AUS DEM LEHRSTUHL FÜR HUMANGENETIK PROF. DR. BERNHARD WEBER DER FAKULTÄT FÜR MEDIZIN DER UNIVERSITÄT REGENSBURG

CHARAKTERISIERUNG PLEIOTROPER GENETISCHER VARIANTEN ANHAND IHRER GENEXPRESSIONSREGULATION

Inaugural – Dissertation zur Erlangung des Doktorgrades der Medizin

der Fakultät für Medizin der Universität Regensburg

> vorgelegt von Simon Stelzl

AUS DEM LEHRSTUHL FÜR HUMANGENETIK PROF. DR. BERNHARD WEBER DER FAKULTÄT FÜR MEDIZIN DER UNIVERSITÄT REGENSBURG

CHARAKTERISIERUNG PLEIOTROPER GENETISCHER VARIANTEN ANHAND IHRER GENEXPRESSIONSREGULATION

Inaugural – Dissertation zur Erlangung des Doktorgrades der Medizin

der Fakultät für Medizin der Universität Regensburg

> vorgelegt von Simon Stelzl

Dekan:	Prof. Dr. Dirk Hellwig
1. Berichterstatter:	Prof. Dr. Bernhard Weber
2. Berichterstatter:	Prof. Dr. Birte Kehr

Tag der mündlichen Prüfung: 20.5.2022

Inhaltsverzeichnis

Zı	Jsamr	nenfassung	4		
1	Ein	leitung	7		
	1.1	Genetik komplexer Merkmale	7		
	1.2	Genomweite Assoziationsstudien (GWAS)	8		
	1.3	Expression Quantitative Trait Loci (eQTL) Studien	10		
	1.4	Altersabhängige Makuladegeneration (AMD)	. 12		
	1.5	Pleiotrope Effekte bei komplexen Erkrankungen und Merkmalen	15		
	1.6	Ziel der Arbeit	16		
2	Mat	terial und Methoden	19		
	2.1	Programme und Datenbanken	. 19		
	2.2	Literaturrecherche	19		
	2.3	Genetische Risiko Score (GRS) Berechnung	20		
	2.4	Expression Quantitative Trait Loci (eQTL) Berechnung	22		
	2.5	Gemeinsame Gene zwischen den Genlisten der Merkmale	22		
3	Erg	ebnisse	26		
	3.1	Literaturrecherche	26		
	3.2	Genetische Risiko Score (GRS) Berechnung	27		
	3.3	Expression Quantitative Trait Loci (eQTL) Berechnung	30		
	3.4	Gemeinsamkeiten zwischen den Genlisten der Merkmale	31		
	3.5	Gemeinsame Gene zwischen den Genlisten der Merkmale	. 34		
4	Dis	kussion	41		
A	Abkürzungsverzeichnis				
A	bildu	ngsverzeichnis	68		
Та	abelle	nverzeichnis	69		
A	nhang	tabellenverzeichnis	70		
Li	teratu	rverzeichnis	71		
A	Anhang				
D	Danksagung 127				
S	Selbstständigkeitserklärung 128				

Zusammenfassung

In den letzten Jahren wurden zu einer Vielzahl an komplexen Merkmalen und Krankheiten genomweite Assoziationsstudien (GWAS) durchgeführt und so eine Fülle an genetischen Varianten identifiziert, welche sich mit den untersuchten Phänotypen assoziiert zeigen. Basierend auf diesen Ergebnissen konnten zudem genetische Korrelationen zwischen diversen komplexen Krankheiten und Merkmalen festgestellt werden. Da sich die Mehrheit der in GWAS gefundenen genetischen Varianten jedoch in intronischen oder intergenischen Genomregionen befindet, bleiben deren funktionellen Eigenschaften und regulatorischen Wirkungen auf die Genexpression zunächst unklar. Infolgedessen sind auch die Ursachen der bereits erforschten genetischen Korrelationen zwischen verschiedenen komplexen Merkmalen meist noch unbekannt. In dieser Arbeit wird eine Vorgehensweise präsentiert, mithilfe von *expression Quantitative Trait Loci* (eQTL) Analysen potenziell pleiotrop wirkende Gene zu identifizieren, um so die Lücke zwischen der genetischen Korrelation komplexer Merkmale, insbesondere mit der altersabhängigen Makuladegeneration (AMD), und den zugrundeliegenden physiologischen Mechanismen zu schließen.

Hierfür wurde zu Beginn eine umfangreiche Literatursuche durchgeführt. Mit Hilfe dieser wurden die aktuell verfügbaren unabhängigen GWAS zu ausgewählten Krankheiten und Merkmalen identifiziert. Der Fokus wurde insbesondere auf Merkmale gelegt, deren genetische Korrelation bereits in vergangenen Forschungsarbeiten beschrieben wurde, um diese Zusammenhänge im Hinblick auf genetische Ursachen und zugrundeliegende Mechanismen vertiefend zu untersuchen. Hierzu wurden neben AMD verschiedene kardiovaskuläre Merkmale, Serumspiegel von Leberenzymen sowie einige Blutlipidwerte gewählt. Die zu den Merkmalen identifizierten GWAS wurden in Bezug auf ausgewählte Qualitätskriterien selektiert und daraus insgesamt 870 genomweit signifikante (p < 5e-08) *Single Nucleotide Polymorphisms* (SNPs) entnommen. Daraufhin wurden die Korrelationen zwischen diesen SNPs und Genexpressionsdaten im Rahmen einer eQTL Analyse untersucht. Mit dieser Methode wurden in ihrer Genexpression beeinflusste Gene identifiziert. Diese Gene ermöglichen wiederum Rückschlüsse auf mögliche Pathomechanismen der jeweiligen betrachteten Erkrankungen oder Merkmale.

