Statistical and experimental analysis of genetic and non-genetic markers associated with risk and progression of age-related macular degeneration (AMD)

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der Fakultät für Biologie und vorklinische Medizin

der Universität Regensburg

vorgelegt von

Felix Graßmann

aus Traunstein

im Jahr 2015

Das Promotionsgesuch wurde eingereicht am: 12.2.2015

Die Arbeit wurde angeleitet von Prof. Dr. Bernhard Weber

Unterschrift

Declaration of personal contribution to submitted/accepted research manuscripts. The results chapters of this thesis are based upon these research publications. All supplementary tables and figures of these publications are shown in the appendix.

1. Grassmann F, Fritsche LG, Keilhauer CN, Heid IM & Weber BHF (2012). **Modelling the genetic risk in age-related macular degeneration.** *PLoS One* **7:** e37979

Personal contribution: Genotyped part of the patient samples. Analyzed the data and developed the algorithms. Designed and realized the figures and tables and wrote a first draft of the manuscript.

2. Grassmann F, Schoenberger PGA, Brandl C, Schick T, Hasler D, Meister G, Fleckenstein M, Lindner M, Helbig H, Fauser S, Weber BHF (2014). **A Circulating MicroRNA Profile Is Associated with Late-Stage Neovascular Age-Related Macular Degeneration.** *PLoS One* **9:** e107461

Personal contribution: Genotyped part of the patient samples. Isolated mirRNA from serum and stabilized whole blood samples from patients and controls. Performed and analyzed RNASeq experiments (discovery study). Performed and anazyled qPCR experiments (replication studies). Performed and analyzed *in vitro* angiogenesis assays. Analyzed the data in the joint study. Designed and realized the figures and tables and wrote a first draft of the manuscript.

3. Grassmann F, Fleckenstein M, Chew EY, Strunz T, Schmitz-Valckenberg S, Göbel AP, Klein ML, Ratnapriya R, Swaroop A, Holz FG, Weber BHF (2015) **Clinical and genetic factors associated with progression of geographic atrophy lesions in age-related macular degeneration.** *PLoS One* in revision

Personal contribution: Genotyped part of the patient samples. Analyzed the data. Designed and realized the figures and tables and wrote a first draft of the manuscript.

4. Grassmann F, Friedrich U, Fauser S, Schick T, Milenkovic A, Schulz HL, von Strachwitz CN, Bettecken T, Lichtner P, Meitinger T, Arend N, Wolf A, Haritoglou C, Rudolph G, Chakravarthy U, Silvestri G, McKay GJ, Freitag-Wolf1 S, Krawczak M, Smith RT, Merriam JC, Merriam JE, Allikmets R, Heid IM, Weber BHF (2015) **A candidate gene association study identifies *DAPL1* as a female-specific susceptibility locus for age-related macular degeneration (AMD).** *Neuromolecular Medicine* in press

Personal contribution: Genotyped part of the patient samples. Analyzed the data. Designed and realized the figures (Figure 1, Figure 2, Figure 4, Supplementary Figure 1, Supplementary Figure 3) and tables (except Supplementary Table 5) and wrote a first draft of the manuscript.

Declaration of personal contribution to submitted/accepted reviews. Figures and Tables in the introduction and discussion of this thesis are based upon these reviews.

1. Grassmann F, Heid IM & Weber BHF (2014) **Genetic risk models in age-related macular degeneration.** *Adv. Exp. Med. Biol.* **801:** 291–300

Personal contribution: Performed comprehensive literature search. Designed and realized the tables and wrote a first draft of the manuscript.

2. Grassmann F, Ach T, Brandl C, Heid IM & Weber BHF (2015) **What does genetics tell us about AMD?** *Ann Rev Vis Sci*, invited review, in revision

Personal contribution: Performed literature research. Designed and realized Table 1, Figure 2 and Figure 3 and wrote a first draft of the manuscript.

3. Grassmann F, Fauser S & Weber BHF (2015) **The genetics of age-related macular degeneration (AMD) and its usability for designing treatment options** *Eur J Pharm Biopharm.* invited review, submitted

Personal contribution: Performed literature research. Designed and realized Table 1 and Figure 2 and wrote a first draft of the manuscript.

Acknowledgement

First and foremost I would like to thank my supervisor Prof. Dr. Bernhard Weber for the interesting tasks, his interest in the progress of this work as well as for his excellent supervision. I have been fortunate that he gave me great latitude to manage my research projects independently and that he was open-minded for new methods and ideas. His insightful comments and constructive criticisms at different stages of my research were thought-provoking and they helped me to pursue my ideas.

I am also grateful to Prof. Dr. Iris Heid for her encouragement and support throughout this work. She has been always there to listen and give advice. I would like to thank her for constructive discussions, her enthusiasm and her encouragement that inspired me and helped me to improve my knowledge in the field of epidemiology.

Thanks are extended to Prof. Dr. Ralph Witzgall for providing critical feedback and advice throughout my thesis. I would also like to thank Prof. Dr. Richard Warth for his willingness to act as chairperson of the examination commitee and Prof. Dr. Rainer Merkl for his willingness to function as third examiner.

I am also indebted to all the patients and controls that participated in the various studies. None of the projects would have been possible without their willingness to participate.

Very special thanks for the great and successful cooperations go to all of the collaborators from the University of Regensburg, the University Hospital of Regensburg, the University Hospital of Cologne, the University of Bonn, the University Eye Hospital Würzburg, the Technical University of Munich, the University Eye Hospital of the Ludwigs-Maximilians-University Munich, the Queen's University of Belfast, the Columbia University New York, the New York University School of Medicine, the Christian-Albrechts University Kiel and the University of Alabama at Birmingham.

I am heartily thankful to all the great people at the Institute of Human Genetics for creating a positive, constructive as well as cooperative working atmosphere.

The joy of discovery is certainly the liveliest that the mind of man can ever feel.

Claude Bernard (1813-78)

Content

1. General introduction

Tables and figures in the introduction are based on these reviews:

1. Grassmann F, Heid IM & Weber BHF (2014) **Genetic risk models in age-related macular degeneration.** *Adv. Exp. Med. Biol.* **801:** 291–300

2. Grassmann F, Fauser S & Weber BHF (2015) **The genetics of age-related macular degeneration (AMD) and its usability for designing treatment options** *Eur J Pharm Biopharm.* invited review. in revision

1.1 Clinical features of age-related macular degeneration (AMD)

According to a recent survey, the majority of people in different ethnic groups agree that "good eye health is important to overall health" [1]. In fact, severe visual impairment and blindness are among the most-feared health conditions, after cancer and cardiovascular disease, mostly due to the loss of independence that comes with it [1].

Age-related macular degeneration is a common, multifactorial, late-onset human disease [2] characterized in its final stages by neurodegeneration of cells in the outer retina, leading to pronounced visual loss and eventually to complete blindness. The affected cell types and functional entities mainly include but are not limited to: (i) the photoreceptors, which are the main sensory cells in the eye, (ii) the retinal pigment epithelium cells (RPE), which among many other functions phagocytose shed photoreceptor outer segments, are responsible for the recycling of metabolites reqired for the visual cylce and provide nurtrients to the photoreceptors [3], (iii) the Bruch's membrane, a five-layered extracellular matrix embedded between the RPE and the choroid, responsible for the regulation of metabolit transport from and to the RPE [4–6] and finally (iv) the choriocapillaris, which delivers oxygen and nutrients to support retinal metabolism [5].

AMD is an age-related disease and thus its phenotype has to be differentiated from age specific changes in the retina and choroid. For example, with increasing age, the RPE cells accumulate (auto)fluorescent material, thus increasing autofluoresence of RPE cells [7–9]. The number of RPE cells and photoreceptor cells, however, remains stable throughout this process [10]. In addition, Bruch's membrane thickens, leading to decreased permeability [5,11], potentially aggravating further accumulation of fluorescent material in the RPE.

Finally, thickness of the choriocapillaris decreases with age, further reducing oxygen and nutrition supply to the eye and thus increasing susceptibility to degenerative processes [5,12].

The diseased retina due to AMD is categorized into two main stages including early AMD and late stage AMD. Early AMD is a heterogenous manifestion, clinically defined by extracellular depositions between Bruch's membrane and the RPE (so-called drusen) and pigmentary changes of the RPE (either hyper- or hypopigmentation). These pathological changes usually do not significantly influence retinal function and thus do not lead to visual impairment, although dark adaptation is delayed in eyes with early AMD [13]. In contrast, late stage AMD can result in severe and irreversible vision loss. Late stage AMD can manifest in two forms, namely as an atrophic, non-exudative ("dry") condition and a neovascular, exudative ("wet") complication. In the neovascular form (NV), growth of blood vessels into the outer retina is stimulated by the release of vascular endothelial growth factor (VEGF) from the RPE [14]. These new vessels extend through Bruch's membrane and can result in profound bleeding, extracellular fluid accumulation and subsequent scarring of the outer retina, thus leading to rapid (within days or weeks) and severe vision loss [15]. The late stage of the non-exudative form is characterized by geographic atrophy (GA) of choriocapillaris, RPE cells and photoreceptors, leading to a slow (within years) but steady progression of vision loss. While neovascular AMD can be treated by anti-VEGF drugs, there is no treatment availble for GA at present [16] .

1.2 Epidemiology of AMD

Per definition, the prevalence of an age-related disease increases with age. The prevalence for late stage AMD in Western societies increases from less than 1% at 55 years of age to more than 15% in individuals aged 85 years and above [17], while the overall prevalence of all AMD types is estimated to be 8.7% [18]. The proportion of GA in late stage AMD is approximately 35-40% [19,20]. While the overall incidence of the neovascular form is more frequent, GA occurs more commonly in individuals over 85 years of age [20]. Conservative estimates project the worldwide number of people with AMD in 2020 to 196 million and in 2040 to 288 million [18]. In the near future, this will dramatically increase the individual as well as the socioeconomic burden of the disease. This greatly emphasizes the impact of late stage AMD on aging populations, and underscores the need for effective treatments to

prevent, slow or cure the disease. To identify potential treatment options, knowledge on genetic and non-genetic factors involved in disease risk and progression are crucial.

1.3 Risk factors associated with AMD

1.3.1 Environmental, life-style and clinical risk factors

Several environmental or life-style factors, like diet, sunlight exposure or allergies [21] have been associated with early and late stage AMD, although only smoking has consistently been found to increase the risk for both stages of AMD [22]. In addition, recent reports indicate a protective effect of nutritional fish oil based fatty acid supplementation (DHA and EPA) [23–25], promising a possible intervention for a delayed disease progression.

Many features of early AMD in the outer retina are indicators for progression to late stage AMD and are thus considered to be clinical and in principle, non-modifiable risk factors. In particular, the presence of large, confluent (soft) drusen and defective dark adaptation are associated with a high risk for progression to late stage AMD [13]. In contrast, pigmentary changes in the RPE and presence of small and non-confluent (hard) drusen are not associated with late stage AMD progression [26,27]. Furthermore, the presence of a late stage form of the disease in one eye is a strong risk factor for the progression of the fellow eye to late stage AMD.

1.3.2 Genetic risk factors associated with late stage AMD

Among the two stages of the disease, early and late, the genetics of late stage AMD has been more extensively studied and is currently better understood. In most genetic association studies, both late stage forms (neovascular and atrophic AMD) are analyzed jointly and the reported genetic variants are risk factors for both forms. The first findings for a genetic contribution to the disease were reported in 1997 by implicating variations in the ATP binding cassette, subfamily A, member 4 (*ABCA4)* gene to be associated with AMD (**Figure 1**) [28]. This association, however, remained controversial since several attempts at replication failed [29,30]. Only recently, it was shown that mono-allelic changes in *ABCA4* are indeed associated with a rare subtype of geographic atrophy, namely fine granular pattern with peripheral punctate spots (GPS) [31]. In 1998, Klaver *et al.* reported a significant association of variations in the apolipoprotein E (*APOE*) with late stage AMD [32].

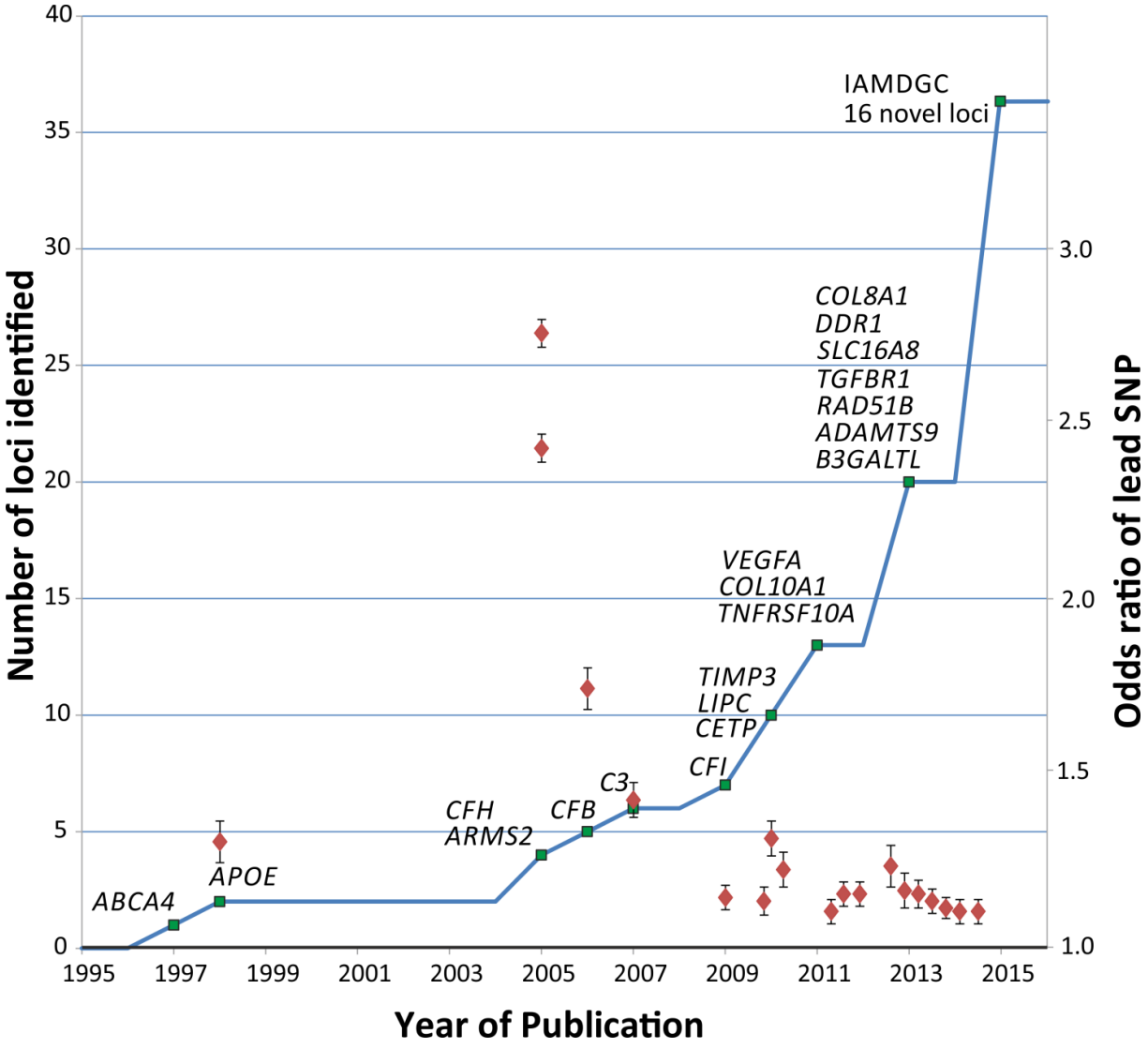


Figure 1. Gene loci associated with AMD. The first genetic locus associated with AMD was *ABCA4* and was implicated in 1997. Since then, the number of associated loci has risen exponentially and as of today comprises 34 independent loci. Included in this count are 16 novel loci to be reported by the International AMD Genomics Consortium (IAMDGC) in 2015. Green squares represent the date of discovery of novel risk loci. Red diamonds represent the odds ratio of the identified top variant in the respective locus and error bars the respective 95% confidence intervals (data based on effect sizes obtained from Fritsche *et al.* 2013 [33]). Figure taken from Grassmann *et al.* 2015, *Eur J Pharm Biopharm*. invited review. in revision

These findings were later replicated in several studies, implicating the contribution of lipid pathways to disease etiology. By conducting the first ever genome-wide association study (GWAS) for a human disease, Klein *et al.* identified in 2005 a common missense mutation (CFH p.Y402H) in the complement factor H gene to be strongly associated with AMD [34]. Later, a similarly strong association was observed in the *ARMS2/HTRA1* locus on 10q26 [35,36]. In the following years, candidate gene studies identified non-synonymous risk variants in other complement genes, e.g. the complement component 3 (C3), complement

component 2 (C2) and complement factor B (CFB) [37,38]. Currently, more than 19 disease associated loci have been implicated with a genome-wide significant p-value below $0.5 * 10^{-8}$ (**Figure 1**) [39,40].

1.3.3 Genetic risk factors associated with early AMD

In contrast to our extensive knowledge on the genetics of late stage AMD, only little is known about the genetic basis of early AMD. The heritability of early AMD has been estimated to be between 35% and 55% and thus slightly lower than the estimate for late stage AMD [41]. Recently, Holliday *et al.* reported a meta-analysis of five genome-wide association studies in early AMD and implicated both, the *ARMS2/HTRA1* as well as the *CFH* locus to be significantly (on a genome-wide level) associated with early AMD. Notably, although late stage AMD risk increasing alleles were also found to be risk increasing alleles for the early stage of the disease, the estimated effect sizes (i.e. the observed odds ratios) for the risk alleles in the *ARMS2/HTRA1* as well as the *CFH* locus were strikingly lower. In the same study, no other variants reached a genome-wide significant p-value threshold.

1.4 (Genetic) Risk models for AMD

AMD is unique among complex diseases as several common variants show strong association with late stage AMD with large effect sizes. This unique feature has led to efforts to predict the risk for AMD based on genetic factors with or without including non-genetic risk factors (**Table 1**). While pure genetic risk models perform reasonably well with area under the receiver operating characteristics curve (AUC) values reported around 0.80, they still fall short of the recommended AUC value of 0.99 needed for presymptomatic DNA testing in the general population [42]. The number of associated variants with average effect sizes necessary to increase the AUC to 0.99 and above will probably range in the hundreds [42,43]. These models have, nevertheless, comparable accuracies to other risk models, e.g. risk models for the prediction of cardiovascular diseases [44] or for the prediction of prostate cancer [45]. However, the validity of AMD risk models in a clinical setting has yet to be shown.

Table 1. Risk models for age-related macular degeneration (from: Grassmann *et al.* 2014 *Adv. Exp. Med. Biol.* 801: 291–300)

Publication	Year	Genetic factors[4]	Non-genetic factors[6]	Clinical factors	Study design (sample size)	Method[1]	Risk Score[2]	disc[5]	Validation[3]	Reported AUC
Maller *et al** [46]	2006	CFH(2), CFB, ARMS2	no	no	case-control (2172)	LR	no	yes	-	-
Gold *et al* [47]	2006	CFH(2), CFB(3)	no	no	case-control (~1500)	DT	no	yes	IS	-
Hughes *et al* [48]	2007	CFH(4), ARMS2	s	no	case-control (667)	LR	yes	yes	-	-
Jakobsdottir *et al* [49]	2008	CFH, ARMS2, C2	no	no	case-control (675)	LR	NA	yes	-	-
Jakobsdottir *et al* [42]	2009	CFH, ARMS2, C2	no	no	case-control (675)	LR	yes	yes	-	0.79
Seddon *et al** [50]	2009	CFH(2), ARMS2, C2, CFB, C3	a, g, e, s, t, b	no	longitudinal (1446)	LR	yes	no	-	0.83
Farwick *et al* [51]	2009	CFH, ARMS2(2), C2, CFB	s	no	longitudinal (913)	LR	NA	no	-	0.81
Gibson *et al* [52]	2010	CFH(3), ARMS2, C3, SERPING	a, g, s	no	case-control (940)	LR	yes	yes	-	0.83
Chen *et al** [53]	2011	CFH(3), C2, CFB, ARMS2(2), C3	s, a, g, b	no	case-control (1844)	LR	yes	no	-	0.82
Hageman *et al* [54]	2011	CFH(4), CFHR4/5(3), FI3B(2), C2, CFB, ARMS2, C3	s	no	case-control (3182)	LR	yes	yes	IS	0.80
Klein *et al** [55]	2011	CFH, ARMS2	a, e, s, **	AREDS Score; late stage AMD	longitudinal (2962)	CR	no	no	IS	0.87
Seddon *et al** [26]	2011	CFH(2), ARMS2, C3, C2, CFB	e, s, b, v	drusen size, late stage AMD	longitudinal (2937)	CR	no	no	CV	0.92
Spencer *et al* [56]	2011	CFH, ARMS2, CFB, C3	a, s	no	case-control (~900)	LR	yes	no	IS	0.84
McCarthy *et al** [57]	2012	CFH, ARMS2	a, g, e, s, b, v	AREDS Score, late stage AMD	longitudinal (2011)	LR	yes	no	IS	0.89
Grassmann *et al* [58]	2012	CFH(3), ARMS2, CFB, C3, C2, APOE(2), CFI, LIPC(2), TIMP3	no	no	case-control (1782)	LR	yes	no	CV	0.82

[1] LR= logistic regression, CR= cox regression, DT= decision tree

[2] Risk score calculated

[3] methods used to avoid inflation of risk: "independent samples" (IS) means independent samples were used to test the models and to obtain AUC values; CV=cross validation

[4] the number in parentheses states the number of risk variants in that locus used in the best performing model. No number indicates only one SNP at given locus

[5] study was used to discover novel variants which were also used in model

[6] non genetic factors are coded: s= smoking, e=education, a=age, b=body mass index, v=vitamins, g=gender, t= AREDS treatment assignment, ** family history of late stage AMD

* ARED study samples

Of note, several risk models report high AUC values around 0.90 [26,55,57]. These models include clinical findings from fundus pictures, e.g. soft drusen area, presence of geographic atrophy in the fellow eye or the AREDS severity score [59] as predictive factors. While these models outperform genetic risk models in their prediction accuracy, these models are aimed at predicting late stage AMD later in life, especially in the case where an individual already has symptoms of (early) AMD. In contrast, genetic risk models predict disease risk at birth and could therefore proof more useful for long term interventions, long before an onset of symptoms.

1.5 Improving genetic risk models and further applications of the genetic risk score

Several venues are currently explored in order to improve existing risk models and to foster our understanding of the genetic basis of the disease:

(1) *Identification of novel genetic variants associated with AMD.* This is primarily achieved by increasing the sample size of case-control studies by combining these studies in large consortia (see also Perspectives). However, due to substantial differences in the composition of case control studies in these consortia, candidate gene studies in single centers are still a promising option to identify novel genetic variations associated with the disease. Furthermore, GWA studies can not readily assess complex variations, which are not easily queried by chip genotyping, e.g. copy number variations or variations in repeat regions. These variations can be assessed with non-highthroughput methods like multiplex ligation-dependent probe amplification (MLPA) and can therefore, at present, only be detected in candidate gene or candidate variant studies.

(2) *Complementing known (common) variants with rare variants by exome sequencing or custom-made variant arrays.* It is estimated that a large proportion of the genetic risk is attributed to rare variants with moderate to high impact on disease risk [60]. However, studies need to be sufficiently large to detect these associations since only few individuals will carry a risk increasing allele, thus decreasing the power to detect the association. This is especially true if such variants are identified in a genome-wide approach since an association signal needs to reach p-values below $0.5 * 10^{-8}$ to adjust for multiple testing/comparisons. Currently, the International AMD Genomics Consortium (IAMDGC) is conducting a genome-wide association study on a custom genotyping platform. This platform specifically includes more than 150,000 coding variants, effectively capturing around 70% of the known coding variants

in the exome. Most of these coding variants are rare and found in less than 1% of the general population.

(3) *Identification of a biomarker for early and late stage AMD.* Although genetic variants can also be considered a "biomarker", true biomarker studies usually aim to find small molecules or cells in the blood/serum of patients and controls (e.g. modified/unmodified proteins [61,62], immune active cells [63], lipids [64], sugars, ribonucleic acids, complement components [65]), that can be used as a diagnostic tool to allow discrimination between affected und unaffected individuals. In addition, those biomarkers can potentially be used as prognostic markers predicting the onset or progression of a disease and levels of a biomarker can correlate with disease severity. Therefore, biomarkers are especially valuable for clinical trials to monitor and evaluate treatment success, especially if they can be analyzed in a non-invasive and inexpensive manner.

The genetic risk score does not only allow to discriminate between cases and controls. It efficiently sums up the known genetic risk for late stage AMD of an individual. There are additional, as of yet unexplored possibilities to further use this information:

(1) *Improving the design for future case-control association studies.* Genotyping patients with arrays or sequencing of specific areas in the genome of patients is still a costly endeavour. Therefore, an algorithm to prioritize samples for such association studies is warranted. Cases with a high genetic risk score are not likely to carry additional (unknown) disease associated variants and thus can be excluded from further studies. The same is true for controls with a very low genetic risk score, since these samples will most likely not carry additional (unknown) protective alleles. In order to identify novel associated variants, cases in the lower risk groups and controls in the higher risk groups as well as samples with an average risk score should be included in future studies.

(2) *AMD risk and influence on monogenic diseasse.* The AMD genetic risk score could also influence disease risk in ocular (monogenic) diseases. For instance, up to 35% of patients affected with autosomal recessive Stargardt disease, a monogenic disease caused by homozygous or compound heterozygous mutations in the *ABCA4* gene, none or only a single disease associated mutation in the *ABCA4* gene could be found [66]. Other genetic variants or environmental factors must therefore play a role in this Mendelian disease. Since Stargardt

disease shows striking similarities with GA AMD, a high risk for AMD might help to explain the observed findings.

1.6 Aims and structure of this thesis

The overall aim of the thesis was to identify and characterize genetic and non-genetic markers involved in AMD risk, progression and severity by analysing a diverse set of epidemiological studies. In addition, I aimed to evaluate the influence of several AMD associated markers on cells and tissues *in vitro*.

Specifically, we wanted to generate a genetic risk model for late stage AMD by calculating a genetic risk score for AMD patients and controls from the *Lower Frankonian AMD* case-control study recruited at the University of Würzburg [35,58]. We calculated the risk score based on 13 common (previously published) variants and then evaluated its impact on disease risk. In addition, we used the genetic risk score to investigate differences in the genetic architechture of the disease in different age-groups, in men and women and in different late stage forms of the disease (e.g. GA or NV AMD).

A second specific aim was to evaluate circulating microRNAs (cmiRNAs) as potential novel biomarkers for AMD. To this end, we recruited AMD cases and AMD-free controls from the eye clinic Regensburg. To validate candidate cmiRNAs in an independent, population based (cross-sectional) study, we analyzed serum samples of patients and controls from the *EUGENDA* study [21] recruited in Cologne in collaboration with Prof. Dr. Fauser from the Department of Ophthalmology at the University Hospital of Cologne. The influence of significantly associated cmiRNAs on neovascularization was assessed *in vitro* with a so called tube formation assay.

The third specific goal was to investigate a potential correlation of common, AMD risk associated variants as well as clinical factors with the growth rate of geographic atrophy lesions. Patients in the initial study were recruited as part of the prospective *Fundus Autofluorescence Imaging in Age-related Macular Degeneration* S*tudy* (FAM) [67] at the Department of Ophthalmology at the University of Bonn. In addition, we used GA progression data from the prospective *Age-Related Eye Disease Study* (AREDS) [22] recruited by the National Institute of Health (NIH) to replicate our findings.

The final objective of this thesis was to find additional common genetic variants associated with AMD in a candidate gene approach in a large collection of cases and controls from

several studies. The initial study consisted of individuals from the *Southern Germany* case-control study recruited at the University Hospitals in Würzburg, München and Tübingen [33]. The samples of the replication study were recruited at the Centre for Public Health at the Queen's University of Belfast, the Department of Ophthalmology at the Columbia University and the Department of Ophthalmology at the University Hospital of Cologne.

2. Modelling the genetic risk in age-related macular degeneration

This chapter is identical to the following publication:

Grassmann F, Fritsche LG, Keilhauer CN, Heid IM & Weber BHF (2012). **Modelling the genetic risk in age-related macular degeneration.** *PLoS One* **7:** e37979

Abstract

Late-stage age-related macular degeneration (AMD) is a common sight-threatening disease of the central retina affecting approximately 1 in 30 Caucasians. Besides age and smoking, genetic variants from several gene loci have reproducibly been associated with this condition and likely explain a large proportion of disease. Here, we developed a genetic risk score (GRS) for AMD based on 13 risk variants from eight gene loci. The model exhibited good discriminative accuracy (area-under-curve (AUC) of the receiver-operating characteristic of 0.820), which was confirmed in a cross-validation approach. Noteworthy, younger AMD patients with an age below 75 years had a significantly higher mean GRS (1.87, 95% CI: 1.69-2.05) than patients aged 75 and above (1.45, 95% CI: 1.36-1.54). Based on five equally sized GRS intervals, we present a risk classification with a relative AMD risk of 64.0 (95% CI: 14.11-1131.96) for individuals in the highest category (GRS 3.44-5.18, 0.5% of the general population) compared to subjects with the most common genetic background (GRS -0.05-1.70, 40.2% of general population). The highest GRS category identifies AMD patients with a sensitivity of 7.9% and a specificity of 99.9% when compared to the four lower categories. Modeling a general population around 85 years of age, 87.4% of individuals in the highest GRS category would be expected to develop AMD by that age. In contrast, only 2.2% of individuals in the two lowest GRS categories which represent almost 50% of the general population are expected to manifest AMD. Our findings underscore the large proportion of AMD cases explained by genetics particularly for younger AMD patients. The five-category risk classification could be useful for therapeutic stratification or for diagnostic testing purposes once preventive treatment is available.

2.1 Introduction

Age-related macular degeneration (AMD) is a common degenerative disease of the central retina and a leading cause of severe vision impairment in Western societies [2]. Advanced forms of AMD (late-stage AMD) are known as geographic atrophy (GA) of the retinal pigment epithelium (RPE) or neovascular (NV) complications with RPE detachment, scar formation, and subretinal hemorrhage [68,69]. To date, effective therapeutic intervention is available for active NV, while GA still remains untreatable [70,71].

AMD is a complex disease influenced by genetic and environmental factors with estimates of heritability varying from 45% to 71% [72]. So far, several AMD susceptibility loci have been identified. Two loci are accounting for an estimated 50% of AMD cases: complement factor H (*CFH)* on 1q32 and age-related maculopathy susceptibility 2 (*ARMS2*) / HtrA serine peptidase 1 (*HTRA1*) on 10q26 [34,73]. Fine-mapping studies and functional analyses at the *CFH* locus indicate at least three independent risk variants [34,37,74–77]. At the *ARMS2/HTRA1* region, a single risk haplotype was found to fully explain the observed association [35].

A crucial role of the complement system in AMD pathogenesis was further supported by subsequent candidate gene studies. These studies identified risk-associated variants in or near three additional complement genes including the complement component 2 (*C2*) / complement factor B (*CFB)*[47], complement component 3 (*C3*) [78,79] and complement factor I (*CFI*) [80]. In addition, variants in genes involved in the cholesterol and lipid metabolism were also implicated in AMD susceptibility [81,82]. Strongest signals peaked near the hepatic lipase gene (*LIPC*) on chromosome 15q22 [81,82], the cholesterylester transfer protein (*CETP*) and the lipoprotein lipase precursor (*LPL*) genes [81]. Also, among the most replicated AMD risk variants are two coding SNPs in the apolipoprotein E (*APOE*) gene [32,83]. A recent genome wide association study established a significant association of AMD with rs9621532, a variant intronic to synapsin III (*SYN3)* and approximately 100kb upstream of the tissue inhibitor of metalloproteinases-3 gene (*TIMP3*) [81]. Finally, common variations near VEGFA and FRK/COL10A1 were associated with AMD, further implicating angiogenesis as well as extracellular matrix metabolism in AMD pathogenesis [84].

To predict the genetic risk in complex diseases, testing of single susceptibility variants is generally of limited value [85]. In contrast, genotyping and evaluating a series of independent

disease associated variants, a process also known as genetic profiling, may be more appropriate [85]. This can be facilitated by a genetic risk score (GRS) which could simply represent the sum of risk associated variants found in each individual. However, such an approach may not be particularly effective in the presence of greatly differing effect sizes of the respective variants [86]. Therefore, an extension to this model weighs each additional risk allele by its effect size. For example, Seddon *et al.* (2009) calculated a risk score for AMD based on 6 known genetic risk variants and additional environmental factors. Their model revealed good discriminatory power with a reported area-under-curve (AUC) of the receiver-operating characteristic of 0.82 [50]. Other studies reporting a GRS [81,82,84] primarily aimed at identifying novel variants without using independent data or a cross-validation approach and are thus likely biased to overestimate the effect of these variants. The quantification of the genetic risk based on frequently replicated AMD loci in a single study which is independent from locus identification is still lacking.

Here, we present a genetic risk model for AMD, specifically the late-stage forms of AMD, based on a large and well characterized AMD case-control study group including 986 cases and 796 controls. We selected 13 genetic variants from eight gene loci that have repeatedly been shown to be associated with AMD and computed a genetic risk score. This was used to establish a classification system that allows for discriminating subjects at high and low genetic risk. Environmental variables such as smoking or diet were not included in the model building.

2.2 Results

SNP selection based on published data and linkage disequilibrium structure

Eight loci (CFH, ARMS2/HTRA1, CFI, CFB, C3, APOE, LIPC and TIMP3) with 13 SNPs and established association with AMD were included into our genetic risk score modeling **(Supplementary Table S1)**. There were three further SNPs with reportedly established association, which we did not select for the model: (i) at the CFH locus, an association of four variants with AMD is known (rs1410996, rs800292, rs1061170, rs6677604); however, rs1410996 is present on two distinct haplotypes, each of which is tagged by rs800292 (correlation r^2=0.473 to rs1410996[87]) or rs6677604 (r^2 =0.283 to rs1410996[87]), respectively [77], while rs800292 and rs667604 are uncorrelated (r^2=0.008 [87]), (ii) among the three highly correlated ARMS/HTRA variants (rs10490924, rs11200638, and c.del443ins54; pairwise r^2=1), rs10490924 was reported to fully capture the disease risk at this

locus [88]. We therefore selected rs1061170, rs800292 and rs667604 at CFH and rs10490924 at the ARMS2/HTRA1 locus yielding the 13 SNPs for model building.

Genotyping of SNPs in the Lower Frankonian AMD case-control study

We genotyped the selected 13 SNPs as well as the three highly correlated SNPs (to validate the correlations) in 986 cases and 796 controls from the Lower Frankonian AMD case-control study (**Table 2**).

Table 2.Summary characteristics of the case-control study

	Cases	**Controls**	**Total**
Subjects	986	796	1782
GA[1]	229	-	
NV[2]	581	-	
Mixed GA+NV[3]	176	-	
Mean Age (S.D.) [in years]	78.7 (6.5)	78.3 (5.1)	78.5 (5.9)
Men [%]	34.1	39.3	36.4
Fraction smoker [%][4,5]	15.9	14.3	

[1] Geographic atrophy

[2] Neovascular AMD

[3] Mixed GA+NV: GA and NV in the same eye or GA in one and NV in the second eye

[4] Smoking was defined as ever smoked more than 20 pack years

[5] This variable was surveyed incompletely in cases and controls and thus was not further considered in the analysis

All variants showed high genotyping quality with an average call rate > 99.5%. With the exception of rs1061170 at CFH, all genotypes were in Hardy-Weinberg equilibrium in controls (HWE, $p>0.04$). The variant rs1061170 was genotyped twice with two independent assays yielding identical genotypes and therefore persistent HWE violation in controls ($p= 0.002$) [89]. There were no missing genotypes at the 13 variants for any individual in the study.

Association of the selected 13 SNPs with AMD

For each SNP, association with AMD was computed using a logistic regression model, unadjusted for age or gender (**Table 3**). Sensitivity analysis additionally adjusting for age and gender yielded similar results. Odds ratio (OR) estimates per AMD risk increasing variant ranged from 1.14 [95% CI: 1.00-1.30] for rs2285714 to 3.13 [95% CI: 2.68-3.68] for rs10490924 and were significantly different from unity for all 13 variants demonstrating

sufficient statistical power in our study (**Table 3**). In a subgroup analysis, AMD cases with GA (n=229) or NV (n=581) or mixed GA+NV in one or both eyes (n=176) were compared to controls using logistic regression for each variant separately (**Supplementary Figure S1)**.

Computing the genetic risk score

Based on the data from the 13 SNPs, we fit a multiple logistic regression model (**Figure 2**). The odds ratios in this model ranged from 1.070 to 4.063. This is, to our knowledge, the first study to report these 13 variants together in one multiple logistic regression accounting for other AMD risk variants. We computed a GRS for each individual as the sum of AMD risk increasing alleles weighted by the relative effect size of each SNP from the logistic model. We added the alpha estimate of -10.13 to center the GRS on zero for our study (see Methods). Cases had a significantly higher mean GRS (1.61, 95% CI: 1.53-1.69) compared to controls (-0.03, 95% CI: -0.12-0.06, p <0.01). The relative risk of AMD per GRS unit approximated by the OR was 2.72 (95% CI: 2.46-3.01). The mean GRS of our controls was slightly lower than the one for the HapMap data representing a general population (0.00, 95% CI: -0.14-0.14), which is in-line with our controls being selected for having no AMD.

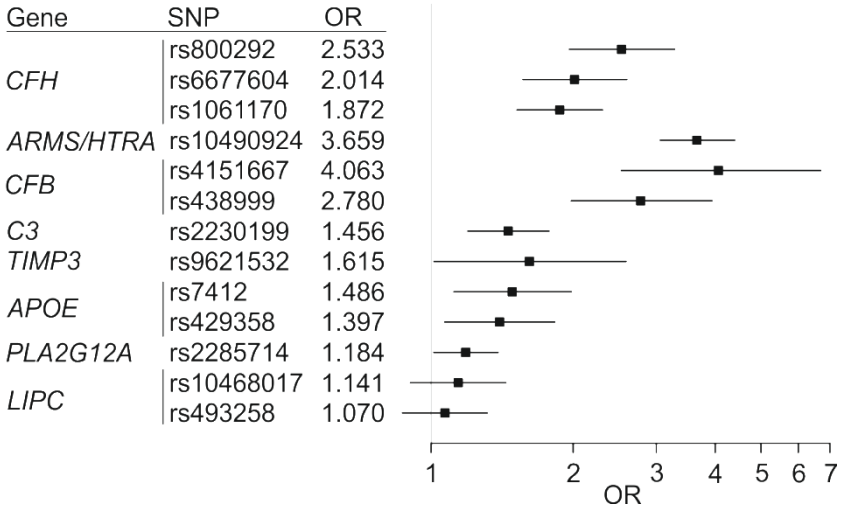


Figure 2. Risk estimates for each of thirteen AMD risk variants from eight gene loci. Odds ratios (OR) per risk allele were derived from multiple logistic regression models. Horizontal lines indicate 95% confidence intervals.

Table 3. Association results for the 13 known AMD associated variants in the lower Frankonian case-control study (986 cases, 796 controls) using single logistic regression

									Frequency of risk allele in			
Nearby gene(s)	**Marker**	**ID**	**Impact/effect of variant**	**Odds ratio**	**95% CI[1]**	**P-value[2]**	**Non risk allele**	**Risk allele[3]**	**Controls (N=796)**	**Cases (N=986)**	**AUC[4] of variant**	**correlation [5]**
CFH	rs1061170	1	p.Y402H	2.74	2.36-3.18	1.66E-45	T	C	0.365	0.600	0.676	
	rs800292	2	p.I62V	2.43	2.02-2.92	6.95E-23	A	G	0.761	0.888	0.606	0.150
	rs6677604	3	proxy for ΔCFHR3/CFHR1	2.19	1.82-2.64	1.42E-17	A	G	0.777	0.884	0.590	0.203
ARMS2	rs10490924	4	p.A69S	3.13	2.68-3.68	7.97E-54	G	T	0.189	0.441	0.684	
CFB	rs4151667	5	p.L9H	2.82	1.90-4.28	1.41E-07	A	T	0.951	0.982	0.530	
	rs438999	6	proxy for rs641153 (p.R32Q)	2.31	1.73-3.11	5.75E-09	C	T	0.915	0.962	0.542	0.01
C3	rs2230199	7	p.R102G	1.52	1.29-1.80	4.71E-07	G	C	0.175	0.245	0.556	
APOE	rs7412	8	p.R158C	1.41	1.12-1.80	0.003613	C	T	0.079	0.107	0.526	
	rs429358	9	p.C112R	1.35	1.09-1.69	0.006812	C	T	0.881	0.908	0.528	0.783
PLA2G12A	rs2285714	10	synonymous exonic, unknown	1.14	1.00-1.30	0.04839	C	T	0.409	0.443	0.523	
LIPC	rs493258	11	intergenic (36kb upstream)	1.18	1.04-1.35	0.01277	T	C	0.538	0.580	0.531	
	rs10468017	12	intergenic (46kb upstream)	1.26	1.08-1.46	0.002992	T	C	0.707	0.751	0.536	0.367
SYN3/TIMP3	rs9621532	13	intronic, unknown	1.58	1.09-2.30	0.01246	C	A	0.96	0.974	0.512	

[1] CI = confidence interval

[2] P-values were derived from a logistic regression model with one SNP as covariate.

