characteristic of
proteins which can be found in the Ingenuity database (Qiagen all right reserved) and
public databases such as DAVID, (https://david.ncifcrf.gov), SWISS-prot
(https://www.ebi.ac.uk/uniprot), Genecards (www.genecard.org) and/or the data
shows that these entries apparently do not have specific characteristics, except perhaps
for their ability to interact with one another, that could explain why they would get
stuck in BrM as a drusen protein. We conclude that, if it is not the proteins that explain
this, it must be the structure of BrM.

Supplementary Table Legends:

Table S1. List of 276 proteins present in the RPE-IVS dataset. For each entry the gene
symbol, Entrez gene name, location and type are provided based on information found
in the Ingenuity knowledge database.

Table S2. List of 170 proteins present in the RPE-ST dataset. For each entry the gene
symbol, Entrez gene name, location and type are provided based on information found
in the Ingenuity knowledge database.

Table S3. List of 412 proteins present in the Pros-EP dataset. For each entry the gene
symbol, Entrez gene name, location and type are provided based on information found
in the Ingenuity knowledge database.

Table S4. List of 995 proteins present in the BLP-SP1 dataset. For each entry the gene
symbol, Entrez gene name, location and type are provided based on information found
in the Ingenuity knowledge database.
Table S5. List of 262 HAP binding proteins in the BL-PHP blood proteome dataset. For each entry the gene symbol, Entrez gene name, location and type are provided based on information found in the Ingenuity knowledge database.

Table S6. List of 754 expressed genes present in the cPR-ET dataset. For each entry the gene symbol, Entrez gene IDs for human and mouse are provided.

Table S7. List of 848 expressed genes present in the cChor-ET dataset. For each entry the gene symbol, Entrez gene IDs for human and mouse are provided.

Table S8. Annotation of 37 drusen proteins (out of 89) that may uniquely originate from the blood. For each entry, the gene symbol. Entrez gene IDs for human and mouse are presented.

Table S8a Functional annotation of 37 drusen proteins that may originate from the blood. Combinations of genes/proteins in this group makes up specific functional categories associated with biological function or disease.

Table S9. Annotation of 23 drusen proteins that may originate either from the neural or from the systemic side, using Ingenuity. For each entry its functional category, specific associated disease or function, p-value, gene names of associated proteins and number of proteins in each category are provided.

Table S9a Functional annotation of 23 drusen proteins that may originate from either the neural or the systemic side of drusen using Ingenuity. Combinations of genes/proteins in this group makes up specific functional categories associated with biological function or disease.

Table S10 Annotation of 19 drusen proteins of unclear origin. Entrez gene IDs for human and mouse are presented.
Table S10a Functional annotation of 19 drusen proteins of unclear origin using Ingenuity. For each entry its functional category, specific associated disease or function, p-value, gene names of associated proteins and number of proteins in each category are provided.

Table S11 List of 64 proteins common to both drusen and atherosclerotic plaques. For each entry, the gene symbol and Entrez Gene IDs for human and mouse and are provided.

Table S11a Functional annotation of 64 proteins common to both drusen and atherosclerotic plaques. For each entry its functional category, specific associated disease or function, p-value, gene names of associated proteins and number of proteins in each category are provided.