Zur eQTL Analyse wurden die Genotypdaten sowie die Genexpressionsdaten der Gewebe Leber und Blut aus der Datenbank des *Genotype-Tissue-Expression* (GTEx) Projekts entnommen. Hinsichtlich des Gewebes Leber konnten so die Genexpressionsinformationen von 131 Spendern zu 26.072 Genen genutzt werden. Zur eQTL Analyse des Gewebes Blut wurden Genexpressionsinformationen von 323 Spendern zu 29.151 Genen verwendet.

Basierend auf den Genexpressionsdaten des Gewebes Leber konnten 268 eQTL identifiziert werden, sowie 194 eQTL mithilfe der Genexpressionsdaten des Gewebes Blut (P < 1e-05). Von den in ihrer Expression beeinflussten Genen sind besonders Gene interessant, deren Expression mit mehreren Merkmalen zugleich in Zusammenhang steht, und folglich als pleiotrope Gene bezeichnet werden. Durch Untersuchung solcher Gene können Erkenntnisse zu überlappenden physiologischen Mechanismen gewonnen werden. Deswegen wurden zu jedem Merkmal und Gewebe die eQTL Gene zwischen den Merkmalen auf gemeinsame Gene verglichen. Insgesamt konnten so bezüglich der beiden untersuchten Gewebe 65 potenziell pleiotrop wirkende Gene identifiziert werden, welche mit unterschiedlichen Merkmalen zugleich in Zusammenhang stehen. Zu einigen dieser Gene wurden die in dieser Arbeit identifizierten Merkmalszusammenhänge bereits in anderen Publikationen eindeutig beschrieben. Ein Beispiel hierfür wäre das Gen signal peptide, CUB domain and EGF like domain containing 2 (SCUBE2), für welches die hier vorliegende Arbeit einen Zusammenhang mit verschiedenen Blutdruckmerkmalen sowie dem Blutspiegel von Lipoproteinen hoher Dichte aufzeigt. Die Tatsache, dass einige Zusammenhänge zwischen pleiotropen Genen und den assoziierten Phänotypen plausibel durch bekannte Publikationen bestätigt werden konnten, weist die Funktionstüchtigkeit des in dieser Arbeit verwendeten Untersuchungsansatzes nach. Zu vielen weiteren hier als pleiotrop identifizierten Genen sind jedoch noch kaum Informationen zu Funktion und assoziierten molekularen Mechanismen bekannt. Solche Gene können im Rahmen weiterer Studien gezielt untersucht werden, um die physiologischen Prozesse aufzudecken, welche den identifizierten pleiotropen Effekten zugrunde liegen. Neben diesen Aspekten kann der in dieser Arbeit vorgestellte Ansatz zur Identifizierung pleiotroper Gene zudem anhand der Genexpressionsdaten zahlreicher weiterer Gewebe und Zellarten angewandt werden, sofern hierzu geeignete Datensätze verfügbar sind. Aufgrund der heterogenen Genexpression der verschiedenen

Körpergewebe und Zelltypen ermöglicht dies die spezifische Untersuchung genetischer Mechanismen sowie hiermit assoziierter Signalwege, mit dem Ziel therapeutisch angreifbare Signalprozesse zu ermitteln.

Zusammengefasst konnten in der vorliegenden Arbeit mithilfe von eQTL Analysen 65 Gene mit potenziell pleiotropen Effekten auf komplexe Merkmale anhand Genexpressionsdaten der Gewebe Leber und Blut identifiziert werden. Bei einem Großteil dieser Gene sind die in dieser Arbeit festgestellten Zusammenhänge jedoch noch unbekannt. Dies zeigt die Notwendigkeit der Identifizierung pleiotrop wirkender Gene, um einerseits Hinweise zu erhalten, welche Gene als Verbindungsglied komplexer Merkmale fungieren und somit als Kandidat weiterer Erforschung in Betracht kommen, und andererseits, um die der Pleiotropie zugrundeliegenden molekularen Prozesse weiter zu beleuchten.

1 Einleitung

1.1 Genetik komplexer Merkmale

Bis vor wenigen Jahren wurde die medizinische Genetik als nahezu reine Domäne rarer Fälle chromosomaler Aberrationen sowie Mendelscher Erkrankungen betrachtet. Die meisten dieser genetischen Erkrankungen sind selten und werden durch Mutationen eines einzigen Gens verursacht und infolgedessen als monogenetisch bezeichnet. So beruht beispielsweise die autosomal rezessiv vererbte Krankheit Zystische Fibrose auf Mutationen im cystic fibrosis transmembrane conductance regulator (*CFTR*) Gen [1]. Solche Zusammenhänge konnten durch Anwendung einer Untersuchungsmethodik namens Kopplungsanalyse (*Linkage Analysis*) aufgedeckt werden. Hierzu wurden von der Krankheit betroffene Familien genotypisiert und genetische Marker definiert. Anschließend wurde verfolgt, wie diese genetischen Marker mit der Krankheit innerhalb der Familien segregieren, wodurch auf die Lokalisation der pathogenen Gene geschlossen werden konnte [2,3].

Im Gegensatz zu monogenetischen Erkrankungen wurde der Genetik sogenannter komplexer Erkrankungen lange keine besondere Beachtung beigemessen. Hauptursächlich dafür war die Tatsache, dass komplexe Merkmale und Erkrankungen durch eine Vielzahl genetischer und umweltbedingter Faktoren beeinflusst werden und Methoden wie die Kopplungsanalyse nicht imstande waren, die entsprechenden Zusammenhänge aufzudecken [4]. Bedeutende Errungenschaften der jüngeren Humangenetik, wie beispielsweise die Entschlüsselung des menschlichen Genoms sowie die Beschreibung ihrer Variation im Rahmen des Humangenomprojekts [5] und des HapMap Projekts [6], ermöglichten zusammen mit Fortschritten in der Gentechnologie jedoch eine Änderung des Fokus der Genomforschung. Infolge der erweiterten Untersuchungsmöglichkeiten wandte sich die Aufmerksamkeit vermehrt der Genetik komplexer Erkrankungen und Merkmale zu [7].