[3] Risk allele is the allele that is associated with increased risk of AMD

[4] AUC = area-under-curve of the receiver-operating characteristic

[5] r^2 values representing the correlation with the first SNP in each gene/locus based on 1000 genomes data (build 1) or HapMap release 22 [87]

Good discriminative ability of the GRS

Computing the area-under-the-curve (AUC) of the receiver-operating characteristic for the 13-SNP GRS, we observed good ability to correctly classify those with and without the disease (AUC=0.820, **Figure 3**). We also computed the AUC per locus demonstrating that the impact by gene varied substantially, as expected. The three SNPs at the CFH locus alone (rs800292, rs1061170, rs6677604) showed the highest classification efficiency (AUC=0.710), followed by rs10490924 at ARMS2/HTRA1 (AUC=0.684), and the remaining variants (AUC from 0.512 to 0.571) (**Figure 3**).

Although we specifically avoided selecting the SNPs based on association in our own data set but rather from the literature, there could be a potential overestimation of the AUC. We estimated the effect sizes per variant from our data and used these as weights for the GRS.

To evaluate this potential over-estimation, we performed a sensitivity analysis via a cross-validation approach by repeated (i=2000) random sub-sampling with 2/3rd of the data for model building and 1/3rd for testing. The cross validated AUC of 0.813 (95% CI: 0.813-0.814) is close to the one described in our initial study (AUC=0.820).

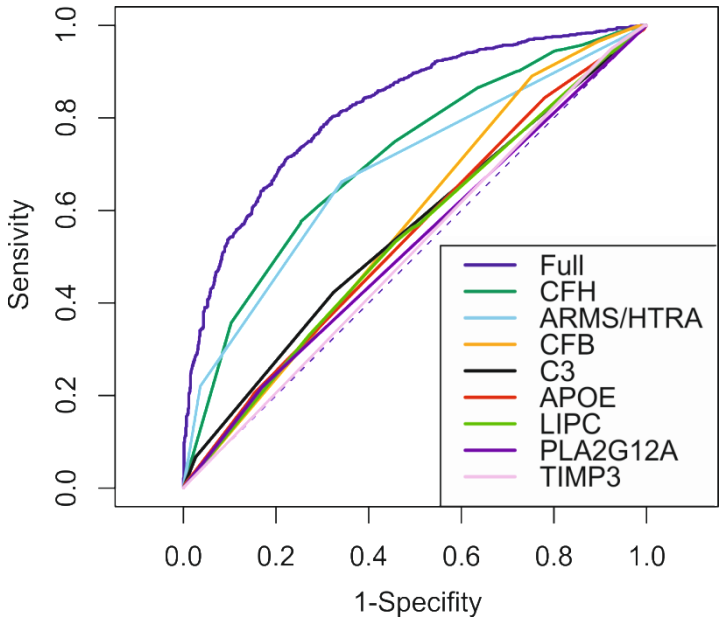


Figure 3. Area-under-the-curve of the receiver operating characteristic for the 13-SNP genetic risk score and by gene locus. Observed AUC was 0.820 and the locus-specific AUCs were 0.513, 0.524, 0.536, 0.547, 0.555, 0.571, 0.686 and 0.710 from bottom to top.

Developing a parsimonious genetic risk score model

We evaluated whether a parsimonious model based on our data could be developed. We thus explored several models by subsequently excluding the loci with the weakest AUC and found a model restricted to 10 variants with equally discriminatory ability (AUC 0.820) and equal model fit (R^2=0.247) (**Table 4**). This model could be of value for translational studies minimizing the genotyping burden. Whether this is specific to our data set or holds true for

other study populations needs to be evaluated further. It should be noted that all further analyses are based on the 13-SNPs-GRS.

Table 4. Model fit and discriminative accuracy of parsimonious models

Model[1]	**Variants**[2]	**R^2**	**AUC**
13-SNP model	*1,2,3,4,5,6,7,8,9,10,11,12,13*	0.2475	0.820
- TIMP3	*1,2,3,4,5,6,7,8,9,10,11,12*	0.2475	0.820
- PLA2G12A	*1,2,3,4,5,6,7,8,9,11,12*	0.2454	0.819
- APOE	*1,2,3,4,5,6,7,10,11,12*	0.2411	0.816
- LIPC[3]	*1,2,3,4,5,6,7,8,9,10*	0.2457	0.820

[1] SNPs from one additional locus at a time were omitted from the 13-SNP model by starting with the locus with the smallest risk

[2] Numbering corresponds to IDs in **Table 3**

[3] This model contained the least number of SNPs without compromising R^2 or AUC values

Distribution of the genetic risk score

The distribution of GRS for cases and controls as observed in our study is given in **Figure 4A**. To provide a more realistic view of the GRS distribution, the proportion of cases were weighted to reflect a general distribution. For this modeling, an AMD prevalence of 15% was assumed as reported for the general population aged >85 years [17,19,90] (**Figure 4B**). The derived GRS is comparable to the distribution estimated from individual HapMap data (**Figure 4B**).

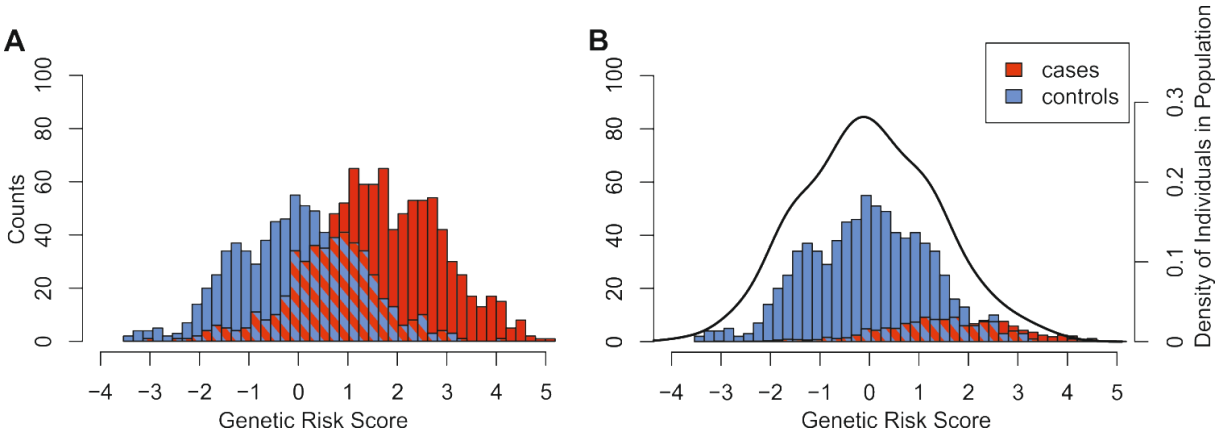


Figure 4. Genetic risk score distribution in the study population and in a modeled population. AMD cases are shown in red, controls in blue, while overlapping bars are shaded blue/red. (A) Genetic risk score distribution for cases (N=986) and controls (N=796) in the present study. (B) Counts of cases in (A) were scaled to represent 15% of the total population (assumed as AMD prevalence of the 85-90 year old general population). The density curve represents the risk score distribution in 381 European ancestry samples available through the 1000 Genomes Project (Release 20110521).

Genetic risk score by age groups, gender and AMD subtype

We further investigated differences of the GRS between age-groups (below or older than 75 years), men and women, or types of AMD (GA, NV, or mixed GA+NV) using a significance level of 0.05/3 to account for the three subgroup tests performed.

Significant differences in mean GRS were found between younger (1.87, 95% CI: 1.69-2.05) and older (1.45, 95% CI: 1.36-1.54) AMD cases ($p=8.7x10^{-5}$), but there was no difference between the age-groups among controls (p= 0.18). The OR per GRS unit was 3.06 (95%CI: 2.64-3.59) for younger and 2.71 (95% CI: 2.44-3.05) for older individuals. We also found that the AUC restricted to the younger subjects (cases and controls) was higher (0.852) than when only older subjects (cases and controls) were included in the calculations (0.809).

Cases with mixed GA+NV had a significantly higher mean GRS (1.87, 95% CI: 1.69-2.04) compared to NV cases (1.44, 95% CI: 1.34-1.55, $p=6.6x10^{-5}$). It was also higher when compared to GA cases (1.65, 95% CI: 1.48-1.83, p=0.03), although the latter was statistically not significant when applying a conservative Bonferroni-adjusted significance level of 0.05/3. The OR per GRS unit was also higher for mixed GA+NV cases (OR=3.79, 95% CI: 3.13-4.67) than for NV cases (OR=3.79, 95% CI: 3.13-4.67) or for GA cases (OR=2.84, 95% CI: 2.44-3.33). This effect appeared to be independent of age, since mean age in GA (78.8 years, 95% CI: 77.9-79.6), NV (78.5 years, 95% CI: 77.9-79.0) and mixed GA+NV (79.4 years,

95% CI: 78.4-80.3) was similar. There was no significant difference in the GRS means between men and women neither among cases nor among controls.

These subgroup analyses demonstrate a higher genetic risk of the younger AMD patients compared to the older patients as well as a higher genetic risk for those with mixed late-stage manifestations (GA+NV) when compared to NV or GA alone.

Genetic risk groups and relative risk estimates

To establish a classification scheme, we formed five equally sized intervals for the GRS spectrum (≤-1.79, (-1.79)-(-0.05), (-0.05)-1.70, 1.60-3.44, and >3.44; **Table 5**). The highest GRS category (no. 5) contained 7.92% of AMD cases, but only 0.13% of controls. In contrast, the two lowest GRS categories (nos. 1+2) jointly contained only 9.8% of cases, but 48.7% of controls. According to the HapMap data reflecting a general population, the proportion of subjects in the two lowest risk groups combined was 48.9% and 0.5% in the highest risk group. This is consistent with the general population being a mixture of mostly controls and only few cases. The relative risks were approximated as ORs for each GRS category using the middle category (no. 3) as reference (**Table 5**). It can be seen that the OR is dramatically increased for category four (OR=5.44, 95% CI=4.03 – 7.46) and even more for category five (OR=64.00, 95% CI=14.11 – 1132.96). The odds ratios are substantially decreased for categories two and one (OR=0.22, CI=0.17-0.29 and OR=0.12, 95% CI=0.05-0.24) compared to the reference. Thus, these GRS categories can effectively describe genetic risk groups for AMD.

Due to the substantial differences found in mean GRS for younger compared to older cases (see above), we derived these ORs also separately by age-group. To avoid scarce data, risk group one and two as well as four and five were combined to a low and a high risk group, respectively (**Table 5**). This highlights the higher genetic relative risk for AMD when restricting the analysis to the younger (OR =12.66, 95% CI: 6.76-25.65) compared to the older (OR=5.18, 95% CI: 3.70-7.38) subjects. Although the 95% confidence intervals overlap slightly, we observed a significant difference (p=0.0194).

Modeled absolute risk for late-stage AMD

To reflect the anticipated situation in the general population and to compute the absolute risk of AMD per GRS group, we computed the fraction of late-stage AMD cases per GRS category by (i) utilizing the fraction of cases and controls as observed in each GRS category (**Table 5**) and (ii) weighting the fraction of cases assuming various AMD prevalences (1%-15%). The fraction of cases and the fraction of subjects of the modeled general population (also for comparison in the HapMap sample) by GRS category are shown in **Table 6**. The fraction of late-stage AMD in the highest GRS group (absolute AMD risk) ranged from 38.6-91.7% depending on the assumed AMD prevalence which were chosen to correspond to the various age-groups as reported [17,19,90]. For example, in a general population with an AMD prevalence of 10% approximately 90% of the persons in the highest GRS group are expected to be affected by late-stage AMD. Consequently, the genetic relative risk for subjects in the highest GRS group (compared to the middle GRS group) is higher for younger compared to the older AMD cases. However, the absolute risk of AMD among subjects in the highest GRS group is higher for the older population due to the higher AMD prevalence among the older persons.

We again adopted the same cross-validation approach to compute absolute risks since the effect sizes of the variants in the GRS model, on which the absolute risk estimates are based, were estimated from our study data. This approach yielded overall similar estimates (**Supplementary Table S2**).

Table 5. Five genetic risk groups and relative risk of AMD (ORs and 95% confidence intervals)

	Genetic risk groups				
GRS category	**1**	**2**	**3**	**4**	**5**
Sample size	**N=63**	**N=417**	**N=761**	**N=450**	**N=79**
GRS interval	≤-1.79	]-1.79,-0.05]	]-0.05,1.70]	]1.70,3.44]	>3.44
Cases [%]	0.81	9.00	42.5	39.7	7.92
Cases <75 years [%]	1.70	6.20	34.0	42.3	15.8
Cases >75 years [%]	0.54	9.96	45.2	38.9	5.38
Controls [%]	6.99	41.7	43.6	7.50	0.13
Frequency in HapMap[1]	8.92	40.2	41.2	9.18	0.53
OR (95% CI)	0.12 (0.05-0.24)	0.22 (0.17-0.29)	reference	5.44 (4.02-7.46)	64.00 (14.11-1131.96)
GRS categories	**low (1+2)**		**3**	**high (4+5)**	
Sample size	**N=480**		**N=761**	**N=529**	
GRS interval	≤-0.05		]-0.05,1.70]	>1.70	
OR (95% CI)	0.21 (0.16-0.27)		reference	6.41 (4.76-8.76)	
<75 years: OR (95% CI)	0.19 (0.10-0.33)		reference	12.66 (6.76-25.65)	
>=75 years: OR (95% CI)	0.22 (0.16-0.29)		reference	5.18 (3.70-7.38)	

[1] Fraction of individuals in 1000 Genome Project European Ancestry Samples residing in risk groups

Table 6.Absolute risks for AMD by modeling a general population for various prevalences of AMD (reflecting various age-groups)

	Modeled prevalence (age-group [yrs])[1]	**Absolute risk of AMD by genetic risk group [%]**				
		1 (low)	**2**	**3**	**4**	**5 (high)**
GRS interval		≤-1.79	]-1.79,-0.05]	]-0.05,1.70]	]1.70,3.44]	>3.44
% cases, modeled general population						
	1% (65-69)	0.12	0.22	0.97	5.08	38.6
	2.5% (70-74)	0.30	0.55	2.44	12.0	61.5
	5% (75-79)	0.61	1.13	4.87	21.8	76.6
	10% (80-84)	1.30	2.40	9.80	37.0	87.4
	15% (>85)	2.00	3.70	14.7	48.3	91.7
% subjects, modeled general population						
	1%	6.84	40.9	43.6	8.31	0.32
	2.5%	6.68	40.1	43.8	8.50	0.34
	5%	6.69	40.1	43.6	9.12	0.52
	10%	6.38	38.5	43.5	10.7	0.91
	15%	6.10	36.8	43.5	12.3	1.30
% subjects, HapMap population[2]		8.92	40.2	41.2	9.18	0.53

[1] Approximate age-groups corresponding to the modeled prevalences for 65 and 79 years [17,19] and for those above 80 years [90].

[2] see **Table 5.**

2.3 Discussion

Based on a genetic risk score including 13 reported SNPs from eight established AMD gene loci, we propose a five-category classification system that effectively differentiates subjects with high or low genetic risk. With this, we extend on earlier efforts to predict the genetic risk for late-stage AMD [42,50,52,56]. Seddon *et al.* described a risk score model for six genetic variants in four loci also including environmental factors like BMI, smoking, age and diet (sample size was 1.446 individuals of which 279 progressed to AMD) [50]. Similarly, a study from Gibson *et al.* included 470 cases and 470 controls and reported an AUC of 0.83 (95% CI 0.81 to 0.86) using six SNPs in four loci and two environmental factors [52]. A study by Spencer *et al.* investigated one variant in each of four loci as well as age and smoking as environmental factors and found an AUC of 0.84 (95% CI: 0.81-0.88) [56]. Jakobsdottir *et al.* reported an AUC of 0.79 based on one SNP in each of three loci [42]. This study consisted of around 1.000 family-based cases and 429 controls as well as a case-control study with 187 cases and 168 controls. We evaluated 13 SNPs from 8 AMD loci in a well characterized and well powered case-control study and observed an AUC of 0.820, which is sufficient to classify AMD patients and controls into high risk and low risk groups [85]. Our study has not contributed to the identification of any of the 13 SNPs as AMD risk-increasing variants and would thus not be subject to winner's curse regarding the effect size. To our knowledge, this is a first study to include most of the currently known genetic loci for their value to predict late-stage AMD risk in a study that is independent of the identification of any of these loci.

Interestingly, we find a higher relative risk of the CFB SNP rs4151667 compared to CFH and ARMS2/HTRA1 risk-increasing SNPs particularly in the multivariable logistic regression model. This can also be seen in a previously published study (Seddon *et al.*, Table 4) [26], although it needs to be noted that the models used in our and the published study differ in the sense that ours considers exclusively genetic factors while the other work largely focused on non-genetic factors. The smaller allele frequency of the CFB SNP (1.8% in our cases, 6.7% in the European ancestry 1000G individuals) compared to SNP frequencies in CFH and ARMS2/HTRA1 results in a reduced power to detect association and may explain why CFB SNP rs4151667 was not among those detected first by AMD GWAS.

As expected, the mean GRS was significantly higher in cases when compared to controls. Importantly, patients with late-stage AMD diagnosed at an earlier age had a significantly higher mean GRS than individuals that developed AMD later in life. This strongly suggests

that genetic predisposition influences disease onset, which is also reflected in the higher relative AMD risk for younger subjects with an OR of 12.66 (95% CI: 6.76-25.65) when compared to older individuals with an OR of 5.18 (95% CI: 3.70-7.38). The mean genetic risk score in our control group was slightly lower but similar to the mean score in the HapMap sample (including a total of 381 European subjects from CEU, GBR, IBS, TSI and FIN, 1000 Genomes Project (Release 20110521, http://www.1000genomes.org, accessed 2 May 2012).). The slight discrepancy would be in-line with the fact that our controls were specifically selected to reveal no signs of early or late-stage AMD.

Limitations of our study for risk prediction should be acknowledged. First, the analysis was based on a case-control study, which has no element of a prospective study or a nested case-control study. The controls were often spouses of AMD patients and thus non-genetic risk factors could not be studied due to the known similarities among spouses regarding life style factors. However, our AMD patients were virtually incident AMD cases and thus the age at study entry is likely the age-at-diagnosis and the best possible proxy for age-of-onset (allowing for a delay of about 1-2 years between onset and diagnosis). In a case-control setting, absolute risk or positive/negative predictive values cannot be derived without making assumptions on the overall AMD prevalence, which a prospective cohort study could estimate directly. Thus, the predictive ability of the risk score groups greatly depends on those assumptions. Second, it might be considered a limitation but also a strength that our study included exclusively late-stage AMD with NV or GA in one or both eyes as well as highly-matched controls with no signs of early or late-stage AMD in any eye. A strength as our data might exhibit less disease misclassification than other studies, but a limitation as the genetic relative risk could be overestimated if the genetic risk is larger for subjects with both eyes affected than for those with only one affected eye. Third, we had no independent and equally well characterized data set available to separate model building from testing although this is also the case for all other studies published on AMD risk score model building [42,50,52]. Only one study [56] reported a small replication study. We avoided selecting SNPs for our model based on association signals in our own data but rather selected SNPs from the literature. However, the SNP-specific effect sizes utilized as weights in the genetic risk score computation were still estimated in our data set. Thus, estimations of AUC or absolute risk in the same data could lead to a slight over-estimation of risk. We therefore adopted a cross-validation approach as sensitivity analysis, which did not provide evidence of remarkable over-estimation.

The highest genetic risk group of our proposed five-category classification scheme can effectively identify subjects at high risk for AMD. The specificity in this risk group was 99.9% (95% CI: 99.3%-100%). For example, our data and model suggest that 87.4% of subjects testing positive at some time in life for a high genetic risk are likely to develop AMD in their mid-eighties (positive predictive value). Thus, this group of individuals could greatly profit from a sight-saving prevention or early intervention program while only 13% of (false-positive) subjects would be alarmed and treated unnecessarily. However, still a large number of cases would be missed if this was established as a screening method (sensitivity 8.0% (95% CI: 6.5%-9.9%), i.e. 92% of all AMD cases would not be found in the highest risk group). Also individuals in the second highest risk group could possibly profit from early intervention, which would increase sensitivity to 47.6% and decrease specificity to 91.2%. However, this would only be acceptable, if the prevention/intervention is not harmful to the 59.9% of subjects treated and alarmed unnecessarily (40.1% positive predictive value). These numbers are well in the range of established screening tests, e.g. for prostate cancer by prostate specific antigen (PSA) (positive predictive value = 25.1%, sensitivity = 72.1%, specificity = 93.2%, [91]), albeit with a higher predictive value at the cost of reduced sensitivity. Abnormal levels of PSA are detected in about 10% of the male population, which is comparable to the coverage of high risk group four and five [91]. Offering an effective prevention program to individuals in the highest AMD risk group (approximately 400,000 individuals in Germany alone), almost 10% of incident late-stage AMD could be avoided. If individuals in risk groups four and five are included (about 10% of the general population), up to 50% of future AMD patients could be addressed.

So far, only the progression of the neovascular complications in AMD can be slowed by treatment [92]. If disease progression to an advanced neovascular form is detected early in high risk patients, immediate intervention might prove essential to sustain full vision for a more extended time. Accordingly, high risk individuals could be advised to seek clinical follow-ups more frequently and could also benefit from dietary recommendations, including the intake of antioxidants [22] or omega-3 fatty acids [93,94]. Identification of individuals at high risk for developing AMD may also help to include defined candidates in clinical AMD trials and thus may allow a better assessment of therapeutic effects.

In conclusion, our study provides a genetic risk score for late-stage AMD from a well characterized case-control study emphasizing the large proportion of disease explained by genetic markers particularly for younger subjects. We propose a classification scheme to

identify subjects at high or low genetic risk that might be suitable for risk stratification in therapy studies or genetic screening once preventive treatment is available.

2.4 Materials and Methods

Ethics statement

This study followed the tenets of the declaration of Helsinki and was approved by the Ethics Review Board at the University of Würzburg, Germany. Informed written consent was obtained from each patient after explanation of the nature and possible consequences of the study.

The study subjects

The case-control sample includes 986 AMD patients and 796 controls recruited from the Lower Frankonian area at the University Eye Clinic of Würzburg, Germany [35]. Controls were often unaffected spouses or nonrelated acquaintances of cases of similar age as the patient. All patients and controls were examined by a trained ophthalmologist (CNK). Stereo fundus photographs were graded according to standardized classification systems as described previously [37,95,96]. Only patients with severe forms of AMD (GA or NV) in at least one eye and signs of early AMD (e.g. large soft drusen) in the other eye were included. The patients were divided into three subgroups according to their type of late-stage AMD: patients with GA in the severe eye, patients with NV in the severe eye and patients that had either GA in one eye and NV in the other eye or that showed both late-stage forms in the same eye (mixed GA+NV). Mean age in cases was 78.7 (±6.5) years and 78.3 (±5.1) in controls. A total of 34.1% of cases and 39.1% of controls were male. Study characteristics are summarized in **Table 2**.

Genotyping

Genomic DNA was extracted from peripheral blood leukocytes according to established protocols. Genotyping of SNPs was achieved by direct sequencing, restriction enzyme digestion of PCR products, TaqMan SNP Genotyping (Applied Biosystems, Foster City, USA) or primer extension of multiplex PCR products with detection of the allele-specific extension products by the matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry method (Sequenom, San Diego, USA) (**Supplementary Table S3**). Direct sequencing was performed with the Big Dye Terminator Cycle Sequencing Kit Version 1.1 (Applied Biosystems, Foster City, USA) according to the manufacturer's instructions.

Reactions were analyzed with an ABI Prism Model 3130xl Sequencer (Applied Biosystems). TaqMan Pre-Designed SNP Genotyping Assays (Applied Biosystems) were performed according to the manufacturer's instructions. Additionally, some variants were genotyped by PCR followed by restriction enzyme digestion (New England Biolabs, Ipswich, USA) and subsequent restriction fragment length analysis. The c.del443ins54 variant in the 3'-region of the *ARMS2* locus was genotyped by a single PCR with oligonucleotide primers 5'-ACTCATCACGTCATCACCAAT-3' and 5'-CTCTCTGCAGCCCTCATTTG-3' resulting in distinct fragment sizes due to the presence or absence of the deletion/insertion polymorphism.

Estimating genetic risk and model fit

Genotypes were coded as the number of AMD risk increasing alleles (0, 1, and 2). Logistic regression analyses were carried out using the R software [97]. Odds ratios (OR) per risk allele and 95% confidence intervals (95% CI) were calculated from the estimated beta-coefficients to derive an approximate relative risk. The goodness-of-fit of each model was assessed by calculating McFaddens pseudo R^2 [98], which however, does not reflect the variance explained by the model [99].

Computing the genetic risk score

Based on the intercept "a" and the single-SNP beta-coefficients estimated using the logistic regression model including all SNPs at once, the genetic risk score (GRS) was calculated as

$$\mathrm{GRS} = a + \sum_{i=1}^{k} b_i * x_i \qquad (1)$$

with k being the number of SNPs in the model and x_i the genotype of the ith SNP. Here, "a" denotes a constant that centers the risk score distribution around zero and b_i relates to the ith variant. The odds ratio of the effect of the ith variant is thus given by $\exp(b_i)$ [50,81,84]. The mean GRS by age-group, sex, or AMD subtype were compared based on the independent samples t test using the R software [97] and differences were considered as significant, if $P < 0.05/3$ accounting for the three comparisons performed.

Assessing the discriminative ability

To estimate the ability of a potential genetic screening test to discriminate between AMD cases and healthy subjects, we computed the receiver-operating-characteristic (ROC) curve. This involves ranking all subjects according to their GRS starting with the smallest, computing sensitivity and specificity at each possible GRS cut-off, and plotting sensitivity

versus 1-specificity. The area-under-the-curve (AUC) is a measure of how well the GRS cut-offs can separate AMD cases from controls. We used the package EPICALC [100] for AUC computations and forest plots were generated with RMETA [101].

Internal validation by cross-validation

Although we have not selected the SNPs into the model based on their association in our data set but rather with information from the literature, there is a potential overestimation of the AUC due to the fact that we used the SNP effect sizes to weigh the risk alleles when computing the GRS. Thus, we conducted a sensitivity analysis using a cross-validation approach to derive AUC estimates that are not subject to this bias to compare with the original data AUC. We randomly assigned 2/3rd of the data to the model building (to compute the effect sizes and thus establish the GRS model) and 1/3rd of the data to testing (to compute the AUC and positive predictive values) [102,103]. We repeated this 2000 times and computed the average AUC as an unbiased estimate.

Modeling of the absolute risk by GRS group

In order to derive the fraction of cases in the five GRS categories as expected in the general population (corresponding to the absolute AMD risk) from the number of cases (N_cases=986) and controls (N_controls=796) in our case-control study, we weighted the number of AMD cases in our study by

$$\text{weight} = \frac{\text{prevalence}*N\_controls}{(1-\text{prevalence})*N\_cases} \qquad (2)$$

where prevalence denotes the fraction of AMD cases in the general population, that we chose to reflect previously reported prevalences of AMD in the various age groups (65-69 years: 1%, 70-74 years: 2.5%, 75-79 years: 5%, 80-84 years: 10% and >85 years: 15%) [17,19,90]. These were also used to compute positive and negative predictive value for the highest GRS category as a screening test for AMD. The cross-validation approach described above was also adopted for a sensitivity analysis to compute unbiased absolute risk.

3. A circulating microRNA profile is associated with late-stage neovascular age-related macular degeneration

This chapter is identical to the following publication:

Grassmann F, Schoenberger PGA, Brandl C, Schick T, Hasler D, Meister G, Fleckenstein M, Lindner M, Helbig H, Fauser S, Weber BHF (2014). **A Circulating MicroRNA Profile Is Associated with Late-Stage Neovascular Age-Related Macular Degeneration.** *PLoS One* **9:** e107461

Abstract

Age-related macular degeneration (AMD) is the leading cause of severe vision impairment in Western populations over 55 years. A growing number of gene variants have been identified which are strongly associated with an altered risk to develop AMD. Nevertheless, gene-based biomarkers which could be dysregulated at defined stages of AMD may point toward key processes in disease mechanism and thus may support efforts to design novel treatment regimens for this blinding disorder. Circulating microRNAs (cmiRNAs) which are carried by nano‑sized exosomes or microvesicles in blood plasma or serum, have been recognized as valuable indicators for various age-related diseases. We therefore aimed to elucidate the role of cmiRNAs in AMD by genome-wide miRNA expression profiling and replication analyses in 147 controls and 129 neovascular AMD patients. We identified three microRNAs differentially secreted in neovascular (NV) AMD (hsa-mir-301a-3p, $p_{corrected} = 5.6*10^{-5}$, hsa-mir-361-5p, $p_{corrected} = 8.0*10^{-4}$ and hsa-mir-424-5p, $p_{corrected} = 9.6*10^{-3}$). A combined profile of the three miRNAs revealed an area under the curve (AUC) value of 0.727 and was highly associated with NV AMD ($p = 1.2*10^{-8}$). To evaluate subtype-specificity, an additional 59 AMD cases with pure unilateral or bilateral geographic atrophy (GA) were analyzed for microRNAs hsa-mir-301a-3p, hsa-mir-361-5p, and hsa-mir-424-5p. While we found no significant differences between GA AMD and controls neither individually nor for a combined microRNAs profile, hsa-mir-424-5p levels remained significantly higher in GA AMD when compared to NV ($p_{corrected} < 0.005$). Pathway enrichment analysis on genes predicted to be regulated by microRNAs hsa-mir-301a-3p, hsa-mir-361-5p, and hsa-mir-424-5p, suggests canonical TGFβ, mTOR and related pathways to be involved in NV AMD. In addition, knockdown of hsa-mir-361-5p resulted in increased neovascularization in an *in vitro* angiogenesis assay.

3.1 Introduction

Age-related macular degeneration (AMD) is a highly prevalent cause of severe vision impairment among people aged 55 years and older [104]. It is a degenerative disorder of the central retina involving predominantly the rod photoreceptors, the retinal pigment epithelium (RPE), Bruchs membrane and the underlying choriocapillaris [40]. The disease aetiology is complex and is influenced by a combination of multiple genetic susceptibility factors and environmental components.

An early sign of AMD is the appearance of drusen, yellowish extracellular deposits of protein and lipid material within and beneath the RPE. Advanced AMD manifests essentially as two distinct late-stage lesions – geographic atrophy (GA) and neovascular (NV) AMD. GA occurs in up to 50% of cases and is clinically defined as a discrete area of RPE atrophy with visible choroidal vessels in the absence of neovascularization in the same eye [2-5]. It may or may not involve the fovea. NV AMD describes the development of new blood vessels beneath and within the retina and is characterized by serous or hemorrhagic detachment of either the RPE or the sensory retina, the presence of subretinal fibrous tissue and eventually widespread RPE atrophy. Progression to visual loss can be rapid in NV AMD [104].

The precise aetiology of AMD is still not fully understood, although risk factors such as age, smoking, and genetic components are known to strongly contribute to disease development [40]. In Western societies, AMD reveals an age-dependent prevalence of almost 1 in 5 people aged 85 and above [17,90,105]. Across a number of epidemiological studies, smoking has consistently been associated with increased risk of developing advanced AMD with an estimated odds ratio of approximately 2 [106]. The exact mechanism, however, by which smoking affects the retina is unknown. Twin studies and familial aggregation studies suggested a significant genetic contribution of up to 70% in disease risk [72]. Subsequently, several genes have been implicated in AMD pathology by candidate gene studies as well as genome wide association studies. Genetic variants in complement factor H (CFH) and ARMS2/HtrA Serine Protease 1 (HTRA1) were found to be strongly associated with odds ratios over 2.5 per risk allele. In addition, multiple medium to low effect size gene variants were discovered in a large number of loci across the genome. A recent meta-analysis of genome wide association studies found a total of 19 independently associated loci by comparing over 17,000 cases and 60,000 controls [33].

The combined effect of the major risk variants on AMD was estimated by modelling risk scores [58]. The multiple logistic regression model was found to have an area under the curve (AUC) of about 82%, which is suitable for classifying individuals in high and low risk groups. Accordingly, roughly 50% of AMD cases and 50% of healthy controls can now reliably be predicted. However, a large proportion of AMD cases do not have the expected genetic risk profile despite their given disease status. Consequently, other components, genetic or environmental, may influence disease development. This makes it crucial to identify these components possibly by defining disease biomarkers correlating with the underlying genetic or environmental factors and eventually reflecting a defined disease stage.

Recently, circulating microRNAs (cmiRNAs) were found in blood plasma or blood serum where they are carried by nano‐sized exosomes or microvesicles [107,108]. Origin and effects of these cmiRNAs are unclear although some studies suggested functional involvement in cell-to-cell signalling [109]. In general, cmiRNAs are potential biomarkers which can be used for diagnostics and prognostics of human diseases [110]. Additionally, synthetic microRNAs in artificial exsosomes could be applicable for therapeutic approaches by modulating cmiRNA levels.

In this study, we aimed to elucidate the role of cmiRNAs in AMD and performed a genome-wide expression profiling in patients affected by late stage neovascular manifestation. Such analyses provide a promising approach to define biomarkers for AMD which could be helpful to identify as of yet unknown gene targets involved in defined aspects of AMD pathology. Such biomarkers could also serve as the long sought-after variable needed to monitor treatment effects in future clinical trials for AMD.

3.2 Results

Study design

We applied a three stage design to identify significantly associated cmiRNAs. First, RNASeq was performed to screen for miRNA candidates in 9 cases and 9 controls from the Regensburg study. The cmiRNAs with a nominal significance of $p > 0.1$ were then validated in an unrelated set of 45 NV cases and 68 controls from the Regensburg study (**Table 7**). Finally, candidate cmiRNAs with a nominal significant association ($p < 0.05$, adjusted or unadjusted for glaucoma) and an odds ratio above 2 or below 0.5 were then replicated in a population based study (Cologne study, **Table 7**) consisting of 75 NV cases and 70 controls. In total, the combined study included 129 patients with NV AMD and 147 AMD-free controls (**Table 7**). Additionally, 59 AMD patients with pure GA were assessed for candidate cmiRNAs to test for specificity of the findings in NV AMD.

Table 7. Summary characteristics of the study

	Regensburg	Bonn	Cologne
Study type	Case/Control	Case/Control	Population based
Number of individuals	131	18	186
Controls	77	0	70
Cases	54	18	116
Geographic atrophy	0	18	41
Neovascular AMD	54	0	75
Mean age cases (S.D.) [years]	75.15 (6.75)	74.60 (8.70)	80.22 (9.24)
Mean age controls (S.D.) [years]	73.26 (8.00)	-	78.44 (8.76)
Female cases [%]	59.3	61.1	56.9
Female controls [%]	54.5	-	55.7
Glaucoma in cases [%]	11.1	5.5	NA
Glaucoma in controls[%]	83.1	-	NA

Identification of cmiRNAs in NV AMD (discovery study)

To search for candidate cmiRNAs, we first performed next-generation sequencing of cmiRNAs extracted from plasma of 9 AMD NV cases and 9 matched controls. Overall, in the 18 samples we identified 203 different cmiRNA species. Of these, 10 cmiRNAs were significantly associated with late-stage NV AMD ($p_{uncorrected} < 0.1$) (**Table 8**).

Table 8.Association of circulating microRNAs with AMD in the Regensburg discovery study (9 NV cases and 9 controls)

microRNA	uncorrected p-value	mean cases (95%CI)*	mean controls (95%CI)*
hsa-miR-142-5p	0.012	1.21 (1.14-1.28)	1.00 (0.93-1.07)
hsa-miR-192-5p	0.010	1.29 (1.20-1.38)	1.00 (0.91-1.09)
hsa-miR-194-5p	0.028	1.28 (1.19-1.38)	1.00 (0.89-1.11)
hsa-miR-26a-5p	0.082	0.90 (0.83-0.96)	1.00 (0.94-1.06)
hsa-miR-301a-3p	0.084	0.83 (0.72-0.93)	1.00 (0.90-1.10)
hsa-miR-335-5p	0.094	1.34 (1.17-1.50)	1.00 (0.84-1.16)
hsa-miR-361-5p	0.056	0.74 (0.56-0.91)	1.00 (0.85-1.15)
hsa-miR-424-5p	0.028	0.52 (0.30-0.73)	1.00 (0.84-1.16)
hsa-miR-4732-5p	0.086	1.24 (1.12-1.36)	1.00 (0.88-1.12)
hsa-miR-505-5p	0.048	1.29 (1.13-1.44)	1.00 (0.85-1.15)

*95% confidence intervals

Circulating miRNAs associated with NV AMD (replication study)

To replicate the initial findings, qRT-PCR was performed for the significant 10 cmiRNAs in 113 samples consisting of 45 NV AMD cases and 68 controls. Three cmiRNAs were identified (hsa-mir-301-3p, hsa-mir-361-5p, and hsa-mir-451a-5p) which showed (1) an association signal in the same direction as in the discovery study, (2) an odds ratio over 2 or under 0.5 and (3) an uncorrected (one-sided) p-value below 0.1. These three cmiRNAs showed reduced levels in the serum of CNV cases compared to AMD free controls. The association was robust also when adjusting for covariates such as age, gender, smoking (measured in packyears), genetic risk score (GRS) or levels of the housekeeping cmiRNA hsa-mir-451a-5p (**Table 9**). Of note, two cmiRNAs (hsa-mir-301-3p and hsa-mir-361-5p) were strongly confounded by glaucoma disease status and showed stronger association signals when adjusting for glaucoma.

Table 9. Sensitivity analysis in the Regensburg study by multiple logistic regression models

covariate	hsa-miR-301a-3p	hsa-miR-361-5p	hsa-miR-424-5p
none	0.31 (0.10-0.86)*	0.50 (0.19-1.27)	0.28 (0.12-0.59)*
age [years]	0.33 (0.13-0.92)*	0.49 (0.19-1.26)	0.27 (0.12-0.59)*
packyears [years]	0.31 (0.10-0.86)*	0.50 (0.19-1.27)	0.27 (0.12-0.59)*
gender	0.29 (0.09-0.82)*	0.48 (0.18-1.23)	0.28 (0.12-0.59)*
genetic risk score	0.38 (0.10-1.33)	0.53 (0.13-2.12)	0.21 (0.07-0.56)*
glaucoma	0.15 (0.04-0.54)*[1]	0.23 (0.06-0.78)*[1]	0.24 (0.08-0.65)*
hsa-mir-451a-5p	0.38 (0.12-1.08)	0.68 (0.24-1.85)	0.35 (0.14-0.78)*

[1] strong increase in association signal by adjusting for glaucoma as a covariate
* statistically significant association ($p < 0.05$)

Circulating miRNAs hsa-mir-301-3p, hsa-mir-361-5p, and hsa-mir-424-5p were then analyzed by qRT-PCR in an additional replication study (Cologne study) consisting of 75 NV cases and 70 controls. In concordance with the Regensburg study, we also found reduced levels of those three cmiRNAs in NV cases compared to controls in the Cologne study. The results of the two replications were pooled and jointly analyzed (**Figure 5**, **Supplementary Table S4**). We found raw (one-sided) p-values of 2.78×10^{-7}, 4.09×10^{-6}, and 4.75×10^{-5} for hsa-mir-301-3p, hsa-mir-361-5p, and hsa-mir-424-5p, respectively. The p-values were adjusted by a conservative Bonferroni correction, assuming 203 statistical tests based on the number of microRNAs detected in the serum of cases and controls. After correction, the p-values for hsa-mir-301-3p, hsa-mir-361-5p, and hsa-mir-424-5p were 5.63×10^{-5}, 8.03×10^{-4}, and 9.64×10^{-3}, respectively. A cmiRNA profile including hsa-mir-301-3p, hsa-mir-361-5p, and hsa-mir-424-5p was significantly associated with AMD in the combined study (129 NV AMD versus 147 controls, $p = 1.17 * 10^{-8}$) as well as in the Cologne study alone (75 NV cases and 70 controls, $p = 2.43 * 10^{-5}$).