Next page: Table 1: 89 drusen proteins used in this study.
<table>
<thead>
<tr>
<th>ID/Symbol</th>
<th>Entrez Gene Name</th>
<th>Location</th>
<th>Type(s)</th>
<th>Source</th>
<th>Human IHC ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTB</td>
<td>Actin beta</td>
<td>Cytoplasm</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>ACTN1</td>
<td>Actinin alpha 1</td>
<td>Cytoplasm</td>
<td>transcription regulator</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>ALB</td>
<td>Albumin</td>
<td>Extracellular Space</td>
<td>transporter</td>
<td>(Crabb et al., 2002)</td>
<td>(Hollyfield et al., 2003)</td>
</tr>
<tr>
<td>ALDH1A1</td>
<td>Aldehyde dehydrogenase family, A1</td>
<td>Cytoplasm</td>
<td>enzyme</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>AMBP</td>
<td>Alpha-1-microglobulin/bikuninulin 5 drusenin precursor</td>
<td>Extracellular Space</td>
<td>transporter</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>ANXA1</td>
<td>Annexin A1</td>
<td>Plasma Membrane</td>
<td>enzyme</td>
<td>(Crabb et al., 2002)</td>
<td>(Rayborn et al., 2006)</td>
</tr>
<tr>
<td>ANXA2</td>
<td>Annexin A2*</td>
<td>Plasma Membrane</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td>(Crabb et al., 2002)*</td>
</tr>
<tr>
<td>ANXA5</td>
<td>Annexin A5</td>
<td>Plasma Membrane</td>
<td>transporter</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>ANXA6</td>
<td>Annexin A6</td>
<td>Plasma Membrane</td>
<td>ion channel</td>
<td>(Crabb et al., 2002)</td>
<td>(Crabb et al., 2002); (Rayborn et al., 2006)</td>
</tr>
<tr>
<td>APCS</td>
<td>Amyloid P component, serum</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td>(Mullins et al., 2000)</td>
</tr>
<tr>
<td>APOA1</td>
<td>Apolipoprotein A1</td>
<td>Extracellular Space</td>
<td>transporter</td>
<td>(Crabb et al., 2002)</td>
<td>(Mullins et al., 2000)</td>
</tr>
<tr>
<td>APOA4</td>
<td>Apolipoprotein A4</td>
<td>Extracellular Space</td>
<td>transporter</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>APOE</td>
<td>Apolipoprotein E</td>
<td>Extracellular Space</td>
<td>transporter</td>
<td>(Crabb et al., 2002)</td>
<td>(Mullins et al., 2000)</td>
</tr>
<tr>
<td>ATP5A1</td>
<td>ATP synth., H+ transp., mitochondrial F1 compl., alpha sub. 1, cardiac muscle</td>
<td>Cytoplasm</td>
<td>transporter</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>ATP5B</td>
<td>ATP synth., H+ transp., mitochondrial F1 compl., beta pp</td>
<td>Cytoplasm</td>
<td>transporter</td>
<td>(Wang et al., 2010)</td>
<td></td>
</tr>
<tr>
<td>BFSP1</td>
<td>Beaded filament structural protein 1</td>
<td>Cytoplasm</td>
<td>enzyme</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>BFSP2</td>
<td>Beaded filament structural protein 2</td>
<td>Cytoplasm</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>BGN</td>
<td>Biglycan</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Protein Name</td>
<td>Location</td>
<td>Function</td>
<td>Reference(s)</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-------------------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>-------------------------</td>
<td></td>
</tr>
<tr>
<td>C7</td>
<td>Complement C7</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>C8A</td>
<td>Complement C8 alpha chain</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td>(Wang et al., 2010)</td>
</tr>
<tr>
<td>C8B</td>
<td>Complement C8 beta chain</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td>(Wang et al., 2010)</td>
</tr>
<tr>
<td>C8G</td>
<td>Complement C8 gamma chain</td>
<td>Extracellular Space</td>
<td>transporter</td>
<td>(Crabb et al., 2002)</td>
<td>(Wang et al., 2010)</td>
</tr>
<tr>
<td>CFH</td>
<td>Complement factor H</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Wang et al., 2010)</td>
<td>(Arya et al., 2018)</td>
</tr>
<tr>
<td>CKB</td>
<td>Creatine kinase B</td>
<td>Cytoplasm</td>
<td>kinase</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>CLU</td>
<td>Clusterin</td>
<td>Cytoplasm</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td>(Sakaguchi et al., 2002)</td>
</tr>
<tr>
<td>COL1A2</td>
<td>Collagen type I alpha 2 chain</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td>(Newsome et al., 1987)</td>