Da komplexe Erkrankungen typischerweise durch die Kombination einer Vielzahl genetischer und umweltbedingter Faktoren verursacht werden, die zudem noch miteinander interagieren können, hat eine einzelne genetische Variante vergleichsweise geringe funktionale Auswirkungen auf die betrachtete Erkrankung [4,8]. Dies steht im Einklang mit der Häufige Erkrankung / Häufige Variante (*Common*

Disease / Common Variant) Hypothese. Sie postuliert, dass häufige beziehungsweise komplexe Erkrankungen aller Wahrscheinlichkeit nach von genetischen Varianten beeinflusst werden, die häufig in der Population vorkommen [9,10]. Falls diese häufigen Varianten das Auftreten von Krankheiten beeinflussen, folgt hieraus, dass die Effektgröße solcher Varianten vergleichsweise klein sein muss. Dies erklärt, wieso die traditionellen, auf Familien basierenden Studienmodelle, welche zur Identifizierung von Genen mit großen Effektgrößen konzipiert wurden, bei komplexen Erkrankungen nicht erfolgversprechend sind und die Untersuchung der Genetik komplexer Erkrankungen neue Studienmethoden verlangt [11,12].

1.2 Genomweite Assoziationsstudien (GWAS)

Ein solcher Ansatz zur Erforschung der Genetik komplexer Erkrankungen oder Merkmale sind genomweite Assoziationsstudien (GWAS). Hierbei wird einerseits eine Fallgruppe, deren Individuen das zu untersuchende Merkmal oder die Erkrankung aufweisen, nicht betroffene Kontrollgruppe genotypisiert. sowie eine Die Genotypinformationen bzw. - frequenzen der beiden Gruppen werden anschließend verglichen und auf Unterschiede untersucht. Falls eine genetische Variante in statistisch signifikant höherer oder niedrigerer Anzahl in der Fallgruppe als in der Kontrollgruppe identifiziert wird, wird von einer Assoziation zwischen dieser genetischen Variante und dem betrachteten Merkmal ausgegangen. Das Allel, welches das Risiko für ein bestimmtes Merkmal oder eine Erkrankung erhöht, wird hierbei als das Risikoallel bezeichnet [13,14]. GWAS ermöglichen es somit, statistische Zusammenhänge zwischen genetischen Varianten und dem jeweiligen Phänotyp aufzuzeigen [15]. Die erste erfolgreiche GWAS wurde im Jahr 2005 zur altersabhängigen Makuladegeneration (AMD) durchgeführt, mit deren Hilfe das complement factor H (CFH) Gen als Risikofaktor identifiziert werden konnte [16].

Eine Form von genetischer Variation stellen sogenannte Einzelnukleotidpolymorphismen (*Single Nucleotide Polymorphisms*, SNPs) dar. SNPs sind einzelne Änderungen von Basenpaaren in der DNA Sequenz, die in hoher Frequenz im menschlichen Genom vorkommen [17]. Aufgrund der Größe des menschlichen Genoms und den bisher über 84,7 Millionen identifizierten SNPs können jedoch nicht alle dieser genetischen Varianten eines Individuums untersucht werden [18]. Um dieses Problem zu umgehen, werden repräsentative SNPs ausgewählt, welche als genetische Marker eine Gruppe von SNPs, einen sogenannten Haplotyp,

repräsentieren. Eine solche Gruppe von SNPs befindet sich hierbei im Kopplungsungleichgewicht (*Linkage Disequilibrium*, LD) zueinander [19,20]. Das LD beschreibt das Ausmaß, in welchem das Allel eines SNPs zusammen mit dem Allel eines anderen SNPs innerhalb einer Population vererbt wird. SNPs werden als im LD zueinanderstehend bezeichnet, wenn die Frequenzen der Allele der beiden SNPs korrelieren, wobei die Frequenz eines SNPs meist durch die Frequenz des weniger häufigen Allels definiert wird (*Minor Allel Frequency*). Um die Korrelation zwischen zwei SNPs quantitativ angeben zu können, wird der Korrelationskoeffizient r² verwendet. Dieser Wert hat eine Schwankungsbreite zwischen 0 und 1, wobei ein Wert von r² = 0 eine nicht vorhandene Korrelation angibt, und r² = 1 eine perfekte Korrelation beschreibt [6,21,22].

Insgesamt bietet die Methodik der GWAS die Möglichkeit einer systematischen und hypothesenfreien Untersuchung der Assoziation von SNPs mit komplexen Merkmalen, wobei keine vorausgehenden Informationen bezüglich Funktion von Genen oder der untersuchten Variablen erforderlich sind. Zudem werden auch potenziell regulatorische Bereiche einbezogen, die sich in intronischen oder intergenischen Bereichen befinden können. Dies ermöglicht einen umfassenden Einblick in die genetischen Einflüsse der untersuchten Erkrankungen und Merkmale [23].

GWAS weisen jedoch auch einige Limitierungen auf. Wenngleich diese Studien zahlreiche SNPs im Genom identifizieren konnten, die mit verschiedensten Krankheiten und Merkmalen assoziiert sind, so sind sie nicht fähig festzustellen, welche Gene und molekularen Mechanismen diesen Assoziationen zugrunde liegen [13]. Oftmals liegen weniger als 10 % der identifizierten SNPs in proteinkodierenden Regionen oder sind in hohem LD mit solchen Bereichen, wodurch die Identifizierung der beeinflussten Gene und deren molekulare Beteiligung am Phänotyp erschwert wird [24,25]. Auch die in GWAS gängige Praxis der Zuordnung von SNPs zum nächstliegenden Genlokus ist wenig aussagekräftig, da die Gene, die sich sehr nahe am assoziierten SNP befinden, oftmals nicht die für die Assoziation kausalen, durch die SNPs beeinflussten Gene darstellen [26,27]. Es konnten beispielsweise funktionelle Verbindungen zwischen Genen und Varianten gefunden werden, die sich über eine Entfernung von mehreren Megabasen (Mbs) erstrecken [28]. Eine weitere Limitation von GWAS ist die Tatsache, dass sie, wie durch die Häufige Erkrankung /

Häufige Variante Hypothese postuliert, typischerweise häufige Varianten mit sehr kleinen Effektgrößen identifizieren, wodurch die Suche nach den beeinflussten Genen zusätzlich erschwert wird [12,29,30].