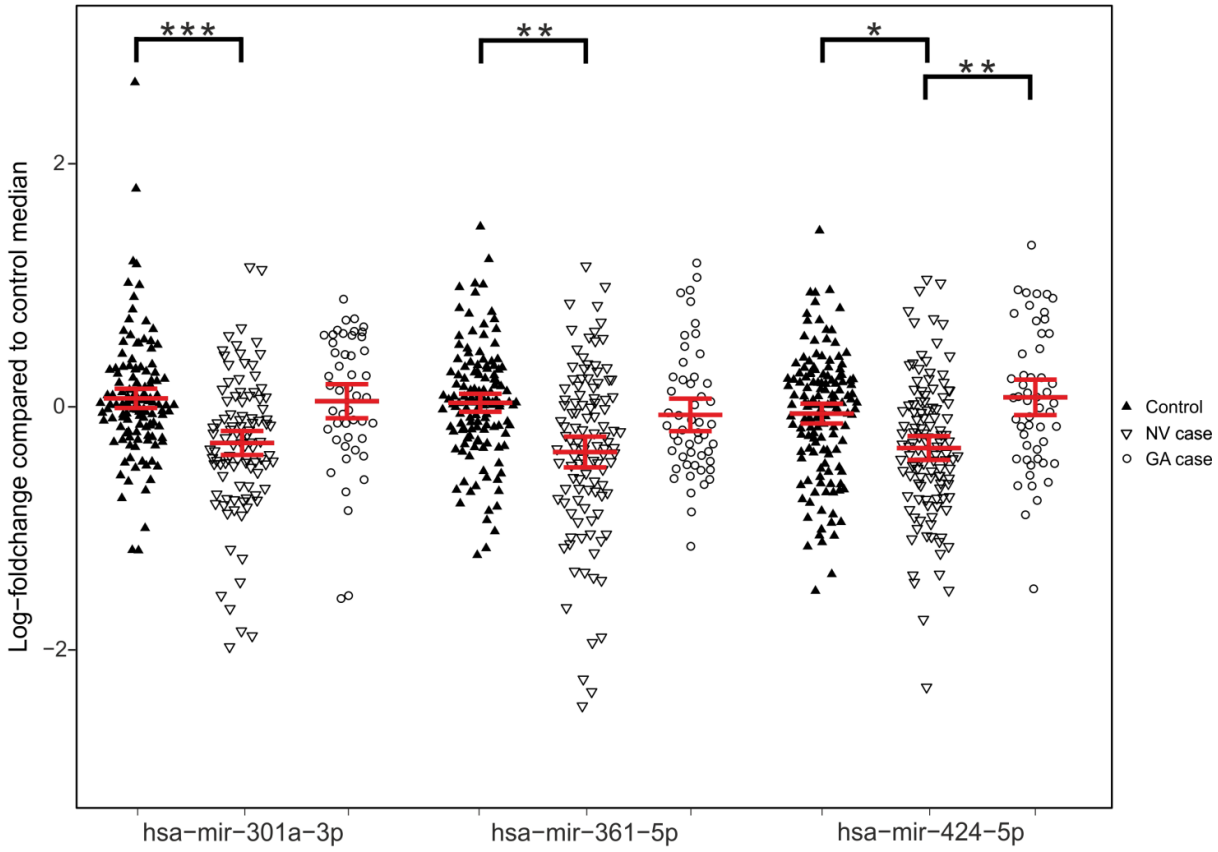


Figure 5. Expression analysis of three cmiRNAs (hsa-mir-301a-3p, hsa-mir-361-5p and hsa-mir-424-5p) in 129 NV AMD cases, 59 GA AMD cases and 147 healthy controls. Expression values for all samples were normalized by the median expression value in controls. Broad horizontal bars represent the mean value in each group (NV cases, GA cases or controls) for each cmiRNA. Smaller horizontal bars represent the 95% confidence intervals for each mean (see **Supplementary Table S4**). Significant differences between means are indicated by asterix. * = $p_{corrected}$< 0.05; ** = $p_{corrected}$< 0.005; *** = $p_{corrected}$< 0.0005

Testing of cmiRNAs specificity in NV and GA AMD

The expression of hsa-mir-301-3p, hsa-mir-361-5p, and hsa-mir-424-5p was analyzed by qRT-PCR in the serum of 59 GA AMD patients from the Cologne and Bonn study and compared to all controls (**Figure 5, Supplementary Table S4**). There was no statistically significant association of cmiRNA levels with GA compared to controls ($p_{corrected}$> 0.05). We also found no significant association of the cmiRNA profile including hsa-mir-301-3p, hsa-mir-361-5p, and hsa-mir-424-5p with GA AMD versus controls (p = 0.084).

Circulating miRNA hsa-mir-424-5p showed significantly higher levels in GA compared to NV ($p_{corrected}$< 0.005), while hsa-mir-301-3p and hsa-mir-361-5p were not significant ($p_{corrected}$>0.05).

Pathway analysis

Pathway enrichment analysis was performed for 3,516 genes predicted by microT-CDS to be regulated by either hsa-mir-301-3p, hsa-mir-361-5p, or hsa-mir-424-5p. A total of 410 genes was predicted to be regulated by at least two of the three cmiRNAs and 35 genes were regulated by the three cmiRNAs jointly (**Supplementary Figure S2**). Evaluation with miRSystem implicated the canonical TGF-β and mTOR pathways as well as related pathways such as WNT signaling, focal adhesion, neutrophin signaling and insulin metabolism as the top regulated pathways. This is in agreement with the results of mirPATH v2.0, which implicated mTOR (KEGG ID: hsa04150, $p < 10^{-13}$) and TGF-β pathways (KEGG ID: hsa04350, $p < 10^{-14}$) as top regulated pathways (**Table 10**).

Table 10. **Pathway enrichment analysis performed with miRSystem and mirPATH2**

	genes observed / genes in pathway		**genetic association reported[1]**
Canonical pathway (ID[2])	miRSystem	mirPATH2	
TGF-β signaling (hsa04350)	25/84	35/80	TGFBR1 [33]
mTOR signaling (hsa04150)	16/52	33/60	VEGFA [84]
Neutrophin signaling (hsa04722)	38/127	-	-
WNT signaling (hsa04310)	48/150	-	
Focal adhesion (hsa04510)	43/199	-	VEGFA [84]
Insulin signaling (hsa04910)	35/137	-	
Melanogenesis (hsa04916)	28/101	-	

[1] genetic associations were reported in or near genes in this pathway by genome wide association studies

[2] KEGG pathway ID (http://www.genome.jp/kegg/)

Functional characterization of candidate miRNAs in human endothelial cells

MicroRNA hsa-mir-361-5p was shown earlier to influence the expression level of VEGFA [111] and thus should also influence angiogenesis. In order to test this hypothesis *in vitro*, we designed antisense oligoribonucleotides against hsa-mir-361-5p but also against hsa-mir-301a-3p and hsa-mir-424-5p and performed tube formation assays with human umbilical vein endothelial cells (HUVEC). We show that a knockdown of hsa-mir-361-5p significantly alters tube formation *in vitro* ($p_{corrected} < 0.05$, **Figure 6, Supplementary Figure S3 and Supplementary Figure S4**). Knockdown of hsa-mir-301a-3p and hsa-mir-424-5p also

showed elevated average tube lengths, however, this was not statistically significant after adjustment for multiple testing ($p_{corrected} > 0.05$).

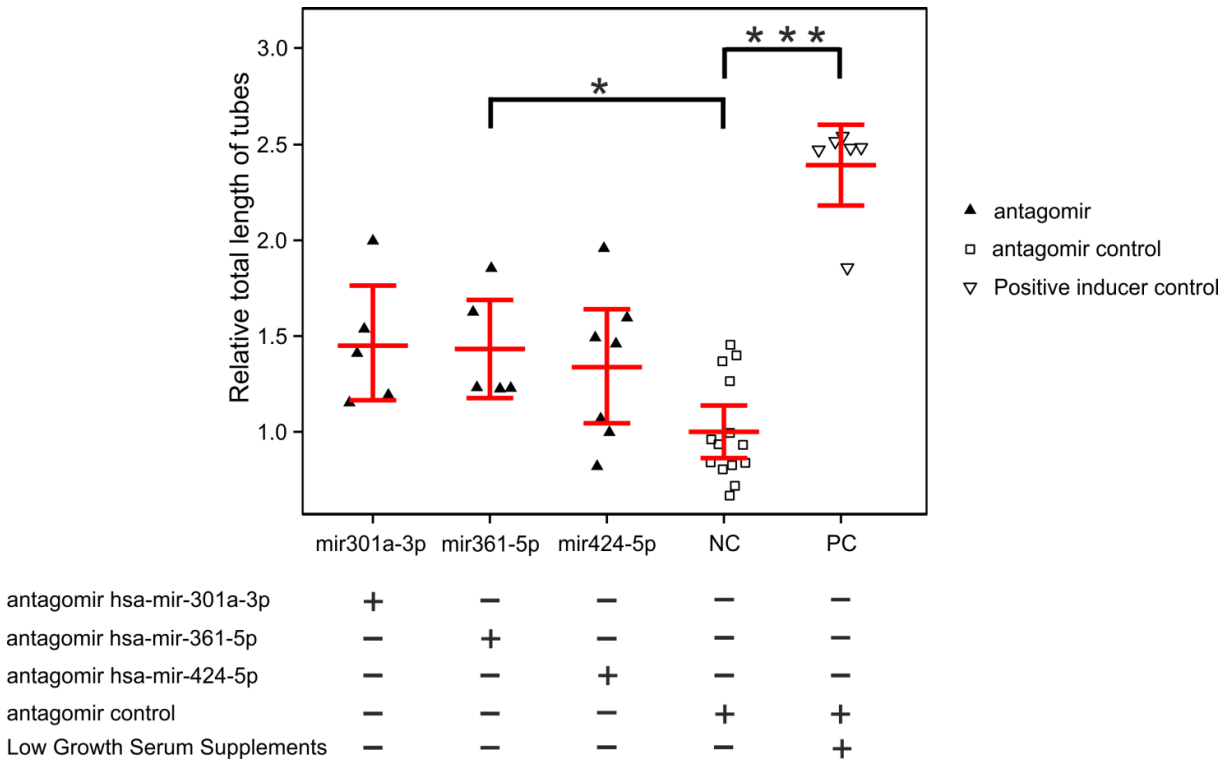


Figure 6. ***In vitro*** **tube formation assays in human endothelial cells.** HUVEC cells were transfected with antagomirs for hsa-mir-301a-3p, hsa-mir-361-5p or hsa-mir-424-5p or with control antagomirs (see **Supplementary Figure S3**) and seeded on Geltrex/Matrigel. Cumulative tube length was quantified with Angiogenesis Analyzer implemented in ImageJ. Each measurement point indicates one independent transfection. Low Serum Growth Supplements (Life) were used as a positive inducer control. Representative images are shown in **Supplementary Figure S4**. Significant differences between means are indicated by asterix. * = $p_{corrected} < 0.05$; *** = $p_{corrected} < 0.0005$.

Classification

The raw AUC value for the cmiRNA profile was 0.727 for NV AMD and controls from the Regensburg study and 0.802 when restricting the analysis to NV AMD and controls from the Cologne study. Additionally, we used the weights obtained from the Regensburg study of each cmiRNA in the profile to predict the outcome (case or no case) in the Cologne study and found an AUC value of 0.722. To estimate non-parametric confidence intervals, we performed a 2,000 fold bootstrap analysis in the pooled study. The bootstrapped AUC value for the profile was 0.730 (95%CI: 0.544-0.877) indicating a good classification accuracy.

3.3 Discussion

To our knowledge, this is the first study to evaluate the relative abundance of cmiRNAs in the serum of late stage AMD patients. We identified three cmiRNAs (hsa-mir-301a-3p, hsa-mir-361-5p, and hsa-mir-424-5p) which were significantly altered in NV AMD patients compared to AMD-free controls. Even when conditioned on covariates such as age, gender, smoking or genetic risk scores computed from known AMD-associated variants, the three cmiRNAs showed little alteration in their association strength, indicating a true association with late stage NV AMD. In contrast, there was no association of cmiRNAshsa-mir-301a-3p, hsa-mir-361-5p, or hsa-mir-424-5p with GA AMD, suggesting subtype-specific cmiRNA profiles for late stage AMD. A global screening strategy similar to the one applied in this study may be suited to eventually characterize a GA AMD specific cmiRNA profile.

Our initial discovery study comprised 9 NV AMD cases and 9 matched controls and identified several cmiRNA candidates with altered expression levels although none reached statistical significance after adjustment for multiple testing (n = 203 equivalent to the discovery of 203 cmiRNAs). A recent study compared cmiRNA levels in long-surviving versus short-surviving patients with lung cancer and found fold changes of significantly altered cmiRNAs between 1.60 and 7.15 [112] and Cohen's effect sizes between 0.92 and 1.54 which are considered to be large [113]. Given the number of samples in our discovery study, we calculated the power to detect comparable effect sizes after adjustment for multiple testing between 4.2% and 33.2%. This would imply a power to identify between 4 and 33 cmiRNAs out of 100 in our discovery study at the assumed effect size or higher. To compensate for lower effect sizes, we increased our sample size to 276 individuals (129 NV cases and 147 AMD-free controls) in the replication and retested individually the top 10 cmiRNAs hits from discovery. This uncovered a statistically significant association of NV AMD with cmiRNAshsa-mir-301a-3p, hsa-mir-361-5p, and hsa-mir-424-5p.

Bioinformatical pathway analysis for genes suggested to be regulated by the NV AMD associated cmiRNAs were performed with two independent programs including the miRSystem and mirPATH v2.0. Both revealed concurring results and implicated the TGF-β and mTOR pathways in neovascular AMD pathology. Interestingly, this is in agreement with a recently published GWAS which also implicated the TGF-β and the mTOR pathways in late stage AMD by identifying risk associated genetic variants near or within the genes encoding the transforming growth factor, beta receptor 1 (*TGFBR1)* and the vascular endothelial growth factor A (*VEGFA)*[33,84]. The TGF-β as well as the mTOR pathway are involved in cellular responses to stress and injury and also regulate angiogenesis. Consequently, we performed *in*

vitro tube formation assays and reduced the levels of hsa-mir-424-5p, hsa-mir-301a-3p, and hsa-mir-361-5p by antisense oligoribonucleotides to evaluate the impact of decreased miRNA levels on angiogenesis. Knockdown efficiency reduced microRNA levels in the test system on average by about two-fold. Antisense treatment of hsa-mir-361-5p lead to a significant increase in tube formation and, thus, angiogenesis in vitro. Results for hsa-mir-424-5p and hsa-mir-301a-3p revealed a similar direction of effect but were not statistically significant due to correction for multiple testing. Together, the data are promising and support our bioinformatical analyses.

Additionally, pathways closely related to the mTOR pathway were implicated by our analysis including WNT signaling, focal adhesion, neutrophin signaling and the insulin pathway. These pathways are involved in (neural) cell survival and therefore are reasonable candidate pathways for the pathogenesis of AMD. However, so far no genetic association with late stage AMD was observed for any genes associated with these signaling pathways. In this context, it should be noted that until now only few studies evaluated a genetic association for progression and severity of AMD [114,115]. These studies mainly focused on strong (and known) signals associated with increased risk for AMD and therefore may have missed possible existing associations. The present study has now identified cmiRNAshsa-mir-301a-3p, hsa-mir-361-5p, and hsa-mir-424-5p as new biomarkers for late stage neovascular AMD. Furthermore, our data show that these biomarkersare not associated with GA AMD implying that different biomarkers and thus different biological pathways are likely involved in subtype-specific manifestations of late stage AMD. If confirmed, this could have major implications for designing treatment regiments for AMD.

A recent study investigated a treatment option for patients with stroke by increasing a disease-related reduction in plasma levels of hsa-mir-424-5p [116]. In an inducible mouse model of acute stroke which also revealed a down-regulation of hsa-mir-424-5p in plasma as well as in brain, lentiviral overexpression of hsa-mir-424-5p in the murine brain prior to induction of ischemic stroke significantly lowered the infarct volume as well as the brain edema levels [116]. A similar approach could be envisioned for treating AMD lesions. The identification of cmiRNAs that are dysregulated in NV AMD patients, now offers a number of novel starting points for therapeutic regimens. For example, such targets could be the genes that are regulated by the cmiRNAs or, alternatively, could directly address the dysregulated cmiRNAs itself. Specifically, the latter approach would initially entail prescreening of patients for altered cmiRNAs levels. Reduced expression of a diagnostic cmiRNA (as pre-microRNA or mature microRNA) could be supplemented by lentiviral transduction, nano-particle aided

transfection or by delivery of the dysregulated cmiRNA via synthetic microRNAs in artificial exsosomes. Therapies to modify up- or down-regulated genes are also conceivable. This could be done by using small molecules to influence gene activity [117], protein activity and stability [118] or by targeting proteins or interacting proteins with specific antibodies [119].

In summary, this study has identified three cmiRNAs with a significantly altered expression profile in the serum of NV AMD patients when compared to AMD-free control individuals. This finding opens up a number of new avenues in understanding disease mechanisms and designing targeted treatment options. Another important aspect of our finding pertains to monitoring treatment effects in clinical trial settings. Although proof of concept is still waranted, measuring drug responses as a means of measuring changes in the cmiRNA profil from blood samples of AMD patients may proof a direct and little invasive approach in the future.

3.4 Materials and Methods

Ethics statement

This study followed the tenets of the declaration of Helsinki and was approved by the Ethics Review Board at the University of Regensburg, Germany (ID: 12-101-0241), University of Bonn, Germany and University of Cologne, Germany. Informed written consent was obtained from each proband after explanation of the nature and possible consequences of the study.

Recruitment of AMD cases and control individuals

The case-control sample included 54 individuals with seemingly non-familial NV AMD and 77 age- and gender-matched AMD-free controls from the Regensburg study, 116 cases and 70 controls from the Cologne study, and 18 GA AMD cases from the Bonn Eye Clinic (**Table 7**). Inclusion and exclusion criteria have been described elsewhere [33,65,67,120].

Genotyping of samples

Genotyping was carried out as described elsewhere [58]. Briefly, genomic DNA was extracted from peripheral blood leukocytes. Ten single nucleotide polymorphisms (SNPs, **Supplementary Table S5**) were genotyped either by direct sequencing, restriction enzyme digestion of PCR products (RFLP) or TaqMan SNP Genotyping (Applied Biosystems, Foster City, USA).

Isolation of cmiRNAs from stabilized blood samples and serum

To reduce degradation of microRNAs and other RNA species [121], for the Regensburg and Bonn samples peripheral venous blood was drawn in PAXgene Blood RNA tubes (PreAnalytiX GmbH, Hombrechtikon, CH) and immediately stored at -80°C. To isolate RNA, tubes were thawed at room temperature on a rocker and centrifuged for 5 minutes at 1500 rcf at 4°C. The RNA isolation was carried out with the mirVANA microRNA isolation kit (Ambion, Austin, TX, USA) as described elsewhere [122]. Briefly, 300 µl of the supernatant were mixed with 600 µl of binding/lysis buffer. Then, 90µl of microRNA homogenate additive was added, thoroughly mixed for 30s and incubated on ice for 10 minutes. An equal amount of acid/phenol/chloroform (Ambion) was then added to each aliquot and vortexed for 1 minute at maximum setting. The solution was spun for 10 minutes at 10,000g at room temperature. The resulting aqueous (upper) phase was mixed with 1.25 volumes of 100% ACS grade ethanol and passed through a mirVANA column in sequential 700µl steps. The columns were then washed according to the manufactures protocol and the RNA was eluted with 50µl nuclease-free water (preheated to 95°C).

For the Cologne samples, RNA isolation from blood serum was carried out with the miRNeasy Serum/Plasma kit (Qiagen) according to the manufacturer's recommendations. Typically, we used 200 ul of serum and eluted the RNA in 24ul of nuclease-free water.

Sequencing of cmiRNAs and data analysis (discovery study)

cDNA libraries were constructed using the Ion Total RNA-Seq v2 kit (Life Technologies) according to the manufacturers recommendations for 9 NV AMD cases and 9 control samples. The resulting cDNA libraries were purified by AMPure beads (Beckman Coulter), and their concentrations and sizes distribution were determined on an Agilent BioAnalyzer DNA high-sensitivity Chip (Agilent Technologies). Emulsion PCR and enrichment of cDNA conjugated particles were performed with an Ion OneTouch 200 Template Kit v2 DL (Life Technologies) according to the manufacturer's instructions. The final particles were loaded on an Ion 316 chip and sequenced on a Personal Genome Machine with 200bp read length (Life Technologies).

The data obtained were analyzed with the mirDEEP2 package [123]. Briefly, all reads were mapped to the human genome. Reads that failed to align were excluded. Remaining reads were then mapped to the pre-microRNA and microRNA sequences obtained from mirbase.org (Release 19, August 2012) and quantified. Reads per microRNA were normalized to the overall number of reads and normalized to 100,000 reads. The data were transformed with the natural logarithm to obtain a normal distribution of expression values. In order to account for

batch effects in the data, we employed an empirical Bayesian batch effect correction algorithm known as ComBat [124]. For each microRNA, mean values of cases were compared to mean values of controls via t-test. Nominal significant associations with a (two-sided) p-value < 0.1 were considered for replication.

Quantitative (q)RT-PCR and data analysis (replication study)

Circulating miRNA was extracted from blood as described above and reverse transcription followed by qRT-PCR was performed according to Hurteau *et al.* [125]. Briefly, 10µl of purified cmiRNA solution were modified by *E. coli* Poly (A) Polymerase I (E-PAP) by the addition of a polyA tail (Ambion, Austin, TX, USA). Reverse transcription was performed with Superscript III reverse transcriptase (Invitrogen Carlsbad, CA) and a Universal RT oligonucleotide primer, which contains a polyT stretch of DNA that binds to the newly synthesized polyA tail (**Supplementary Table S6**). The RT solution was diluted 1:50, of which 4µl were used per qRT-PCR reaction. Each qRT-PCR master mix was prepared according to the protocol of the Power SYBR Green Master Mix (Applied Biosystems, Foster, CA, USA) and run on an ABI Viia-7 (Applied Biosystems, Paisley, UK). Each microRNA was assayed in triplicates. Primers that performed poorly (<50% qRT-PCR efficiency) were excluded from further analysis. We further excluded measurements with a standard deviation greater than 0.4 Ct values in the triplicates. In order to normalize the Ct-values according to the amount of isolated RNA and reverse transcription efficiency, we used hsa-mir-451-5p as a housekeeping cmiRNA. This microRNA showed the least variance between cases and controls and within each group in our discovery study and was therefore regarded suitable as a housekeeper. The normalized Ct values of each individual were then normalized versus the median of the Ct values of the controls. We considered associated cmiRNAs with an odds ratio greater than 2 or lower than 0.5 for further replication.

The standard student's t-tests was applied to evaluate a statistically significant association as implemented in R [97]. In the final dataset, we adjusted the observed raw p-values ($p_{uncorrected}$) by a conservative Bonferroni correction ($p_{corrected}$). Adjusted p-values below 0.05 were considered significant. Sensitivity analysis was carried out by fitting logistic regression models adjusted for possible confounding variables.

Target prediction for cmiRNAs

We used miRSystem[126] and DIANA mirPATH v2.0 [127] to identify canonical pathways involved in AMD pathogenesis based on differentially regulated microRNAs. We used the default settings in miRSystem to identify target genes and to find canonical KEGG pathways.

With mirPATH v2.0, targets predicted by microT-CDS were selected with a threshold of 0.7. The intersection of pathways which showed an involvement of all investigated microRNAs (p-value threshold: 0.005, with Conservative Stats) was considered. We excluded KEGG pathways with more than 200 genes to increase specificity and to exclude pathways considered to be too general. Furthermore, we excluded validated cmiRNA targets as well as cancer pathways such as prostate cancer (hsa05215) or glioma (hsa05214), as the majority of the cmiRNA work has been in the field of oncology and thus cancer pathways are expected by design to be among the top findings.

Classification of cases and controls

Area under the curve (AUC) measurements were carried out with the function lroc from the package "epicalc" [100]. We used a bootstrap (n=2000) approach to calculate robust mean and confidence interval estimates for the AUC measurements by randomly selecting half of the cases and half of the controls (with replacement) and calculating the risk model with this sub-sample (training data). A randomly selected sample of half of the cases and half of the controls (with replacement) was then used to calculated the AUC (test data).

In vitro angiogenesis assay

Pooled human umbilical vein endothelial cells (HUVECs) were purchased from Life Technologies and cultured in Medium 200PRF with Low Serum Growth Supplement and Gentamicin/Amphotericin Solution (Life Technologies). Transfection of HUVECs was carried out as described in Bonauer *et al.* 2009 [128]. Briefly, cells were subcultured to passage 3 and grown until 70% confluent. 2`O-methyl antisense oligoribonucleotides against hsa-mir-424-5p (5'-UUCAAAACAUGAAUUGCUGCUG-3`), hsa-mir-301a-3p (5'-GCUUUGACAAUACUAUUGCACUG-3') or hsa-mir-361-5p (5'-ACAGGCCGGGACAAGUGCAAUA-3') or GFP (5'-AAGGCAAGCUGACCCUGAAGUU-3`) were synthesized by VBC Biotech and 50 nM were transfected with GeneTrans II (MoBiTec) according to the manufacturer's protocol. After 24h the medium was changed to full growth medium with supplements and antibiotics. 48h after transfection, $3.5x10^4$ HUVECs of each transfection were sown onto one well of a 24 well plate coated with 150µl Geltrex (Life Technologies). As a positive inducer control, cells were cultured in full growth medium with supplements. Total tube length was quantified after 24 hours by measuring the cumulative tube length in four random fields (area in each field: 2.25 mm^2) using the Angiogenesis Analyzer in ImageJ [129]. In total, we performed between 5 and 14 independent transfections for each knockdown or control experiment. In order to

assess the transfection efficiency, miRNAs were isolated with the mirVANA microRNA isolation kit (Ambion, Austin, TX, USA) according to the manufacturer's protocol. cDNA synthesis and qRT-PCR was carried out as described above.

4. Clinical and genetic factors associated with progression of geographic atrophy lesions in age-related macular degeneration.

This chapter is identical to the following publication currently in revision:

Grassmann F, Fleckenstein M, Chew EY, Strunz T, Schmitz-Valckenberg S, Göbel AP, Klein ML, Ratnapriya R, Swaroop A, Holz FG, Weber BHF (2015) **Clinical and genetic factors associated with progression of geographic atrophy lesions in age-related macular degeneration.** *PLoS One* in revision

Abstract

Worldwide, age-related macular degeneration (AMD) is a serious threat to vision loss in individuals over 50 years of age with a pooled prevalence of approximately 9 %. For 2020, the number of people afflicted with this condition is estimated to reach 200 million. While AMD lesions presenting as geographic atrophy (GA) show high inter-individual variability, only little is known about prognostic factors. Here, we aimed to elucidate the contribution of clinical, demographic and genetic factors on GA progression. Analyzing the currently largest dataset on GA lesion growth (N = 388), our findings suggest a significant and independent contribution of three factors on GA lesion growth including at least two genetic factors (ARMS2_rs10490924 [P < 0.00088] and C3_rs2230199 [P < 0.00015]) as well as one clinical component (presence of GA in the fellow eye [P < 0.00023]). These correlations jointly explain up to 7.2% of the observed inter-individual variance in GA lesion progression and should be considered in strategy planning of interventional clinical trials aimed at evaluating novel treatment options in advanced GA due to AMD.

4.1 Introduction

Age-related macular degeneration (AMD) is a common cause of blindness in Western societies with an estimated prevalence for all forms of the disease to reach almost 8.7 % and prevalence rates higher in Europeans than Asians or Africans [18]. Well-founded estimates project the worldwide number of people with AMD in 2020 to 196 million and in 2040 to 288 million [18]. In the near future, this will dramatically increase the individual as well as the socioeconomic burden of the disease.

AMD can progress in a succession of stages from an early to an intermediate and finally to a late form, where atrophic and neovascular subtypes are distinguished [130]. The early form is characterized by abnormalities at the level of the retinal pigment epithelium (RPE) and depositions of extracellular material located predominantly between Bruch's membrane and the RPE [131]. As this material accumulates, it becomes recognized clinically as individual drusen deposits. When drusen progress in size to greater than 125 microns in diameter, they are designated as large drusen and the eye is classified as having intermediate AMD. Drusen may become confluent and be associated with RPE hyper- and hypopigmentation. The presence of large drusen is a strong indicator of increased risk to develop a late form of the disease [132] which can manifest as geographic atrophy (GA), involving a gradual degeneration and disappearance of RPE, photoreceptor cells and the choriocapillaris layer of the choroid in the central retina. Another form of late stage manifestation is the exudative or neovascular phenotype, which is accompanied by choroidal neovascularisation (CNV) with subpigmentepithelial, subretinal and/or intraretinal extracellular fluid accumulation evolving to retinal scarring if left untreated [133]. Both late stage disease manifestations can exist at the same time in the same eye or mixed with neovascular AMD in one eye and GA in the fellow eye. While CNV development may be associated with rapid functional impairment, GA typically progresses slowly and eventually may extend beyond the macular area of the retina [133]. While intravitreal administration of anti-vascular endothelial growth factor agents are beneficial for the treatment of neovascular AMD [134], there is no proven therapy for GA.

The proportion of GA in late stage AMD is approximately 35-40% [19,20]. While the overall incidence of the neovascular form is more frequent, GA occurs more common in individuals over 85 years of age [20]. This further emphasizes the impact of GA on ageing populations, and underscores the need of effective treatment to prevent, slow or cure the disease.

GA lesions usually expand with an average growth rate of about 1.3 to 2.6 mm^2 per year [115,135–137] and may ultimately result in severe central vision loss [138]. While meta-analysis of genome-wide association studies for advanced stage AMD has identified at least 19 loci [33] and biological pathways underlying AMD are slowly getting recognized [40], limited information is available about the influence of genetic, demographic and clinical factors on GA growth. One study reported a significant contribution of a common, AMD risk associated haplotype in 10q26 (*ARMS2/HTRA1* locus) on GA progression as measured by lesion size [115], although this correlation was not replicated in two subsequent studies of similar sample size [139,140]. This inconsistency can be ascribed to a number of specifics in

the respective studies, e.g. related to imaging (color fundus photographs vs fundus autofluorescence), correction for initial lesion size [141] or different summarization of obtained growth rates [67]. Also, the effect of the presence of GA in the fellow eye was found to be a significant modulator of GA growth [137].

Design and evaluation of state-of-the-art clinical trials in GA require knowledge on factors that contribute to the progression of atrophic lesions. To validate previous findings [115,137] and to identify novel factors correlated with GA lesion growth, we analyzed the currently largest dataset on GA and GA lesion growth by combining two available studies: (i) the Fundus Autofluorescence in Age-related Macular Degeneration Study (FAM) [142], a multicenter study conducted in Germany and (ii) the Age-Related Eye Disease Study (AREDS) conducted in the United States [135]. We validated earlier findings for variant ARMS2_rs10490924 and for the presence of bilateral GA. We also expanded the analysis and searched for novel genetic and demographic factors correlated with GA lesion growth. Taken together, our data provide evidence for significant correlations between GA lesion growth and ARMS2_rs10490924, C3_rs2230199 and the presence of GA in the fellow eye, respectively. These correlations are independent of each other and jointly explain up to 7.2% of the observed inter-individual variance of GA growth.

4.2 Results

Study design

Overall, we investigated the effect of five clinical/demographic variables as well as ten genetic factors on GA lesion growth in 529 eyes from 388 individuals in a three stage study design (**Table 11**). Each area measurement was root transformed to eliminate the dependence of the growth rate on the initial lesion size [141]. A single growth rate per individual was calculated by taking the means of all calculated growth rates for each individual. In a first step, we aimed to replicated the findings from an earlier study for risk haplotype on 10q26 (*ARMS2/HTRA1* locus) [115]. This study showed that the risk increasing allele at ARMS2_rs10490924 increased GA lesion growth rate. We then searched for significant correlations between GA growth and novel clinical and genetic factors in a discovery study including 86 randomly selected individuals (*FAM study – discovery*) and considered factors which showed a nominally significant correlation ($P_{raw} < 0.05$) for further replication. Significant findings were replicated in two additional studies (*FAM study – replication* and *AREDS – replication,* N=302 individuals) and a meta-analysis was conducted to combine the effect sizes and standard errors from each individual study assuming a random effects model. The final P-values were adjusted by a conservative Bonferroni correction assuming 16 independent statistical tests and corrected P-values ($P_{corrected}$) below 0.05 were considered significant. Lastly, we fitted a multivariate linear regression model to evaluate the independence of significantly correlated factors.

Table 11. Summary characteristics of participating study populations

	FAM - discovery	FAM - replication	ARED replication	combined
Imaging technique	FAF	FAF	color fundus	mixed
Number of individuals	86	48	254	388
Mean follow-up time (S.D.) [years]	3.19 (1.97)	2.77 (1.66)	5.21 (3.00)	4.46 (2.85)
Mean interval between examinations (S.D.) [years]	1.37 (0.86)	1.56 (1.22)	1.13 (0.37)	1.24 (0.68)
Mean number of examinations (S.D.)	3.95 (2.86)	2.85 (1.61)	4.72 (2.83)	4.31 (2.78)
Mean age (S.D.) [years]	75.47 (7.37)	76.77 (5.90)	70.27 (5.07)	72.22 (6.36)
Mean growth [mm²/year] (S.D.)	1.62 (0.96)	1.34 (0.92)	1.55 (1.74)	1.54 (1.51)
Mean √growth [mm/year] (S.D.)	0.28 (0.14)	0.28 (0.14)	0.32 (0.30)	0.30 (0.25)
Mean ln (√growth [mm/year]) (S.D.)	-1.40 (0.60)	-1.43 (0.58)	-1.54 (1.03)	-1.50 (0.91)
Patients with bilateral GA [%]	67.4	64.6	29.1	40.72
Mean initial size (S.D.) [mm²]	6.53 (4.4)	5.03 (5.02)	2.96 (4.24)	4.01 (4.62)
Fraction male [%]	39.5	31.3	43.3	41.0

FAF = fundus autofluorescence

Table 12. Correlation between genetic, clinical and demographic factors and GA growth

	FAM - discovery			FAM - replication			AREDS - replication			Combined (random effects model)				
	Effect size	**95% CI [2]**	**P [3]**	**Effect size**	**95% CI [2]**	**P [3]**	**Effect size**	**95% CI [2]**	**P [3]**	**Effect size**	**95% CI [2]**	**P [3]**	**Pcorr[4]**	**Pheter [5]**
Gender	0,036	-0.228 - 0.300	0,788	-	-	-	-	-	-	-	-	-	-	-
Age [years]	0,001	-0.016 - 0.019	0,866	-	-	-	-	-	-	-	-	-	-	-
Initial size [mm²]	0,015	-0.015 - 0.044	0,325	-	-	-	-	-	-	-	-	-	-	-
Genetic Risk Score[1]	0,037	-0.057 - 0.132	0,434	-	-	-	-	-	-	-	-	-	-	-
CFH_rs1061170	-0,042	-0.233 - 0.148	0,660	-	-	-	-	-	-	-	-	-	-	-
CFH_rs6677604	-0,181	-0.551 - 0.189	0,334	-	-	-	-	-	-	-	-	-	-	-
CFH_rs800292	0,243	-0.048 - 0.534	0,100	-	-	-	-	-	-	-	-	-	-	-
C3_rs2230199	**-0,262**	**-0.455 - -0.070**	**0,008**	**-0,358**	**-0.635 - -0.081**	**0,012**	**-0,141**	**-0.340 - 0.057**	**0,162**	**-0,24**	**-0.358 - -0.114**	**1,50E-04**	**0,0024**	0,4186
ARMS2_rs10490924[6]	**0,149**	**-0.019 - 0.317**	**0,082**	**0,193**	**-0.016 - 0.402**	**0,069**	**0,194**	**0.017 - 0.370**	**0,032**	**0,176**	**0.072-0.280**	**8,80E-04**	**0,0141**	0,919
CFB_rs438999	0,054	-0.452 - 0.561	0,832	-	-	-	-	-	-	-	-	-	-	-
CFB_rs4151667	0,113	-0.743 - 0.969	0,793	-	-	-	-	-	-	-	-	-	-	-
APOE_rs7412	-0,195	-0.500 - 0.111	0,209	-	-	-	-	-	-	-	-	-	-	-
APOE_rs429358	0,086	-0.216 - 0.388	0,574	-	-	-	-	-	-	-	-	-	-	-
CFI_rs2285714	-0,146	-0.354 - 0.062	0,167	-	-	-	-	-	-	-	-	-	-	-
bilateral GA	**0,322**	**0.054 - 0.590**	**0,019**	**0,278**	**-0.09 - 0.647**	**0,135**	**0,333**	**0.055 - 0.611**	**0,019**	**0.317**	**0.148 - 0.485**	**2,30E-04**	**0,0037**	0,9704
No. of exams	0,006	-0.051 - 0.039	0,796	-	-	-	-	-	-	-	-	-	-	-

[1] GRS computed with reduced (10 SNPs) set according to Grassmann *et al.* 2012

[2] 95% confidence intervals

[3] P value from linear regression model without covariates

[4] P value adjusted for multiple testing (Bonferroni correction) assuming 16 tests performed

[5] P value for evidence of heterogeneity from random effects model

[6] previously been shown to influence GA growth. We used the FAM study to replicate this finding

Influence of ARMS2_rs10490924 risk allele on GA growth

Klein *et al.* [115] found a significant contribution of the ARMS2_rs10490924 risk allele on GA lesion growth in 114 individuals from AREDS. We expanded these analyses by calculating the correlation of the number of risk alleles at this variant with GA growth in an extended AREDS panel of 254 patients (**Table 12**). We found a nominally significant correlation (P = 0.032) and a positive slope (0.194, 95% CI: 0.017 - 0.370) consistent with the previous study [115]. The findings were replicated in both FAM studies (*FAM - discovery* and *FAM - replication*) and the obtained slopes and standard errors were pooled using a random effects model. The combined slope was 0.176 (95CI: 0.072-0.280) with a raw P-value of 0.00088 (**Table 12** and **Figure 7**). This correlation remained statistically significant after correction for multiple testing ($P_{corrected}$ = 0.0141) and no evidence for heterogeneity between the studies was found (P_{het} = 0.919). The correlation was visualized for the combined study (N = 388) by stratifying the patients according to their genotype (homozygous non-risk, heterozygous risk and homozygous risk) and plotting the growth rate for each individual (**Figure 8**).

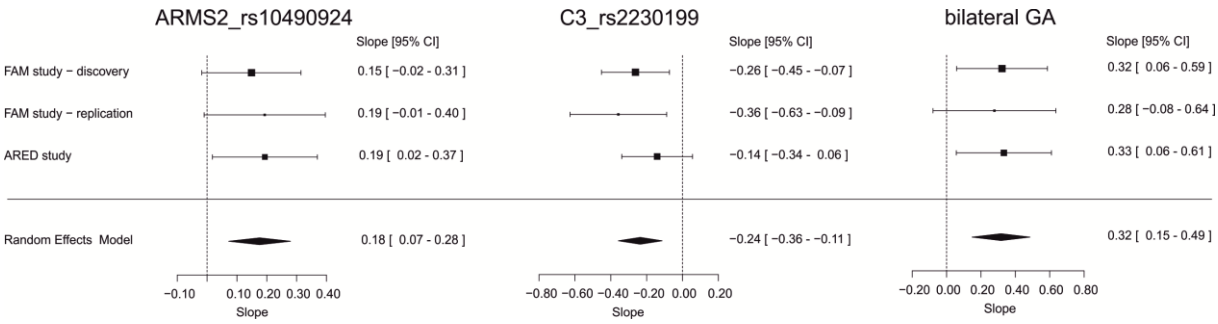


Figure 7. **Forestplot representations of univariate linear regression models.** Univariate linear regression models were fitted for variables ARMS2_rs10490924, C3_rs2230199 and bilateral GA for each study separately. Slope and standard errors obtained from the models of each study were combined by performing a meta-analysis assuming a random effects model. The combined estimates for slope and 95% confidence intervals (CI) were computed from the random effects model. In all analyses, no evidence was found for heterogeneity ($P_{heterogeneity}$> 0.05).

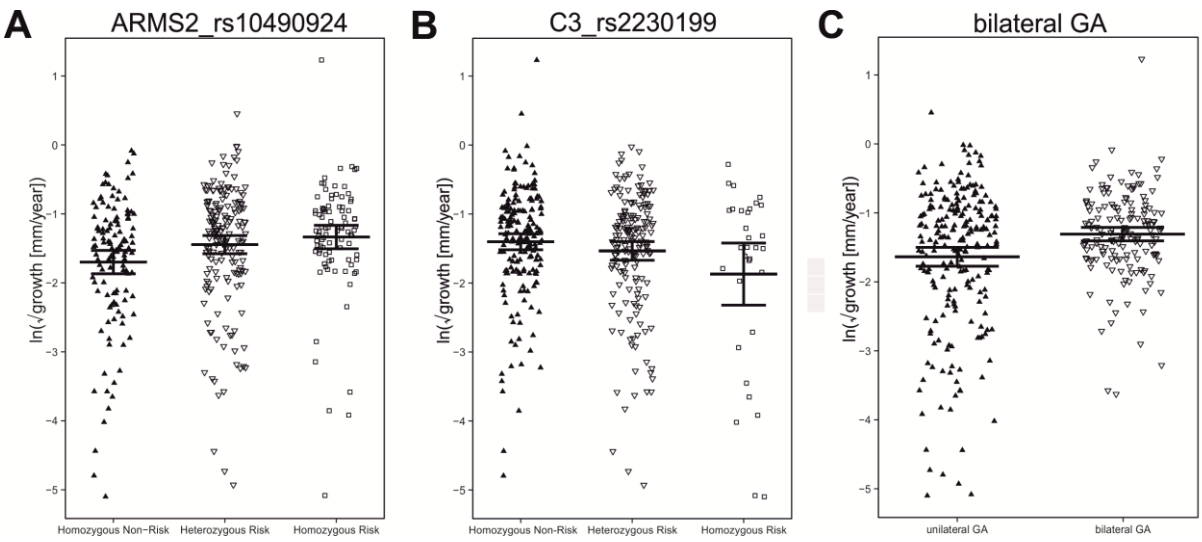


Figure 8. **GA lesion growth rates for each individual in the combined study.** The measured area of GA was square-root transformed. From the transformed area the growth rate was calculated per year in [mm/year]. Growth rates from each individual were then obtained by calculating the mean of all growth rates of the individual. If both eyes were affected, the mean of both eyes were calculated resulting in a single growth variable per individual. These individual growth rates were further transformed by the natural logarithm (ln) and were stratified either by **A** the genotype at ARMS2_rs10490924 or **B** the genotype at C3_rs2230199 or **C** the presence or absence of bilateral GA.

Correlation of genetic factors with GA growth

To estimate the contribution of additional genetic factors to GA lesion growth, we genotyped nine common AMD associated variants at 5 loci in FAM - discovery (**Table 12** and **Supplementary Table S7**). The frequencies of the variants in FAM - discovery were comparable to those observed in AMD cases in other studies which have been shown to significantly and independently influence the risk for AMD [58]. In addition, we calculated a genetic risk score (GRS) to summarize the genetic risk of the individual participants. The average GRS in the discovery study was 1.96 (S.D. = 1.37) and thus slightly higher than the observed values for GA patients in a previous study [58].