</tr>
<tr>
<td>COL6A1</td>
<td>Collagen type VI alpha 1 chain</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>COL6A2</td>
<td>Collagen type VI alpha 2 chain</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>COL8A1</td>
<td>Collagen type VIII alpha 1 chain</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>CRYAB</td>
<td>Crystallin alpha B*</td>
<td>Nucleus</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td>(De et al., 2007)*</td>
</tr>
<tr>
<td>CRYBA1</td>
<td>Crystallin beta A1</td>
<td>Other</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>CRYBA4</td>
<td>Crystallin beta A4</td>
<td>Other</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>CRYBB1</td>
<td>Crystallin beta B1</td>
<td>Other</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>CRYBB2</td>
<td>Crystallin beta B2</td>
<td>Other</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>CRYGB</td>
<td>Crystallin gamma B</td>
<td>Nucleus</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>CRYGC</td>
<td>Crystallin gamma C</td>
<td>Cytoplasm</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>CRYGD</td>
<td>Crystallin gamma D</td>
<td>Cytoplasm</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>CRYGS</td>
<td>Crystallin gamma S</td>
<td>Other</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>CTSD</td>
<td>Cathepsin D*</td>
<td>Cytoplasm</td>
<td>peptidase</td>
<td>(Crabb et al., 2002)</td>
<td>(Rakoczy et al., 1999)*</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>Location</td>
<td>Type</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------------------------------</td>
<td>------------------</td>
<td>-----------------</td>
<td>-----------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>DIP2C</td>
<td>Disco interacting protein 2 homolog C</td>
<td>Other</td>
<td>Other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>EFEMP1</td>
<td>EGF containing fibulin like ECM protein 1</td>
<td>Extracellular</td>
<td>enzyme</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>ELN</td>
<td>Elastin*</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002)*</td>
<td></td>
</tr>
<tr>
<td>ENO2</td>
<td>Enolase 2</td>
<td>Cytoplasm</td>
<td>enzyme</td>
<td>(Wang et al., 2010)</td>
<td></td>
</tr>
<tr>
<td>FBLN5</td>
<td>Fibulin 5</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002) (Mullins et al., 2000)</td>
<td></td>
</tr>
<tr>
<td>FGG</td>
<td>Fbrinogen gamma chain</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002) (Mullins et al., 2000)</td>
<td></td>
</tr>
<tr>
<td>FHAD1</td>
<td>Forkhead ass.phosphopept.bind. dom. 1</td>
<td>Other</td>
<td>other</td>
<td>(Wang et al., 2010)</td>
<td></td>
</tr>
<tr>
<td>FN1</td>
<td>Fibronecit 1</td>
<td>Extracellular</td>
<td>enzyme</td>
<td>(Crabb et al., 2002) (Newsome et al., 1987)</td>
<td></td>
</tr>
<tr>
<td>FRZB</td>
<td>Frizzled-related protein</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
<td>Cytoplasm</td>
<td>enzyme</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>GPNMB</td>
<td>Glycoprotein nmb</td>
<td>Plasma Membrane</td>
<td>enzyme</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>HIST1H1E</td>
<td>Histone cluster 1 H1 family member e</td>
<td>Nucleus</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>HIST1H2BJ</td>
<td>Histone cluster 1 H2B family member j</td>
<td>Nucleus</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>HIST1H2B</td>
<td>Histone cluster 1 H2B family member l</td>
<td>Nucleus</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>HIST2H2B</td>
<td>Histone cluster 2 H2B family member e</td>
<td>Nucleus</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>HLA-DRA</td>
<td>Major histocompatibility complex, class II, DR alpha</td>
<td>Plasma Membrane</td>
<td>transmembrane receptor</td>
<td>(Wang et al., 2010)</td>
<td></td>
</tr>
<tr>
<td>HRG</td>
<td>Histidine rich glycoprotein</td>
<td>Extracellular</td>
<td>other</td>
<td>(Kobayashi et al., 2014)**</td>
<td></td>
</tr>
<tr>
<td>LAMB2</td>
<td>Laminin subunit beta 2</td>
<td>Extracellular</td>
<td>enzyme</td>
<td>(Crabb et al., 2002) (Newsome et al., 1987)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>LTF</td>
<td>Lactotransferrin</td>