Diese Einschränkungen der GWAS erklären, wieso weiterführende Methoden erforderlich sind, um die funktionellen Auswirkungen der durch GWAS identifizierten krankheitsassoziierten Varianten auf Gene gezielt zu untersuchen.

1.3 Expression Quantitative Trait Loci (eQTL) Studien

Eine Möglichkeit, die Grenzen von GWAS bezüglich der Identifizierung der beeinflussten Gene zu überwinden und die Risikoallele in physiologische Einsichten zu überführen, ist neben der Betrachtung der Genotypinformationen die zusätzliche Untersuchung von Genexpressionsdaten.

Das Expressionsniveau eines Individuums, also die Menge an mRNA der transkribierten Gene in der untersuchten Probe, stellt einen intermediären Phänotypen dar und steht physiologisch somit zwischen genomischer DNA und der Ausprägung des komplexen Phänotypen auf Ebene der Zellen, Organe und des Organismus [31]. Das Prinzip der kombinierten Analyse von Genotyp und Genexpressionsdaten wurde 2001 erstmal von Jansen und Nap unter der Bezeichnung *Genetical Genomics* vorgestellt und wird in der Gegenwart unter dem Begriff *expression Quantitative Trait Loci* (eQTL) Analyse angewandt [32]. Hierbei wird die Korrelation zwischen SNPs und dem Genexpressionsniveau in einem Gewebe betrachtet, um in ihrer Expression beeinflusste Gene zu identifizieren [33,34].

Da die Genexpression stark gewebespezifisch ist, sind neben entsprechenden Genotypinformationen die Expressionsdaten verschiedener Gewebetypen erforderlich [35]. Eine Datenbank, welche sowohl Genotyp- als auch Genexpressionsdaten zahlreicher Gewebe beinhaltet, wird durch das *Genotype-Tissue-Expression* (GTEx) Projekt bereitgestellt. Dieser Datensatz enthält Genotyp- und Genexpressionsinformationen von über 600 Spendern zu 48 Geweben, darunter beispielsweise Lebergewebe, Blut oder Nierengewebe, wobei nicht zu jedem Spender Proben aller 48 Gewebe zur Verfügung stehen [36]. Die durch diese Datenbank verfügbaren Informationen ermöglichen es, die eQTL in verschiedenen Geweben

gezielt zu untersuchen und die erlangten Ergebnisse zwischen den Geweben zu vergleichen.

Bei der Studie von eQTL kann zwischen sogenannten cis und trans eQTL unterschieden werden (siehe Abbildung 1). Cis eQTL wirken über lokale Effekte, wobei sich der SNP direkt im oder in direkter Nachbarschaft zum beeinflussten Gen befindet. So können genetische Varianten beispielsweise zu einer veränderten Bindungsfähigkeit von Transkriptionsfaktoren führen oder die Methylierung der DNA beeinflussen und somit eine veränderte Expression des Gens verursachen. Falls diese Änderungen der Genexpression auch einen Einfluss auf andere, weiter entfernt liegende Gene ausüben, so werden die mit diesen Effekten assoziierten Loci als trans eQTL bezeichnet. Diese Effekte werden typischerweise durch Intermediärfaktoren vermittelt, welche beispielsweise als Transkriptionsfaktoren auf das beeinflusste Gen wirken [37].





Abbildung 1: eQTL und ihre Wirkungsweisen.

Bei cis eQTL befindet sich der eQTL SNP in der Nähe (<1Mb) des eQTL Gens und beeinflusst die Genexpression über lokale, direkte Effekte. Durch solche Änderungen der Genexpression und nachfolgender Beeinflussung des zugehörigen Proteinniveaus kann auch eine Wirkung auf die Expression weiter entfernt liegender Gene ausgeübt werden. Diese indirekten, durch Intermediärfaktoren vermittelten Effekte des SNPs werden als trans eQTL bezeichnet. Cis eQTL zeigen häufig ausgeprägtere Unterschiede in den Expressionsniveaus als trans eQTL. (modifizierte Abbildung nach Westra et al. (2014)[37])

Im Rahmen einer eQTL Analyse können somit die durch GWAS identifizierten SNPs, welche mit verschiedenen komplexen Merkmalen oder Erkrankungen assoziiert sind, vertieft untersucht werden. Dies ermöglicht es, in ihrer Expression beeinflusste Gene zu identifizieren, um damit Rückschlüsse auf die potenziellen Pathomechanismen der betrachteten Erkrankungen oder Merkmale zu ziehen.

1.4 Altersabhängige Makuladegeneration (AMD)

Ein Beispiel einer komplexen Erkrankung ist die AMD. Mit 196 Millionen Betroffenen im Jahre 2020 stellt sie eine der häufigsten Ursachen für Erblindung in westlichen Gesellschaften dar [38,39].