In the discovery study, neither the genetic risk score nor most single genetic variants analyzed revealed a significant influence on GA growth rate ($P_{raw} > 0.05$, **Table 12**) with the exception of variant C3_rs2230199 ($P_{raw} < 0.05$). Interestingly, the risk increasing allele (C) at C3_rs2230199 reduced the growth rate with a slope of -0.262 (95% CI: -0.455 - -0.070). Thus, individuals who have an increased risk to develop AMD due to the risk increasing allele at C3_rs2230199 show a reduced rate of GA lesion growth when compared to individuals who do not carry risk increasing alleles. To further validate this finding, we replicated this correlation in *FAM - replication* and *AREDS - replication*. We found similar negative slopes

in both replication cohorts and pooled the findings in a mixed effects model (**Figure 7**). The random effects model had a slope of -0.236 (95% CI: -0.358 - -0.114) and demonstrated a highly significant correlation of the number of C3_rs2230199 risk alleles with GA growth ($P_{raw} = 1.5x10^{-4}$), which remained statistically significant after adjustment for multiple testing ($P_{corrected} = 0.0024$). Additionally, we found no significant evidence for heterogeneity between the studies ($P_{het} = 0.4186$).

Correlation of clinical and demographic factors with GA growth

After root transforming the measured area prior to the calculation of the growth rates, a significant correlation between the initial lesion size (in [mm²]) or the root transformed initial lesion size (in [mm]) and the growth rate of the lesion was not observed (**Table 12**). We also found no significant correlation to gender, age or the number of examinations to determine the GA growth rate. However, we found a strong correlation between the presence of GA in the fellow eye (bilateral GA) and GA growth (**Table 12**), in agreement with a previous report [137]. The findings indicate a significant increase in GA growth in cases where GA is present in both eyes of an individual. The slope of the regression model was estimated to be 0.322 (95% CI: 0.054 – 0.590) in the discovery study. To validate this finding, the effect of bilateral GA was investigated in the two replication cohorts. In each replication study, we found similar effect sizes in the same direction as observed in the discovery study. The findings were summarized in a meta-analysis of slopes and standard errors obtained from the individual studies (**Figure 7**). The random effects model showed a highly significant correlation (slope: 0.317, 95% CI: 0.148-0.485) of the presence of bilateral GA with GA growth ($P_{raw} = 0.00023$), which remained significant after adjustment for multiple testing ($P_{corrected}= 0.0037$, **Table 12**). Again, no significant evidence for heterogeneity between the studies was found ($P_{het} = 0.9704$). The growth rates for each individual were stratified according to the presence or absence of bilateral GA and were visualized in a jitterplot (**Figure 8)**.

Multivariate linear regression models

After correlating risk alleles at ARMS2_rs10490924 and presence of bilateral GA with GA lesion growth and identification of a novel genetic variant (C3_rs2230199) in the univariate regression analysis, we further evaluated the possibility that one of these factors influences or confounds the correlation with the other factors. We therefore fitted a multivariate linear regression model for each of the three studies and all three studies combined with the three identified factors in the same model (**Table 13**).

Table 13. Multivariate linear regression analysis of factors significantly correlated to GA growth

	Factors included in multivariate linear regression models			
	ARMS2_rs10490924	C3_rs2230199	bilateral GA	P[1]
FAM discovery	0.131 (-0.033 - 0.295)	-0.239 (-0.429 - -0.049)	0.341 (0.084 - 0.598)	0.0014
FAM replication	0.200 (-0.008 - 0.407)	-0.401 (-0.708 - -0.094)	0.244 (-0.097 - 0.585)	0.0089
ARED replication	0.205 (0.031 - 0.380)	-0.168 (-0.363 - 0.028)	0.362 (0.087 -0.639)	0.0039
All combined[2]	0.174 (0.072 - 0.276)	-0.232 (-0.355 - -0.109)	0.326 (0.163 - 0.488)	5.83e-08

[1] P value of linear regression model vs. null model

[2] combined effect sizes were estimated from random effects model (meta-analysis)

No significant evidence for a confounding effect of these factors on the correlation of the other factors was found ($P_{corrected} < 0.05$). The adjusted slopes in the multivariate analyses differed on average by 0.048 (S.D. = 0.086) for ARMS2_rs10490924, 0.0946 (S.D. = 0.0913) for C3_rs2230199 and 0.034 (S.D. = 0.027) for bilateral GA when compared to the slopes in the univariate analysis. The combined regression model showed a highly significant fit ($P_{raw} = 5.83 \times 10^{-8}$) and an adjusted R^2 of 0.072, thus explaining up to 7.2% of the variation in the GA growth rate.

4.3 Discussion

Here, we aimed to further elucidate the contribution of genetic as well as clinical and demographic factors on the rate of AMD GA lesion enlargement. We extended previous efforts [115,139,140] and now provide data on the largest available dataset on GA lesion progression so far. If both eyes were affected, the previous studies either included only one eye at random [115,140] or pooled the data for both eyes [139]. Here, we chose to combine the growth rates observed for the two eyes of an individual to a single variable under the assumption that a germline genetic variant should influence lesion growth in a similar fashion in both eyes. And indeed, we and others report a high degree of concordance between GA growth rate in the two affected eyes of the patient [135–137,143,144]. Of noted, the ARED study reveals a lower occurence of bilateral GA which can be explained mainly by two findings. Firstly, the ascertainment strategy was different for AREDS and FAM. While the ARED study specifically recruited patients with unilateral late stage AMD at baseline, the FAM study protocol included both uni- and bilateral patients. Secondly, we observed a significantly lower mean age in the ARED study compared to the FAM study. As individuals with unilateral GA are much more likely to progress to advanced stages in the fellow eye than individuals without late AMD manifestations [145,146], we expect the number of bilateral GA cases to increase in the ARED study over time.

A number of earlier studies have not reported a significant influence of genetic factors on GA lesion growth after adjustment for multiple testing [115,139,140]. Nevertheless, the correlation for C3_rs2230199 and ARMS2_rs10490924 presented in this report, was also suggested in a previous study which was based on a subsample of the present patient cohort [139]. However, the correlation in the previous work was not statistically significant after adjustment for multiple testing. Furthermore, in the FAM study several eyes were excluded from the current anaylsis which exhibited FAF phenotypes reminiscent of monogenetic diseases (e.g. GPS-FAF pattern [31] or CACD-FAF phenotypes [147]). In addition, we note that previous studies usually did not account for the large influence of the initial size of the lesion on the rate of progression possibly confounding the analyses. By root transforming the measured GA areas, we eliminated this problem. Furthermore, the present study included more than twice the number of participants than each single previous study and, thus, had a much higher power to detect significant correlations.

Our data confirm a significant correlation between GA lesion size and the number of risk alleles at ARMS2_rs10490924, a variant which represents the AMD risk haplotype at the *ARMS2*/*HTRA1* locus [148]. As the functional gene at this locus has not yet been identified [149], it is inappropriate to speculate about the mechanisms involved by which this locus may contribute to both the development as well as the progression of the disease.

At the C3 locus, the risk increasing allele at C3_rs2230199 reduced the growth rate of GA lesions. As the risk increasing allele at C3 results in reduced levels of CFH binding and thus in overly active complement [150], the observed statistical correlation between lesion size and C3 risk variant is inverse to findings for AMD risk and C3 risk variant and thus rather counter-intuitive. Complement activation is generally thought to be associated with increased inflammation due to priming [151] and activation of microglia [152] and as such would be expected to be associated with faster disease progression. On the other hand, several studies demonstrated low levels of inflammation to be beneficial: (i) *C3* knockout mice ($C3^{-/-}$) implicated the complement system in neurogenesis [153] by showing that increased levels of complement activation promoted neurogenesis in healthy and diseased neuronal tissue; (ii) in the presence of immune cells low levels of complement activation revealed neuroprotective properties [154] and (iii) properdin, the only positive regulator of the alternative complement pathway, was found to be a protective factor in inflammatory diseases, thus shedding new light on the complement activation in neurodegenerative and inflammatory diseases [155]. Taking these findings into account, active neurogenesis and neuroprotection due to increased complement activation could counteract neurodegenerative activities in GA and thus could reduce lesion growth [156,157].

Our study replicated an earlier report on the influence of disease status of the fellow eye on GA lesion growth [137]. We demonstrate that (i) the slope, and thus the effect size, for bilateral GA is virtually the same in all cohorts analyzed in this study and (ii) that this correlation is independent of C3_rs2230199 and ARMS2_rs10490924. The *ARMS2*_rs10490924 variant increases the risk of progression from early to late AMD [158] as well as the growth rate of GA lesions. Thus, this variant may explain the increased growth in bilateral GA patients by (i) promoting the onset of GA in the fellow eye and (ii) increasing the growth rate of the lesion. However, in the multivariate model, both variables (*ARMS2* variant and bilateral GA) show an independent correlation with GA growth. Additionally, no significant association of *ARMS*_rs10490924 is detected with the presence of bilateral GA in the combined dataset (OR_{Allele}: 0.810, 95% CI: 0.606 – 1.078) or in any individual study

analyzed. Furthermore, the risk increasing alleles at C3_rs2230199 are independently correlated with a decreased progression rate. This finding argues strongly against a confounding effect of risk increasing genetic factors on the increased growth rate observed in bilateral GA patients, specifically, as risk increasing alleles at C3_rs2230199 reduce GA lesion growth. As no significant association is observed between C3_rs2230199 risk alleles and the presence of bilateral GA (OR_{Allele} = 1.087, 95% CI: 0.788 – 1.498), other factors must influence the onset of GA in the fellow eye and explain the increased growth rate in bilateral GA.

In summary, our findings reveal a significant and independent contribution of at least two genetic factors and the presence of GA in the fellow eye to GA lesion growth. Our studies have an impact on design and evaluation of future clinical trials aimed at testing novel treatment approaches for GA. These factors can greatly confound the outcome, particularly when GA growth rate is the main outcome parameter. As the area and growth rate of GA lesions do not necessarily correlate with visual acuity due to the frequently observed phenomenon of foveal sparing over a long period of time [138], it is difficult to draw any conclusion from our data on how long the patient's vision can be retained. Future studies will be needed to evaluate the impact of genetic and clinical factors on visual perception.

4.4 Material and Methods

Ethics Statement

The study followed the tenets of the Declaration of Helsinki and was approved by the local Ethics Review Board at the University of Bonn (ID: 082/04) and the NIH (IRB operates under FWA00005897 and the IRB Blue Panel, IRB00005894). Informed written consent was obtained from each patient after explanation of the nature and possible consequences of the study.

Study characteristics

The study characteristics are summarized in **Table 11**. For screening (FAM - discovery) and initial validation (FAM - replication), we included a total of 201 eyes from 134 patients from the Fundus Autofluorescence in Age-Related Macular Degeneration (FAM) study (www.clinicaltrials.gov: NCT00393692) [139]. To validate the findings (AREDS - replication), we included 328 eyes from 254 patients from the Age-Related Eye Disease Study (AREDS) [115].

Classification of geographic atrophy

From the FAM-study, eyes with central (within a 500 µm radius of the foveal center) and non-central GA were included into the analysis and were classified as 'GA' eyes. In general, GA is funduscopically defined as one or more well-circumscribed, usually more or less circular patch of partial or complete depigmentation of the RPE, typically with exposure of underlying large choroidal blood vessels [135]. GA due to AMD is further defined as sharply demarcated lesion with clearly reduced FAF of an extend of $\geq$ 0.05 mm² (approximately 178µm in diameter) that does not correspond to exudative retinal changes (e.g. bleeding, exudates, fibrous scar) in an eye with funduscopically visible soft drusen and/or retinal pigment abnormalities consistent with AMD [137].

For the AREDS Study, GA associated with AMD was defined on stereoscopic color fundus photographs as sharply circumscribed areas of RPE depigmentation occurring in the macular area, generally considered to be circular in shape, with obvious visualization of the underlying choroidal blood vessels. The size must be as large as 1/8 disk diameter. An area of RPE atrophy within or adjacent to fibrosis or other features of neovascularization is not considered GA. Central GA is defined as the involvement of the center of the fovea, which was deteremined by retinal vascular configuration and pigment change. The digitized images were evaulated for GA area (mm^2) using computerized planimetry [115].

Fundus autofluorescence (FAF) measurements and calculation of GA growth rate

In the FAM study, FAF was measured and the lesion area was determined as previously reported [139] while AREDS analyzed the area of GA by color fundus photographs [115]. To eliminate the dependency of growth rates on baseline lesion size measurements, the individual area measurement was square root transformed (√area [mm]) [141]. We calculated lesion growth per examination interval by dividing the root transformed area by the time between examination points (in [years]), yielding a linear growth rate of the lesion (in [mm/year]). In case a growth rate was negative due to measurement imprecision, the growth rate for this interval was set at zero. The resulting growth rates per examination interval (√growth [mm/year]) were summarized for each individual by computing the mean of all growth rates for one individual. If data were available for both eyes of a patient, the mean of all computed growth rates from both eyes per individual were used. To generate normal distributed data and to reduce the bias from outliers, the growth rates per individual were log transformed with the natural logarithm (ln). This resulted in a single log-square-root growth rate variable per individual (ln(√growth [mm/year])), which was used for all subsequent analyses.

Clinical and demographic variables

The age at first examination (in [years]), the gender and lesion sizes in one eye (unilateral GA) or both eyes (bilateral GA) at the last examination as well as the mean follow-up time of our patients and the mean interval between examinations (in [years], **Table 11**) were recorded. Additionally, we report the mean number of examinations each individual received (minimum: 2, maximum: 21) in order to exclude a confounding effect of the number of examinations on GA growth (**Table 11**).

Genotyping and genetic risk score (GRS)

Genotyping was performed as described [58,115]. Briefly, genomic DNA was extracted from peripheral blood leukocytes by established methods. Genotyping was performed by TaqMan SNP genotyping (Applied Biosystems, Foster City, USA) or by PCR followed by restriction enzyme digestion (New England Biolabs, Ipswich, USA) and subsequent restriction fragment length analysis (RFLP).The resulting genotypes were coded as the number of AMD risk increasing alleles (0, 1 or 2), i.e. alleles which are more frequent in cases than in controls (**Supplementary Table S7**) [58]. These variants were used to compute the genetic risk score according to Grassmann *et al.* 2012 with weights obtained from the parsimonious model based on 10 SNPs (**Supplementary Table S7**).

Statistical analyses and visualization

To visualize the raw data, we computed means and 95% confidence intervals of growth rates in different subgroups and used the *ggplot* function from the *ggplot2* [159] package in R [97] for drawing jitterplots. Linear regression was done to evaluate correlations between clinical and genetic variables with GA growth and computed P-values and confidence intervals as implemented in R. Furthermore, obtained P-values were adjusted by a conservative Bonferroni correction multiplying the P-values with the number of (independent) tests performed. Since we conducted the study in a three stage setup, we subsequently combined the obtained slopes and standard errors from the three independent studies by using the function *rma* from the packages *metafor* [160] and conducted the meta-analysis assuming a random effects model. This approach also allowed an assessment of heterogeneity between the estimates from each study.

5. A candidate gene association study identifies *DAPL1* as a female-specific susceptibility locus for age-related macular degeneration (AMD).

This chapter is identical to the following publication currently in press:

Grassmann F, Friedrich U, Fauser S, Schick T, Milenkovic A, Schulz HL, von Strachwitz CN, Bettecken T, Lichtner P, Meitinger T, Arend N, Wolf A, Haritoglou C, Rudolph G, Chakravarthy U, Silvestri G, McKay GJ, Freitag-Wolf1 S, Krawczak M, Smith RT, Merriam JC, Merriam JE, Allikmets R, Heid IM, Weber BHF (2015) **A candidate gene association study identifies DAPL1 as a female-specific susceptibility locus for age-related macular degeneration (AMD).** *Neuromolecular Medicine,* in press

Abstract

Age-related macular degeneration (AMD) is the leading cause of blindness among white caucasians over the age of 50 years with a prevalence rate expected to increase markedly with an anticipated increase in the life span of the world population. To further expand our knowledge of the genetic architecture of the disease, we pursued a candidate gene approach assessing 25 genes and a total of 109 variants. Of these, synonymous single nucleotide polymorphism (SNP) rs17810398 located in *DAPL1* (death associated protein-like 1) was found to be associated with AMD in a joint analysis of 3,229 cases and 2,835 controls from five studies (combined $P_{ADJ} = 1.15\times10^{-6}$, OR=1.332 [1.187-1.496]). This association was characterised by a highly significant sex difference ($P_{diff} = 0.0032$) in that it was clearly confined to females with genome wide significance ($P_{ADJ} = 2.62\times10^{-8}$, OR = 1.541 [1.324-1.796]; males: $P_{ADJ} = 0.382$, OR = 1.084 [0.905-1.298]). By targeted resequencing of risk and non-risk associated haplotypes in the *DAPL1* locus, we identified additional potentially functional risk variants, namely a common 897bp deletion and a SNP predicted to affect a putative binding site of an exonic splicing enhancer. We show that the risk haplotype correlates with a reduced retinal transcript level of two, less frequent, non-canoncical *DAPL1* isoforms. *DAPL1* plays a role in epithelial differentiation and may be involved in apoptotic processes thereby suggesting a possible novel pathway in AMD pathogenesis.

5.1 Introduction

Age-related macular degeneration (AMD) is a common condition of complex aetiology with major risk factors including age, gender, smoking, ethnicity and genetics [161]. While AMD ultimately represents the primary cause of blindness in developed countries [162], its early

form is less severe and characterized by the mere presence of drusen and pigmentary abnormalities in the macular area of the retina [163]. Late stage AMD manifests as choroidal neovascularization and/or geographic atrophy and is associated with irreversible central visual loss [59,161].

Genetic predisposition plays an important role in AMD and is estimated to contribute up to 70% of the disease risk [72]. To date, two major and several minor to moderate AMD susceptibility loci have been identified with per allele odds ratios ranging from 1.3 to 3.4 [33]. Of note, many of these loci suggest an involvement of inflammatory processes and impaired complement activation in AMD pathogenesis [34,47,75,78,80], a fact that has raised major interest in novel therapeutic approaches to address progression of the disease [164].

Genetic variants associated with complex diseases are usually identified by high-throughput genome-wide association studies of large numbers of cases and controls [165]. However, candidate gene studies with similar sample sizes normally have greater statistical power to detect genetic disease associations [166], especially for genes not covered efficiently by commercially available genotyping platforms [167].

In this study, we aimed to expand our current knowledge of the genetic architecture of AMD pathogenesis, following a candidate gene approach. In a well-powered case-control study, we screened 109 haplotype tagging variants in 25 genes for an association with late stage AMD. Attempts to replicate any positive findings in over 4,000 individuals from four previous studies revealed that variation in the *DAPL1* gene is significantly associated with AMD. Importantly, this association is restricted to females and the variants of interest correlate with altered transcription levels of specific retinal isoforms of the *DAPL1* gene.

5.2 Results

Association of 109 SNPs in 25 candidate genes with late stage AMD

We first selected 25 genes and 109 haplotype tagging single-nucleotide polymorphisms (SNPs) for an initial analysis of 710 late stage AMD cases and 612 controls (GER1) (**Table 14** and **Supplementary Table S8** and **Supplementary Table S9**). Criteria for candidate gene selection included one or a combination of the following: (i) causative involvement of the gene in phenotypically related retinopathies, (ii) known gene function compatible with suspected AMD pathogenesis, (iii) specific or predominant gene expression in cellular sites of

primary AMD pathology, i.e. the photoreceptor/retinal pigment epithelium (RPE)/choroid complex. All SNPs were tested for a significant deviation from Hardy-Weinberg equilibrium ($p < 0.05$) in all controls and in female and male controls separately. This identified three SNPs (*RGR:*rs2279227, rs4620343 and *TRPM3:*rs3812532) which were subsequently excluded from further analyses. Association tests adjusted for age and sex revealed a nominally significant association using logistic regression between AMD and three SNPs (*DAPL1*: rs17810398:C>T, P=0.016; *RP1*: rs9643828:T>C, P=0.037; *CST3*: rs2424577:C>T, P=0.028) (**Supplementary Table S9**).

Replication of three nominally significant AMD associated candidate gene variants

The three SNPs with a nominally significant AMD association were genotyped in an independent German replication sample of 996 late stage AMD cases and 645 controls (GER2). The disease association could be confirmed only for rs17810398 (P=0.0014), a synonymous SNP in the coding sequence of the death-associated protein-like 1 (*DAPL1*) gene (**Table 15**). Analysis of this SNP in three other studies (681/367 late stage cases/controls from US, 300/183 from UK and 542/1028 from Cologne, **Table 14**) yielded consistent results (**Table 15**). Combined analyses of the 3,229 cases and 2,835 controls yielded a P value of 1.15×10^{-6} after adjustment for age, sex and study (**Table 15**). Given that 106 tests were performed, this result is significant at a significance level of 1.20×10^{-4} after Bonferroni correction. The risk allele frequencies were similar in all four studies (13.3-14.3 % in cases; 10.3-12.4 % in controls) and the per-allele odds ratios (OR) was consistent in direction and magnitude ($1.177 \leq OR \leq 1.530$) with no indication for heterogeneity ($I^2 = 0$).

Imputation and replication of genetic variants at the DAPL1 locus

We next imputed the genotypes of 20,422 additional SNPs around rs17810398 in GER1 study based on 8 tag SNPs in the DAPL1 locus. After quality control, 517 SNPs remained for analysis and we obtained association signals ($P_{adj} < 0.05$) that were confined to a 154 kb region devoid of any gene other than *DAPL1* (**Figure 9** and **Supplementary Table S10)**.

Table 14. Summary characteristics of participating study populations

		Number of individuals							Mean age (SD) [years] in		
Stage	**Study**	**Cases**	**GA [1]**	**NV [2]**	**GA&NV [3]**	**Controls**	**Total**	**Study type**	**cases**	**controls**	**Fraction male [%]**
1	GER1	710	161	423	126	612	1322	Case/Control	78.81 (6.64)	78.21 (5.28)	36.99
2	GER2	996	216	535	245	646	1642	Case/Control	76.15 (7.32)	73.05 (8.34)	37.72
2	US	681	165	516	0	367	1048	Case/Control	79.08 (8.48)	74.57 (7.10)	39.79
2	UK	300	38	252	10	183	483	Case/Control	78.45 (9.75)	74.53 (8.91)	34.78
2	COL	542	55	459	28	1028	1570	Population based	75.49 (7.11)	69.51 (5.82)	42.93
1+2	ALL	3229	635	2185	409	2835	6064	mixed	77.68 (7.78)	73.14 (7.49)	39.36

[1] Geographic Atrophy
[2] Neovascular AMD
[3] Individuals with both GA and NV in either the same eye or in different eyes

Table 15. Association between AMD and rs17810398:C>T and rs17810816:A>G in five independent studies computed by logistic regression adjusted for covariates (N=3229 cases and 2835 controls)

	Number of individuals		All				Females				Males			
	(Females/Males)		MAF[1]				MAF[1]				MAF[1]			
Sample	**Cases**	**Controls**	**Cases**	**Controls**	P_{ADJ} *	**OR (95% CI) [2] ***	**Cases**	**Controls**	P_{ADJ} **	**OR (95% CI) [2] ****	**Cases**	**Controls**	P_{ADJ} d	**OR (95% CI) [2] ****
rs17810398														
GER1	710(455/255)	612(378/234)	0.133	0.106	0.028	1.311 (1.031-1.673)	0.134	0.098	0.026	1.415 (1.046-1.930)	0.131	0.118	0.501	1.147 (0.770-1.714)
GER2	996(670/326)	645(352/293)	0.143	0.109	1.39E-03	1.448 (1.157-1.823)	0.150	0.090	5.64E-05	1.891 (1.396-2.598)	0.127	0.132	0.957	1.010 (0.714-1.433)
US	681(428/253)	367(203/164)	0.141	0.109	0.037	1.366 (1.024-1.838)	0.149	0.094	3.58E-03	1.838 (1.233-2.804)	0.127	0.128	0.869	0.964 (0.626-1.497)
UK	300(193/107)	183(122/61)	0.145	0.104	0.047	1.530 (1.013-2.352)	0.155	0.102	0.044	1.701 (1.026-2.900)	0.126	0.107	0.560	1.238 (0.613-2.611)
COL	542(314/228)	1028(582/446)	0.137	0.124	0.199	1.177 (0.917-1.508)	0.131	0.126	0.234	1.235 (0.870-1.746)	0.145	0.122	0.472	1.139 (0.796-1.623)
ALL* **	**3229(2060/1169)**	**2835(1637/1198)**	**0.140**	**0.113**	**1.15E-06**	**1.332 (1.187-1.496)**	**0.144**	**0.105**	**2.62E-08**	**1.541 (1.324-1.796)**	**0.131**	**0.124**	**0.382**	**1.084 (0.905-1.298)**
rs17810816														
GER1	710(455/255)	612(378/234)	0.184	0.134	8.49E-04	1.435 (1.163-1.778)	0.183	0.124	0.0021	1.526 (1.169-2.008)	0.185	0.152	0.147	1.292 (0.916-1.832)
GER2	996(670/326)	645(352/293)	0.179	0.145	4.92E-03	1.335 (1.093-1.636)	0.184	0.122	1.47E-04	1.704 (1.300-2.256)	0.170	0.172	0.854	0.972 (0.716-1.322)
US	681(428/253)	367(203/164)	0.180	0.144	0.021	1.370 (1.051-1.797)	0.187	0.132	0.0085	1.622 (1.139-2.346)	0.167	0.159	0.683	1.089 (0.725-1.646)
UK	300(193/107)	183(122/61)	0.166	0.149	0.497	1.144 (0.779-1.693)	0.168	0.145	0.390	1.231 (0.770-1.998)	0.162	0.158	0.944	0.976 (0.503-1.938)
COL	542(314/228)	1028(582/446)	0.188	0.159	0.072	1.223 (0.981-1.522)	0.175	0.156	0.139	1.257 (0.926-1.701)	0.205	0.162	0.242	1.209 (0.879-1.660)
ALL* **	**3229(2060/1169)**	**2835(1637/1198)**	**0.180**	**0.148**	**1.76E-07**	**1.318 (1.188-1.462)**	**0.181**	**0.137**	**2.68E-08**	**1.471 (1.285-1.687)**	**0.179**	**0.162**	**0.141**	**1.129 (0.961-1.326)**

[1] minor allele frequency

[2] odds ratio (OR) and 95% confidence intervals (95% CI)

* adjusted for age and sex

** adjusted for age

*** analyses were additionally adjusted for study

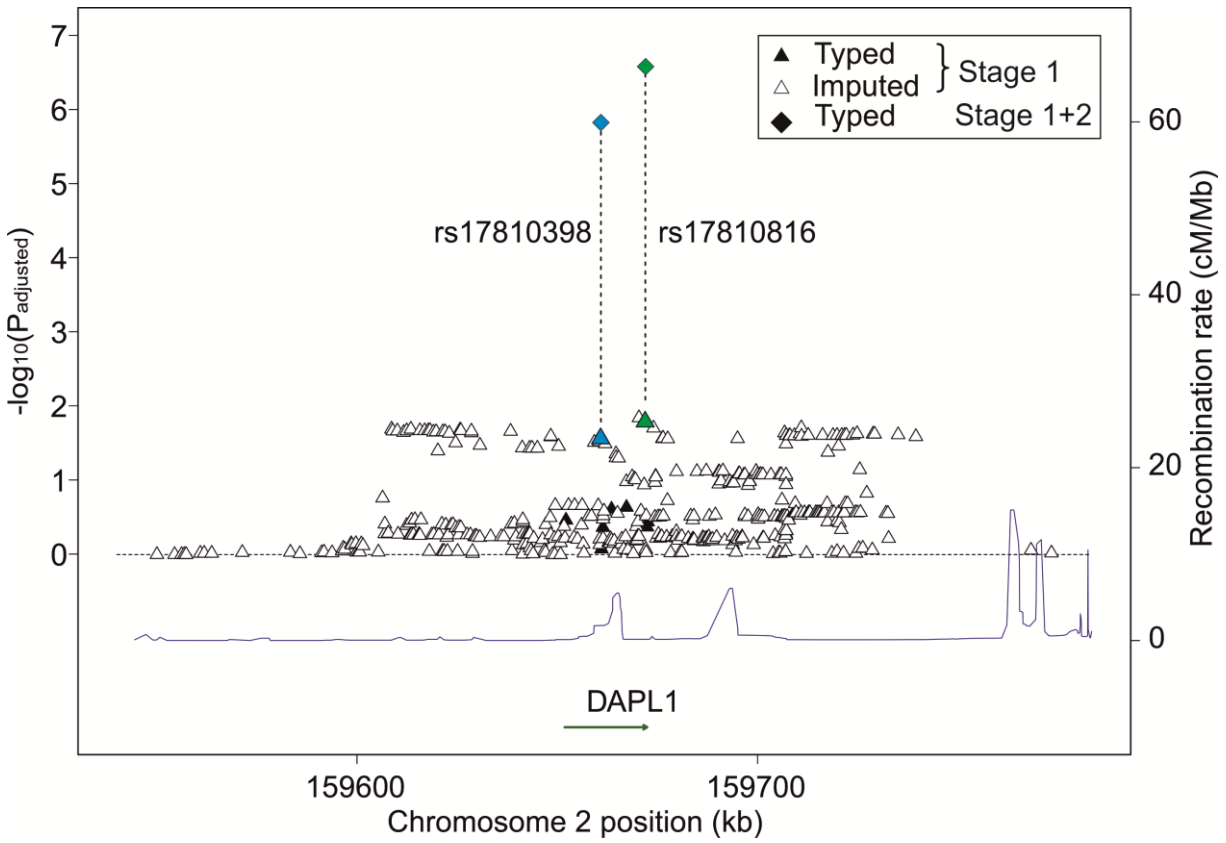


Figure 9. Association with AMD of imputed and typed variants at the *DAPL1* locus. Association signals of markers are shown by the log-P value from a logistic regression model (additive model adjusted for age and sex; y-axis) and are plotted against their physical position (x-axis). Stage 1 results (GER1 sample) are marked by filled (genotyped) and open triangles (imputed). Association signals of rs17810398 and rs17810816 in the pooled samples (3053 cases and 2737 controls) are indicated by blue and green diamonds respectively.

Forty-eight imputed SNPs revealed a more significant association than rs17810398 (P_{ADJ}=0.027), with a minimum P_{ADJ} of 0.014 at rs74923781 and its perfect proxy rs17810816:A>G (r^2 = 1 based on 1000 Genomes CEU samples) in *DAPL1* intron 4. De novo genotyping of rs17810816 gave consistent associations in terms of its direction in all four replication studies with a nominal significance ($P_{adj} \leq 0.05$) attained in GER2, US and COL. There was no indication of heterogeneity (I^2 = 0). The combined analysis yielded a P_{adj} of 1.76×10^{-7} after adjustment for age, sex and study (**Table 15** and **Figure 9**). Linkage disequilibrium (LD) and haplotype analysis in the GER1 study revealed rs17810398 and rs17810816 to be in moderate LD in controls (r^2 = 0.55, **Supplementary Figure S5**).

Variants rs1710398 and rs17810816 show a female specific association

Stratifying the combined analysis by phenotype, including AMD subtype and age-group, revealed no subgroup-specific association for rs17810398 or rs17810816 (**Figure 10**).

However, stratification by sex revealed that the association signals of both SNPs were confined to females with genome-wide significance (rs1710398: P_{ADJ}= 2.62 × 10^{-8}, rs17810816: P_{ADJ}= 2.68 × 10^{-8}). No AMD association was evident in males (rs1710398: P_{ADJ}= 0.382, rs17810816: P_{ADJ}= 0.141; **Figure 10**, **Supplementary Figure S6, Table 15**). The difference between sex-specific ORs was statistically significant (rs17810398: P_{diff} = 0.0034, rs17810816: P_{diff} = 0.014) at the 5% level and was observed in all studies analyzed (**Supplementary Figure S6**, **Table 15**). In the combined study, the minor allele frequency (MAF) of rs17810398 was lower in female controls than in male controls and higher in female cases than in male cases. A similar, albeit less pronounced effect was seen for variant rs17810816.

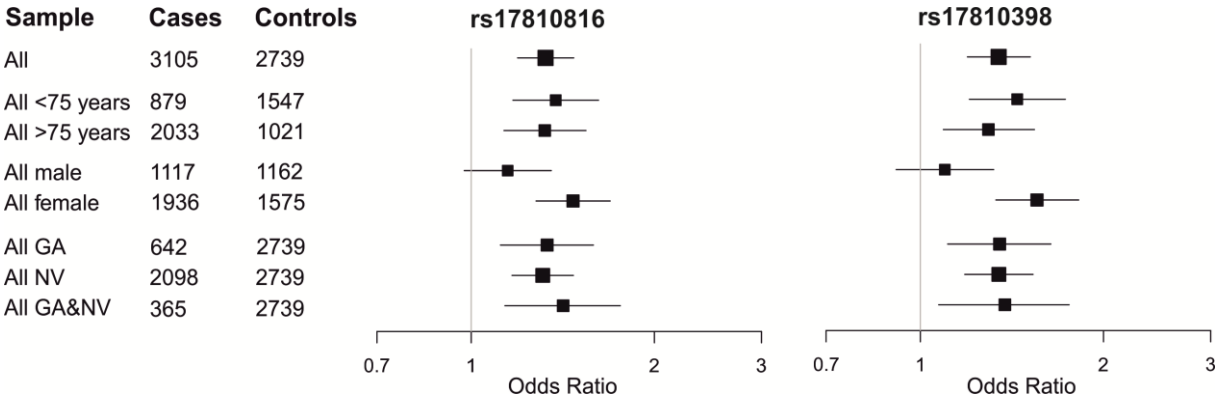


Figure 10. Subgroup analysis in the combined study of candidate SNPs rs17810398 and rs17810816 in the DAPL1 gene. Odds ratios and corresponding 95% confidence intervals are given with the size of each rectangles representing the respective number of cases. AMD phenotypic subgroups comprise patients with geographic atrophy (GA), and neovascular AMD (NV) and both latestage forms (GA&NV).

DAPL1 encodes four isoforms in retina/RPE

For expression analysis of the *DAPL1* locus, we scrutinized expressed sequence tags (EST) and identified three entries (GenBank accession numbers: DA417123 [thalamus], BG818506 [oligodendroglioma], BI016096 [lung tumor]) that suggested alternative splicing of *DAPl1* gene products. A potential correlation between rs1710398 and rs17810816 genotype and the occurrence of *DAPL1* isoforms was investigated by 3'-RACE experiments on four unrelated RPE/retina tissue samples, two of which were homozygous (ID_16 and ID_17) for the non-risk alleles and two of which were heterozygous (ID_13 and ID_14) for the risk alleles. After plasmid cloning of PCR products, we sequenced 1,200 cDNA clones and identified a total of 24 specific DAPL1 isoforms four of which (referred to as isoforms 1 to 4, **Supplementary Figure S7**) were consistently found in all samples. The most abundant isoform 1 (65–77 %

over all samples) corresponded to the *DAPL1* reference sequence (NM_001017920). Isoform 2 (6–12 %) and 3 (3–6 %) had not been reported before, whereas isoform 4 (3–6 %) matched EST BG818506 (**Figure 11A** and **Figure 11B**). Sequences corresponding to DA417123 and BI016096 were not detected in the RPE/retina RNA samples. RT-PCR analysis confirmed the expression of isoforms 1 to 4 in human tissues with isoform 4 likely being specific for RPE/retina (**Supplementary Figure S8**).

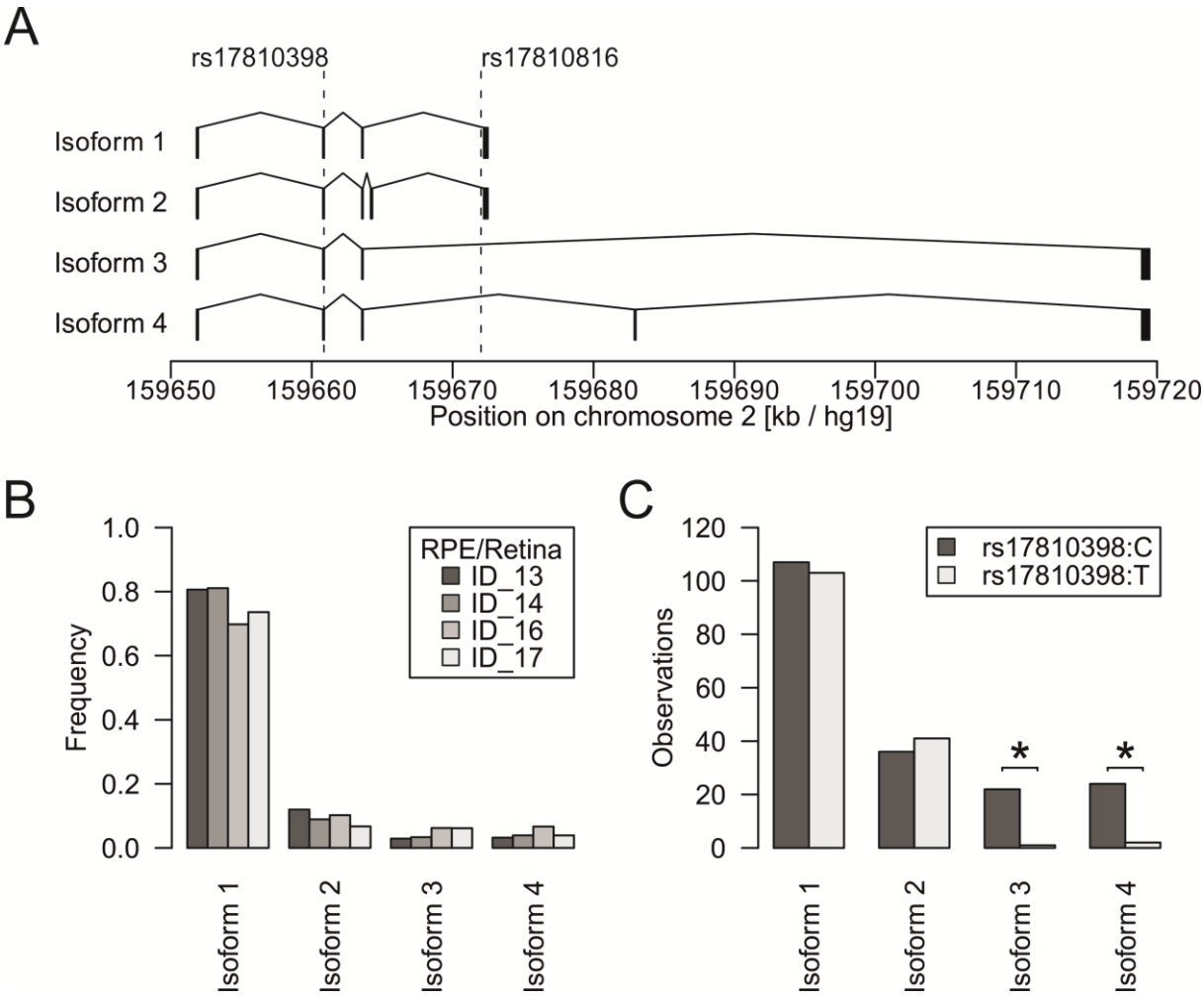


Figure 11. Functional consequences for isoform expression of DAPL1 variants. A. Exon/intron structure of four frequent DAPL1 isoforms (gene orientation is from left to right). SNP positions are marked by vertical dotted lines. B. Frequency of the four isoforms as determined from sequencing 1,200 cDNA clones that were obtained after 3'-RACE of four unrelated RPE/retina tissue samples either homozygous (ID_16, ID_17) or heterozygous (ID_13, ID_14) for the non-risk alleles of rs17810398 and rs17810816. C. Distribution of rs17810398 alleles in heterozygous RPE/retina tissue samples ID_13 and ID_14 (**Supplementary Table S11**). Statistically significant deviations from the reference transcript (i.e. isoform 1) are indicated by asterisks (P<0.0001).

Resequencing of candidate regions at the DAPL1 locus

In a search for additional risk variants at the extended *DAPL1* locus, we re-sequenced over 10 kb of intronic/exonic sequences in each of 12 probands homozygous for AMD risk alleles

rs17810398:T and rs17810816:G and eight probands homozygous for AMD non-risk alleles rs17810398:C and rs17810816:A (**Supplementary Figure S7**). Due to its extensive saturation with repeat structures, resequencing of the genomic region around *DAPL1*exon 4 of HQ179937 (isoform 4) was carried out for three individuals following sub-cloning of PCR fragments. In total, we detected 33 sequence variants (**Supplementary Table S12**, **Supplementary Table S13**), three of which (rs75277023:G>A, rs6146986, and rs144087548:A>T) were in strong LD with rs17810398 and rs17810816 (r^2 in controls > 0.9, **Supplementary Figure S5**). Variants rs6146986 and rs144087548 were of particular interest because the minor allele of the former represents a common 878 bp deletion in *DAPL1* intron 2 and the latter was predicted to affect a putative binding site of an exonic splicing enhancer (serine/arginine-rich splicing factor 1, SRSF1), 24-bp upstream of the most 3' exon shared by isoforms 3 and 4 (**Supplementary Figure S7**).Genotyping of rs6146986 and rs144087548 in the GER1 study confirmed their strong LD with rs17810398 (rs6146986: r^2 = 0.93) and rs17810816 (rs144087548: r^2 = 0.77) (**Supplementary Figure S5**, **Table 16**). Additional cDNA resequencing of eight RPE/retina tissues heterozygous for rs17810398 did not reveal additional coding variants (**Supplementary Table S14**).

Table 16. Association results in the GER1 study for four functional candidate SNPs in *DAPL1*.