<td>Extracellular Space</td>
<td>peptidase</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>LUM</td>
<td>Lumican</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>MFAP4</td>
<td>Microfibril associated protein 4</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>MYH9</td>
<td>Myosin heavy chain 9</td>
<td>Cytoplasm</td>
<td>enzyme</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>OGN</td>
<td>Osteoglycin</td>
<td>Extracellular Space</td>
<td>growth factor</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>ORM1</td>
<td>Orosomucoid 1</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>PLG</td>
<td>Plasminogen</td>
<td>Extracellular Space</td>
<td>peptidase</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>PRDX1</td>
<td>peroxiredoxin 1</td>
<td>Cytoplasm</td>
<td>enzyme</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>PLEP</td>
<td>Pro, Arg rich end Leu rich repeat protein</td>
<td>Extracellular Space</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>PSMB5</td>
<td>Proteasome subunit beta 5</td>
<td>Cytoplasm</td>
<td>peptidase</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>RBP3</td>
<td>Retinol binding protein 3</td>
<td>Extracellular Space</td>
<td>transporter</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>RDH5</td>
<td>Retinol dehydrogenase 5</td>
<td>Cytoplasm</td>
<td>enzyme</td>
<td>(Wang et al., 2010)</td>
<td></td>
</tr>
<tr>
<td>RGR</td>
<td>Retinal G protein coupled receptor</td>
<td>Plasma Membrane</td>
<td>G-protein coupled recept.</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>RNASE4</td>
<td>Ribonuclease A family member 4</td>
<td>Extracellular Space</td>
<td>enzyme</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>S100A7</td>
<td>S100 calcium binding proteA7</td>
<td>Cytoplasm</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>S100A8</td>
<td>S100 calcium binding protein A8</td>
<td>Cytoplasm</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>S100A9</td>
<td>S100 calcium binding protein A9</td>
<td>Cytoplasm</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>SAA1</td>
<td>Serum amyloid A1</td>
<td>Extracellular Space</td>
<td>transporter</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>SCARB2</td>
<td>Scavenger receptor class B member 2</td>
<td>Plasma Membrane</td>
<td>other</td>
<td>(Wang et al., 2010)</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>Description</td>
<td>Location</td>
<td>Function</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------</td>
<td>----------------</td>
<td>--------------</td>
<td>-------------------------------</td>
<td></td>
</tr>
<tr>
<td>SEMA3B</td>
<td>Semaphorin 3B</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>SERPINA1</td>
<td>Serpin family A member 1</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>SERPINA3</td>
<td>Serpin family A member 3</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>SERPINF1</td>
<td>Serpin family F member 1</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>SPP2</td>
<td>Secreted phosphoprotein 2</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>SPTAN1</td>
<td>Spectrin alpha, non-erythrocytic 1</td>
<td>Plasma Membrane</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>THBS4</td>
<td>Thrombospondin 4</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>TIMP3</td>
<td>TIMP metallopeptidase inhibitor 3</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002) (Kamei and Hollyfield, 1999)*</td>
<td></td>
</tr>
<tr>
<td>TNC</td>
<td>Tenascin C</td>
<td>Extracellular</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>TUBA1C</td>
<td>Tubulin alpha 1c</td>
<td>Cytoplasm</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>TUBB3</td>
<td>Tubulin beta 3 class III</td>
<td>Cytoplasm</td>
<td>other</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>TYRP1</td>
<td>Tyrosinase related protein 1</td>
<td>Cytoplasm</td>
<td>enzyme</td>
<td>(Crabb et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>VIM</td>
<td>Vimentin*</td>
<td>Cytoplasm</td>
<td>other</td>
<td>(Crabb et al., 2002) (Johnson et al., 2003)*</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2C
Figure 2D
Database curation examples