AMD wird in verschiedene Progressionsstadien eingeteilt, die typischerweise mit einem zunehmenden Verlust des zentralen Sehens einhergehen. So wird nach aktueller klinischer Klassifikation zwischen einer Frühform, einer Intermediärform und zwei Spätformen der AMD unterschieden [40,41]. Die Früh- und Intermediärformen sind symptomatisch weniger auffällig und zeichnen sich durch Pigmentveränderungen sowie durch Ablagerungen von extrazellulärem Material, genannt Drusen, unter dem retinalen Pigmentepithel (RPE) vor allem im Bereich der Makula Lutea, der zentralen Stelle der Netzhaut, aus, Zusätzlich zu diesen Ansammlungen kommt es zu lokalen Entzündungsreaktionen mit nachfolgenden neurodegenerativen Prozessen. Die Spätform der AMD wird gewöhnlich in eine trockene Form und eine feuchte Form unterteilt. Die trockene Form zeichnet sich hierbei durch lokale Atrophien, also einen Gewebeschwund, des RPEs aus und wird in seiner Endphase als geographische Bei der feuchten Form kommt es zu pathologischen Atrophie bezeichnet. Gefäßneubildungen der Choroidea, den sogenannten choroidalen Neovaskularisationen, welche die Retina durch das RPE infiltrieren. Die Ursache für den Visusverlust liegt bei den Spätformen in der progressiven Schädigung der Photorezeptoren durch Dysfunktionen des RPEs, einem beeinträchtigten Transport von Stoffwechselprodukten und Nährstoffen zwischen Gefäßen und den äußeren retinalen Zellen, sowie Leckagen der choroidalen Kapillaren beziehungsweise der choroidalen Neovaskularisationen [40,41].

Wie für eine komplexe Erkrankung typisch, wird die AMD von einem komplizierten Zusammenspiel aus zahlreichen genetischen Faktoren, Lebensstileinflüssen und Umweltbedingungen beeinflusst [42–46].

Umweltfaktoren und Lebensstileinflüsse, die das Auftreten dieser Erkrankung beeinflussen sind unter anderem das Lebensalter sowie der Raucherstatus [43,44,47]. Die Auswirkungen des Rauchens auf die AMD wird vor allem den fortschreitenden Schäden aufgrund von oxidativem Stress zugeschrieben. Auch altersabhängigen Schäden der mitochondrialen DNA wird ein Beitrag zur AMD Pathogenese attestiert [47–49]. Zudem wird von Unterschieden im AMD Risiko zwischen verschiedenen Ethnien berichtet, wobei für hellhäutige Populationen ein höheres AMD Risiko im Vergleich zu dunkelhäutigeren Populationen beobachtet werden konnte [45]. Es wird vermutet, dass dies unter anderem auf Unterschieden in der Verteilung von Varianten risikoassoziierter Gene zwischen den Bevölkerungsgruppen basiert [50].

Bezüglich der genetischen Faktoren, die zur Pathogenese der AMD beitragen, zeigen Zwillingsstudien größere Konkordanzen bei monozygoten als dizygoten Zwillingen [51]. Auch eine Häufung von AMD Fällen in Familien unterstreicht die genetische Basis der AMD, wobei gezeigt werden konnte, dass Personen mit einem AMD erkrankten Verwandten 1. Grades ein bis zu 27-fach erhöhtes Risiko aufweisen, ebenfalls an AMD zu erkranken [52]. Mithilfe von GWAS konnten bisher 34 mit AMD assoziierte Loci identifiziert werden, welche 52 unabhängige Signale enthalten (siehe **Abbildung 2**) [42]. Bis zu 50 % der Erblichkeit von AMD basiert auf 2 Loci, welche kodierende und nicht kodierende Varianten auf Chromosom 1 (*CFH*) und Chromosom 10 (*ARMS2/HTRA1*) aufweisen [16,53–55].



Abbildung 2: Manhattan Plot der mit AMD assoziierten Loci. Jeder Punkt repräsentiert einen SNP und dessen P-Wert der AMD Assoziation. Auf der Y-Achse wurden die P-Werte der Assoziation und auf der X-Achse die genomische Position aufgetragen. Die rote Linie zeigt die P-Wert Schwelle für genomweite Signifikanz (P < 5e-08). In blau sind genomweit signifikante AMD Loci dargestellt. (Modifiziert nach einer Abbildung von Fritsche et al. (2016) [42])

Verschiedene Studien legen nahe, dass AMD trotz der lokalen Augen-assoziierten Pathologien eher als eine systemische Erkrankung begriffen werden könnte, da neben Retina-assoziierten Körperprozessen auch Zusammenhänge der AMD mit fundamental verschiedenen molekularen Vorgängen in zahlreichen Geweben aufgezeigt wurden. Es werden beispielsweise Beteiligungen von Immun- und Entzündungsprozessen wie der Komplementkaskade vermutet, wobei anhand von Untersuchungen der mit AMD assoziierten Drusen und den der Drusen umgebenden RPE Zellen sowohl Komponenten als auch Regulatoren des Komplementsystems festgestellt werden konnten [56,57]. Einige Studien zeigen ebenfalls eine Korrelation von AMD mit dem Metabolismus und den Serumspiegeln von Lipiden wie Cholesterol und Triglyceriden sowie deren Metaboliten [58–61]. Lipide erfüllen im menschlichen Körper wichtige Funktionen wie Energiespeicherung, Signalgebung zwischen Zellen oder als Baustoff von Zellmembranen [62,63]. Der Lipidmetabolismus kann somit ebenfalls als ubiquitär im Organismus vorkommender Prozess betrachtet werden, welcher wahrscheinlich die AMD Pathogenese beeinflusst.

Kiel et al. (2017) konnten die Komplexität der mit AMD assoziierten Mechanismen darlegen, indem sie im Rahmen einer Netzwerkanalyse vermutliche AMD Risikogene aus bekannten Datenbanken und Publikationen extrahierten und die mit diesen Genen assoziierten Prozesse aufzeigten. So wurde beobachtet, dass nur 7 % der

vermutlichen AMD Risikogene direkt mit dem Sehvorgang assoziiert sind, während der Großteil der Gene auf molekulare Prozesse verschiedenster Gewebe und Kompartimente einwirkt, wie beispielsweise Beteiligungen am Aufbau der extrazellulären Matrix, Regulation von oxidativem Stress sowie den bereits genannten Entzündungsprozessen [64]. Die genannten Beispiele und Studien untermauern die These, dass die AMD aufgrund ihrer engen Korrelation mit anderen komplexen und nicht Retina assoziierten Phänotypen und Körperprozessen auch als systemische Erkrankung betrachtet werden sollte.