SNP	**Position on chr 2 [bp / hg19]**	**Major allele**	**Minor allele**	**MAF[2]**		**Odds Ratio (95% CI)**	***P*[1]**	**R2[3]**
				Cases	**Controls**			
rs17810398	159,660,870	C	T	13.3%	10.5%	1.314 (1.033-1.680)	0.027	Ref.
rs6146986	159,661,997-159,662,874	-	878 bp deletion	13.5%	10.7%	1.310 (1.033-1.667)	0.027	NA
rs17810816	159,671,992	A	G	18.4%	13.4%	1.435 (1.163-1.778)	8.49×10^{-4}	0.866
rs144087548	159,718,894	A	T	17.4%	13.8%	1.296 (1.047-1.609)	0.018	0.735

[1]*P* from logistic regression adjusted for age and sex.
[2] MAF: minor allele frequency calculated in 710 cases or 612 controls
[3] R^2 to top variant based on CEU samples from the 1000 genomes project

AMD associated variants are correlated to a differential expression of DAPL1 isoforms

Samples heterozygous for rs17810398 (ID_13 and ID_14) were characterized by a significantly different abundance of non-risk and risk isoforms 3 and 4 ($P < 10^{-4}$). This was not the case for isoforms 1 and 2 (**Figure 11C**, **Supplementary Table S11**). *DAPL1* isoform expression in RPE/retina was further evaluated in vivo by semi-quantitative cDNA sequencing (**Figure 12**). Of 39 unrelated RPE/retina tissues available, seven were heterozygous for AMD associated variants rs17810398, rs6146986, rs17810816, and rs144087548. In agreement with our 3'-RACE data, these samples revealed differential expression of isoforms 3 and 4 but not isoforms 1 and 2 (**Figure 12**). Interestingly, sample ID_11 was heterozygous for rs17810398, but homozygous for the non-risk alleles of rs6146986, rs17810816, and rs144087548. In this sample, expression intensities of isoform 3 and 4 alleles were equal excluding rs17810398 as a functional variant involved in the differential expression of isoforms 3 and 4. This leaves rs6146986, rs17810816, rs144087548 or an as yet unknown but correlated variant as the truly functional risk variant at the *DAPL1* locus.

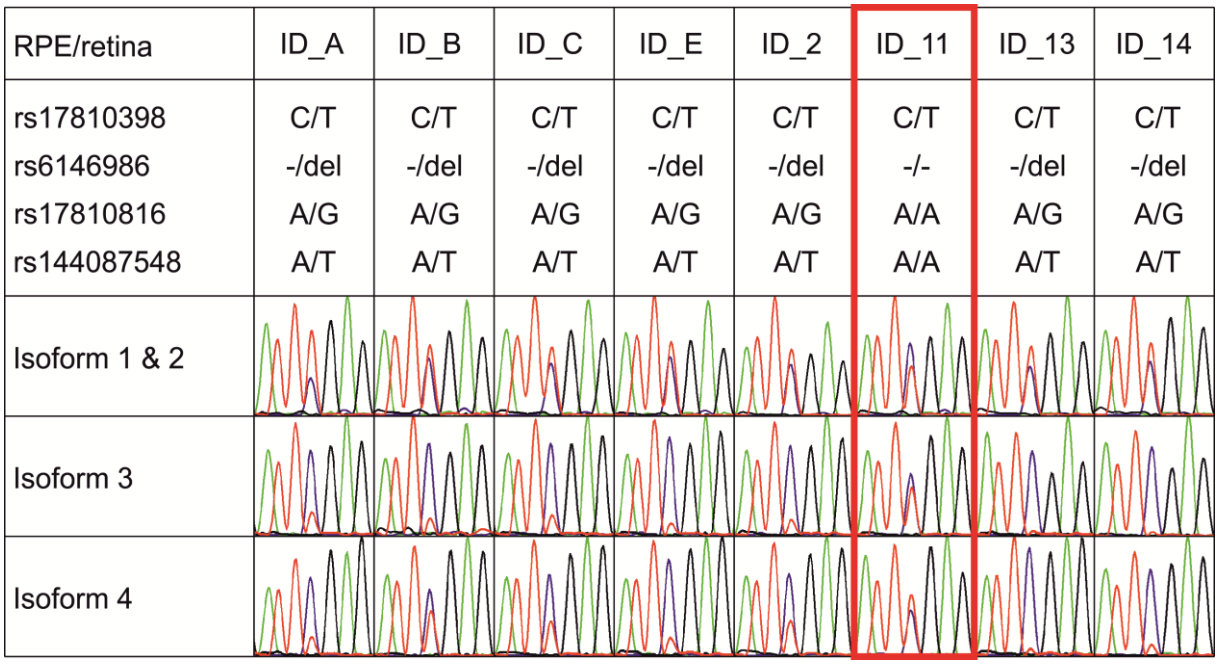


Figure 12. Semi-quantitative cDNA sequencing of eight RPE/retina tissue samples heterozygous for synonymous coding SNP rs17810398:C>T. The chromatograms of the variant nucleotide at rs17810398 flanked by +/- 3 bp are shown for isoforms 1&2, 3, and 4. Isoforms were specifically amplified by three different exon-spanning primer combinations. Genotypes of the four candidate variants are given above the chromatograms. The sample heterozygous for rs17810398 but homozygous for the non-risk alleles of rs6416986, rs17810816, and rs144087548 is highlighted in red.

5.3 Discussion

Here, we provide evidence that *DAPL1* is an AMD-associated gene and that its disease association is female-specific. To our knowledge, this is a first study reporting a sex-specific genetic association with AMD at a genome wide significance level. Although lead SNPs rs17810398 and rs17810816 have been imputed into large GWAS data sets, neither variant has been identified before as AMD-associated [33]. This is likely due to the female specificity of the association as male/female ratios in multi-center GWAS tend to differ greatly between cohorts thereby potentially leading to reduced power. Behrens *et al*. [168] have methodologically shown that gender-stratified analyses greatly increase power to detect gender-specific effects. Additionally, we have also observed the female specific association in our population-based sample (COL study) only by adjusting the analysis for age. This further indicates that different study types increase the heterogeneity and therefore may lead to decreased power to detect this association.

We also considered the possibility that age is a confounding factor in our study since (i) females are slightly older than males and (ii) cases are older than controls. If any of the variants would be correlated with longevity, this could potentially confound our analysis. However, we found no evidence for a correlation between either SNP or age, neither in cases, controls, females or males separately or analyzed jointly ($P > 0.05$). Additionally, we note that logistic regression analyses were adjusted for age in our analyses. Furthermore, two SNPs at the *DAPL1* locus (rs9869 and rs10497199) were investigated in a recent study by [169]. For these two variants, the authors found no association with longevity ($P > 0.5$). Variant rs9869 is weakly linked to markers rs17810398 ($r^2 = 0.13$) and rs17810816 ($r^2 = 0.1$) while rs10497199 is independent of either variant ($r^2 < 0.1$). Taken together, these findings have led us to exclude a confounding effect of age in our analysis.

The observed association in the present study could eventually be explained by a population substructure in our cases or controls from the UK or US. Although we cannot definitely exclude such a possibility, it is of note that frequency and effect sizes of the two DAPL1 risk variants rs17810398 and rs17810816 (in males, females and jointly) observed in the UK and US study are similar to the frequencies observed in the combined German cohort (**Table 15**). The German samples derive from a genetically homogenous population from a small area in southern Germany. Homogeneity was estimated previously from genome-wide data available for a subset of cases [33]. From these data, we conclude that the US and UK study primarily

consists of Caucasians which are genetically similar to the German cohort and, if at all, population substructure may only exert a minor effect on study outcome.

Although *DAPL1* is evolutionarily conserved, only little is known about its function. It has been shown to be abundantly expressed in the retina/RPE transcriptome [170] as well as in epidermis, oesophageal epithelium, and tongue epithelium where it appears to be involved in the early stages of stratified epithelial differentiation [171]. Based upon strong amino acid sequence similarities, *DAPL1* has also been connected to the death-associated protein (DAP), a basic, proline-rich protein of 15-kD molecular weight that acts as a positive mediator of programmed cell death upon induction by interferon-gamma [172]. Clarification of the cellular function of DAPL1 in the RPE/retina is required to allow more detailed insight into this novel pathway of AMD pathogenesis.

We have shown that *DAPL1* is present in a multitude of correctly spliced isoforms, two of which, isoforms 3 and 4, were specifically down-regulated in the presence of AMD associated alleles. Although we could not identify the causative variant at the *DAPL1* locus, we excluded lead SNP rs17810398 as the presence of the T-risk allele in one patient (ID_11) had no influence on the transcript levels of isoform 3 or 4. Notably, the unique C-terminus of isoform 4 encodes two potential transmembrane domains with significant homology to the rhodopsin-like G-protein coupled receptor (GPCR) family. Another member of the GPCR family, the G protein-coupled estrogen receptor 1 (GPER), plays a role in intracellular signaling following estrogen binding and could provide a useful lead when searching for factors involved in sex-dependent AMD risk. While at present we cannot explain the gender-specificity of the association with *DAPL1*, our results provide a starting point at a molecular level to investigate why AMD is more frequent in women than in men [173].

Taken together, we investigated 25 gene loci of interest to AMD pathology and excluded all but one from being disease-associated. Our data implicate *DAPL1* as a novel gene involved in AMD pathology although the cellular functions of this gene and of its various differentially spliced transcripts remain elusive. Our study revealed a correlation between risk variants at rs17810398 and rs17810816 on the one hand and expression levels of *DAPL1* isoforms 3 and 4 on the other, the latter being specifically expressed in RPE/retina tissue. We also reported a significant sex difference of the effect of *DAPL1* where only females showed an association signal at this locus. Although speculative at present, this sex difference may be explained by a role of *DAPL1* variants in sex-specific signaling processes. Our findings add another piece to

the puzzle of the genetic architecture of AMD, which, once completed, should allow refined identification of individuals at risk for this disease.

5.4 Materials and Methods

Subjects

Five independent studies were included in our study comprising a total of 3,229 unrelated Caucasian patients with clinically documented AMD (cases) and 2,835 unrelated individuals with comparable age range and ethnicity without signs of macular disease (controls) (**Table 14**). All data were available for analysis at the analysis center in Regensburg.

Discovery study GER1 (stage 1) included 710 AMD patients and 612 controls from the University Eye Clinic of Würzburg (Germany). The four replication studies (Stage 2) comprised (i) 996 AMD patients and 645 controls from the University Eye Clinics in München, Tübingen and Würzburg (Germany) (GER2); (ii) 681 AMD patients and 367 controls from Columbia University (New York, USA) (US); (iii) 300 AMD patients and 183 controls from the Royal Victoria Hospital (Belfast, UK) (UK); and (iv) 542 AMD patients and 1,028 controls from the Department of Ophthalmology at the University Hospital Cologne, Germany (COL). Cases and controls were examined by trained ophthalmologists. Stereo fundus photographs were graded according to standardized classification systems as described previously [58]. The study was conducted at all sites in strict adherence to the tenets of the Declaration of Helsinki and was approved by the respective Ethics Committees at the University Eye Clinics of Würzburg, München and Tübingen, by the Institutional Review Board at Columbia University, by the Research Ethics Committee of Queen's University Belfast and by the local Ethics Committee in Cologne.

Genotyping

Genomic DNA was extracted from peripheral blood leukocytes according to established protocols. Genotyping of SNPs was carried out by direct sequencing, TaqMan SNP genotyping (Applied Biosystems, Foster City, USA) or by primer extension of multiplex PCR products and subsequent allele detection by matrix-assisted laser desorption/ionization time of flight (MALDI-TOF; Sequenom, San Diego, USA). Direct sequencing was performed with the Big Dye Terminator Cycle sequencing kit version 1.1 (Applied Biosystems, Foster City, U.S.A.) according to the manufacturer’s instructions. Reactions were analyzed with an ABI Prism 3130xl sequencer (Applied Biosystems). TaqMan pre-designed SNP genotyping assays (Applied Biosystems) were used according to the manufacturer’s instructions. The

rs144087548 variant was genotyped by polymerase chain reaction (forward primer: 5'-CGC AGA CAT GAT GCT GGG GGT-3'; reverse primer: 5'-ACA TGC AAG ACG GGG AAT TGA-3') followed by *HpyCH4III* digestion (New England Biolabs, Ipswich, USA) and restriction fragment length analysis. All SNPs showed high genotyping quality with an average call rate > 98% in each of the five case-control samples.

Statistical methods

Discovery study: We excluded three SNPs (rs2279227 [*RGR*], rs4620343 and rs3812532 [*TRPM3*]), each with significant deviation from Hardy-Weinberg equilibrium (HWE, $P \leq 0.05$) in the control group of the discovery sample. SNP association analysis was carried out by logistic regression adjusted for age and sex. All analyses modeled an additive genetic effect and the genotype was coded as the number of alleles present at a given variant (i.e. 0, 1 or 2).

Replication studies and combined analysis. All SNPs were in HWE ($P > 0.05$). We used the same tests for SNP association analysis as in the discovery study. We also combined the individual data from all five studies and also adjusted the respective analyses by study center (coded as factors). The I^2 measure was computed to measure between-study heterogeneity. We also conducted sex-stratified analyses for each study separately and for all study samples combined. Sex differences were asses for statistical significance using a t-test derived from sex-specific beta estimates and corresponding standard errors.

All reported P values were two-sided except where noted otherwise. All SNP association analyses were carried out with R (v3.0.1, http://R-Forge.R-project.org/). To allow a more detailed inspection of the genomic region of interest, measures of linkage disequilibrium (LD) were calculated using R package snp.plotter [174].

Imputation of SNPs

Prior to imputation, 8 tag SNPs in *DAPL1* were phased in the GER1 study individuals using SHAPEIT2 [175]. Then, untyped SNPs were imputed with IMPUTE2 [176] using the 1000 Genomes Phase I integrated haplotypes (release 20110521) as reference panel. After the exclusion of SNPs with imputation quality ("info") < 0.5, the genotype probabilities (dosages) of the remaining SNPs were also analyzed by logistic regression in R, using an additive model adjusted for age and sex.

Genomic resequencing

Genomic resequencing was done for regions of interest defined by the presence of certain gene elements (putative promoter, coding exons of transcripts NM_001017920.2, HQ179935, HQ179936, and HQ179937) or conserved elements based upon the "46-Way Most Cons" track of the UCSC genome browser, NCBI Build 37/hg19]. Regions within extensive repeat structures were excluded (**Supplementary Figure S7**). Resequencing primers are listed in **Supplementary Table S11**.

Prediction of functional impact of risk variants

The functional impact AMD associated SNPs (with known dbSNP ID) on RNA processing as well as protein sequence, structure and function was predicted using the web-based "SNP Function Prediction" tool implemented in the "SNPinfo Web Server" (http://snpinfo.niehs.nih.gov/index.html) [177]. For newly identified SNPs, we used ESEfinder 3.0 to predict the effect of a given SNP allele on putative exonic splicing enhancers **(**http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi) [178].

Characterization of major splice variants of DAPL1 in human retina/RPE

To determine major splice variants and functional polyadenylation sites, 3' rapid amplification of cDNA ends (3'-RACE) experiments were conducted. RNAs from RPE/retina tissues that were either heterozygous (ID_13 and ID_14) or homozygous (ID_16 and ID_17) for the non-risk rs17810398:C allele were isolated by RNeasy Mini Kit followed by DNAseI treatment (QIAGEN, Hilden, Germany). 3'-RACE was conducted with the FirstChoice RLM-RACE Kit (Applied Biosystems/Ambion, Austin, USA) according to the manufacturer's instructions. Forward primers for first and second (nested) PCR were 5'-GCA CTG GCA CACG CTA TG-3' and 5'-TTG GCA CCT TGG AAA GAC ATA CC-3', respectively. Amplified RACE products were ligated into the pGEM-T vector (Promega, Madison, USA). PCR products were obtained with M13 forward and M13 reverse primers from a total of 1,200 clones. Of these, 597 clones were sequenced; the remaining 603 could unequivocally be assigned to *DAPL1* isoform 1 (NM_001017920.2, HQ179934) by visual gel inspection. The sequences of isoforms 2 to 6 were submitted to GenBank (HQ179935, HQ179936, HQ179937, HQ179938, HQ179939).

Expression analysis and semi-quantitative resequencing

Eight RPE/retina tissues with risk variant genotypes as given in **Figure 11** and **Supplementary Table S13** were used as templates to amplify isoform-specific PCR products

with forward primer 5'-GCA CTG GCA CAC GCT ATG-3' and the isoform-specific reverse primers 5'-CGA GGC TGC TGA ATA ATG TAG-3' (isoform 1 & 2), 5'-TCT GGA TCC TCT GAG CTT CTT CTC-3' (isoform 3) or 5'-CTG GAT CCT CTG AGC TTC TTG TGT-3' (isoform 4), followed by sequencing with the forward primer. Primers for the *GUSB* gene were 5'-ACT ATC GCC ATC AAC AAC ACA CTC ACC-3' and 5'-GTG ACG GTG ATG TCA TCG AT-3'. For tissue samples, sex was determined with fluorescence-based PCR analysis of the homologous, X- and Y-linked genes *AMELX* and *AMELY* as described in Sullivan *et al.* [179].

6. General Discussion

Tables and figures in the discussion are based on the following review articles:

1. Grassmann F, Ach T, Brandl C, Heid IM & Weber BHF (2015) **What does genetics tell us about AMD?** *Ann Rev Vis Sci,* in revision.

2. Grassmann F, Fauser S & Weber BHF (2015) **The genetics of age-related macular degeneration (AMD) and its usability for designing treatment options** *Eur J Pharm Biopharm.,* in revision

The main aim of this thesis was to investigate genetic as well as non-genetic markers involved in AMD pathogenesis. Specifically, this thesis focused on a number of questions related to AMD pathogenesis:

(i) is it possible to quantify the cummulative genetic risk for AMD of an individual and project absolute risk estimates based on a genetic risk score (GRS) calculated from common and previously published genetic variants? Are there additional features in AMD pathogenesis that can be evaluated based on such a genetic risk score?

(ii) are (circulating) microRNAs involved in neovascular AMD pathogenesis and can they be used as a specific biomarker for this late stage disease? What pathways are regulated by identified circulating microRNAs and how do these microRNAs influence angiogenesis? Are other forms of late stage disease such as atrophic AMD also associated with neovascular AMD related cmiRNAs?

(iii) are genetic variants or non-genetic factors correlated with progression of geographic atrophy lesions and are these correlations independent of each other?

(iv) are there additional common genetic variants associated with AMD, which have not been detected by meta analyses of individual study data in the AMD Genetics Consortium [33]?

6.1 Genetic variants are among the strongest factors influencing AMD risk

The heritability of late stage AMD is estimated to be between 45% to 71% [72] and therefore comparable to the estimated heritabilities of other complex diseases e.g. Alzheimer's disease [180] or Psoriasis [181]. However, the genetics of AMD is unique among complex diseases as several loci with strong effect sizes of common variants have been identified (e.g. variants in

ARMS2/HTRA1, *CFH*, *C3* or *CFB* **Figure 13**). These variants exceed the expectation of effect size vs. their respective effect allele frequency usually observed on complex diseases (**Figure 13**) [33,182].

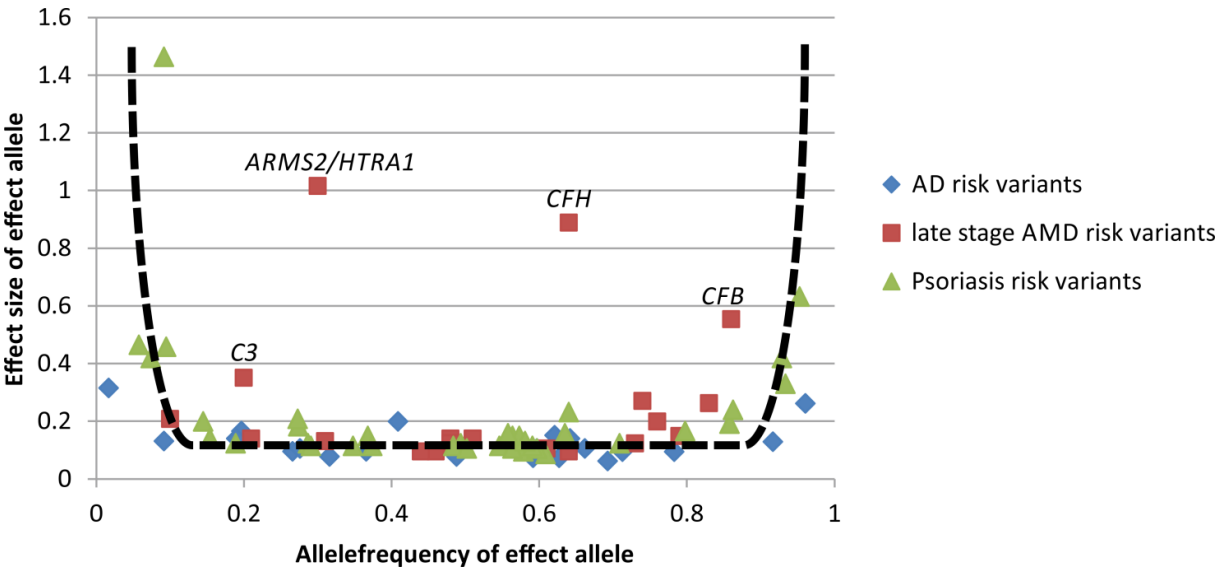


Figure 13. Plot of effect size versus effect-allele frequency of AMD, Psioriasis and Alzheimer's disease (AD) risk variants. For complex diseases, the observed effect sizes are usually dependent on the frequency of the variant. Common variants have generally low effect sizes while rare variants have increased effect sizes. For AMD associated variants, several loci exceed this expectation (ARMS2/HTRA1, CFH, CFB and C3). Data based on Fritsche *et al.* 2013, Lambert *et al.* 2013, Tsoi *et al.* 2012 [33,181,182]. Figure taken from Grassmann *et al.* 2015, *Ann Rev Vis Sci.*

The cumulative impact of known variants on disease risk can also be expressed as the fraction of disease variance being explained by known factors. This estimate usually varies between 1% and 20% for most complex diseases [183]. However, these fractions are still considered minor by recent reports on genomics and genetic studies in popular media, which have proclaimed a "failure of the human genome" [184]. While the explained fraction of disease variance of known genetic factors for diseases with a heritability smaller than 30%, like type 2 diabetes [185], is indeed low (around 2%), they can still account for a higher disease variabilty than individual environmental risk factors. For instance, smoking explains about 1% of disease variance in type 2 diabetes, while the body mass index explains around 9% [186]. In diseases with a stronger genetic contribution, known genetic risk factors account for a higher explained disease variance. For Crohn's disease, variations identified in genome-wide association studies explain up to 12% of the disease variance [187], while the strongest measurable non-genetic predictor (smoking) only accounts for 3% of the variance [188].

Depending on the overall prevalence assumed for late stage AMD, the variance jointly explained by 19 disease associated variants was estimated to be between 10% and 30% [33], further highlighting the unique properties of AMD genetics among complex diseases.

Although we can not explain a large proportion of the variance in most complex diseases, it is not essential to fully explain all of the disease variance to provide valuable insights into disease mechanisms. Even small effects can implicate important pathways, e.g. the previously underappreciated importance of the autophagy pathway in Crohn's disease pathology [189].

6.2 Novel genetic factors associated with AMD risk

In order to advance our understanding of genes and pathways involved in AMD beyond GWAS-related approaches, we conducted a candidate gene based association study including more than 109 candidate variants in 25 candidate genes (see chapter 5) and found a significant association for variants in the *DAPL1* gene. This association was primarily confined to females. By adding both associated top variants (rs17810398 and rs17810816) to the genetic risk score model proposed in reference [58], the resulting AUC of the model increased by 0.01 to 0.83, thus slightly increasing classification accuracy. This increase in the AUC value, may, however, be inflated. Due to the "winner's curse", the odds ratio estimates of the risk variants are larger in the study that identified the variant. Since all individuals contributing to the calculation of the risk model were also part of the study identifying the DAPL1 risk variants, the impact of those variants is likely overestimated. The rather small contribution to the risk prediction is, however, not surpising and was predicted previously [42]. Even variants with large odds ratios above 2.00 can only be expected to increase the prediction accuracies by small margins, because the genetic risk for cases and controls overlaps significantly. For instance, a risk allele that increases the odds for the disease by two fold (i.e. has an odds ratio of 2.00) and an allele frequency frequency of 20% in cases, would have an allele frequency of 11% in controls. Consequently, more than 20% of all controls would carry at least one risk allele. Therefore, it is difficult to distinguish cases and controls from such associated variants reliably, unless the effects of several variants are analysed jointly. Even then, a substantial overlap of the genetic risk between cases and controls is evident (see **Figure 4**).

By performing statistical power analyses, we estimated that our candidate gene approach allowed us to find a significant association with AMD ($p < 0.05$) in the discovery study in case the risk variant is associated with AMD with an odds ratio greater than 1.40. Therefore, we cannot entirely exclude the tested variants and consequently the analyzed genes from

being relevant for AMD, since our study may have missed weakly associated variants (OR < 1.40). Hence, large scale studies with more participants are necessary to provide a definite answer on the involvement of the remaining candidate genes in AMD.

6.3 Genetics of AMD and implications for treatment options

Several studies estimated the contribution of known risk variants to treatment response in anti-VEGFA treatment for neovascular AMD. Variants in CFH, ARMS2 and VEGFA have been implicated with moderate effect sizes to be involved in treatment success, although there are conflicting data [190–194]. Even though neovascular AMD can be reliably treated and vision can be stabilized for some time, eventually patients with the neovascular form of AMD show a degeneration of the RPE and photoreceptors reminiscent of geographic atrophy [195]. Therefore, efforts are currently made to explore novel treatment options for neovascular AMD as well as for the so far untreatable atrophic form (GA) [16]. While 9 out of 10 clinical trials aimed at treating neovascular AMD are based on anti-VEGFA treatment and most of these report a successful impediment of neovascularisation and leakage, many different targets are currently being considered and tested in preclinical and clinical trials in an effort to treat atrophic AMD (**Table 17**) [16]. These latter trials either use neuroprotective agents, anti-complement antibodies or anti-inflammatory substances, many of which are based on antibody approaches. So far, none of the studies treating atrophic AMD, however, report a significant treatment effect (**Table 17**) and this can, in part, be attributed to our limited knowledge on factors involved in disease progression and severity. Genetic association studies have a tremendous success to identify variants and thus genes involved in AMD. The identified genes are, however, per definition involved in disease risk (i.e. risk for incident AMD) and not necessarily in disease severity or progression and thus provide little insight into processess that can be targeted in AMD patients currently suffering from the disease (i.e. risk for prevalent AMD). Rather, they provide insight into processess involved before the onset of disease and as such are valuable targets for risk prediction before the onset of symptoms and consequently, for preventive treatment. In fact, little is known on the impact of genetic variants or non-genetic factors on disease severity, e.g. growth rates of GA lesions, area of neovascularisation or scarring, cummulative drusen area or similar measurements.

Table 17. Clinical trials for early AMD and geographic atrophy (adopted from Holz et al 2014), taken from Grassmann *et al.* 2015 *Eur J Pharm Biopharm.* in revision

Clinical trial registration (sponsor)	AMD type	Phase	No. of patients	Drug/Treatment	Mechanism	Current status	primary end point	Result
NCT01802866 (Acucela Inc.)	GA	2/3	about 440	Emixustat (RPE65 inhibitor) vs. placebo	Visual cycle inhibitor	Recruiting	Change in GA area	Pending
NCT00429936 (Sirion Therapeutics Inc.)	GA	2	246	Fenretinide (retinol analogon) vs. placebo	Visual cycle inhibitor	Completed	Change in GA area	No significant effect
NCT00695318 (Alimera Sciences)	GA	2	40	Fluocinolone acetonide (corticosteroid) vs. sham	Suppresion of inflammation	Ongoing, not recruiting	Change in GA area	Pending
NCT00766649 (NEI)	GA	1/2	11	Rapamycin (mTor pathway inhibitor)	Suppresion of inflammation/ Immunisuppresant	Recruiting	Change in GA area and visual acuity	No significant effect
NCT01229215 (Genentech)	GA	2	143	Lampalizumab (antibody against CFD) vs. sham	Suppression of complement activation	Completed	Change in GA area	Pending
NCT00935883 (Alexion Pharmaceuticals)	GA	2	30	Eculizumab (antibody against C5) vs. sham	Suppression of complement activation	Completed	Change in GA area	No significant effect
NCT01527500 (Novartis)	GA	2	about 120	LFG316 (antibody against C5) vs. sham	Suppression of complement activation	Recruiting	Change in GA area	Pending
NCT00447954 (Neurotech Pharmaceuticals)	GA	2	48	CNTF (neural growth factor) vs. sham	Neuroprotection	Completed	Visual acuity	No significant effect
NCT00890097 (Alcon Research)	GA	3	772	Tandospirone (serotonin 1A receptor agonist) vs. placebo	Neuroprotection	Completed	Change in GA area	No significant effect
NCT00658619 (Allergan)	GA	2	119	Brimonidine tartrate (α-2 adrenergic receptor antagonist) vs. sham	Neuroprotection	Completed	Change in GA area	Pending
NCT01344993 (Advanced Cell Technology)	GA	1/2	16	Stem cell derived human RPE	Transplanation of RPE cells	Recruiting	Safety/tolerability	Pending
NCT01782989 (MEDARVA Foundation)	GA	2/3	286	Doxycyclin (antibiotic, inhibitor of metalloproteases) vs. placebo	Inhibition of metalloproteases/ suppresion of inflammation	Recruiting	Change in GA area	Pending
ACTRN12612000704897 (Ellex)	early	2	360	Low-energy laser (2RT laser)	Stimulating wound healing	Recruiting	Progression to advanced AMD	Pending
NCT00951288 (Catholic University)	early	1/2	30	Saffron supplementation (contains antioxidant carotids) vs. placebo	Neuroprotection	Recruiting	Visual acuity and function	Improved visual function
NCT01528605 (Peking University)	early	2	168	Lutein and zeaxanthin (antioxidant carotids) supplementation vs. placebo	Neuroprotection	Completed	Macular pigment optical density (MPOD)	Improved MPOD

6.4 Factors associated with disease severity and prevelant AMD

A recent study excluded AMD associated risk variants in *CFH, ARMS2/HTRA1* and *C3* from being associated with GA lesion growth [139], while others report a significant association for some of those variants [115,140]. By analyzing the largest dataset available on GA lesion growth, we showed a significant contribution of two genetic variants to GA growth (C3_rs2230199 and ARMS2_rs10490924). Noteworthy, GA in both eyes also significantly increased the rate of progression of GA lesions, with effect sizes comparable to the genetic variants. Taken together, the genetic variants explained about 4% of the inter-individual variability of GA lesion growth, while the addition of the disease status of the fellow eye (presence of GA or absence of GA in the fellow eye) increased the explained variance to 7.2%. These findings have direct implications for the design and interpretation of the success of treatment trials. Although the risk increasing allele in C3_rs2330199 results in a more active form of complement component 3, thus increasing risk for AMD, we could show that the same risk increasing allele reduces GA growth significantly. Inhibition of the complement cascade with anti-complement antibodies might therefore not be promising venues since low levels of inflammation seem to be beneficial in neurodegenerative diseases [154,156,157].

We also explored non-genetic (epigenetic) markers which are associated with prevalent AMD. Circulating microRNAs in serum and plasma have been shown to be dysregulated in a variety of age related degenerative diseases. We therefore aimed to further our knowledge on dysregulated microRNAs in prevalent AMD and showed that, indeed, several microRNAs are associated with prevalent neovascular AMD. These microRNAs are predicted to regulate target genes in inflammatory as well as angiogenic pathways. The precise targets of these microRNAs can be best evaluated *in vitro* and *in vivo* with microRNA mimics and anti-microRNAs. Such agents are potential treatment options in itself. Future studies tackling the role of circulating microRNAs in AMD pathology are underway (see Perspectives)

6.5 Understanding genes and pathways involved in disease risk

Although genome-wide association studies showed a huge success to identify disease associated variants for virtually all complex diseases [196], care has to be taken in the interpretation of these results. Currently, several pitfalls for the interpretation of GWAS data exist and are outlined below.

6.5.1 Clinical heterogeneity of AMD and implications for associated variants

In AMD, most association studies investigated the association of variants with early AMD or late stage AMD, both of which are clinically highly heterogenous groups. Profound clinical heterogeneity exists for the early stages of the disease, and the late stage AMD is comprised of at least two major forms of the disease, atrophic and neovascular AMD [197]. The reasons for jointly investigating distinct disease manifestations are manifold: (i) the more patients a study includes, the more statistical power this study has to identify a significant assocation; (ii) both late stage forms are sometimes not easily distinguished, for instance in case where an individual is affected with GA and NV in the same eye; (iii) individuals suffering from GA may spontaneously develop NV AMD and vice versa, further increasing classification difficulties; (iv) in early AMD, various phenotypical manifestations of the disease exist (e.g. different drusen characteristics, changes in pigmentation), which often occur jointly in the same eye. It is therefore difficult to delineate distinct sub phenotypes of early AMD. Taken together, any association found so far is a genetic associations common to jointly investigated phenotypes (i.e. GA and NV or all stages or types of early AMD). Furthermore, clinical heterogeneity is probably one of several possible factors reponsible for the reduced effect sizes of late stage AMD associated risk variations in early AMD. Only recently, a sufficient amount of patients and controls have been recruited to further investigate the genetic basis of different AMD subtypes (see Perspectives).

6.5.2 The search for the causative variant, affected gene and pathway

More than 19 disease associated loci have been identified in a recent meta analysis [40] and an additional 16 loci are expected to be reported soon by the International AMD Genomics Consortium (IAMDGC) (**Figure 1**, **Figure 14**). In a first attempt to annotate and prioritize risk variants for further research, we have used the {FunciSNP} pipeline implemented in R to find variants correlated ($R^2 > 0.7$) with the SNP showing strongest association at each locus and created a circos plot of all annotated variants. For example, among the available data sets of the 19 loci [40], we found coding variants in six loci and several 3'-UTR variants in one locus. Additionally, in one locus we found variants annotated as expression quantitative trait loci (eQTLs), potentially altering the expression of a gene. In order to find tissue specific exons, we conducted next generation sequencing of RNA libraries (RNASeq) derived from several retinal cell types like photoreceptors, RPE and choroid cells [198]. In two loci, we found variants in novel (non-canonical) exons, representing the most likely candidates in

these loci. In the remaining nine loci, the most likely functional variant is still elusive (**Figure 14**) and identification could be hindered by complex haplotype structures with long stretches of linkage disequilibrium, multiple signals at each locus, insufficient genotype information, particularly for copy number variants (e.g. deletions or insertions) and lack of affordable high-throughput methods to characterize many candidate variants especially in the appropriate tissue and cell types [199].

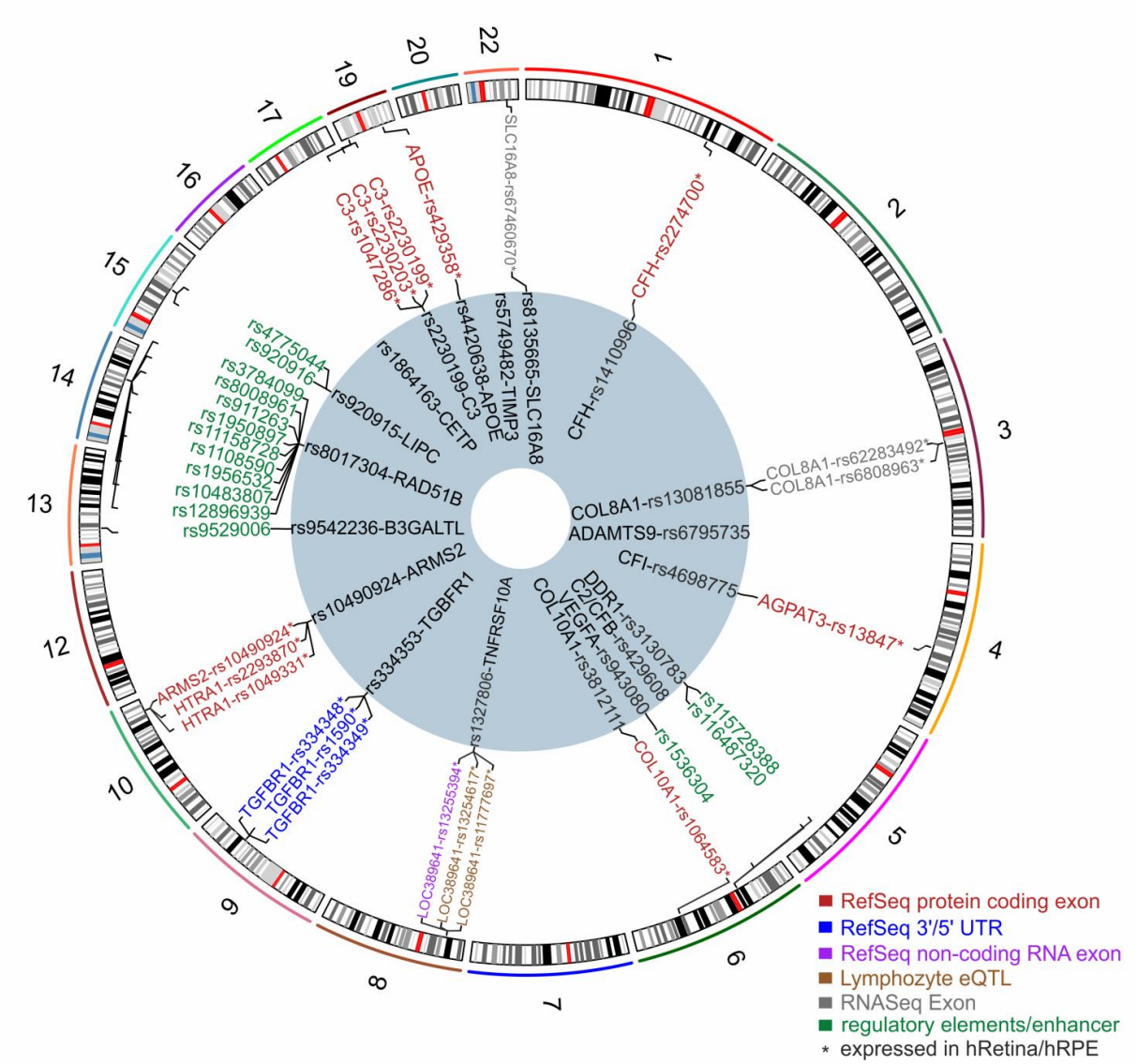


Figure 14. Functional annotation of AMD associated candidate variants. Top variants as well as correlated variants with $R^2 > 0.7$ to the top variant in nineteen AMD associated loci were annotated based on biofeatures implemented in package {FunciSNP} and a detailed map of the location and annotation of AMD associated candidate variants and regions in the genome was plotted. Candidate variants are color-coded by functional category. Transcripts expressed in the human retina, human RPE or human RPE derived from induced pluriponent stem cells (iPSC) are marked with an asterisk *. Several, previously published functional variants are not shown as their correlation with the top SNP is low ($R^2 < 0.7$), thus these variants are not detected by simple correlation analyses. Figure taken from Grassmann *et al.* 2015, *Ann Rev Vis Sci*

In order to explore novel ways to prevent or delay the onset of AMD, further molecular and statistical research is required to identify the responsible variants and affected genes and pathways [199].

6.6 Perspectives

Despite many efforts and huge successes of genome-wide association studies, the fundamental challenges in AMD and in other diseases still await to be addressed [200], namely finding appropriate animal models and treatment options based on the knowledge of genetic risk factors. For several diseases, findings of genome-wide association studies translated into identification of appropriate mouse models that mimic the genetic risk of the disease [201,202] and help to further shed light on processes involved in disease pathology. So far, no association study has led to the direct identifcation of a treatment option, although it is now possible for some diseases to adjust treatment recommendations according to the genotype of affected individuals [200].

In AMD, we have a good understanding of a large proportion of the disease genetics, yet there is still no suitable mouse model [203], nor is there a treatment option based on findings from genetic association studies [16]. This can be attributed to the possible pitfalls of current study designs discussed above, i.e. clinical heterogeneity of AMD subtypes and lack of knowledge on the true associated causal variant and the affected gene and pathway.

Here, I want to highlight two approaches currently underway to further our knowledge on disease risk and factors involved in AMD pathogenesis.

6.6.1 The International AMD Genomics Consortium

The International AMD Genomics Consortium (IAMDGC), an association of more than 26 research groups worldwide, has collected over 50,000 samples comprising healthy and diseased individuals. Traditionally, individual participating study data are pooled by combining the effect size estimates from each separate study group. In IAMDGC, all of the samples were genotyped on the same genotyping platform in the same genotyping center, thus greatly reducing inter-study variability which should increase the number of loci significantly implicated in AMD. Furthermore, in the IAMDGC all genotypes are fully accessible for

analysis, thus facilitating dissection of complex haplotype structures and identification of additional, independent signals. In addition, the large number of individuals now allows investigations into genetic factors associated with AMD sub-phenotypes like pure geographic atrophy or distinct manifestations of early AMD. First results of genome-wide association analyses are expected to be published in 2015.

6.6.2 Circulating microRNAs

The identification of significantly disregulated microRNAs in AMD patients compared to controls points to genes and pathways involved in prevalent AMD. In order to find the targets (genes) influenced by the significantly associated circulating microRNAs, several methods have been proposed [204,205]. The most straightforward methods include the modulation of those microRNAs in a variety of cell types e.g. RPE cells or endothelial cells. After knockdown or overexpression of a microRNA species, the expression profile of all genes can readily be evaluated with next generation sequencing. In addition, several *in vitro* assays can be used to assess the influence of these microRNAs on cells (e.g. tube formation assay [128], cell migration assay [109], measuring oxidative stress or immune responses [151]). Another promising approach is to modulate multiple microRNAs at the same time, thus mimicing the disease associated microRNA profile.