A

strip and merge

Σ

B

RPE (top 10%) enriched

select

Choroid (top) 10%

RPE (top 10%) highly enriched

C

substract

X-Y
Apical source of drusen -omics database

Curated database drusen proteomics

Basal source of drusen -omics database

Blood –omics (BL-SP1; BL-SP2)

Curated Chor/blood –omics (cChor-ET)

RPE –omics (RPE-ST; RPE-IVS)

Curated Phot/RPE –omics (cPRos-EP; cPR-ET)

Non-Curated Phot/RPE –omics (PRos-EP; PR-ET)

Non-Curated Chor/blood –omics (Chor-ET)

Figure 4
Figure 8
Figure 9A

Figure 9B

Figure 9C
### Top Canonical Pathways

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute Phase Response Signaling</td>
<td>7.51E-15</td>
<td>8.2 %</td>
</tr>
<tr>
<td>LXR/RXR Activation</td>
<td>2.29E-12</td>
<td>9.1 %</td>
</tr>
<tr>
<td>FXR/RXR Activation</td>
<td>9.39E-11</td>
<td>7.9 %</td>
</tr>
<tr>
<td>Atherosclerosis Signaling</td>
<td>2.35E-09</td>
<td>7.1 %</td>
</tr>
<tr>
<td>IL-12 Signaling and Production in Macrophages</td>
<td>1.42E-07</td>
<td>5.5 %</td>
</tr>
</tbody>
</table>

### Top Diseases and Bio Functions

#### Diseases and Disorders

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>#Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hereditary Disorder</td>
<td>1.04E-04 - 3,14E-19</td>
<td>51</td>
</tr>
<tr>
<td>Ophthalmic Disease</td>
<td>1.04E-04 - 3,14E-19</td>
<td>37</td>
</tr>
<tr>
<td>Organismal Injury and Abnormalities</td>
<td>1.19E-04 - 3,14E-19</td>
<td>88</td>
</tr>
<tr>
<td>Metabolic Disease</td>
<td>9.39E-05 - 4,75E-14</td>
<td>47</td>
</tr>
<tr>
<td>Developmental Disorder</td>
<td>1.04E-04 - 9,87E-14</td>
<td>34</td>
</tr>
</tbody>
</table>

### Molecular and Cellular Functions

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>#Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular Movement</td>
<td>1.18E-04 - 2,86E-16</td>
<td>44</td>
</tr>
</tbody>
</table>
### Top Networks

<table>
<thead>
<tr>
<th>ID</th>
<th>Associated Network Functions</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Cancer, Connective Tissue Disorders, Organismal Injury and Abnormalities</td>
<td>44</td>
</tr>
<tr>
<td>2.</td>
<td>Developmental Disorder, Ophthalmic Disease, Organismal Injury and Abnormalities</td>
<td>30</td>
</tr>
<tr>
<td>3.</td>
<td>Cell-To-Cell Signaling and Interaction, Hematological System Development and Function, Lipid Metabolism</td>
<td>27</td>
</tr>
<tr>
<td>4.</td>
<td>Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Neurological Disease</td>
<td>25</td>
</tr>
<tr>
<td>5.</td>
<td>Neurological Disease, Infectious Diseases, Respiratory Disease</td>
<td>22</td>
</tr>
</tbody>
</table>

### Physiological System Development and Function

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>#Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embryonic Development</td>
<td>1.04E-04 - 3.33E-17</td>
<td>30</td>
</tr>
<tr>
<td>Nervous System Development and Function</td>
<td>1.04E-04 - 3.33E-17</td>
<td>31</td>
</tr>
<tr>
<td>Organ Development</td>
<td>1.04E-04 - 3.33E-17</td>
<td>25</td>
</tr>
<tr>
<td>Organismal Development</td>
<td>1.04E-04 - 3.33E-17</td>
<td>53</td>
</tr>
<tr>
<td>Tissue Development</td>
<td>1.08E-04 - 3.33E-17</td>
<td>52</td>
</tr>
</tbody>
</table>

### Summary of Analysis - Table 2 Final annotation 89 drusen proteins 2018 - 2018-05-02 09:05 AM

- **Cell-To-Cell Signaling and Interaction**: 1.18E-04 - 1.74E-12, 43 molecules
- **Lipid Metabolism**: 9.71E-05 - 1.02E-10, 27 molecules
- **Molecular Transport**: 1.09E-04 - 1.02E-10, 33 molecules
- **Small Molecule Biochemistry**: 9.71E-05 - 1.02E-10, 27 molecules

**Top Networks**

- 1. Cancer, Connective Tissue Disorders, Organismal Injury and Abnormalities
- 2. Developmental Disorder, Ophthalmic Disease, Organismal Injury and Abnormalities
- 3. Cell-To-Cell Signaling and Interaction, Hematological System Development and Function, Lipid Metabolism
- 4. Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Neurological Disease
- 5. Neurological Disease, Infectious Diseases, Respiratory Disease
### Ingenuity Pathway Analysis (IPA). Table 4.

Analysis Name: Table 4 Summary photoreceptor core annotation analysis (745); Bergen et al.