1.5 Pleiotrope Effekte bei komplexen Erkrankungen und Merkmalen

Neben der Tatsache, dass komplexe Erkrankungen wie die AMD durch eine Vielzahl von Varianten und Genen beeinflusst werden, also polygenetisch sind, existieren auch zahlreiche Gene, die auf mehrere komplexe Erkrankungen zugleich einwirken. Falls ein einzelnes Gen oder ein Genlokus mehrere unterschiedliche Phänotypen beeinflusst, so wird dieses Phänomen als Pleiotropie bezeichnet [65,66]. Ein Gen, welches Effekte auf mehrere Phänotypen zugleich ausübt, wird folglich pleiotropes Gen genannt.

Pleiotrope Effekte sind hierbei im Genom als häufig zu betrachten. Viele der bei GWAS identifizierten SNPs zeigen sich mit mehr als nur einem Phänotyp assoziiert und wirken folglich pleiotrop. Eine Analyse genomweit signifikanter SNPs des GWAS Catalog des *National Human Genome Research Institutes* (NHGRI) konnte beispielsweise zu durchschnittlich 4,6 % aller untersuchter GWAS SNPs pleiotrope Effekte aufzeigen [67]. Cotsapas et al. (2011) schätzen, dass zwischen Autoimmunerkrankungen der Anteil an pleiotropen SNPs bei mindestens 44 % liegt [68].

Basierend auf den durch GWAS identifizierten Risikovarianten für komplexe Erkrankungen und Merkmale zeigten Grassmann et al. (2017) zudem, dass zahlreiche Korrelationen zwischen der Genetik komplexer Merkmale bestehen. So zeigen sich zum Beispiel Korrelationen in der Genetik unterschiedlicher Autoimmunerkrankungen, sowie zwischen verschiedenen Blutlipiden und kardiovaskulären Merkmalen. Auch für AMD wurde eine Korrelation der Genetik mit unterschiedlichsten komplexen Erkrankungen und Merkmalen gezeigt, darunter unter anderem kardiovaskuläre

Erkrankungen wie koronare Herzkrankheit (CAD) und verschiedene Blutlipidspiegel [69].

Studien wie die oben genannte Publikation konnten die genetischen Zusammenhänge zwischen komplexen Erkrankungen und Merkmalen bereits beschreiben. Die hierfür ursächlichen pleiotropen Gene sowie die mit diesen Genen assoziierten und somit zwischen den Merkmalen überlappenden Pathomechanismen lassen sich aufgrund des fehlenden Einbezugs von Genexpressionsdaten durch diesen Ansatz jedoch nicht eindeutig bestimmen und erfordern den Einsatz weiterführender Methoden.

1.6 Ziel der Arbeit

In den letzten Jahren wurden zu einer Vielzahl an komplexen Krankheiten und Merkmalen GWAS durchgeführt und so eine Fülle genetischer Varianten identifiziert, welche sich mit den untersuchten Phänotypen hochsignifikant assoziiert zeigten. Basierend auf diesen Daten konnten zudem genetische Korrelationen zwischen diversen komplexen Krankheiten und Merkmalen festgestellt werden. Da sich die Mehrheit der in GWAS gefundenen genetischen Varianten jedoch in intronischen oder intergenischen Genomregionen befindet, bleiben deren funktionelle Eigenschaften unklar. Infolgedessen sind auch die Ursachen der bereits erforschten genetischen Korrelationen zwischen verschiedenen komplexen Merkmalen meist noch unbekannt. Ziel dieser Arbeit ist es daher, mithilfe von eQTL Analysen pleiotrop wirkende Gene zu identifizieren, um so die Lücke zwischen der genetischen Korrelation komplexer Merkmale, insbesondere mit der AMD, und den zugrundeliegenden physiologischen Mechanismen zu schließen.

Um dieses Ziel zu erreichen, wurde zuerst eine umfassende Literatursuche durchgeführt, um GWAS zu ausgewählten Krankheiten und Merkmalen zu identifizieren. Die GWAS wurden in Bezug auf ausgewählte Qualitätskriterien selektiert und daraus genomweit signifikante SNPs extrahiert. Bei der Auswahl der untersuchten Merkmale wurde der Fokus insbesondere auf Merkmale gelegt, deren genetische Zusammenhänge bereits in vergangenen Publikationen beschrieben wurden, wie beispielsweise zwischen dem Blutlipidstatus und der AMD [69]. Die Korrelationen zwischen SNPs und Genexpressionsdaten wurden anschließend im Rahmen einer eQTL Analyse untersucht. Dies sollte in ihrer Genexpression durch genetische

Varianten beeinflusste Gene identifizieren, um letztlich Rückschlüsse auf die Pathomechanismen der jeweiligen betrachteten Erkrankungen oder Merkmale ziehen zu können. Von besonderem Interesse waren hierbei pleiotrope Gene, deren Expressionen mit mehreren Merkmalen in Zusammenhang stehen, um dadurch Erkenntnisse zu überlappenden physiologischen Mechanismen zu gewinnen. Der Zusammenhang zwischen den komplexen Merkmalen, den hiermit assoziierten SNPs und den in ihrer Genexpression beeinflussten pleiotropen Genen ist in **Abbildung 3** schematisch dargestellt.