We could show that a specific microRNA profile is associated specifically with neovascular AMD. Consequently, the next step is to find a microRNA profile for geographic atrophy. The current study included 59 cases with (pure) geographic atrophy. Unless the effect sizes for dysregulated candidate microRNAs are larger, we would need more cases affected with GA to identify a microRNA profile. The recruitment of these cases is currently in progress at the Department of Ophthalmology at the University of Bonn.

Further studies also include the evaluation of circulating microRNAs in mouse models of AMD. We are currently working on a mouse model mimicing disease progression of geographic atrophy by inducing atrophic lesions in the RPE with oxidized lipid metabolites [203]. In addition, we are collaborating with Prof. Langmann from the Department of Experimental Immunology of the Eye from the Eye Clinic Cologne to evaluate cmiRNAs in a laser induced mouse model of neovascular AMD [206]. These experiments will be supplemented by *in vitro* analyses of identified microRNAs.

7. References

1. Jorkasky J (2014) Attitudinal Survey of Minority Populations on Eye and Vision Health and Research. Natl Public Opin Poll: 17.

2. Congdon N, O'Colmain B, Klaver CCW, Klein R, Muñoz B, et al. (2004) Causes and prevalence of visual impairment among adults in the United States. Arch Ophthalmol 122: 477–485.

3. Strauss O (2005) The retinal pigment epithelium in visual function. Physiol Rev 85: 845–881.

4. Sivaprasad S, Bailey TA, Chong VNH (2005) Bruch's membrane and the vascular intima: is there a common basis for age-related changes and disease? Clin Experiment Ophthalmol 33: 518–523.

5. Ramrattan RS, van der Schaft TL, Mooy CM, de Bruijn WC, Mulder PG, et al. (1994) Morphometric analysis of Bruch's membrane, the choriocapillaris, and the choroid in aging. Invest Ophthalmol Vis Sci 35: 2857–2864.

6. Booij JC, Baas DC, Beisekeeva J, Gorgels TGMF, Bergen AAB (2010) The dynamic nature of Bruch's membrane. Prog Retin Eye Res 29: 1–18.

7. Feeney L (1978) Lipofuscin and melanin of human retinal pigment epithelium. Fluorescence, enzyme cytochemical, and ultrastructural studies. Invest Ophthalmol Vis Sci 17: 583–600.

8. Feeney L (1978) Lipofuscin and melanin of human retinal pigment epithelium. Fluorescence, enzyme cytochemical, and ultrastructural studies. Invest Ophthalmol Vis Sci 17: 583–600.

9. Feeney-Burns L, Hilderbrand ES, Eldridge S (1984) Aging human RPE: morphometric analysis of macular, equatorial, and peripheral cells. Invest Ophthalmol Vis Sci 25: 195–200.

10. Ach T, Huisingh C, McGwin G, Messinger JD, Zhang T, et al. (2014) Quantitative autofluorescence and cell density maps of the human retinal pigment epithelium. Invest Ophthalmol Vis Sci 55: 4832–4841.

11. Feeney-Burns L, Ellersieck MR (1985) Age-related changes in the ultrastructure of Bruch's membrane. Am J Ophthalmol 100: 686–697.

12. Spaide RF (2009) Age-related choroidal atrophy. Am J Ophthalmol 147: 801–810.

13. Jackson GR, Owsley C, Curcio CA (2002) Photoreceptor degeneration and dysfunction in aging and age-related maculopathy. Ageing Res Rev 1: 381–396.

14. Pe'er J, Shweiki D, Itin A, Hemo I, Gnessin H, et al. (1995) Hypoxia-induced expression of vascular endothelial growth factor by retinal cells is a common factor in neovascularizing ocular diseases. Lab Invest 72: 638–645.

15. Spaide RF, Armstrong D, Browne R (2003) Continuing medical education review: choroidal neovascularization in age-related macular degeneration--what is the cause? Retina 23: 595–614.

16. Holz FG, Schmitz-Valckenberg S, Fleckenstein M (2014) Recent developments in the treatment of age-related macular degeneration. J Clin Invest 124: 1430–1438.

17. Friedman DS, O'Colmain BJ, Muñoz B, Tomany SC, McCarty C, et al. (2004) Prevalence of age-related macular degeneration in the United States. Arch Ophthalmol 122: 564–572.

18. Wong WL, Su X, Li X, Cheung CMG, Klein R, et al. (2014) Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. lancet Glob Heal 2: e106–16.

19. Augood C a, Vingerling JR, de Jong PTVM, Chakravarthy U, Seland J, et al. (2006) Prevalence of age-related maculopathy in older Europeans: the European Eye Study (EUREYE). Arch Ophthalmol 124: 529–535.

20. Klein R, Klein BEK, Knudtson MD, Meuer SM, Swift M, et al. (2007) Fifteen-year cumulative incidence of age-related macular degeneration: the Beaver Dam Eye Study. Ophthalmology 114: 253–262.

21. Ristau T, Ersoy L, Lechanteur Y, den Hollander AI, Daha MR, et al. (2014) Allergy is a protective factor against age-related macular degeneration. Invest Ophthalmol Vis Sci 55: 210–214.

22. Age-Related Eye Disease Study Research Group (2001) A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS report no. 8. Arch Ophthalmol 119: 1417–1436.

23. Age-Related Eye Disease Study 2 Research Group. (2013) Lutein + zeaxanthin and omega-3 fatty acids for age-related macular degeneration: the Age-Related Eye Disease Study 2 (AREDS2) randomized clinical trial. JAMA 309: 2005–2015.

24. Reynolds R, Rosner B, Seddon JM (2013) Dietary omega-3 fatty acids, other fat intake, genetic susceptibility, and progression to incident geographic atrophy. Ophthalmology 120: 1020–1028.

25. Tan JSL, Wang JJ, Flood V, Mitchell P (2009) Dietary fatty acids and the 10-year incidence of age-related macular degeneration: the Blue Mountains Eye Study. Arch Ophthalmol 127: 656–665.

26. Seddon JM, Reynolds R, Yu Y, Daly MJ, Rosner B (2011) Risk models for progression to advanced age-related macular degeneration using demographic, environmental, genetic, and ocular factors. Ophthalmology 118: 2203–2211.

27. Buitendijk GHS, Rochtchina E, Myers C, van Duijn CM, Lee KE, et al. (2013) Prediction of age-related macular degeneration in the general population: the Three Continent AMD Consortium. Ophthalmology 120: 2644–2655.

28. Allikmets R, Shroyer NF, Singh N, Seddon JM, Lewis RA, et al. (1997) Mutation of the Stargardt disease gene (ABCR) in age-related macular degeneration. Science 277: 1805–1807.

29. Rivera A, White K, Stöhr H, Steiner K, Hemmrich N, et al. (2000) A comprehensive survey of sequence variation in the ABCA4 (ABCR) gene in Stargardt disease and age-related macular degeneration. Am J Hum Genet 67: 800–813.

30. Stone EM, Webster AR, Vandenburgh K, Streb LM, Hockey RR, et al. (1998) Allelic variation in ABCR associated with Stargardt disease but not age-related macular degeneration. Nat Genet 20: 328–329.

31. Fritsche LG, Fleckenstein M, Fiebig BS, Schmitz-Valckenberg S, Bindewald-Wittich A, et al. (2012) A subgroup of age-related macular degeneration is associated with mono-allelic sequence variants in the ABCA4 gene. Invest Ophthalmol Vis Sci 53: 2112–2118.

32. Klaver CCW, Kliffen M, Duijn CM Van, Hofman A, Cruts M, et al. (1998) Genetic Association of Apolipoprotein E with Age-Related Macular Degeneration. Ophthalmic Res: 200–206.

33. Fritsche LG, Chen W, Schu M, Yaspan BL, Yu Y, et al. (2013) Seven new loci associated with age-related macular degeneration. Nat Genet 45: 433–9, 439e1–2.

34. Klein RJ, Zeiss C, Chew EY, Tsai J, Sackler RS, et al. (2005) Complement factor H polymorphism in age-related macular degeneration. Science 308: 385–389.

35. Rivera A, Fisher SA, Fritsche LG, Keilhauer CN, Lichtner P, et al. (2005) Hypothetical LOC387715 is a second major susceptibility gene for age-related macular degeneration, contributing independently of complement factor H to disease risk. Hum Mol Genet 14: 3227–3236.

36. Jakobsdottir J, Conley YP, Weeks DE, Mah TS, Ferrell RE, et al. (2005) Susceptibility genes for age-related maculopathy on chromosome 10q26. Am J Hum Genet 77: 389–407.

37. Hageman GS, Anderson DH, Johnson L V, Hancox LS, Taiber AJ, et al. (2005) A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. Proc Natl Acad Sci U S A 102: 7227–7232.

38. Spencer KL, Olson LM, Anderson BM, Schnetz-Boutaud N, Scott WK, et al. (2008) C3 R102G polymorphism increases risk of age-related macular degeneration. Hum Mol Genet 17: 1821–1824.

39. Gorin MB (2012) Genetic insights into age-related macular degeneration: controversies addressing risk, causality, and therapeutics. Mol Aspects Med 33: 467–486.

40. Fritsche LG, Fariss RN, Stambolian D, Abecasis GR, Curcio CA, et al. (2014) Age-related macular degeneration: genetics and biology coming together. Annu Rev Genomics Hum Genet 15: 151–171.

41. Hammond CJ, Webster AR, Snieder H, Bird AC, Gilbert CE, et al. (2002) Genetic influence on early age-related maculopathy: a twin study. Ophthalmology 109: 730–736.

42. Jakobsdottir J, Gorin MB, Conley YP, Ferrell RE, Weeks DE (2009) Interpretation of genetic association studies: markers with replicated highly significant odds ratios may be poor classifiers. PLoS Genet 5: e1000337.

43. So H-C, Li M, Sham PC (2011) Uncovering the total heritability explained by all true susceptibility variants in a genome-wide association study. Genet Epidemiol 35: 447–456.

44. Siontis GCM, Tzoulaki I, Siontis KC, Ioannidis JPA (2012) Comparisons of established risk prediction models for cardiovascular disease: systematic review. BMJ 344: e3318.

45. Louie KS, Seigneurin A, Cathcart P, Sasieni P (2014) Do prostate cancer risk models improve the predictive accuracy of PSA screening? A meta-analysis. Ann Oncol.

46. Maller J, George S, Purcell S, Fagerness J, Altshuler D, et al. (2006) Common variation in three genes, including a noncoding variant in CFH, strongly influences risk of age-related macular degeneration. Nat Genet 38: 1055–1059.

47. Gold B, Merriam JE, Zernant J, Hancox LS, Taiber AJ, et al. (2006) Variation in factor B (BF) and complement component 2 (C2) genes is associated with age-related macular degeneration. Nat Genet 38: 458–462.

48. Hughes AE, Orr N, Patterson C, Esfandiary H, Hogg R, et al. (2007) Neovascular age-related macular degeneration risk based on CFH, LOC387715/HTRA1, and smoking. PLoS Med 4: e355.

49. Jakobsdottir J, Conley YP, Weeks DE, Ferrell RE, Gorin MB (2008) C2 and CFB genes in age-related maculopathy and joint action with CFH and LOC387715 genes. PLoS One 3: e2199.

50. Seddon JM, Reynolds R, Maller J, Fagerness J a, Daly MJ, et al. (2009) Prediction model for prevalence and incidence of advanced age-related macular degeneration based on genetic, demographic, and environmental variables. Invest Ophthalmol Vis Sci 50: 2044–2053.

51. Farwick A, Dasch B, Weber BHF, Pauleikhoff D, Stoll M, et al. (2009) Variations in five genes and the severity of age-related macular degeneration: results from the Muenster aging and retina study. Eye (Lond) 23: 2238–2244.

52. Gibson J, Cree A, Collins A, Lotery A, Ennis S (2010) Determination of a gene and environment risk model for age-related macular degeneration. Br J Ophthalmol 94: 1382–1387.

53. Chen Y, Zeng J, Zhao C, Wang K, Trood E, et al. (2011) Assessing susceptibility to age-related macular degeneration with genetic markers and environmental factors. Arch Ophthalmol 129: 344–351.

54. Hageman GS, Gehrs K, Lejnine S, Bansal AT, Deangelis MM, et al. (2011) Clinical validation of a genetic model to estimate the risk of developing choroidal neovascular age-related macular degeneration. Hum Genomics 5: 420–440.

55. Klein ML, Francis PJ, Ferris FL, Hamon SC, Clemons TE (2011) Risk Assessment Model for Development of Advanced Age-Related Macular Degeneration. Arch Ophthalmol.

56. Spencer KL, Olson LM, Schnetz-Boutaud N, Gallins P, Agarwal A, et al. (2011) Using genetic variation and environmental risk factor data to identify individuals at high risk for age-related macular degeneration. PLoS One 6: e17784.

57. McCarthy LC, Newcombe PJ, Whittaker JC, Wurzelmann JI, Fries M a, et al. (2012) Predictive models of choroidal neovascularization and geographic atrophy incidence applied to clinical trial design. Am J Ophthalmol 154: 568–578.e12.

58. Grassmann F, Fritsche LG, Keilhauer CN, Heid IM, Weber BHF (2012) Modelling the genetic risk in age-related macular degeneration. PLoS One 7: e37979.

59. Ferris FL, Davis MD, Clemons TE, Lee L-Y, Chew EY, et al. (2005) A simplified severity scale for age-related macular degeneration: AREDS Report No. 18. Arch Ophthalmol 123: 1570–1574.

60. Hoffmann TJ, Marini NJ, Witte JS (2010) Comprehensive Approach to Analyzing Rare Genetic Variants. PLoS One 5: e13584.

61. Uehara H, Mamalis C, McFadden M, Taggart M, Stagg B, et al. (2015) The reduction of serum soluble flt-1 in patients with neovascular age-related macular degeneration. Am J Ophthalmol 159: 92–100.e2.

62. Kim H-J, Woo SJ, Suh EJ, Ahn J, Park JH, et al. (2014) Identification of vinculin as a potential plasma marker for age-related macular degeneration. Invest Ophthalmol Vis Sci 55: 7166–7176.

63. Ilhan N, Daglioglu MC, Ilhan O, Coskun M, Tuzcu EA, et al. (2014) Assessment of Neutrophil/Lymphocyte Ratio in Patients with Age-related Macular Degeneration. Ocul Immunol Inflamm: 1–4.

64. Jonasson F, Fisher DE, Eiriksdottir G, Sigurdsson S, Klein R, et al. (2014) Five-year incidence, progression, and risk factors for age-related macular degeneration: the age, gene/environment susceptibility study. Ophthalmology 121: 1766–1772.

65. Ristau T, Paun C, Ersoy L, Hahn M, Lechanteur Y, et al. (2014) Impact of the Common Genetic Associations of Age-Related Macular Degeneration upon Systemic Complement Component C3d Levels. PLoS One 9: e93459.

66. Zernant J, Xie YA, Ayuso C, Riveiro-Alvarez R, Lopez-Martinez M-A, et al. (2014) Analysis of the ABCA4 genomic locus in Stargardt disease. Hum Mol Genet 23: 6797–6806.

67. Dreyhaupt J, Mansmann U, Pritsch M, Dolar-Szczasny J, Bindewald A, et al. (2005) Modelling the natural history of geographic atrophy in patients with age-related macular degeneration. Ophthalmic Epidemiol 12: 353–362.

68. De Jong PTVM (2006) Age-related macular degeneration. N Engl J Med 355: 1474–1485.

69. Jager RD, Mieler WF, Miller JW, Progress M (2008) Age-related macular degeneration. N Engl J Med 358: 2606–2617.

70. Rosenfeld PJ, Brown DM, Heier JS, Boyer DS, Kaiser PK, et al. (2006) Ranibizumab for neovascular age-related macular degeneration. N Engl J Med 355: 1419–1431.

71. Brown DM, Kaiser PK, Michels M, Soubrane G, Heier JS, et al. (2006) Ranibizumab versus verteporfin for neovascular age-related macular degeneration. N Engl J Med 355: 1432–1444.

72. Seddon JM, Cote J, Page WF, Aggen SH, Neale MC (2005) The US twin study of age-related macular degeneration: relative roles of genetic and environmental influences. Arch Ophthalmol 123: 321–327.

73. Fisher SA, Abecasis GR, Yashar BM, Zareparsi S, Swaroop A, et al. (2005) Meta-analysis of genome scans of age-related macular degeneration. Hum Mol Genet 14: 2257–2264.

74. Li M, Atmaca-Sonmez P, Othman M, Branham KEH, Khanna R, et al. (2006) CFH haplotypes without the Y402H coding variant show strong association with susceptibility to age-related macular degeneration. Nat Genet 38: 1049–1054.

75. Hughes AE, Orr N, Esfandiary H, Diaz-torres M, Goodship T, et al. (2007) A common CFH haplotype , with deletion of CFHR1 and CFHR3 , is associated with lower risk of age-related macular degeneration. Nat Genet 38: 1173–1178.

76. Schmid-kubista KE, Tosakulwong N, Wu Y, Ryu E, Hecker LA, et al. (2009) Contribution of Copy Number Variation in the Regulation of Complement Activation Locus to Development of Age-Related Macular Degeneration. Invest Ophthalmol: 5070–5079.

77. Fritsche LG, Lauer N, Hartmann A, Stippa S, Keilhauer CN, et al. (2010) An imbalance of human complement regulatory proteins CFHR1, CFHR3 and factor H influences risk for age-related macular degeneration (AMD). Hum Mol Genet 19: 4694–4704.

78. Yates JRW, Sepp T, Matharu BK, Khan JC, Thurlby DA, et al. (2007) Complement C3 variant and the risk of age-related macular degeneration. N Engl J Med 357: 553–561.

79. Maller JB, Fagerness JA, Reynolds RC, Neale BM, Daly MJ, et al. (2007) Variation in complement factor 3 is associated with risk of age-related macular degeneration. Nat Genet 39: 1200–1201.

80. Fagerness J a, Maller JB, Neale BM, Reynolds RC, Daly MJ, et al. (2009) Variation near complement factor I is associated with risk of advanced AMD. Eur J Hum Genet 17: 100–104.

81. Chen W, Stambolian D, Edwards AO, Branham KE, Othman M, et al. (2010) Genetic variants near TIMP3 and high-density lipoprotein-associated loci influence susceptibility to age-related macular degeneration. Proc Natl Acad Sci U S A 107: 7401–7406.

82. Neale BM, Fagerness J, Reynolds R, Sobrin L, Parker M, et al. (2010) Genome-wide association study of advanced age-related macular degeneration identifies a role of the hepatic lipase gene (LIPC). Proc Natl Acad Sci U S A 107: 7395–7400.

83. Fritsche LG, Freitag-Wolf S, Bettecken T, Meitinger T, Keilhauer CN, et al. (2009) Age-related macular degeneration and functional promoter and coding variants of the apolipoprotein E gene. Hum Mutat 30: 1048–1053.

84. Yu Y, Bhangale TR, Fagerness J, Ripke S, Thorleifsson G, et al. (2011) Common variants near FRK/COL10A1 and VEGFA are associated with advanced age-related macular degeneration. Hum Mol Genet 20: 3699–3709.

85. Janssens ACJW, Aulchenko YS, Elefante S, Borsboom GJJM (2006) Predictive testing for complex diseases using multiple genes : Fact or fiction ? Genet Med 8: 395–400.

86. Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, et al. (2008) Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med 359: 2208–2219.

87. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, et al. (2008) SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 24: 2938–2939.

88. Friedrich U, Myers C a., Fritsche LG, Milenkovich a., Wolf a., et al. (2011) Risk and non risk associated variants at the 10q26 AMD locus influence ARMS2 mRNA expression but exclude pathogenic effects due to protein deficiency. Hum Mol Genet 20: 1387–1399.

89. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575.

90. Jonasson F, Arnarsson A, Eiríksdottir G, Harris TB, Launer LJ, et al. (2011) Prevalence of age-related macular degeneration in old persons: Age, Gene/environment Susceptibility Reykjavik Study. Ophthalmology 118: 825–830.

91. Mistry K, Cable G (2003) Meta-Analysis of Prostate-Specific Antigen and Digital Rectal Examination as Screening Tests for Prostate Carcinoma. J Am Board Fam Med 16: 95–101.

92. Miller JW (2010) Treatment of age-related macular degeneration: beyond VEGF. Jpn J Ophthalmol 54: 523–528.

93. SanGiovanni JP, Chew EY, Clemons TE, Davis MD, Ferris FL, et al. (2007) The relationship of dietary lipid intake and age-related macular degeneration in a case-control study: AREDS Report No. 20. Arch Ophthalmol 125: 671–679.

94. SanGiovanni JP, Chew EY, Clemons TE, Ferris FL, Gensler G, et al. (2007) The relationship of dietary carotenoid and vitamin A, E, and C intake with age-related macular degeneration in a case-control study: AREDS Report No. 22. Arch Ophthalmol 125: 1225–1232.

95. Klein R, Davis MD, Magli YL, Segal P, Klein BE, et al. (1991) The Wisconsin age-related maculopathy grading system. Ophthalmology 98: 1128–1134.

96. Bird a C, Bressler NM, Bressler SB, Chisholm IH, Coscas G, et al. (1995) An international classification and grading system for age-related maculopathy and age-related macular degeneration. The International ARM Epidemiological Study Group. Surv Ophthalmol 39: 367–374.

97. R Development Core Team (2010) R: A Language and Environment for Statistical Computing.

98. McFadden D (1974) Conditional Logit Analysis of Qualitative Choice Behavior. Front Econom: 105–142.

99. Hu B, Shao J, Palta M (2006) PSEUDO-R 2 IN LOGISTIC REGRESSION MODEL. Stat Sin 16: 847–860.

100. Chongsuvivatwong V (2010) epicalc: Epidemiological calculator.

101. Lumley T (2009) rmeta: Meta-analysis. R package version 2.16.

102. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, et al. (2001) Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 54: 774–781.

103. Liu Q, Sung AH, Chen Z, Liu J, Huang X, et al. (2009) Feature selection and classification of MAQC-II breast cancer and multiple myeloma microarray gene expression data. PLoS One 4: e8250.

104. Swaroop A, Branham KE, Chen W, Abecasis G (2007) Genetic susceptibility to age-related macular degeneration: a paradigm for dissecting complex disease traits. Hum Mol Genet 16 Spec No: R174–82.

105. VanNewkirk MR, Nanjan MB, Wang JJ, Mitchell P, Taylor HR, et al. (2000) The prevalence of age-related maculopathy: the visual impairment project. Ophthalmology 107: 1593–1600.

106. Tomany SC, Wang JJ, Van Leeuwen R, Klein R, Mitchell P, et al. (2004) Risk factors for incident age-related macular degeneration: pooled findings from 3 continents. Ophthalmology 111: 1280–1287.

107. Gallo A, Tandon M, Alevizos I, Illei GG (2012) The majority of microRNAs detectable in serum and saliva is concentrated in exosomes. PLoS One 7: e30679.

108. Turchinovich A, Weiz L, Burwinkel B (2012) Extracellular miRNAs: the mystery of their origin and function. Trends Biochem Sci 37: 460–465.

109. Zhang Y, Liu D, Chen X, Li J, Li L, et al. (2010) Secreted monocytic miR-150 enhances targeted endothelial cell migration. Mol Cell 39: 133–144.

110. Lässer C (2012) Exosomal RNA as biomarkers and the therapeutic potential of exosome vectors. Expert Opin Biol Ther 12 Suppl 1: S189–97.

111. Kanitz A, Imig J, Dziunycz PJ, Primorac A, Galgano A, et al. (2012) The expression levels of microRNA-361-5p and its target VEGFA are inversely correlated in human cutaneous squamous cell carcinoma. PLoS One 7: e49568.

112. Hu Z, Chen X, Zhao Y, Tian T, Jin G, et al. (2010) Serum microRNA signatures identified in a genome-wide serum

microRNA expression profiling predict survival of non-small-cell lung cancer. J Clin Oncol 28: 1721–1726.

113. Cohen J (1988) Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

114. Leveziel N, Puche N, Richard F, Somner JEA, Zerbib J, et al. (2010) Genotypic influences on severity of exudative age-related macular degeneration. Invest Ophthalmol Vis Sci 51: 2620–2625.

115. Klein ML, Ferris FL, Francis PJ, Lindblad AS, Chew EY, et al. (2010) Progression of geographic atrophy and genotype in age-related macular degeneration. Ophthalmology 117: 1554–9, 1559.e1.

116. Zhao H, Wang J, Gao L, Wang R, Liu X, et al. (2013) MiRNA-424 Protects Against Permanent Focal Cerebral Ischemia Injury in Mice Involving Suppressing Microglia Activation. Stroke 44: 1706–1713.

117. Koh JT, Zheng J (2007) The new biomimetic chemistry: artificial transcription factors. ACS Chem Biol 2: 599–601.

118. Hagan EL, Banaszynski LA, Chen L, Maynard-Smith LA, Wandless TJ (2009) Regulating protein stability in mammalian cells using small molecules. Cold Spring Harb Protoc 2009: pdb.prot5172.

119. Scott AW, Bressler SB (2013) Long-term follow-up of vascular endothelial growth factor inhibitor therapy for neovascular age-related macular degeneration. Curr Opin Ophthalmol 24: 190–196.

120. Ferris FL, Davis MD, Clemons TE, Lee L-Y, Chew EY, et al. (2005) A simplified severity scale for age-related macular degeneration: AREDS Report No. 18. Arch Ophthalmol 123: 1570–1574.

121. Köberle V, Pleli T, Schmithals C, Augusto Alonso E, Haupenthal J, et al. (2013) Differential stability of cell-free circulating microRNAs: implications for their utilization as biomarkers. PLoS One 8: e75184.

122. Kosaka N, Izumi H, Sekine K, Ochiya T (2010) microRNA as a new immune-regulatory agent in breast milk. Silence 1: 7.

123. Friedlaender MR, Chen W, Adamidi C, Maaskola J, Einspanier R, et al. (2008) Discovering microRNAs from deep sequencing data using miRDeep. Nat Biotechnol 26: 407–415.

124. Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8: 118–127.

125. Hurteau GJ, Spivack SD, Brock GJ (2006) Potential mRNA degradation targets of hsa-miR-200c, identified using informatics and qRT-PCR. Cell Cycle 5: 1951–1956.

126. Lu T-P, Lee C-Y, Tsai M-H, Chiu Y-C, Hsiao CK, et al. (2012) miRSystem: an integrated system for characterizing enriched functions and pathways of microRNA targets. PLoS One 7: e42390.

127. Vlachos IS, Kostoulas N, Vergoulis T, Georgakilas G, Reczko M, et al. (2012) DIANA miRPath v.2.0: investigating the combinatorial effect of microRNAs in pathways. Nucleic Acids Res 40: W498–504.

128. Bonauer A, Carmona G, Iwasaki M, Mione M, Koyanagi M, et al. (2009) MicroRNA-92a controls angiogenesis and functional recovery of ischemic tissues in mice. Science 324: 1710–1713.

129. Carpentier G (2012) Angiogenesis Analyzer for ImageJ.

130. Sallo FB, Peto T, Leung I, Xing W, Bunce C, et al. (2009) The International Classification system and the progression of age-related macular degeneration. Curr Eye Res 34: 238–240.

131. Kinnunen K, Petrovski G, Moe MC, Berta A, Kaarniranta K (2012) Molecular mechanisms of retinal pigment epithelium damage and development of age-related macular degeneration. Acta Ophthalmol 90: 299–309.

132. Abdelsalam A, Del Priore L, Zarbin MA (n.d.) Drusen in age-related macular degeneration: pathogenesis, natural course, and laser photocoagulation-induced regression. Surv Ophthalmol 44: 1–29.

133. Lim LS, Mitchell P, Seddon JM, Holz FG, Wong TY (2012) Age-related macular degeneration. Lancet 379: 1728–1738.

134. Schmidt-Erfurth U, Chong V, Loewenstein A, Larsen M, Souied E, et al. (2014) Guidelines for the management of neovascular age-related macular degeneration by the European Society of Retina Specialists (EURETINA). Br J Ophthalmol 98: 1144–1167.

135. Lindblad AS, Lloyd PC, Clemons TE, Gensler GR, Ferris FL, et al. (2009) Change in area of geographic atrophy in the Age-Related Eye Disease Study: AREDS report number 26. Arch Ophthalmol 127: 1168–1174.

136. Fleckenstein M, Adrion C, Schmitz-Valckenberg S, Göbel AP, Bindewald-Wittich A, et al. (2010) Concordance of disease progression in bilateral geographic atrophy due to AMD. Invest Ophthalmol Vis Sci 51: 637–642.

137. Fleckenstein M, Schmitz-Valckenberg S, Adrion C, Visvalingam S, Göbel AP, et al. (2011) Progression of age-related geographic atrophy: role of the fellow eye. Invest Ophthalmol Vis Sci 52: 6552–6557.

138. Sunness JS, Gonzalez-Baron J, Applegate CA, Bressler NM, Tian Y, et al. (1999) Enlargement of atrophy and visual acuity loss in the geographic atrophy form of age-related macular degeneration. Ophthalmology 106: 1768–1779.

139. Scholl HPN, Fleckenstein M, Fritsche LG, Schmitz-Valckenberg S, Göbel A, et al. (2009) CFH, C3 and ARMS2 are significant risk loci for susceptibility but not for disease progression of geographic atrophy due to AMD. PLoS One 4: e7418.

140. Caire J, Recalde S, Velazquez-Villoria A, Garcia-Garcia L, Reiter N, et al. (2014) Growth of Geographic Atrophy on Fundus Autofluorescence and Polymorphisms of CFH, CFB, C3, FHR1-3, and ARMS2 in Age-Related Macular Degeneration. JAMA Ophthalmol.

141. Feuer WJ, Yehoshua Z, Gregori G, Penha FM, Chew EY, et al. (2013) Square root transformation of geographic atrophy area measurements to eliminate dependence of growth rates on baseline lesion measurements: a reanalysis of age-related eye disease study report no. 26. JAMA Ophthalmol 131: 110–111.

142. Holz FG, Bindewald-Wittich A, Fleckenstein M, Dreyhaupt J, Scholl HPN, et al. (2007) Progression of geographic atrophy and impact of fundus autofluorescence patterns in age-related macular degeneration. Am J Ophthalmol 143: 463–472.

143. Klein R, Meuer SM, Knudtson MD, Klein BEK (2008) The epidemiology of progression of pure geographic atrophy: the Beaver Dam Eye Study. Am J Ophthalmol 146: 692–699.

144. Sunness JS, Margalit E, Srikumaran D, Applegate CA, Tian Y, et al. (2007) The long-term natural history of geographic atrophy from age-related macular degeneration: enlargement of atrophy and implications for interventional clinical trials. Ophthalmology 114: 271–277.

145. Seddon JM, Reynolds R, Yu Y, Rosner B (2014) Three new genetic loci (R1210C in CFH, variants in COL8A1 and RAD51B) are independently related to progression to advanced macular degeneration. PLoS One 9: e87047.

146. Klein ML, Francis PJ, Ferris FL, Hamon SC, Clemons TE (2011) Risk Assessment Model for Development of Advanced Age-Related Macular Degeneration. Arch Ophthalmol 129: 1543–1550.

147. Smailhodzic D, Fleckenstein M, Theelen T, Boon CJF, van Huet RAC, et al. (2011) Central areolar choroidal dystrophy (CACD) and age-related macular degeneration (AMD): differentiating characteristics in multimodal imaging. Invest Ophthalmol Vis Sci 52: 8908–8918.

148. Fritsche LG, Loenhardt T, Janssen A, Fisher S a, Rivera A, et al. (2008) Age-related macular degeneration is associated with an unstable ARMS2 (LOC387715) mRNA. Nat Genet 40: 892–896.

149. Friedrich U, Myers C a., Fritsche LG, Milenkovich A, Wolf A, et al. (2011) Risk- and non-risk-associated variants at the 10q26 AMD locus influence ARMS2 mRNA expression but exclude pathogenic effects due to protein deficiency. Hum Mol Genet 20: 1387–1399.

150. Heurich M, Martínez-Barricarte R, Francis NJ, Roberts DL, Rodríguez de Córdoba S, et al. (2011) Common polymorphisms in C3, factor B, and factor H collaborate to determine systemic complement activity and disease risk. Proc Natl Acad Sci U S A 108: 8761–8766.

151. Ramaglia V, Hughes TR, Donev RM, Ruseva MM, Wu X, et al. (2012) C3-dependent mechanism of microglial priming relevant to multiple sclerosis. Proc Natl Acad Sci U S A 109: 965–970.

152. Griffiths MR, Neal JW, Fontaine M, Das T, Gasque P (2009) Complement factor H, a marker of self protects against experimental autoimmune encephalomyelitis. J Immunol 182: 4368–4377.

153. Rahpeymai Y, Hietala MA, Wilhelmsson U, Fotheringham A, Davies I, et al. (2006) Complement: a novel factor in basal and ischemia-induced neurogenesis. EMBO J 25: 1364–1374.

154. Van Beek J, Nicole O, Ali C, Ischenko A, MacKenzie ET, et al. (2001) Complement anaphylatoxin C3a is selectively protective against NMDA-induced neuronal cell death. Neuroreport 12: 289–293.

155. Steiner T, Francescut L, Byrne S, Hughes T, Jayanthi A, et al. (2014) Protective role for properdin in progression of experimental murine atherosclerosis. PLoS One 9: e92404.

156. Steiner B, Wolf S, Kempermann G (2006) Adult neurogenesis and neurodegenerative disease. Regen Med 1: 15–28.

157. Pekny M, Wilhelmsson U, Bogestål YR, Pekna M (2007) The role of astrocytes and complement system in neural plasticity. Int Rev Neurobiol 82: 95–111.

158. Klein ML, Francis PJ, Rosner B, Reynolds R, Hamon SC, et al. (2008) CFH and LOC387715/ARMS2 genotypes and treatment with antioxidants and zinc for age-related macular degeneration. Ophthalmology 115: 1019–1025.

159. Wickham H (2009) ggplot2: Elegant Graphics for Data Analysis. Springer New York.

160. Viechtbauer W (2010) Conducting Meta-Analyses in R with the metafor Package. J Stat Softw 36: 1–48.

161. Zarbin MA, Casaroli-Marano RP, Rosenfeld PJ (2014) Age-related macular degeneration: clinical findings, histopathology and imaging techniques. Dev Ophthalmol 53: 1–32.

162. Resnikoff S, Pascolini D, Etya'ale D, Kocur I, Pararajasegaram R, et al. (2004) Global data on visual impairment in the year 2002. Bull World Health Organ 82: 844–851.

163. Sarks SH, Arnold JJ, Killingsworth MC, Sarks JP (1999) Early drusen formation in the normal and aging eye and their relation to age related maculopathy: a clinicopathological study. Br J Ophthalmol 83: 358–368.

164. Troutbeck R, Al-Qureshi S, Guymer RH (n.d.) Therapeutic targeting of the complement system in age-related macular degeneration: a review. Clin Experiment Ophthalmol 40: 18–26.

165. Fu W, O'Connor TD, Akey JM (2013) Genetic architecture of quantitative traits and complex diseases. Curr Opin Genet Dev 23: 678–683.

166. Amos W, Driscoll E, Hoffman JI (2011) Candidate genes versus genome-wide associations: which are better for detecting genetic susceptibility to infectious disease? Proc Biol Sci 278: 1183–1188.

167. Wilkening S, Chen B, Bermejo JL, Canzian F (2009) Is there still a need for candidate gene approaches in the era of genome-wide association studies? Genomics 93: 415–419.

168. Behrens G, Winkler TW, Gorski M, Leitzmann MF, Heid IM (2011) To stratify or not to stratify: power considerations for population-based genome-wide association studies of quantitative traits. Genet Epidemiol 35: 867–879.

169. Flachsbart F, Franke A, Kleindorp R, Caliebe A, Blanché H, et al. (2010) Investigation of genetic susceptibility factors for human longevity - a targeted nonsynonymous SNP study. Mutat Res 694: 13–19.

170. Schulz HL, Goetz T, Kaschkoetoe J, Weber BHF (2004) The Retinome - defining a reference transcriptome of the adult mammalian retina/retinal pigment epithelium. BMC Genomics 5: 50.

171. Sun L, Ryan DG, Zhou M, Sun T-T, Lavker RM (2006) EEDA: a protein associated with an early stage of stratified epithelial differentiation. J Cell Physiol 206: 103–111.

172. Deiss LP, Feinstein E, Berissi H, Cohen O, Kimchi A (1995) Identification of a novel serine/threonine kinase and a novel 15-kD protein as potential mediators of the gamma interferon-induced cell death. Genes Dev 9: 15–30.

173. Owen CG, Jarrar Z, Wormald R, Cook DG, Fletcher AE, et al. (2012) The estimated prevalence and incidence of late stage age related macular degeneration in the UK. Br J Ophthalmol 96: 752–756.

174. Luna A, Nicodemus KK (2007) snp.plotter: an R-based SNP/haplotype association and linkage disequilibrium plotting package. Bioinformatics 23: 774–776.

175. Delaneau O, Zagury J-F, Marchini J (2013) Improved whole-chromosome phasing for disease and population genetic studies. Nat Methods 10: 5–6.

176. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 5: e1000529.

177. Xu Z, Taylor JA (2009) SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. Nucleic Acids Res 37: W600–5.

178. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR (2003) ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res 31: 3568–3571.

179. Sullivan KM, Mannucci A, Kimpton CP, Gill P (1993) A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. Biotechniques 15: 636–638, 640–641.

180. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, et al. (2006) Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry 63: 168–174.

181. Tsoi LC, Spain SL, Knight J, Ellinghaus E, Stuart PE, et al. (2012) Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. Nat Genet 44: 1341–1348.

182. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, et al. (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 45: 1452–1458.

183. So H-C, Gui AHS, Cherny SS, Sham PC (2011) Evaluating the heritability explained by known susceptibility variants: a survey of ten complex diseases. Genet Epidemiol 35: 310–317.

184. Latham J (2011) The failure of the genome. Guard.

185. Poulsen P, Kyvik KO, Vaag A, Beck-Nielsen H (1999) Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. Diabetologia 42: 139–145.

186. Willi C, Bodenmann P, Ghali WA, Faris PD, Cornuz J (2007) Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis. JAMA 298: 2654–2664.

187. Franke A, McGovern DPB, Barrett JC, Wang K, Radford-Smith GL, et al. (2010) Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 42: 1118–1125.

188. Lewis CM, Whitwell SCL, Forbes A, Sanderson J, Mathew CG, et al. (2007) Estimating risks of common complex diseases across genetic and environmental factors: the example of Crohn disease. J Med Genet 44: 689–694.

189. Fritz T, Niederreiter L, Adolph T, Blumberg RS, Kaser A (2011) Crohn's disease: NOD2, autophagy and ER stress converge. Gut 60: 1580–1588.

190. Hagstrom SA, Ying G, Pauer GJT, Sturgill-Short GM, Huang J, et al. (2014) VEGFA and VEGFR2 gene polymorphisms and response to anti-vascular endothelial growth factor therapy: comparison of age-related macular degeneration treatments trials (CATT). JAMA Ophthalmol 132: 521–527.

191. Cruz-Gonzalez F, Cabrillo-Estévez L, López-Valverde G, Cieza-Borrella C, Hernández-Galilea E, et al. (2014) Predictive value of VEGF A and VEGFR2 polymorphisms in the response to intravitreal ranibizumab treatment for wet AMD. Graefes Arch Clin Exp Ophthalmol 252: 469–475.

192. Dikmetas O, Kadayıfcılar S, Eldem B (2013) The effect of CFH polymorphisms on the response to the treatment of age-related macular degeneration (AMD) with intravitreal ranibizumab. Mol Vis 19: 2571–2578.

193. Yuan D, Yuan D, Liu X, Yuan S, Xie P, et al. (2013) Genetic association with response to intravitreal ranibizumab for neovascular age-related macular degeneration in the Han Chinese population. Ophthalmologica 230: 227–232.

194. Lotery AJ, Gibson J, Cree AJ, Downes SM, Harding SP, et al. (2013) Pharmacogenetic associations with vascular endothelial growth factor inhibition in participants with neovascular age-related macular degeneration in the IVAN Study. Ophthalmology 120: 2637–2643.

195. Rofagha S, Bhisitkul RB, Boyer DS, Sadda SR, Zhang K (2013) Seven-year outcomes in ranibizumab-treated patients in ANCHOR, MARINA, and HORIZON: a multicenter cohort study (SEVEN-UP). Ophthalmology 120: 2292–2299.

196. Welter D, MacArthur J, Morales J, Burdett T, Hall P, et al. (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 42: D1001–6.

197. Ferris FL, Wilkinson CP, Bird A, Chakravarthy U, Chew E, et al. (2013) Clinical classification of age-related macular degeneration. Ophthalmology 120: 844–851.

198. Brandl C, Grassmann F, Riolfi J, Weber B (2015) Tapping Stem Cells to Target AMD: Challenges and Prospects. J Clin Med 4: 282–303.

199. Edwards SL, Beesley J, French JD, Dunning AM (2013) Beyond GWASs: illuminating the dark road from association to function. Am J Hum Genet 93: 779–797.

200. Manolio TA (2013) Bringing genome-wide association findings into clinical use. Nat Rev Genet 14: 549–558.

201. Kitsios GD, Tangri N, Castaldi PJ, Ioannidis JPA (2010) Laboratory mouse models for the human genome-wide associations. PLoS One 5: e13782.

202. Welch CL (2012) Beyond genome-wide association studies: the usefulness of mouse genetics in understanding the complex etiology of atherosclerosis. Arterioscler Thromb Vasc Biol 32: 207–215.