2018-05-07 10:07 PM Analysis Creation Date: 2018-05-07

Build version: 470319M

Content version: 43605602 (Release Date: 2018-03-28)

#### Top Canonical Pathways

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phototransduction Pathway</td>
<td>7.98E-11</td>
<td>28,3 %</td>
</tr>
<tr>
<td>Huntington's Disease Signaling</td>
<td>2.04E-04</td>
<td>8,0 %</td>
</tr>
<tr>
<td>Glutamate Receptor Signaling</td>
<td>4.79E-04</td>
<td>14,0 %</td>
</tr>
<tr>
<td>Superpathway of Cholesterol Biosynthesis</td>
<td>1.87E-03</td>
<td>17,9 %</td>
</tr>
<tr>
<td>Wnt/Ca+ pathway</td>
<td>4.20E-03</td>
<td>11,1 %</td>
</tr>
</tbody>
</table>

#### Top Diseases and Bio Functions

##### Diseases and Disorders

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>#Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>1.31E-02 - 6.87E-28</td>
<td>680</td>
</tr>
<tr>
<td>Organismal Injury and Abnormalities</td>
<td>1.33E-02 - 6.87E-28</td>
<td>685</td>
</tr>
<tr>
<td>Gastrointestinal Disease</td>
<td>1.21E-02 - 4.26E-22</td>
<td>629</td>
</tr>
<tr>
<td>Hepatic System Disease</td>
<td>2.47E-03 - 8.92E-15</td>
<td>474</td>
</tr>
<tr>
<td>Reproductive System Disease</td>
<td>1.27E-02 - 3.21E-08</td>
<td>420</td>
</tr>
</tbody>
</table>

##### Molecular and Cellular Functions

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>#Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular Assembly and Organization</td>
<td>1.28E-02 - 1.06E-08</td>
<td>165</td>
</tr>
</tbody>
</table>
Summary of Analysis - photoreceptor core annotation analysis; Bergen et al. (745) - 2018-05-07

| Cellular Function and Maintenance | 1.31E-02 - 1.06E-08 | 185 |
| Cell Death and Survival | 1.32E-02 - 3.71E-06 | 240 |
| Cell Morphology | 1.31E-02 - 5.38E-06 | 149 |
| Cell-To-Cell Signaling and Interaction | 1.16E-02 - 8.21E-06 | 61 |

**Physiological System Development and Function**

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>#Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organ Development</td>
<td>1.25E-02 - 3.12E-06</td>
<td>88</td>
</tr>
<tr>
<td>Tissue Development</td>
<td>1.28E-02 - 3.12E-06</td>
<td>129</td>
</tr>
<tr>
<td>Visual System Development and Function</td>
<td>1.16E-02 - 8.07E-06</td>
<td>30</td>
</tr>
<tr>
<td>Nervous System Development and Function</td>
<td>1.28E-02 - 9.43E-06</td>
<td>92</td>
</tr>
</tbody>
</table>

**Top Networks**

<table>
<thead>
<tr>
<th>ID</th>
<th>Associated Network Functions</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Cellular Assembly and Organization, Cellular Function and Maintenance, Molecular Transport</td>
<td>50</td>
</tr>
<tr>
<td>2.</td>
<td>Molecular Transport, RNA Trafficking, Behavior</td>
<td>47</td>
</tr>
<tr>
<td>3.</td>
<td>Developmental Disorder, Neurological Disease, Cellular Assembly and Organization</td>
<td>47</td>
</tr>
<tr>
<td>4.</td>
<td>Molecular Transport, RNA Trafficking, Connective Tissue Development and Function</td>
<td>42</td>
</tr>
<tr>
<td>5.</td>
<td>Developmental Disorder, Hereditary Disorder, Organismal Injury and Abnormalities</td>
<td>42</td>
</tr>
</tbody>
</table>
**Ingenuity Pathway Analysis (IPA). Table 5.**

Analysis Name: Table 5 Summary Functional Annotation Choroid- Bergen et al
Analysis Creation Date: 2018-05-07
Build version: 470319M
Content version: 43605602 (Release Date: 2018-03-28)