Es sind die Zusammenhänge zwischen komplexen Merkmalen, den mit den Merkmalen assoziierten SNPs und den in ihrer Expression beeinflussten pleiotropen Genen dargestellt. Korrelationen sind durch weiße Doppelpfeile dargestellt, wobei die Methoden zur Identifizierung der Korrelationen in den Doppelpfeilen genannt sind. Die blauen gestrichelten Pfeile verdeutlichen die Zusammenhänge, welche in dieser Arbeit vertieft untersucht werden sollen. Zu komplexen Erkrankungen oder Merkmalen wurden GWAS durchgeführt, um SNPs zu identifizieren, welche Assoziationen mit den jeweiligen Krankheiten oder Merkmalen aufweisen. Teilweise werden so SNPs identifiziert, welche, wie hier dargestellt, Zusammenhänge mit mehreren komplexen Merkmalen zugleich zeigen. Im Rahmen einer eQTL Analyse können anschließend die Korrelationen dieser SNPs mit dem Genexpressionsniveau in Geweben untersucht werden. Dies erfolgt in der Absicht, auf pleiotrop wirkende Gene zu schließen, die den genetischen Zusammenhängen der betrachteten komplexen Merkmale ursächlich sind. Die dargestellte Methodik wurde anhand der Gewebe Leber und Blut etabliert und kann in zukünftigen Forschungsarbeiten ebenso mithilfe der Genexpressionsdaten zahlreicher weiterer Gewebe und Zelltypen angewendet werden und ermöglicht so eine gewebe- und zellspezifische Untersuchung.

Der präsentierte Ansatz kann durch die Identifizierung potenziell pleiotroper Gene somit zu einem erweiterten Verständnis bezüglich der Ursachen des genetischen Zusammenhangs zwischen Merkmalen und Erkrankungen beitragen. Neben der Ermittlung potenzieller Kandidatengene und deren Signalwegen zur Untersuchung in zukünftigen Studien könnten die so erlangten Informationen Hinweise bei der Entwicklung zielgerichteter Therapien bereitstellen und so beispielsweise mögliche Angriffspunkte neuer Medikamente aufzeigen.

2 Material und Methoden

2.1 Programme und Datenbanken

Tabelle 1: Programme

Programm	Referenz/URL
Corel Draw	http://www.coreldraw.com
Gephi	http://gephi.org [70]
Microsoft Office	http://www.office.com
Paint.NET	http://www.getpaint.net
R Version 3.6.0	http://www.R-project.org [71]

Tabelle 2: Datenbanken

Datenbank	Referenz/URL
Ensembl grch37 Genome Browser	http://grch37.ensembl.org/Homo_sapiens [72]
GTEx	http://gtexportal.org [36]
GWAS Catalog	http://www.ebi.ac.uk/gwas/home [73]
LDLink	http://ldlink.nci.nih.gov [74]
Mouse Genome Database	http://www.informatics.jax.org [75]
OMIM	http://www.omim.org [76]
PubMed	http://www.ncbi.nlm.nih.gov/pubmed
PubMed Gene Database	http://www.ncbi.nlm.nih.gov/gene [77]
UniProt	http://www.uniprot.org [78]

2.2 Literaturrecherche

Zu Beginn wurde im Rahmen einer Literaturrecherche mithilfe von PubMed sowie dem GWAS Catalog [73] nach aktuellen GWAS-Daten gesucht, um hieraus SNPs zu entnehmen, welche Assoziationen mit den in dieser Arbeit betrachteten Merkmalen aufweisen (Stand November 2018). Es wurden nur SNPs zur Analyse verwendet, welche folgende Kriterien erfüllen: 1) Vorhandene Angaben zu P-Werten der Assoziationen der jeweiligen SNPs mit dem Merkmal 2) Vorhandene Informationen zu den jeweiligen Effektgrößen und Effektallelen 3) Genomweite Signifikanz (GWS; p < 5e-08) entweder in europäischer Bevölkerung oder in gemischter Bevölkerung mit mehrheitlich europäischen Probanden 4) Nicht auf X-Chromosom gelegen. Die genetischen Varianten wurden anschließend mithilfe der LDLink Webanwendung [74] auf LD in europäischen Populationen getestet, um nur SNPs zur weiteren Analyse zuzulassen, welche nicht übermäßig miteinander korreliert sind ($r^2 < 0.5$).

2.3 Genetische Risiko Score (GRS) Berechnung

Die mit den Merkmalen assoziierten SNPs wurden zur Berechnung der Genetischen Risiko Scores (GRSs) verwendet, um so potenzielle genetische Korrelationen zwischen den einzelnen Merkmalen zu beschreiben. Dazu wurden die Genotypdaten von 635 Individuen des GTEx Projekts (release v7) verwendet [36]. Die Genotypdaten wurden auf europäische Individuen gefiltert, wodurch die Daten von 527 Individuen verblieben. Bei den aus den GWAS entnommenen SNPs wurde die Orientierung angepasst, falls die jeweiligen Allele auf gegenüberliegenden DNA-Strängen angegeben waren. Bei manchen SNPs waren zu einigen Individuen die entsprechenden Informationen zum Genotyp nicht im GTEx Datensatz vorhanden. Diese fehlenden Genotypdaten wurden durch den durchschnittlichen Genotyp dieses SNPs ersetzt, solange der Gesamtanteil fehlender Daten 5 % nicht überstieg. Wurde diese Schwelle überschritten, wurde der entsprechende SNP von der weiteren Analyse ausgeschlossen. Im Rahmen einer Qualitätskontrolle wurden die Allelfrequenzen der SNPs berechnet und sowohl vor als auch nach Ausrichtung auf das risikoerhöhende Allel den in den GWAS publizierten Allelfrequenzen mithilfe von Streudiagrammen gegenübergestellt.