203. Pennesi ME, Neuringer M, Courtney RJ (2012) Animal models of age related macular degeneration. Mol Aspects Med 33: 487–509.

204. Gennarino VA, D'Angelo G, Dharmalingam G, Fernandez S, Russolillo G, et al. (2012) Identification of microRNA-regulated gene networks by expression analysis of target genes. Genome Res 22: 1163–1172.

205. Liu H, Qin C, Chen Z, Zuo T, Yang X, et al. (2014) Identification of miRNAs and their target genes in developing maize ears by combined small RNA and degradome sequencing. BMC Genomics 15: 25.

206. Lambert V, Lecomte J, Hansen S, Blacher S, Gonzalez M-LA, et al. (2013) Laser-induced choroidal neovascularization model to study age-related macular degeneration in mice. Nat Protoc 8: 2197–2211.

207. Dewan A, Liu M, Hartman S, Zhang SS-M, Liu DTL, et al. (2006) HTRA1 promoter polymorphism in wet age-related macular degeneration. Science 314: 989–992.

8. Zusammenfassung

Die altersabhängige Makuladegeneration (AMD) ist eine degenerative Erkrankung der zentralen Retina und eine der Hauptursachen von Erblindungen in höherem Alter. AMD ist eine komplexe Krankheit ausgelöst durch genetische und umweltbedingte Faktoren. Die Frühform der AMD führt in der Regel nicht zu Einschränkungen der Sehfähigkeit. Im Gegensatz dazu kann die Spätform der AMD zu einem starken Verlust der Sehschärfe und schließlich zur Erblindung führen. In der Literatur werden zwei verschiedene Erscheinungsformen der Spätform unterschieden: Geographische Atrophie (GA) und neovaskuläre AMD (NV).

In der vorliegenden Arbeit wurden vier Hauptprojekte zusammengefaßt:

(i) Im ersten Projekt sollte ein genetisches Risikomodell für AMD berechnet und bewertet werden. Das Risikomodell basiert auf einem genetischen Risiko Score (GRS), welcher in einer großen Fall-Kontroll Studie aus den Genotypen häufig vorkommender, AMD assoziierter Risikovarianten berechnet wurde. Für das Risikomodell wurde ein AUC Wert (*area under the receiver operating characteristics curve* Wert) von 0.82 berechnet, welcher als ausreichend angesehen wird um Personen in Hochrisiko- bzw. Niedrigrisikogruppen einzuteilen. Zudem haben wir die absoluten Risiken verschiedener Risikogruppen berechnet und konnten so das Krankheitsrisiko von Personen in diesen Gruppen innerhalb der nächsten 5 Jahre vorhersagen. Desweiteren zeigten unsere Analysen, dass AMD Patienten vor dem 75. Lebensjahr einen statistisch signifikanten höheren genetischen Risiko Score haben als Patienten, die nach dem 75. Lebensjahr erkrankten.

(ii) Um die Klassifikationseffizienz des Risikomodells eventuell zu verbessern, wurden zirkulierende microRNAs als mögliche Biomarker für AMD untersucht. Wir konnten zeigen, dass sowohl drei zirkulierende microRNAs individuell (hsa-mir-301a-3p, hsa-mir-424-5p and hsa-mir-361-5p), sowie auch ein, auf den drei microRNAs gemeinsam basierendes Profil, signifikant mit neovaskulärer AMD assoziiert sind. Mit verschiedenen Vorhersageprogrammen wurden für diese microRNAs mögliche Zielgene berechnet. Die so identifizierten Zielgene lassen sich hauptsächlich inflammatorischen und angiogenetischen Signalwegen zuordnen. Zudem konnten wir für die microRNA hsa-mir-361-5p in einer *in vitro* Angiogeneseuntersuchung zeigen, dass diese microRNA das Wachstum von neuen Blutgefäßen beeinflusst.

.

(iii) Im dritten Projekt untersuchten wir den Einfluss von einzelnen genetischen Varianten, den daraus berechneten genetischen Risiko Score (GRS), sowie klinischen Variablen auf die Wachstumsrate der geographischen Atrophie. Wir konnten diese Untersuchungen in dem bisher größten Patientenkollektiv dieser Art durchführen und zeigen, dass zwei genetische Risikovarianten (C3_rs2230199 and ARMS2_rs10490924) signifikant mit der Progression der geographischen Atrophie korreliert sind. Zudem fanden wir, dass Patienten mit bilateral auftretender geographischer Atrophie eine signifikant größere Wachstumsrate aufwiesen als Patienten, die nur in einem Auge an GA erkrankt waren. Um mögliche Korrelationen zwischen den identifzierten Variablen auszuschließen, wurden multivariate lineare Regressionsmodelle erstellt. Damit konnten wir zeigen, dass die Variablen sowohl unabhängig voneinander mit der Progression von GA korreliert waren, als auch dass sie etwa 7% der Wachstumsratenvariabilität erklären.

(iv) Um neue AMD Risikovarianten zu identifizieren, die aufgrund von Überlegungen zur statistischen Power eher nicht über GWAS Ansätze zu entdecken sind, wurde eine Kandidatengen-Assoziationsstudie durchgeführt. In dieser Studie fanden wir eine statisch signifikante Assoziation von zwei häufigen Varianten im *DAPL1* Gen mit der Spätform der AMD. Interessanterweise zeigte sich diese Assoziation nur in weiblichen Fällen und Kontrollen und nicht in männlichen Individuen. Der *DAPL1* Lokus ist daher der erste Genlokus für AMD, für den eine geschlechterspezifische Assoziation beschrieben wurde. Durch weitere *in vitro* Untersuchungen der Risikovarianten und deren Effekten auf die Expression von *DAPL1* konnten wir zeigen, dass die Risikoallele der Varianten mit einer reduzierten Expression von Retina- und RPE spezifischen Transkripten korrelieren

Zusammenfassend konnte ich im rahmen meiner Dissertationsarbeit den Einfluss neuer Faktoren auf die Pathologie der AMD nachweisen aufgrund folgender Beobachtungen: (i) Die Identifzierung neuer genetischer Varianten, die mit dem Risiko an AMD zu erkranken assoziiert sind; (ii) Die Berechnung eines genetischen Risikomodells für AMD, das es erlaubt das genetische Risiko einer Person an AMD zu berechnen; (iii) Die Identifizierung genetischer und klinischer Variablen, die mit dem Wachstum der geographischen Atrophie und damit der Schwere der Krankheit korrelieren; (iv) Und zuletzt durch die Identifzierung und Charakterisierung eines Biomarkers basierend auf zirkulierenden microRNAs für die neovaskuläre AMD.

9. Summary

Age-related macular degeneration (AMD) is a degenerative disease of the central retina and a leading cause of severe vision impairment in industrialized countries. AMD is a complex disease caused by genetic predisposition and environmental factors. Early stages of the disease usually do not lead to visual impairment. In contrast, the late stage forms of AMD can result in profound vision loss, eventually leading to blindness. Two distinct forms of late stage AMD have been described: geographic atrophy (GA) and neovascular AMD (NV).

In this thesis, four main projects were pursued:

(i) In the first project, a genetic risk model for AMD should be computed and evaluated. This model was based on a genetic risk score (GRS) calculated from twelve known, common genetic risk factors in a large cohort of late stage AMD cases and AMD-free controls. The computed model presented to have a decent classification accuracy with an area under the receiver operating characteristics curve (AUC) value of 0.82, which is sufficient to classify individuals as having a high or low risk for AMD. Furthermore, we computed absolute risk estimates for several risk groups to develop AMD in the next five years. In addition, we showed that individuals that were affected with AMD before the age of 75 had a statistically significantly higher genetic risk score than those individuals that got AMD after the age of 75.

(ii) In order to improve the classification scheme, circulating microRNAs were investigated as potential biomarkers for AMD. We showed that three circulating microRNAs (hsa-mir-301a-3p, hsa-mir-424-5p and hsa-mir-361-5p) as well as a combined profile are significantly associated with neovascular AMD. With pathway enrichment analyses performed on predicted target genes of these microRNAs, inflammatory as well as angiogenc pathways were implicated in AMD. In addition, hsa-mir-361-5p was shown to influence the rate of neovascularization in an *in vitro* angiogenesis assay.

(iii) The third project investigated a potential involvement of the genetic risk score and single genetic variants as well as clinical factors on the growth rate of geographic atrophy lesions. By analyzing the largest dataset available on geographic atrophy lesion growth, we showed that two genetic risk variants (C3_rs2230199 and ARMS2_rs10490924) are significanlty correlated with the growth rate of GA lesions. Furthermore, the presence of GA in the fellow eye (bilateral GA) was shown to be significantly correlated with increased GA growth. We computed multivariate linear regression models and showed that these factors are independently correlated with GA growth and jointly explain around 7% of the disease variability.

(iv) In a candidate gene approach, novel genetic risk variants associated with AMD risk were investigated. We found a statistically significant association of two variants in the *DAPL1* gene with late stage AMD. Interestingly, the observed associations were confined to females, thus identifying the first gender specific locus associated with AMD. By further functional characterization of these risk variants and *DAPL1* transcript levels, we showed that the presence of risk alleles correlated with reduced levels of retinal and RPE specific isoforms of *DAPL1*.

In conclusion, this thesis provided further insights into AMD pathology by: (i) implicating novel genetic factors associated with AMD risk; (ii) computing a genetic risk model, effectively summarizing the genetic risk for AMD of an individual; (iii) implicating factors correlated with geographic atrophy lesion growth and thus disease severity and (iv) identification of a biomarker for neovascular AMD based on circulating microRNAs.

10. Appendix – Supplementary Material

10.1 Publication List

10.1.1 First author publications

1. Grassmann F, Fritsche LG, Keilhauer CN, Heid IM & Weber BHF (2012). **Modelling the genetic risk in age-related macular degeneration.** *PLoS One* **7:** e37979

2. Grassmann F, Schoenberger PGA, Brandl C, Schick T, Hasler D, Meister G, Fleckenstein M, Lindner M, Helbig H, Fauser S, Weber BHF (2014). **A Circulating MicroRNA Profile Is Associated with Late-Stage Neovascular Age-Related Macular Degeneration.** *PLoS One* **9:** e107461

3. Grassmann F, Fleckenstein M, Chew EY, Strunz T, Schmitz-Valckenberg S, Göbel AP, Klein ML, Ratnapriya R, Swaroop A, Holz FG, Weber BHF (2015) **Clinical and genetic factors associated with progression of geographic atrophy lesions in age-related macular degeneration.** *PLoS One*, in revision

4. Grassmann F, Friedrich U, Fauser S, Schick T, Milenkovic A, Schulz HL, von Strachwitz CN, Bettecken T, Lichtner P, Meitinger T, Arend N, Wolf A, Haritoglou C, Rudolph G, Chakravarthy U, Silvestri G, McKay GJ, Freitag-Wolf1 S, Krawczak M, Smith RT, Merriam JC, Merriam JE, Allikmets R, Heid IM, Weber BHF (2015) **A candidate gene association study identifies DAPL1 as a female-specific susceptibility locus for age-related macular degeneration (AMD).** *Neuromolecular Medicine*, in press

5. Grassmann F, Heid IM & Weber BHF (2014) **Genetic risk models in age-related macular degeneration.** *Adv. Exp. Med. Biol.* **801:** 291–300

6. Grassmann F, Ach T, Brandl C, Heid IM & Weber BHF (2015) **What does genetics tell us about AMD?** *Ann Rev Vis Sci*, in revision

7. Grassmann F, Fauser S & Weber BHF (2015) **The genetics of age-related macular degeneration (AMD) and its usability for designing treatment options** *Eur J Pharm Biopharm.*, in revision

8. Grassmann F, Bergholz R, Jaegle H, Maendl J, Rüther K, Weber BFH (2015) **Common synonymous variants in ABCA4 are protective for chloroquine induced maculopathy (toxic maculopathy)** *BMC Ophthalmology*, in revision

10.1.2 Co-author publications

1. Ratnapriya R, Zhan X, Fariss RN, Branham KE, Zipprer D, Chakarova CF, Sergeev YV, Campos MM, Othman M, Friedman JS, Maminishkis A, Waseem NH, Brooks M, Rajasimha HK, Edwards AO, Lotery A, Klein BE, Truitt BJ, Li B, Schaumberg DA, Morgan DJ, Morrison MA, Souied E, Tsironi EE, Grassmann F, Fishman GA, Silvestri G, Scholl HP, Kim IK, Ramke J, Tuo J, Merriam JE, Merriam JC, Park KH, Olson LM, Farrer LA, Johnson MP, Peachey NS, Lathrop M, Baron RV, Igo RP Jr, Klein R, Hagstrom SA, Kamatani Y, Martin TM, Jiang Y, Conley Y, Sahel JA, Zack DJ, Chan CC, Pericak-Vance MA, Jacobson SG, Gorin MB, Klein ML, Allikmets R, Iyengar SK, Weber BH, Haines JL, Léveillard T, Deangelis MM, Stambolian D, Weeks DE, Bhattacharya SS, Chew EY, Heckenlively JR, Abecasis GR, Swaroop A (2014) **Rare and common variants in extracellular matrix gene Fibrillin 2 (FBN2) are associated with macular degeneration.** *Hum Mol Genet*. 2014 Nov 1;23(21).

2. Brandl C, Zimmermann SJ, Milenkovic VM, Rosendahl SM, Grassmann F, Milenkovic A, Hehr U, Federlin M, Wetzel CH, Helbig H, Weber BH (2014) **In-depth characterisation of Retinal Pigment Epithelium (RPE) cells derived from human induced pluripotent stem cells (hiPSC).** *Neuromolecular Medicine* 2014 Sep;16(3):551-64.

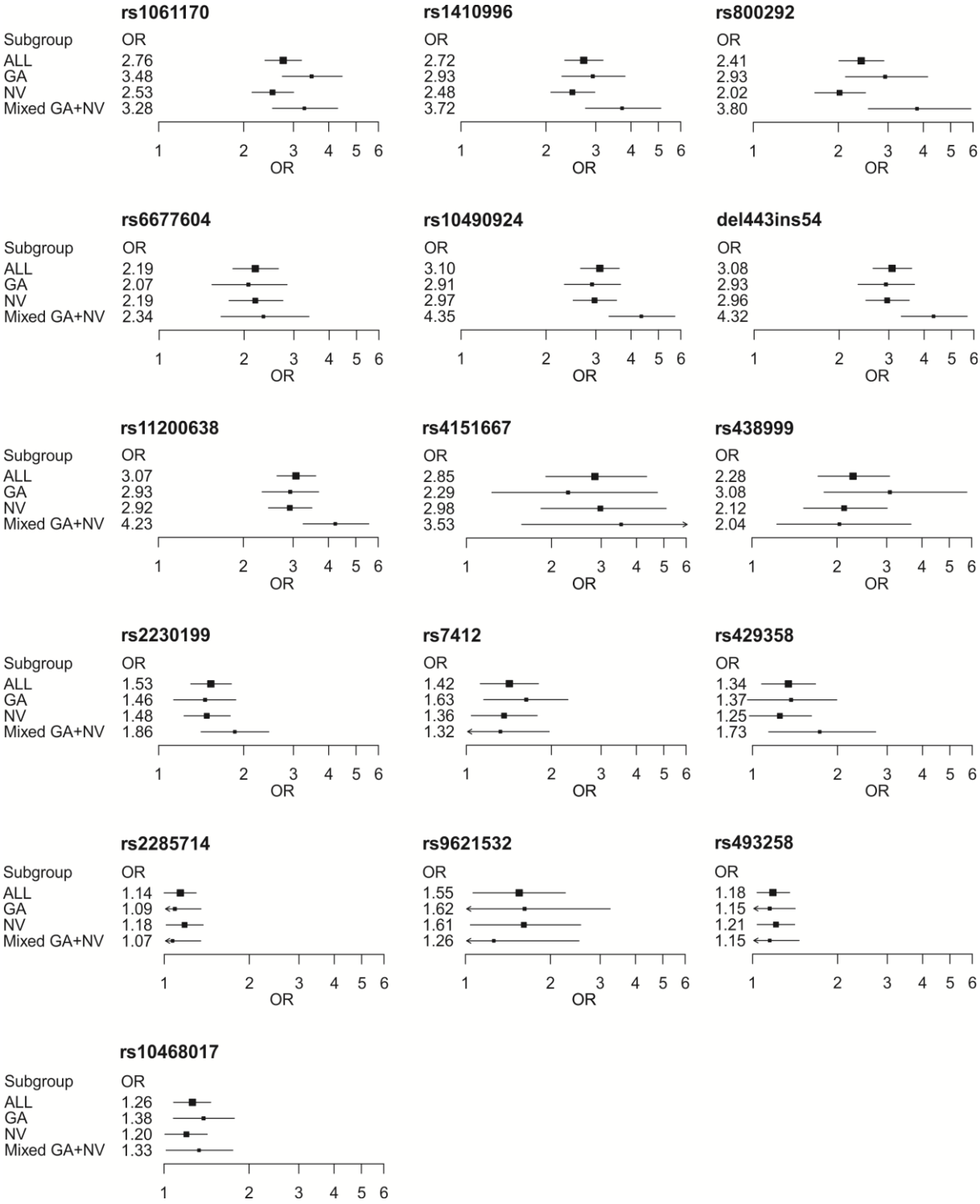
3. Weber BH, Charbel Issa P, Pauly D, Herrmann P, Grassmann F, Holz FG (2014) **The role of the complement system in age-related macular degeneration.** *Dtsch Arztebl Int.* 2014 Feb 21;111(8):133-8

4. Zhan X, Larson DE, Wang C, Koboldt DC, Sergeev YV, Fulton RS, Fulton LL, Fronick CC, Branham KE, Bragg-Gresham J, Jun G, Hu Y, Kang HM, Liu D, Othman M, Brooks M, Ratnapriya R, Boleda A, Grassmann F, von Strachwitz C, Olson LM, Buitendijk GH, Hofman A, van Duijn CM, Cipriani V, Moore AT, Shahid H, Jiang Y, Conley YP, Morgan DJ, Kim IK, Johnson MP, Cantsilieris S, Richardson AJ, Guymer RH, Luo H, Ouyang H, Licht C, Pluthero FG, Zhang MM, Zhang K, Baird PN, Blangero J, Klein ML, Farrer LA, DeAngelis MM, Weeks DE, Gorin MB, Yates JR, Klaver CC, Pericak-Vance MA, Haines JL, Weber BH, Wilson RK, Heckenlively JR, Chew EY, Stambolian D, Mardis ER, Swaroop A, Abecasis GR. (2013) **Identification of a rare coding variant in complement 3 associated with age-related macular degeneration.** *Nat Genet.* 2013 Nov;45(11):1375-9

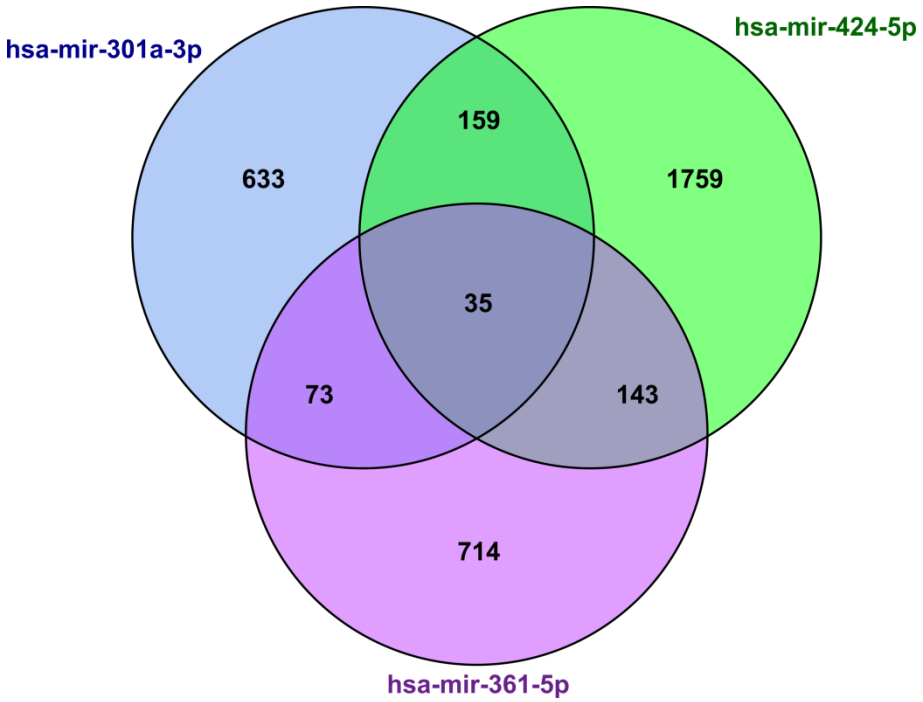
5. Zach F, Grassmann F, Langmann T, Sorusch N, Wolfrum U, Stöhr H. **The retinitis pigmentosa 28 protein FAM161A is a novel ciliary protein involved in intermolecular protein interaction and microtubule association.** *Hum Mol Genet.* 2012 Nov 1;21(21):4573-86

6. Brandl, C.; Grassmann, F.; Riolfi, J.; Weber, BH. **Tapping Stem Cells to Target AMD: Challenges and Prospect**s. *J. Clin. Med*. 2015, 4, 282-303.

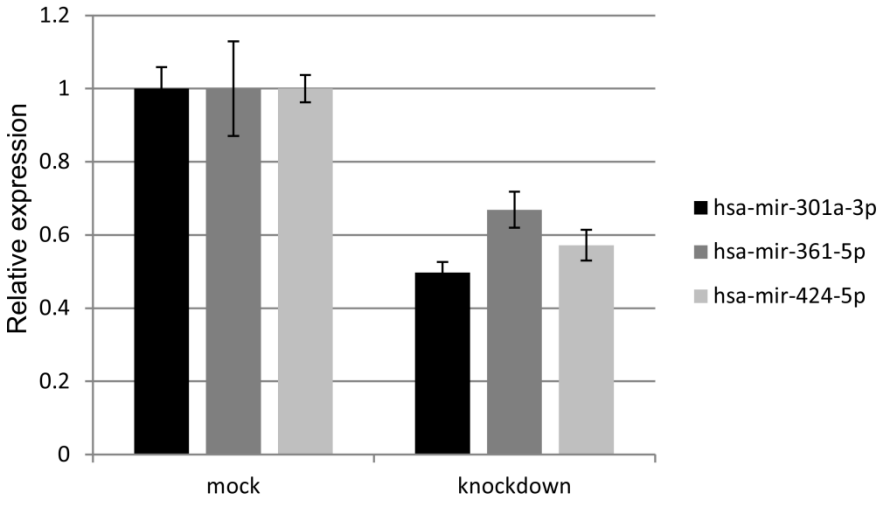
10.2 Supplementary Figures



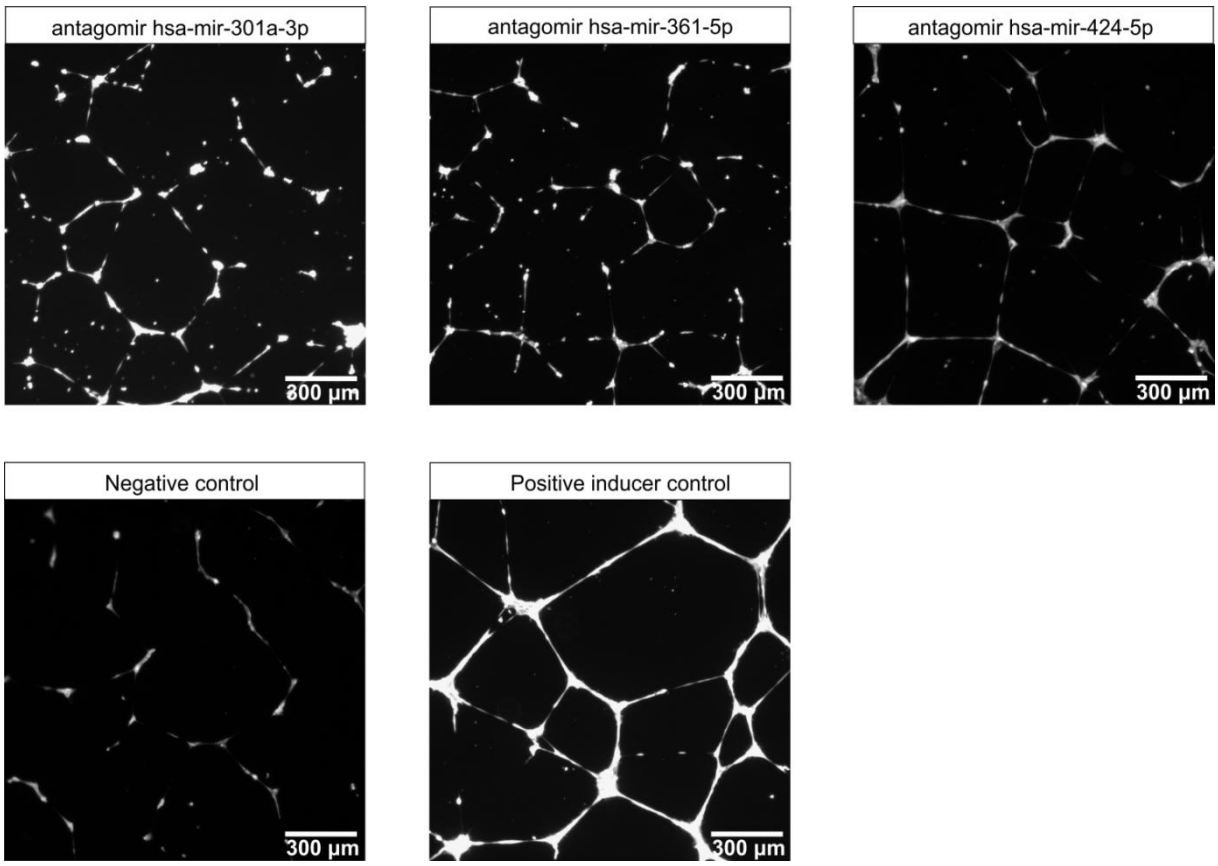
Supplementary Figure S1. Risk estimates for 16 AMD associated variants by disease subtypes. Logistic regression models were fitted with all patients (N=986), GA cases only (N=229), NV cases only (N=581) or mixed GA+NV cases (N=176) versus controls (N=796). Odds ratio estimates (OR) are given per risk allele; horizontal bars indicate 95% confidence intervals and the arrow indicates that the boundary extends below 1 or above 6.



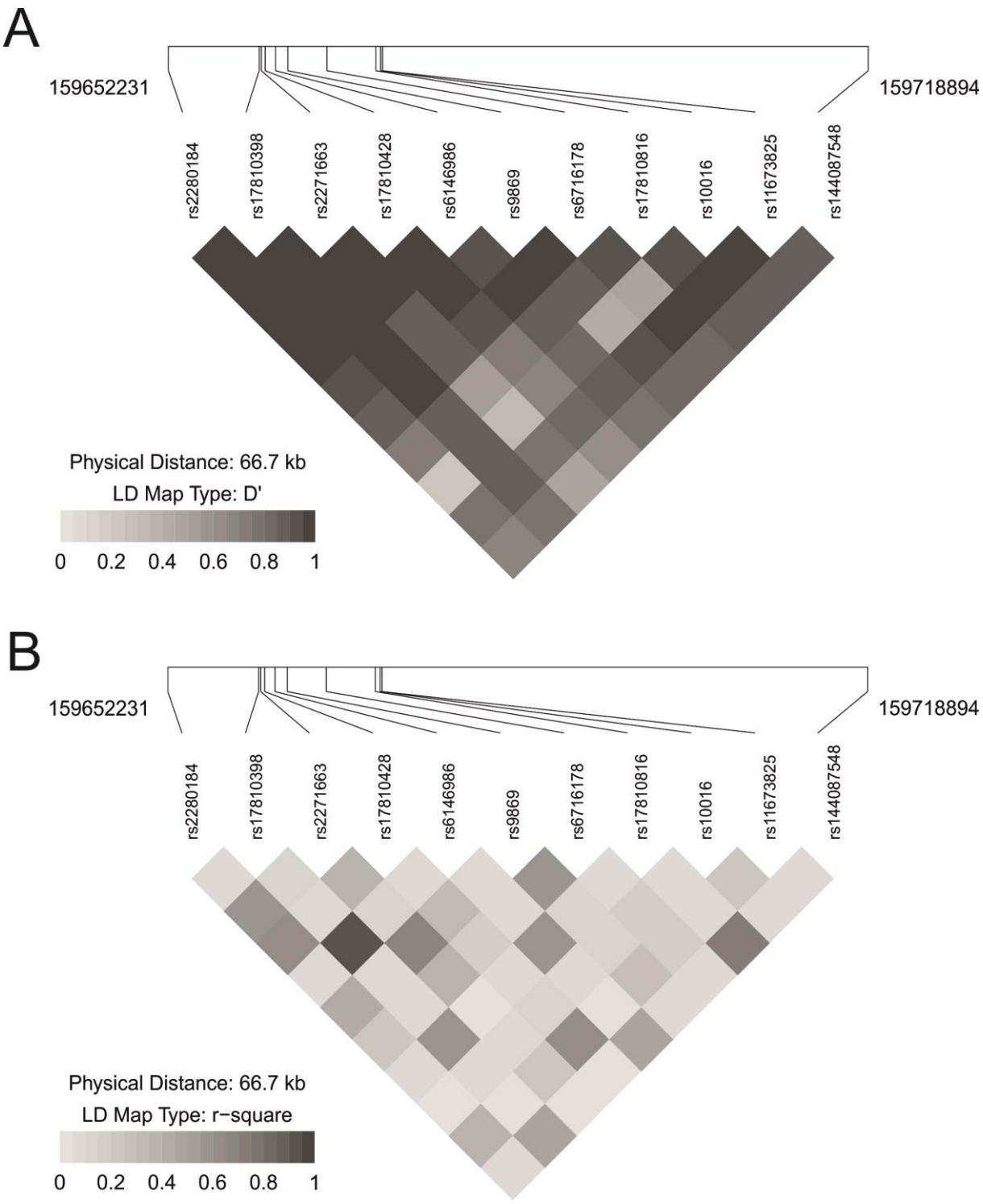
Supplementary Figure S2. Venn diagram of target genes predicted by microT-CDS. Target genes were predicted with microT-CDS with a microT threshold of 0.7. In total, 3,516 target genes were predicted.



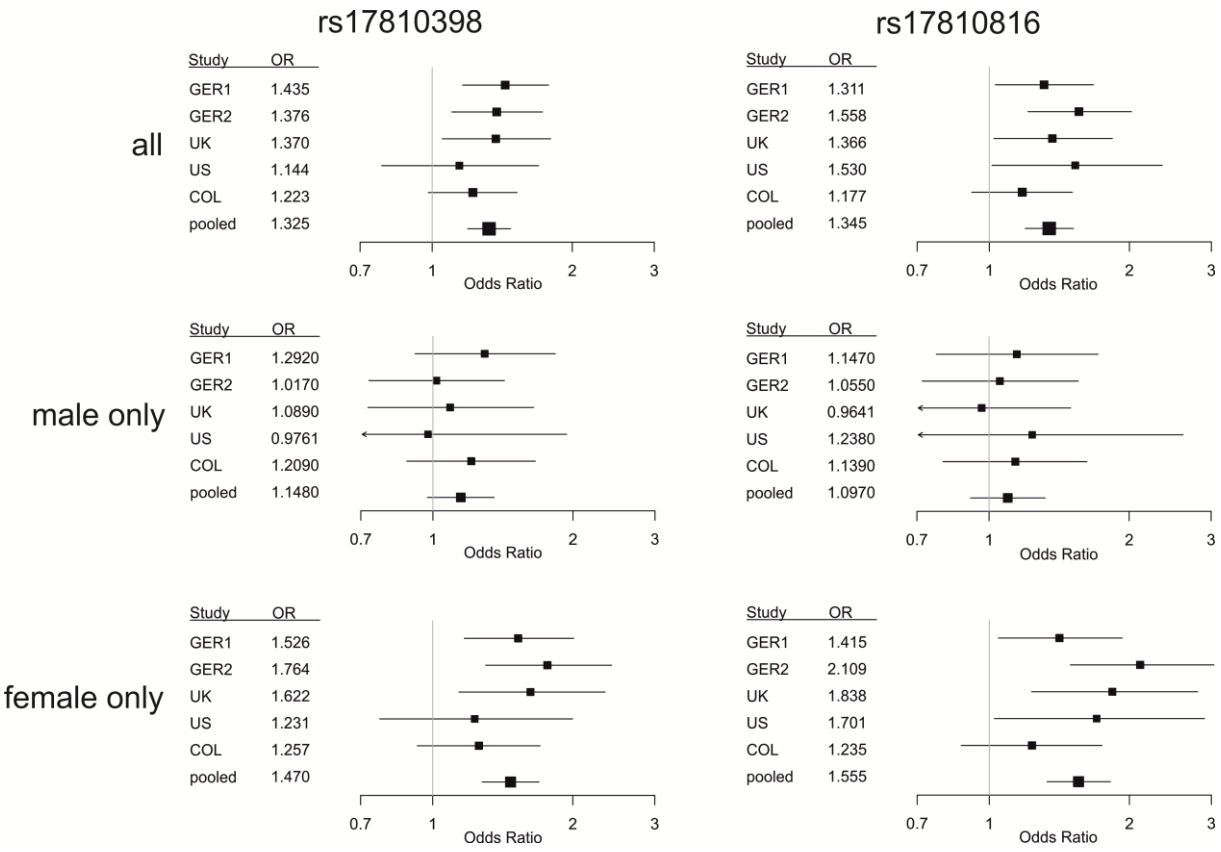
Supplementary Figure S3. Knockdown of candidate miRNAs in human endothelial cells. Mean relative reduction in miRNA levels compared to control antagomir (mock). Whiskers represent the standard error of the mean.



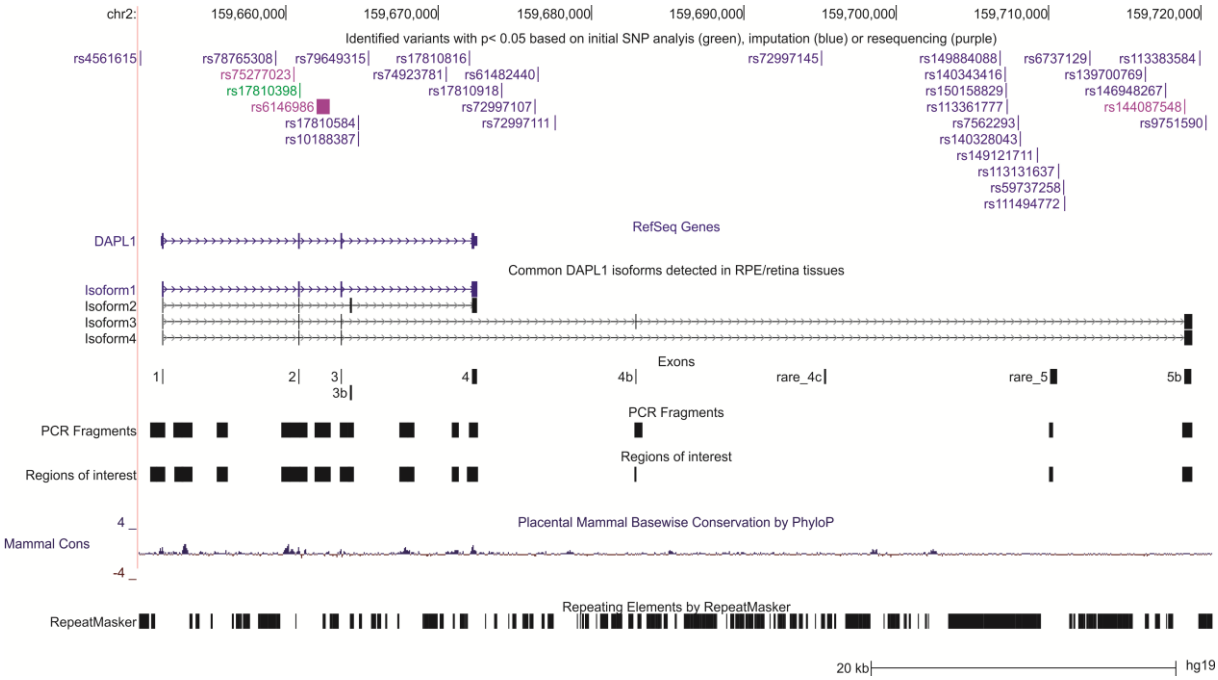
Supplementary Figure S4. Representative images of *in vitro* tube formation assays in human endothelial cells. The measured cumulative tube length in each image was close to the mean cumulative tube length measured in all images of the respective treatment.



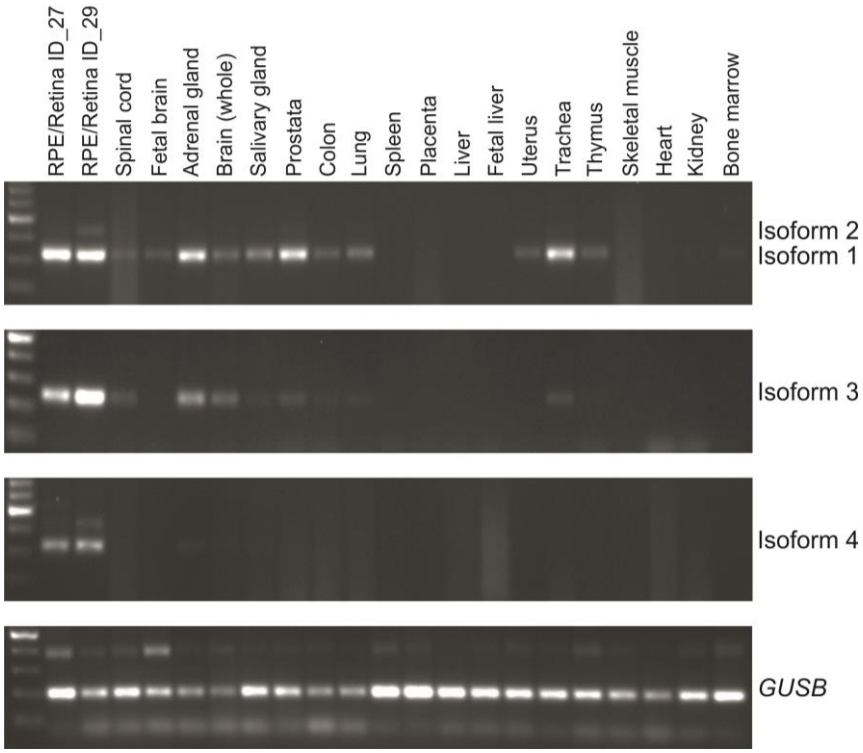
Supplementary Figure S5. Linkage disequilibrium (LD) map of the *DAPL1* gene locus. SNP positions are indicated by vertical/diagonal lines. **A.** Values of r^2 are indicated by coloring (white, low r^2; black, high r^2). **B.** Values of D' are indicated by coloring (white, low D'; black, high D').



Supplementary Figure S6. Sex-specific analysis in the combined study of candidate SNPs rs17810398 and rs17810816 in the *DAPL1* gene. Odds ratios and corresponding 95% confidence intervals are given with the size of each rectangle representing the respective number of cases. Late stage sub-phenotypes were combined.



Supplementary Figure S7. Re-sequencing strategy of the *DAPL1* locus. A screenshot of the UCSC genome browser with default and custom tracks at position hg19:chr2: 159,650,435-159,721,003 is shown (http://genome.ucsc.edu). Track names are given above each track. From top to bottom, tracks are as follows: (1) Identified risk variants based upon discovery study (green), imputation analysis (blue; **Supplementary Table S10**) or resequencing (purple; **Supplementary Table S13**); (2) RefSeq sequence of *DAPL1* (NM_001017920.2); (3) Common *DAPL1* transcripts as identified in four RPE/retina tissue samples (isoform 1: HQ179934/NM001017920.2, isoform 2: HQ179935, isoform 3: HQ179936, isoform 4: HQ179937); (4) Exons identified in common and rare isoforms of *DAPL1*; (5) Resequenced PCR fragments (**Supplementary Table S11**); (6) Regions of interest based upon exon structure and conservation; (7) "46-Way Most Cons" track of the UCSC genome browser(14.03.2014); (8) UCSC RepeatMasker track (14.03.2014).



Supplementary Figure S8. RT-PCR expression analysis of *DAPL1* isoforms. All forward and reverse primers used were exon-spanning to avoid amplification of traces of genomic contamination in the mRNA preparations. Expected and observed fragment sizes were as follows: 329 bp (isoform 1), 448 bp (isoform2), 340 bp (isoform 3) and 331 bp (isoform 4). Expression analysis of housekeeping gene β-glucuronidase (GUSB; 197 bp) served as a control for first-strand cDNA integrity.