### Top Canonical Pathways

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antigen Presentation Pathway</td>
<td>7.58E-12</td>
<td>36.8 %</td>
</tr>
<tr>
<td>Atherosclerosis Signaling</td>
<td>3.26E-11</td>
<td>18.0 %</td>
</tr>
<tr>
<td>Hepatic Fibrosis / Hepatic Stellate Cell Activation</td>
<td>3.88E-11</td>
<td>14.7 %</td>
</tr>
<tr>
<td>Acute Phase Response Signaling</td>
<td>5.19E-10</td>
<td>14.5 %</td>
</tr>
<tr>
<td>Complement System</td>
<td>1.90E-09</td>
<td>31.6 %</td>
</tr>
</tbody>
</table>

### Top Diseases and Bio Functions

#### Diseases and Disorders

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>#Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>1.44E-06 - 4.74E-32</td>
<td>730</td>
</tr>
<tr>
<td>Organismal Injury and Abnormalities</td>
<td>1.48E-06 - 4.74E-32</td>
<td>746</td>
</tr>
<tr>
<td>Inflammatory Response</td>
<td>4.1E-06 - 1.87E-21</td>
<td>263</td>
</tr>
<tr>
<td>Connective Tissue Disorders</td>
<td>4.1E-06 - 1.46E-17</td>
<td>192</td>
</tr>
<tr>
<td>Inflammatory Disease</td>
<td>4.57E-08 - 1.46E-17</td>
<td>174</td>
</tr>
</tbody>
</table>

#### Molecular and Cellular Functions

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>#Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular Movement</td>
<td>1.26E-06 - 1.81E-32</td>
<td>257</td>
</tr>
</tbody>
</table>
Summary of Analysis -

Cell Death and Survival 1.47E-06 - 4.17E-21 323
Cell-To-Cell Signaling and Interaction 1.50E-06 - 9.12E-15 218
Cellular Development 1.49E-06 - 1.14E-14 326
Cellular Function and Maintenance 8.75E-07 - 3.23E-12 274

Physiological System Development and Function

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>#Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiovascular System Development and Function</td>
<td>9.30E-07 - 1.46E-28</td>
<td>192</td>
</tr>
<tr>
<td>Organismal Development</td>
<td>1.47E-06 - 1.38E-24</td>
<td>325</td>
</tr>
<tr>
<td>Immune Cell Trafficking</td>
<td>1.29E-06 - 1.10E-23</td>
<td>158</td>
</tr>
<tr>
<td>Hematological System Development and Function</td>
<td>1.29E-06 - 3.90E-23</td>
<td>236</td>
</tr>
<tr>
<td>Organismal Survival</td>
<td>5.54E-08 - 1.79E-22</td>
<td>244</td>
</tr>
</tbody>
</table>

Top Networks

<table>
<thead>
<tr>
<th>Id</th>
<th>Associated Network Functions</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Organ Morphology, Organismal Injury and Abnormalities, Renal Atrophy</td>
<td>49</td>
</tr>
<tr>
<td>2.</td>
<td>Organismal Injury and Abnormalities, Skeletal and Muscular Disorders, Developmental Disorder</td>
<td>41</td>
</tr>
<tr>
<td>3.</td>
<td>Molecular Transport, Nucleic Acid Metabolism, Small Molecule Biochemistry</td>
<td>38</td>
</tr>
<tr>
<td>4.</td>
<td>Tissue Development, Cellular Movement, Hair and Skin Development and Function</td>
<td>37</td>
</tr>
<tr>
<td>5.</td>
<td>Cell Cycle, Gene Expression, Cellular Growth and Proliferation</td>
<td>36</td>
</tr>
</tbody>
</table>
**Authors' statement**

All authors have seen and approved the final version of the manuscript being submitted. They warrant that the article is the authors' original work, hasn't received prior publication and isn't under consideration for publication elsewhere.

None of the authors has a financial or personal conflicts of interest defined as a set of conditions in which professional judgment concerning a primary interest, such as the validity of research, may be influenced by a secondary interest, such as financial gain.

The work on this manuscript is approximately divided as follows:

AAB 28 %, SA 5 %, CK 5 %, MP 5 %, DW 4 %, PvdS 4%, SMH 4 %, CJFB 4 %, EE 5 %, AJS 18 %, IL 18 %