10.3 Supplementary Tables

Supplementary Table S1. Published genetic variations associated with AMD

Variant	Gene	Year	Ref.
rs1061170	*CFH*	2005	[34]
rs2274700 (proxy: rs1410996)	*CFH*	2006	[74]
rs800292	*CFH*	2005	[37]
ΔCFHR3/CFHR1 (proxy: rs6677604)	*CFH/CFHR*	2007	[75]
rs10490924	*ARMS2*	2005	[73]
c.del443ins54	*ARMS2*	2008	[148]
rs11200638	*HTRA1*	2006	[207]
rs4151667	*CFB*	2006	[47]
rs547154 (proxy: rs438999)	*CFB*	2006	[47]
rs2230199	*C3*	2008	[38]
rs7412	*APOE*	1998	[32]
rs429358	*APOE*	1998	[32]
rs2285714	*CFI/PLA2G12A*	2009	[80]
rs493258	*LIPC*	2010	[81]
rs10468017	*LIPC*	2010	[82]
rs9621532	*SYN3/TIMP3*	2010	[81]

Supplementary Table S2. Cross validated absolute risks for late stage AMD in different risk groups in the modeled population

	Modeled prevalence (age group in yrs)[1]	**Positive predictive value in risk group [%]**				
		1 (low)	**2**	**3**	**4**	**5 (high)**
Fraction of cases in the modeled general population (absolute risk) [%]						
	1% (65-69 yrs)	0.13	0.22	0.91	4.22	46.8
	2.5% (70-74)	0.35	0.56	2.26	10.0	58.1
	5% (75-79)	0.70	1.13	4.54	18.5	68.4
	10% (80-84)	1.46	2.39	9.13	32.4	80.0
	15% (>85)	2.20	3.72	13.7	42.7	85.6
Fraction of cases and controls in our study[2]						
Cases [%]		1.00	8.50	38.7	40.8	11.0
Controls [%]		7.90	38.2	43.5	10.2	0.40

[1] Approximate age groups corresponding to prevalences according to [17,19] for prevalences between 65 and 79 years and [90] for prevalences above 80 years

[2] Averaged fractions of controls and cases observed in 2000 test sets in each risk group

Supplementary Table S3. Primers and methods used for genotyping

Nearby gene(s)	Marker	Method	Forward primer (5‘-3‘)	Rreverse primer (5‘-3‘)	Universal Extension primer (MALDI-TOF)
CFH	rs1061170	Sequencing	TTGCACACAAGATGGATGGT	GCATCTGGGAGTAGGAGACC	
		RFLP (Tsp509I)[1]	CCTTTGTTAGTAACTTTAGTTCGTCTT	CCAAAAACTAAATAGGTCCATTGGT	
	rs1410996	MALDI-TOF	ACGTTGGATGCCGTCAATGAGATTTACGTC	ACGTTGGATGCCTCTACATCAGTGGTATAG	ACGCAGTCCCTGACTACCTCATG
	rs800292	Taqman (C___2530382_10)			
		MALDI-TOF	ACGTTGGATGAAATGCCGCCCTGGATATAG	ACGTTGGATGTAAGAGCAACCCATTCTCCC	CCTGGATATAGATCTCTTGGAAAT
	rs6677604	Sequencing	TTGCTTGAGGAAAGTTGTGC	TTCCCATCTCCTGTCACAAA	
		MALDI-TOF	ACGTTGGATGCCACCAAAGCACAATACCTC	ACGTTGGATGCATGACTTACATAGTTGCCC	TAGCTGTGAGTCCTTTCC
ARMS2/HTRA	rs10490924	Sequencing	GCCTGAGATGGCAAGTCTGT	TGTAGCAGGTGCATTGGAAG	
	del443ins54	PCR[2]	ACTCATCACGTCATCACCAAT	CTCTCTGCAGCCCTCATTTG[3]	
	rs11200638	RFLP (EagI)[1]	ATGCCACCCACAACAACTTT	GGTTCTCTCGCTGAGATTCG	
CFB	rs4151667	Sequencing	GGTCTGGAGTTTCAGCTTGG	TCTTGGAGAAGTCGGAAGGA	
		MALDI-TOF	ACGTTGGATGCTTCTCTCCTGCCTTCCAAC	ACGTTGGATGCAAGAGGCCCAAGATAAAGG	ATCTCAGCCCCCAAC
	rs438999	Sequencing	GAGGTCAGGGGTCATGAGAA	AGACAGGGATTCATGGGATG	
		MALDI-TOF	ACGTTGGATGACAACCTCCTTGTCTCTTCG	ACGTTGGATGAAGAGTCACCTGGCCAGAAG	GGGGACTTATGGGGAAATCCAACTC
C3	rs2230199	RFLP (HhaI)[1]	GTGGTTGACGGTGAAGATCC	CAAGATCCGGAAGCTGGAC	
		MALDI-TOF	ACGTTGGATGCAACAGGGAGTTCAAGTCAG	ACGTTGGATGTCCACCACTTGGGTCCCGAA	AGTGAGTTCAAGTCAGAAAAGGGG
APOE	rs7412	Sequencing	GATGGACGAGACCATGAAGG	CTCGAACCAGCTCTTGAGG	
	rs429358	Sequencing	GATGGACGAGACCATGAAGG	CTCGAACCAGCTCTTGAGG	
PLA2G12A	rs2285714	Sequencing	CAAGCCACCAGATCATCCTT	CAAATGCCTTTTGCAGCTTA	
		MALDI-TOF	ACGTTGGATGCCTGACAAAGTGTTGCAACC	ACGTTGGATGCAGTCATTCTTGCTTTTGCC	GGGTGACACGACAGGTGCTATGA
SYN3	rs9621532	RFLP (EcoO109I)	GGTTCTACTGGCTGGGTGAA	TACCCCCACTACCCCTAGTT	
		MALDI-TOF	ACGTTGGATGTGAAAGGGATTGAAAGCAGG	ACGTTGGATGTCTGGGCAGCCTGAAAACTC	GGGATTGAAAGCAGGTCATTA
LIPC	rs10468017	RFLP (SspI)	TTTACGGTCTCCAAGACTGCT	CCAAGTTCATTCACAGGGACT	
	rs493258	Taqman (C___1929355_10)			
		Sequencing	AGACCAGCAGGCATCACC	CCAGAAACAAACAAGTGGAGTG	

[1] Restriction enzyme used for RFLP assay
[2] PCR with these primers yielded two distinct PCR products of different lengths corresponding to each allele.
[3] SNPs were found in primer binding site

Supplementary Table S4. **Mean and 95% confidence intervals of log transformed fold changes of cmiRNA levels in the combined study**

	NV cases	GA cases	controls
Number of individuals	129	59	147
hsa-mir-301a-3p	-0.318 (-0.416 - -0.220)	0.030 (-0.110 - 0.170)	0.055 (-0.025 - 0.134)
hsa-mir-361-5p	-0.373 (-0.500 - -0.245)	-0.069 (-0.203 - 0.065)	-0.002 (-0.076 - 0.072)
hsa-mir-424-5p	-0.338 (-0.437 - -0.240)	0.077 (-0.068 - 0.223)	-0.071 (-0.152 - 0.011)

Supplementary Table S5. Previously published associated variations used to calculate the genetic risk score

									Frequency of risk allele in	
Nearby gene(s)	**Marker**	**ID**	**Impact/effect of variant**	**Odds ratio**	**95% CI[1]**	**P-value**	**Non risk allele**	**Risk allele[2]**	**Cases (N=72)**	**Controls (N=77)**
CFH	rs1061170	1	p.Y402H	2.18	1.33-3.68	0.001762	T	C	0.59	0.36
	rs800292	2	p.I62V	1.76	0.79-4.06	0.1662	A	G	0.89	0.83
	rs6677604	3	proxy for ΔCFHR3/CFHR1	1.96	0.94-4.28	0.0742	A	G	0.87	0.79
ARMS2	rs10490924	4	p.A69S	5.62	2.90-12.01	2.63E-08	G	T	0.55	0.19
CFB	rs4151667	5	p.L9H	8.47	1.48-160.00	0.01319	A	T	0.99	0.93
	rs438999	6	proxy for rs641153 (p.R32Q)	2.29	0.77-7.73	0.1374	C	T	0.95	0.91
C3	rs2230199	7	p.R102G	1.73	0.95-3.25	0.07395	G	C	0.28	0.18
APOE	rs7412	8	p.R158C	0.70	0.30-1.57	0.3924	C	T	0.08	0.12
	rs429358	9	p.C112R	1.80	0.76-4.54	0.1829	C	T	0.92	0.86
PLA2G12A	rs2285714	10	synonymous exonic, unknown	1.19	0.72-1.99	0.492	C	T	0.45	0.40

[1] 95% confidence intervals

[2] Risk allele refers to AMD risk increasing allele

Supplementary Table S6. Primers and mature microRNA sequences

primer name	mature miRNA sequence	primer sequence
hsa-miR-142-RT	CAUAAAGUAGAAAGCACUACU	CATAAAGTAGAAAGCACTACT
hsa-miR-361-RT	UUAUCAGAAUCUCCAGGGGUAC	TTATCAGAATCTCCAGGGGTA
hsa-miR-424-RT	CAGCAGCAAUUCAUGUUUUGAA	CAGCAGCAATTCATGTTTTGAA
hsa-miR-4732-5p_RT	UGUAGAGCAGGGAGCAGGAAGCU	TGTAGAGCAGGGAGCAGGAAGCT
hsa-miR-451a_RT	AAACCGUUACCAUUACUGAGUU	AAACCGTTACCATTACTGAGTT
hsa-miR-192-RT	CUGACCUAUGAAUUGACAGCC	CTGACCTATGAATTGACAGCC
hsa-miR-26a-RT	UUCAAGUAAUCCAGGAUAGGCU	TTCAAGTAATCCAGGATAGGCT
hsa-miR-505-RT	GGGAGCCAGGAAGUAUUGAUGU	GGGAGCCAGGAAGTATTGAT
hsa-miR-335-5p_RT	UCAAGAGCAAUAACGAAAAAUGU	TCAAGAGCAATAACGAAAAATGT
hsa-miR-301a-3p_RT	CAGUGCAAUAGUAUUGUCAAAGC	CAGTGCAATAGTATTGTCAAAGC
hsa-miR-30b-5p_RT	UGUAAACAUCCUACACUCAGCU	TGTAAACATCCTACACTCAGCT
hsa-miR-194-5p_RT	UGUAACAGCAACUCCAUGUGGA	TGTAACAGCAACTCCATGTGGA
hsa-miR-4732-5p_RT	UGUAGAGCAGGGAGCAGGAAGCU	TGTAGAGCAGGGAGCAGGAAGCT
Univeral_PCR_Primer	-	AACGAGACGACGACAGACTTT
URT_Primer	-	AACGAGACGACGACAGACTTTTTTTTTTTTTTV

Supplementary Table S7. Allele frequencies of evaluated genetic variants in this study

			FAM study - discovery	FAM study - replication	ARED study	combined
Variant	Risk allele*	Weight**	Allele frequency	Allele frequency	Allele frequency	Allele frequency
CFH_rs1061170	C	0.629	0.669	-	-	-
CFH_rs6677604	G	0.702	0.930	-	-	-
CFH_rs800292	G	0.932	0.913	-	-	-
C3_rs2230199	C	0.377	0.326	0.229	0.305	0.300
ARMS2_rs10490924	T	1.301	0.483	0.479	0.451	0.461
CFB_rs438999	T	1.026	0.965	-	-	-
CFB_rs4151667	T	1.406	0.988	-	-	-
APOE_rs7412	T	0.397	0.093	-	-	-
APOE_rs429358	T	0.335	0.901	-	-	-
CFI_rs2285714	T	0.169	0.413	-	-	-

* Risk allele refers to the risk increasing allele, i.e. the allele that is more frequent in cases than in controls

** Weights were obtained from the multiple logistic regression model in Grassmann *et al.* 2012 and used to calculate the genetic risk score (GRS)

Supplementary Table S8. Candidate gene based analyses of 106 SNPs. Table spans more than one page

Gene name	Symbol	Selection criteria*		Chromosome band	# SNPs analyzed	Discovery Sample GER1	Replication Sample GER2
		Protein Function	Retinal / RPE expression			*P*min**	*P****
SH3-domain GRB2-like (endophilin) interacting protein 1	*SGIP1*	-	X	1p31.3	3	0.696	…
Retinal pigment epithelium-specific protein 65kDa	*RPE65*	X	X	1p31	4	0.059	…
ATP-binding cassette, sub-family A (ABC1), member 4	*ABCA4*	X	X	1p22	17	0.183	…
Cathepsin S	*CTSS*	X	-	1q21	2	0.742	…
G protein-coupled receptor 75	*GPR75*	-	X	2p16	2	0.780	…
c-mer proto-oncogene tyrosine kinase	*MERTK*	X	X	2q14.1	3	0.342	…
Death associated protein-like 1	*DAPL1*	-	X	2q24	8	**0.016**	**0.002**
Membrane protein, palmitoylated 4	*MPP4*	-	X	2q33.2	5	0.430	…
Retinol binding protein 1, cellular	*RBP1*	-	X	3q21-q23	4	0.444	…
Succinate receptor 1	*SUCNR1*	X	-	3q25.1	6	0.067	…
WD repeat domain 17	*WDR17*	-	X	4q34	3	0.356	…
Neuropeptide VF precursor	*NPVF*	-	X	7p21-p15	2	0.164	…
Retinitis pigmentosa 1 (autosomal dominant)	*RP1*	X	X	8q12.1	3	**0.037**	0.44
Transient receptor potential cation channel, subfamily M, member 3	*TRPM3*	-	X	9q21.11	10	0.141	…
Cadherin-related family member 1	*CDHR1*	X	X	10q23.1	6	0.197	…
Retinal G protein coupled receptor	*RGR*	X	X	10q23	2	0.096	…
Cathepsin D	*CTSD*	-	X	11p15.5	5	0.355	…
Fatty acid desaturase 3	*FADS3*	X	-	11q12-q13.1	3	0.496	…

Supplementary Table S8. continued

Bestrophin 1	*BEST1*	X	X	11q12	3	0.070	…
Chromosome 11 open reading frame 48	*C11orf48*	-	X	11q12.3	2	0.149	…
Retinol dehydrogenase 12 (all-trans/9-cis/11-cis)	*RDH12*	X	X	14q24.1	1	0.247	…
Retinaldehyde binding protein 1	*RLBP1*	X	X	15q26.1	2	0.488	…
ATPase, Na+/K+ transporting, beta 2 polypeptide	*ATP1B2*	-	X	17p13.1	2	0.212	…
Neuropilin (NRP) and tolloid (TLL)-like 1	*NETO1*	-	X	18q22.2	6	0.256	…
Cystatin C	*CST3*	X	X	20p11.2	2	**0.028**	1.00

* X: present; -: not present

** Pmin: minimum *P*valuefrom logistic regression model adjusted for age and gender for all SNPs analyzed in this region

*** *P*value from logistic regression model adjusted for age and gender

Supplementary Table S9. Association results for the 106 SNPs analyzed in the discovery sample. Table spans more than one page

								Minor Allele Frequency in		Number of non-missing genotypes in		
Gene	SNP (dbSNP-ID)	Chr.	Position [hg19]	Major allele	Minor allele	Odds ratio	95% Confidence Intervals	Controls	Cases	Controls	Cases	P-Value*
SGIP1	rs1373909	1	67040875	A	G	1.018	0.872-1.188	0.420	0.424	611	710	0.821
SGIP1	rs1536112	1	67107225	A	G	1.019	0.869-1.194	0.435	0.438	608	707	0.817
SGIP1	rs6588216	1	67145847	A	G	1.010	0.850-1.200	0.277	0.280	610	708	0.910
RPE65	rs3118415	1	68894296	A	G	1.041	0.850-1.276	0.171	0.176	588	696	0.696
RPE65	rs3125895	1	68901735	G	T	1.002	0.860-1.167	0.410	0.410	608	702	0.980
RPE65	rs3790472	1	68910999	G	T	1.160	0.995-1.354	0.441	0.478	608	703	0.059
RPE65	rs436070	1	68918895	G	A	0.866	0.716-1.048	0.209	0.186	608	706	0.139
ABCA4	rs1800555	1	94463617	G	A	0.872	0.397-1.915	0.011	0.009	612	710	0.729
ABCA4	rs7537325	1	94469631	A	G	1.007	0.840-1.207	0.225	0.227	604	704	0.939
ABCA4	rs1800553	1	94473807	G	A	1.559	0.534-5.107	0.004	0.006	612	710	0.429
ABCA4	rs2275033	1	94480037	G	A	0.937	0.804-1.092	0.434	0.418	604	687	0.404
ABCA4	rs1932014	1	94488497	T	C	1.014	0.872-1.180	0.453	0.457	606	708	0.853
ABCA4	rs3789395	1	94501594	G	T	0.990	0.849-1.154	0.461	0.459	609	704	0.894
ABCA4	rs11165069	1	94504545	G	A	0.972	0.807-1.172	0.212	0.206	604	703	0.768
ABCA4	rs497511	1	94523113	T	C	1.063	0.910-1.242	0.471	0.485	611	705	0.438
ABCA4	rs549114	1	94534354	C	T	1.055	0.893-1.248	0.321	0.332	605	698	0.529
ABCA4	rs4147827	1	94548080	G	C	0.920	0.762-1.112	0.209	0.195	609	704	0.390
ABCA4	rs952499	1	94558425	A	G	0.964	0.825-1.127	0.498	0.488	611	702	0.647
ABCA4	rs950283	1	94567223	A	G	1.049	0.896-1.227	0.360	0.373	612	710	0.555
ABCA4	rs1209515	1	94571335	G	A	1.091	0.920-1.295	0.268	0.284	607	704	0.318
ABCA4	rs4147815	1	94574808	C	T	0.944	0.780-1.141	0.223	0.213	611	692	0.550
ABCA4	rs2297634	1	94576968	A	G	1.103	0.944-1.290	0.469	0.491	611	706	0.219

Supplementary Table S9. continued

ABCA4	rs2184339	1	94585331	A	G	0.891	0.752-1.056	0.307	0.286	608	704	0.183
ABCA4	rs3761911	1	94588754	A	T	1.044	0.890-1.226	0.376	0.385	605	703	0.597
CTSS	rs1136774	1	150738197	A	G	1.001	0.858-1.167	0.471	0.472	611	708	0.991
CTSS	rs3754212	1	150738200	T	C	1.028	0.874-1.208	0.360	0.366	607	707	0.742
GPR75	rs805368	2	54083420	C	T	1.025	0.861-1.221	0.272	0.276	612	708	0.780
GPR75	rs805373	2	54086595	T	C	0.983	0.840-1.150	0.452	0.446	612	708	0.829
MERTK	rs7604639	2	112751928	A	G	0.930	0.792-1.092	0.395	0.378	611	709	0.373
MERTK	rs3811634	2	112754943	C	T	1.036	0.877-1.224	0.300	0.308	612	710	0.677
MERTK	rs55812028	2	112780969	A	G	0.915	0.760-1.100	0.234	0.219	612	708	0.342
DAPL1	rs2280184	2	159652231	T	G	0.924	0.790-1.081	0.416	0.399	607	706	0.322
DAPL1	**rs17810398**	**2**	**159660870**	**C**	**T**	**1.348**	**1.058-1.722**	**0.103**	**0.133**	**611**	**710**	**0.016**
DAPL1	rs2271663	2	159661077	C	T	0.983	0.843-1.147	0.465	0.460	611	695	0.832
DAPL1	rs17810428	2	159661451	G	A	0.951	0.808-1.119	0.321	0.312	611	703	0.544
DAPL1	rs9869	2	159663599	T	C	0.912	0.782-1.063	0.444	0.419	608	706	0.240
DAPL1	rs6716178	2	159667328	G	A	0.894	0.757-1.056	0.327	0.302	608	707	0.187
DAPL1	rs10016	2	159672442	A	G	0.937	0.800-1.096	0.370	0.353	610	706	0.414
DAPL1	rs11673825	2	159672626	T	C	0.921	0.782-1.085	0.300	0.284	609	707	0.327
MPP4	rs3754932	2	202509814	G	A	1.064	0.913-1.240	0.488	0.502	608	708	0.430
MPP4	rs2597900	2	202523347	G	A	1.036	0.888-1.208	0.428	0.437	608	702	0.655
MPP4	rs1208083	2	202532447	C	T	1.006	0.837-1.211	0.223	0.224	611	708	0.947
MPP4	rs1914267	2	202542836	G	A	0.954	0.816-1.115	0.435	0.424	607	707	0.554
MPP4	rs888012	2	202561571	C	T	0.958	0.808-1.137	0.294	0.286	608	704	0.625
RBP1	rs211585	3	139235564	G	A	1.021	0.876-1.189	0.476	0.481	610	708	0.792
RBP1	rs2071388	3	139236683	A	G	0.947	0.808-1.109	0.390	0.378	604	708	0.499
RBP1	rs10935331	3	139256732	G	C	1.062	0.911-1.237	0.416	0.430	611	708	0.444
RBP1	rs9862672	3	139262793	G	A	1.045	0.875-1.249	0.236	0.244	609	707	0.625
SUCNR1	rs6763405	3	151589418	G	A	1.034	0.775-1.382	0.076	0.079	594	699	0.821
SUCNR1	rs1402012	3	151589783	T	C	0.949	0.763-1.181	0.151	0.146	598	694	0.639

Supplementary Table S9. continued

SUCNR1	rs1445359	3	151591741	T	C	1.261	0.985-1.618	0.102	0.124	596	691	0.067
SUCNR1	rs1445358	3	151593218	C	T	1.105	0.897-1.365	0.155	0.169	590	680	0.350
SUCNR1	rs13315275	3	151597310	G	A	1.031	0.877-1.212	0.418	0.426	596	684	0.712
SUCNR1	rs13079080	3	151599393	C	T	1.133	0.972-1.321	0.461	0.494	590	681	0.111
WDR17	rs17062505	4	177019413	T	C	0.961	0.814-1.136	0.307	0.300	610	704	0.643
WDR17	rs17625943	4	177098285	G	A	0.952	0.792-1.144	0.254	0.245	601	678	0.599
WDR17	rs11736872	4	177100644	G	A	0.924	0.782-1.092	0.292	0.276	610	708	0.356
NPVF	rs739749	7	25262594	A	G	1.026	0.880-1.196	0.482	0.489	601	689	0.743
NPVF	rs2074423	7	25263948	C	T	1.149	0.945-1.400	0.182	0.203	601	688	0.164
RP1	**rs9643828**	**8**	**55529073**	**T**	**C**	**0.835**	**0.704-0.989**	**0.295**	**0.260**	**612**	**710**	**0.037**
RP1	rs2293869	8	55539395	A	T	1.060	0.91-1.2340	0.426	0.439	605	702	0.458
RP1	rs446227	8	55541450	G	A	1.133	0.949-1.354	0.248	0.272	612	710	0.168
TRPM3	rs1889915	9	73164712	G	A	1.069	0.916-1.247	0.474	0.489	609	708	0.398
TRPM3	rs11142497	9	73198353	G	A	0.953	0.815-1.116	0.443	0.431	609	709	0.552
TRPM3	rs11142503	9	73218892	G	A	0.923	0.784-1.087	0.350	0.333	611	707	0.336
TRPM3	rs1538670	9	73255337	G	T	0.953	0.817-1.111	0.421	0.409	607	704	0.538
TRPM3	rs10123161	9	73296400	G	A	0.966	0.822-1.135	0.328	0.320	612	710	0.670
TRPM3	rs7031754	9	73311986	T	C	0.902	0.772-1.054	0.442	0.417	611	707	0.194
TRPM3	rs564929	9	73434585	G	A	1.079	0.922-1.263	0.361	0.378	612	707	0.342
TRPM3	rs579587	9	73438011	G	A	0.910	0.779-1.064	0.439	0.416	612	708	0.237
TRPM3	rs1337029	9	73459960	T	C	1.050	0.886-1.246	0.263	0.272	611	707	0.573
TRPM3	rs2152757	9	73478555	G	A	0.868	0.719-1.048	0.213	0.190	609	707	0.141
CDHR1	rs7099098	10	85950406	C	T	0.959	0.815-1.129	0.346	0.337	612	701	0.615
CDHR1	rs11200915	10	85957681	A	C	0.994	0.839-1.177	0.295	0.292	607	706	0.939
CDHR1	rs11200920	10	85960274	A	C	1.092	0.908-1.315	0.224	0.238	612	710	0.350
CDHR1	rs4933975	10	85960395	G	C	0.955	0.819-1.113	0.467	0.454	602	706	0.553
CDHR1	rs4933978	10	85971347	G	A	0.893	0.752-1.060	0.281	0.257	612	703	0.197
CDHR1	rs3814213	10	85974236	C	T	1.052	0.903-1.225	0.471	0.485	612	702	0.518

Supplementary Table S9. continued

RGR	rs1042454	10	86012713	C	T	1.054	0.896-1.240	0.345	0.360	612	710	0.524
RGR	rs11200947	10	86020637	A	G	1.179	0.972-1.432	0.193	0.220	611	710	0.096
CTSD	rs2334411	11	1773477	A	C	0.944	0.789-1.129	0.260	0.247	603	699	0.528
CTSD	rs8839	11	1774136	A	C	1.071	0.867-1.324	0.152	0.161	610	706	0.528
CTSD	rs55923455	11	1777866	A	G	0.955	0.807-1.130	0.325	0.315	599	696	0.593
CTSD	rs1317356	11	1779138	G	A	1.076	0.921-1.257	0.489	0.508	609	707	0.355
CTSD	rs7122341	11	1781790	A	G	0.949	0.792-1.136	0.240	0.229	608	709	0.568
FADS3	rs174626	11	61637057	T	C	1.054	0.907-1.224	0.453	0.466	604	710	0.496
FADS3	rs174634	11	61647387	G	C	1.044	0.873-1.249	0.243	0.251	612	710	0.638
FADS3	rs174468	11	61663691	C	T	0.969	0.828-1.134	0.412	0.405	611	707	0.697
BEST1	rs149698	11	61730036	G	A	0.936	0.795-1.101	0.317	0.302	609	704	0.424
BEST1	rs1800008	11	61730183	C	T	1.105	0.920-1.328	0.234	0.250	610	705	0.288
BEST1	rs1800009	11	61730234	T	C	1.163	0.988-1.371	0.334	0.366	609	707	0.070
C11orf48	rs7386	11	62430335	C	T	0.908	0.782-1.053	0.482	0.455	611	708	0.201
C11orf48	rs17637597	11	62438750	A	G	1.135	0.956-1.348	0.257	0.283	611	709	0.149
RDH12	rs718212	14	68196636	T	C	0.910	0.776-1.067	0.400	0.381	608	704	0.247
RLBP1	rs2710	15	89753220	G	A	1.021	0.874-1.193	0.399	0.403	602	703	0.791
RLBP1	rs3825991	15	89761664	G	T	1.055	0.908-1.225	0.489	0.503	601	701	0.488
ATP1B2	rs1642764	17	7557834	T	C	1.032	0.884-1.205	0.456	0.462	608	707	0.688

Supplementary Table S9. continued

ATP1B2	rs55831773	17	7559037	C	T	1.132	0.932-1.375	0.182	0.201	611	708	0.212
NETO1	rs11872857	18	70416119	A	G	1.020	0.809-1.288	0.119	0.122	611	710	0.868
NETO1	rs2000809	18	70435137	T	C	1.063	0.892-1.266	0.254	0.265	601	684	0.496
NETO1	rs1032102	18	70453405	G	A	1.035	0.876-1.224	0.299	0.305	612	704	0.683
NETO1	rs753147	18	70461059	G	A	0.912	0.779-1.069	0.425	0.405	612	708	0.256
NETO1	rs753744	18	70527649	G	A	0.915	0.766-1.094	0.274	0.256	612	708	0.331
NETO1	rs10164255	18	70532840	C	A	1.089	0.933-1.271	0.400	0.423	611	710	0.282
CST3	**rs2424577**	**20**	**23613750**	**C**	**T**	**1.246**	**1.025-1.517**	**0.181**	**0.215**	**611**	**707**	**0.028**
CST3	rs3787499	20	23616807	T	C	1.153	0.983-1.354	0.377	0.410	612	710	0.081

*from logistic regression model adjusted for age and gender

Supplementary Table S10. Association results for imputed SNPs with significant association (p < 0.05) in the discovery sample (710 late stage AMD patients/612 controls). Table spans more than one page

				Minor Allele Frequency in				
SNP (dbSNP-ID)	Odds ratio	95% Confidence Intervals	P-Value*	Cases	Controls	Minor allele	Position [hg19]	imputation quality
rs74923781	1.399	1.072-1.835	0.014	0.135	0.107	G	159670440	0.722
rs17810816	1.3801	1.064-1.799	0.016	0.140	0.112	G	159671992	0.765
rs59737258	1.3398	1.05-1.716	0.019	0.183	0.154	A	159710981	0.717
rs17810918	1.338	1.049-1.712	0.020	0.151	0.122	A	159674084	0.832
rs79029713	1.4177	1.058-1.909	0.020	0.144	0.120	T	159608659	0.609
rs76276633	1.4136	1.058-1.899	0.020	0.144	0.120	T	159613499	0.617
rs16843331	1.4102	1.056-1.892	0.021	0.146	0.122	G	159625777	0.621
rs80284529	1.4092	1.055-1.892	0.021	0.142	0.118	T	159625985	0.633
rs60864902	1.4101	1.055-1.894	0.021	0.147	0.123	C	159618280	0.614
rs75962772	1.4073	1.054-1.887	0.021	0.148	0.124	G	159615755	0.61
rs80335441	1.4071	1.054-1.887	0.021	0.148	0.124	A	159617834	0.611
rs10497202	1.407	1.054-1.887	0.021	0.148	0.124	C	159616934	0.611
rs76502338	1.407	1.054-1.887	0.021	0.148	0.124	A	159612450	0.611
rs10497200	1.408	1.054-1.89	0.021	0.148	0.125	A	159610120	0.606
rs60324897	1.4079	1.054-1.89	0.021	0.148	0.125	A	159608966	0.606
rs76719581	1.4065	1.054-1.886	0.022	0.144	0.120	T	159628604	0.633
rs76721474	1.4	1.053-1.871	0.022	0.112	0.088	C	159638393	0.818
rs76686128	1.4056	1.052-1.886	0.022	0.145	0.121	A	159622055	0.626
rs139858290	1.406	1.052-1.887	0.022	0.144	0.121	A	159619820	0.625
rs148057425	1.4059	1.052-1.887	0.022	0.144	0.121	A	159621470	0.625
rs56894726	1.4054	1.052-1.886	0.022	0.145	0.121	C	159619146	0.625
rs142646819	1.4025	1.051-1.88	0.022	0.147	0.123	C	159611690	0.617

Supplementary Table S10. continued

rs77664069	1.4014	1.05-1.878	0.023	0.142	0.118	T	159628595	0.643
rs149884088	1.3406	1.043-1.729	0.023	0.153	0.126	G	159706790	0.774
rs145657727	1.4164	1.051-1.917	0.023	0.141	0.118	T	159623085	0.604
rs113143827	1.3574	1.044-1.772	0.023	0.146	0.120	C	159724896	0.733
rs113131637	1.3416	1.042-1.733	0.023	0.159	0.131	T	159710656	0.732
rs112605456	1.3592	1.044-1.777	0.024	0.145	0.119	A	159728893	0.729
rs190418388	1.3592	1.044-1.777	0.024	0.145	0.119	C	159729290	0.729
rs62185469	1.4572	1.052-2.026	0.024	0.130	0.110	C	159735154	0.525
rs35901244	1.3429	1.041-1.74	0.024	0.152	0.125	T	159723905	0.752
rs6737129	1.338	1.04-1.728	0.025	0.152	0.125	A	159712710	0.769
rs139700769	1.3379	1.04-1.728	0.025	0.152	0.125	T	159716306	0.769
rs113383584	1.338	1.04-1.728	0.025	0.152	0.125	G	159719907	0.769
rs112202751	1.338	1.04-1.728	0.025	0.152	0.125	G	159722218	0.769
rs113711908	1.338	1.04-1.728	0.025	0.152	0.125	T	159721786	0.769
rs113340479	1.3379	1.04-1.728	0.025	0.152	0.125	T	159721099	0.769
rs150158829	1.3377	1.039-1.728	0.025	0.151	0.124	A	159707210	0.774
rs113361777	1.3377	1.039-1.728	0.025	0.151	0.124	T	159707249	0.774
rs144087548	1.3393	1.04-1.732	0.025	0.151	0.125	T	159718894	0.765
rs149121711	1.33	1.038-1.709	0.025	0.176	0.149	C	159709248	0.651
rs140328043	1.339	1.039-1.732	0.025	0.151	0.124	T	159708160	0.767
rs7562293	1.339	1.039-1.732	0.025	0.151	0.124	T	159707998	0.767
rs6729542	1.3507	1.04-1.762	0.025	0.125	0.099	C	159648421	0.885
rs111494772	1.34	1.039-1.735	0.025	0.152	0.125	T	159711036	0.757
rs62185474	1.4558	1.049-2.029	0.026	0.134	0.114	C	159739588	0.503
rs148268005	1.348	1.039-1.757	0.026	0.127	0.102	C	159648405	0.875
rs61482440	1.3199	1.034-1.691	0.027	0.171	0.144	A	159676470	0.742
rs17810398	1.3135	1.033-1.676	0.027	0.133	0.105	T	159660870	1
rs72997145	1.3273	1.034-1.71	0.027	0.170	0.143	T	159695064	0.717

Supplementary Table S10. continued

rs72997107	1.319	1.033-1.69	0.027	0.171	0.144	T	159676285	0.741
rs72997111	1.319	1.033-1.69	0.027	0.170	0.143	T	159677635	0.743
rs75277023	1.3081	1.028-1.67	0.030	0.132	0.105	A	159660494	0.999
rs78765308	1.307	1.027-1.669	0.031	0.132	0.105	A	159659291	0.998
rs55989586	1.3745	1.031-1.84	0.031	0.158	0.136	C	159624664	0.581
esv2676504	1.3093	1.025-1.678	0.032	0.134	0.108	T	159661994	0.951
rs140343416	1.3348	1.026-1.742	0.032	0.201	0.176	A	159707144	0.542
rs79632155	1.3612	1.026-1.813	0.033	0.139	0.116	A	159630744	0.679
rs9751590	1.3172	1.022-1.703	0.034	0.180	0.155	C	159720291	0.664
rs4561615	1.2987	1.021-1.658	0.035	0.136	0.109	C	159650386	0.963
rs79614044	1.319	1.019-1.713	0.036	0.124	0.100	A	159641297	0.91
rs17204323	1.3181	1.019-1.712	0.037	0.123	0.099	G	159645078	0.918
rs147441846	1.3181	1.019-1.712	0.037	0.123	0.099	A	159643141	0.918
rs79281718	1.3181	1.019-1.712	0.037	0.123	0.099	A	159643742	0.918
rs11422552	1.3543	1.016-1.812	0.040	0.183	0.162	TA	159620245	0.519
rs146948267	1.339	1.013-1.777	0.041	0.122	0.100	A	159717629	0.755
rs17810584	1.294	1.009-1.666	0.044	0.132	0.108	T	159664703	0.918
rs10188387	1.2867	1.002-1.658	0.049	0.145	0.121	T	159664713	0.825

Supplementary Table S11. Isoform-specific allele frequencies of SNP rs17810398 in 351 resequenced cDNA clones from two heterozygous RPE/retina tissue samples (ID_13, ID_14).

	rs17810398 alleles		***P****
	C	**T**	
Isoform 1	107	103	Reference
Isoform 2	36	41	0.60
Isoform 3	22	1	1.2×10^{-5}
Isoform 4	24	2	3.3×10^{-5}

* obtained from Fisher's exact test to test for deviations from the distribution of the reference transcript (i.e. isoform 1)

Supplementary Table S12. Analyzed genomic regions and corresponding primers for product amplification with PCR.

Fragment	Position on chromosome 2 [hg19]	Fragment size [bp]	Forward primer (5' -> 3')	Reverse Primer (5' -> 3')
Promoter	159,651,088-159,651,537	450	CAATAATGGCTAAGCTTAAAGTTGT	GTGCACCTGCAATGTAACAG
Exon 1	159,651,457-159,652,045	589	AGATGTCCTGAAGCAAAGAGC	GGTAGACAGAGAAGCTTCCGT
Intron 1.1	159,652,693-159,653,272	580	CTTCCAACTGCCTTCCTTCTT	TCATTCGAACTTTCATTTGGTT
Intron 1.2	159,653,108-159,653,867	760	AGGCTCCTTACATAGACAAAGACA	TACCGGGCCTGCATTTTA
Intron 1.3	159,655,462-159,656,171	710	TGCAGGCTTATGTTTATCTTCACT	GAAGCGGAAGATAATATCCCTAGA
Intron 1.4	159,659,652-159,660,267	616	ATGTGTAGAATAACTGTGGATGA	ATACCCACCTTCCATCCAGA
Exon 2	159,660,178-159,660,965	788	GGCAGGATTATTAATGCAGTTT	GCCTTGGTTTGTCCTAAAATAACT
Intron 2.1	159,660,875-159,661,384	510	AAACAAGGTAGGGACTCTTAATTTT	CTCGGTACCTATGTGTTTCTGG
Intron 2.2*	159,661,876-159,662,944	1069 (deletion absent) 191 (deletion present)	CATGACAAAAGGGAAGCTGTGC	GCTATCCTCATTTATAGCCCCAAAGA
Exon 3	159,663,504-159,664,070	567	ATGTGGCCTTTTCAGTGGGA	TTCCTCCAACCCTAGACCAT
Exon 3b	159,664,059-159,664,411	353	GGGTTGGAGGAAGGCTGGGTACC	TCGAGGTGGCCGGATGGCCTG
Intron 3.1	159,667,434-159,668,027	594	CTGGGATTTGTGCTCGTTCA	TTATGACTAGAGAGCCCTGAAAT
Intron 3.2	159,667,947-159,668,395	449	ATACTACCCAACTTCAGGCTCTAA	CAGCTTTCCAGAGTATAAGCTACA
Intron 3.3	159,670,892-159,671,351	460	GTTGGCAGAGATAATCCTTGG	GGCCAGCAAGGAAGGAAT
Intron 3.4	159,671,834-159,672,072	239	TGAGGAGTTTTTCGTCCCCACTCT	TTTCTAGCCCTAAACCCATGGGAGT
Exon 4	159,672,011-159,672,594	584	AAGGACAGTAAATGAGAGAAGAC	AGGTATTCTGGATGCTTCACTT
Exon 4b	159,682,865-159,682,995	537	CCCTTTCCTTCCTTAGACCTTCC	TTGGAAGAAGATTCCTCAACTACAC
Exon rare_5	159,710,049-159,710,318	272	CCAGCATTAAGGGCTGGTTTT	GGTTGGAGAGGAAGAAGTACACG
Exon 5b	159,718,792-159,719,425	634	CGCAGACATGATGCTGGGGGT	ACATGCAAGACGGGGAATTGA

* Fragment was not resequenced but visually inspected by gel electrophoresis

Supplementary Table S13. Resequencing results of selected genomic regions (see Supplementary Table S12 and Supplementary Figure S7).

dbSNP ID / GeneBank ID (Position)	Position on chromosome 2 [hg19]	Allele 1	Allele 2	Individuals homozygous for non-risk alleles	Individuals homozygous for either risk allele
				N(Allele 2) / N(chromosomes)	N(Allele 2) / N(chromosomes)
rs925781	159,651,734	G	A	10 / 16	0 / 24
rs1356173	159,651,973	G	A	4 / 16	0 / 24
rs12463934	159,652,852	C	T	4 / 16	0 / 24
rs7578195	159,653,250	T	C	1 / 16	0 / 24
rs71421092	159,653,626	A	G	0 / 16	5 / 22
rs12996550	159,653,721	C	G	3 / 16	0 / 22
rs6437202	159,656,114	G	C	8 / 14	0 / 16
rs6756488	159,660,277	T	A	4 / 16	0 / 24
rs75277023	159,660,494	G	A	0 / 16	24 / 24
rs11297084	159,660,646	-	A	6 / 16	0 / 24
rs17810398	159,660,870	C	T	0 / 16	24 / 24
rs2271663	159,661,077	C	T	6 / 16	0 / 24
rs147412692	159,661,111	A	G	0 / 16	9 / 24
rs11684616	159,661,318	A	G	0 / 14	0 / 22
rs6146986	159,661,997-159,662,874	-	878 bp deletion	0 / 16	24 / 24
rs9869	159,663,599	T	C	4 / 14	2 / 24
rs10497199	159,663,616	G	A	4 / 14	0 / 24
rs75859613	159,664,016	T	C	1 / 16	0 / 24
rs62183691	159,664,285	G	A	6 / 16	0 / 24
NM_001017920.2_c.207+3984 / HQ220184 (77)	159,667,611	G	A	1 / 16	0 / 24
rs908402	159,667,731	G	A	6 / 16	0 / 24
rs71421093	159,668,241	A	G	4 / 16	0 / 24
rs11693126	159,670,945	T	C	0 / 16	0 / 22
NM_001017920.2_c.208-1029 / HQ220186 (225)	159,671,188	C	T	0 / 16	1 / 22
rs17810816	159,671,992	A	G	0 / 16	24 / 24
rs1515922	159,672,168	G	A	1 / 16	0 / 24
rs61740878	159,672,252	G	A	1 / 14	0 / 24
rs10016	159,672,442	A	G	9 / 14	0 / 24
rs35919361	159,683,271	C	A	2-4 / 4*	1-2 / 2*
rs9750235	159,710,189	A	T	11 / 16	1 / 24
rs144087548	159,718,894	A	T	0 / 16	23 / 24
rs34995873	159,719,031	A	G	1 / 16	0 / 24

* Due to an extensively long AT/T-stretch, only three individuals could be resequenced after subcloning the obtained stutter bands.

Supplementary Table S14. Variants detected by cDNA re-sequencing of eight RPE/retina tissue samples heterozygous for rs17810398. Alleles are given for the + strand.

dbSNP ID	RPE/retina tissue sample							
	ID_A	**ID_B**	**ID_C**	**ID_E**	**ID_2**	**ID_11**	**ID_13**	**ID_14**
Gender	Female	Male	Male	Male	Male	Female	Male	Male
rs17810398	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T
rs9869	C/T	T/T	T/T	T/T	T/T	C/T	C/T	C/T
rs10497199	G/G	A/G	A/G	G/G	A/G	A/G	G/G	G/G
rs34995873	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/G

10.4 List of illustrations and tables