In einer GWAS wird der Effekt einer Variante auf ein Merkmal durch die Odds Ratio (OR) angegeben. Die Berechnung der OR erfolgt hierbei entsprechend der Formel:

 $OR = \frac{F\ddot{a}lle (Risikoallel) / Kontrollen (Risikoallel)}{F\ddot{a}lle (Nichtrisikoallel) / Kontrollen (Nichtrisikoallel)}$

In der Formel bezeichnet *Fälle (Risikoallel)* die Frequenz der Risikoallele in der Fallgruppe, *Kontrollen (Risikoallel)* die Frequenz der Risikoallele in der Kontrollgruppe, *Fälle (Nichtrisikoallel)* die Frequenz der Nichtrisikoallele in der Fallgruppe und *Kontrollen (Nichtrisikoallel)* die Frequenz der Nichtrisikoallele in der Kontrollgruppe. Ein Allel mit einer OR >1 hat somit einen risikoerhöhenden Effekt auf die betrachtete Erkrankung. Im Gegensatz dazu gibt eine OR <1 einen risikomindernden Effekt an.

Im Rahmen der GRS Berechnung wurde zu jedem Merkmal zunächst je Individuum die Anzahl an Risikoallelen (A_n) mit der logarithmierten OR des jeweiligen SNPs multipliziert und die Ergebnisse addiert [79]. Hierdurch wird der kombinierte Effekt aller mit dem jeweiligen Merkmal assoziierten SNPs angegeben, gewichtet nach der jeweiligen Effektgröße eines jeden SNPs. Im nächsten Schritt wurde diese Summe durch die durchschnittliche logarithmierte OR aller Varianten dividiert. Dies erlaubt eine anschauliche Interpretation der Ergebnisse, da eine Erhöhung des resultierenden GRS um eine Einheit somit einem zusätzlichen Risikoallel von durchschnittlicher Stärke entspricht. Die Berechnung wird in folgender Formel zusammengefasst, wobei n die Anzahl an verwendeten Varianten angibt:

$$GRS = \frac{\sum_{i=1}^{n} A_n \times \ln(OR_n)}{\frac{1}{n} \sum_{i=1}^{n} \ln(OR_n)}$$

Nach der Berechnung wurde je GRS ein Histogramm erstellt, welches die Übereinstimmung mit einer Normalverteilung überprüft. Anschließend wurde die Korrelation der GRSs untereinander untersucht, um die Qualität der GRSs und der dafür verwendeten SNPs zu kontrollieren. Hierfür wurde die lineare Regression zwischen den einzelnen GRSs betrachtet und die P-Werte mithilfe der Falscherkennungsrate (*false discovery rate*, FDR) für mehrfaches Testen korrigiert [80]. Um die berechneten GRS Korrelationen darzustellen, wurde eine Heatmap mithilfe der Funktion heatmap2 der R Erweiterung gplots erstellt [81]. Die verwendete Farbpalette wurde hierbei unter Verwendung der R Erweiterung RColorBrewer erzeugt [82].

2.4 Expression Quantitative Trait Loci (eQTL) Berechnung

Mithilfe der bei der GRS Berechnung verwendeten SNPs wurde eine eQTL Analyse durchgeführt, um so den Einfluss dieser Varianten auf die Genexpression zu untersuchen. Dies sollte Rückschlüsse auf die Gene erlauben, die dem Zusammenhang zwischen den SNPs und den assoziierten komplexen Merkmalen ursächlich sind.

Zur eQTL Analyse wurden die Genotypdaten verwendet, welche bereits bei der GRS Berechnung genutzt wurden. Die Genexpressionsdaten der Gewebe Leber und Blut wurden der Datenbank des GTEx Projekts (release v7) entnommen [36]. Die Datenbank enthält hinsichtlich des Gewebes Leber die Genexpressionsinformationen 26072 Genen. von 131 Spendern zu Zum Gewebe Blut sind Genexpressionsinformationen von 323 Spendern zu 29151 Genen enthalten. Die Genexpressionsdaten wurden auf proteinkodierende Gene gefiltert, um nur Gene zu untersuchen, deren physiologische Relevanz anhand ihrer Proteine nachvollzogen werden kann. Somit verblieben zum Gewebe Leber 15939 proteinkodierende Gene und zum Gewebe Blut 16258 proteinkodierende Gene. Die Berechnung der eQTL erfolgte mithilfe der R Implementierung Matrix eQTL, welche die Genotyp- und Genexpressionsdaten mithilfe linearer Regression auf Assoziationen testete [83]. Angaben zu Geschlecht und Alter sowie die ersten fünf Hauptkomponenten der Genotyp-Hauptkomponentenanalyse wurden als Kovariaten verwendet. Es wurden trans eQTL berechnet, wobei nachfolgend die Entfernung der eQTL Varianten zum entsprechenden eQTL Gen untersucht wurde. eQTL, bei denen sich die eQTL Variante maximal 1 Mb upstream oder downstream zum zugehörigen Gen befand, wurden als cis eQTL definiert. Die eQTL Ergebnisse wurden anschließend auf den P-Wert 1e-05 gefiltert, um nicht signifikante sowie falsch positive Ergebnissen innerhalb der Listen zu adjustieren.

2.5 Gemeinsame Gene zwischen den Genlisten der Merkmale

Die nachfolgenden Analysen wurden jeweils separat mit den eQTL Ergebnissen der Gewebe Leber und Blut durchgeführt. Aus den auf den P-Wert 1e-05 gefilterten eQTL Ergebnissen jedes Phänotyps wurden die in ihrer Genexpression beeinflussten Gene extrahiert und bezüglich jedes Merkmals in einer Genliste aufgetragen. Hierdurch sollten Überschneidungen zwischen den Genlisten der Merkmale aufgedeckt werden, um so auf möglicherweise pleiotrop wirkende Gene zu schließen. Hierfür wurde zuerst eine Berechnung aller möglichen Merkmalspaare durchgeführt und anschließend die Genlisten der Merkmalspaare in einer 2 x 2 Kontingenztabelle gegenübergestellt. **Tabelle 3** zeigt eine Musterkontingenztabelle.

Tabelle 3: Musterkontingenztabelle zu den Überschneidungen der Genlisten.

Jede Tabelle zeigt die Überschneidungen zwischen den Genlisten zweier Merkmale eines Gewebes. "Ja" impliziert hierbei, dass die betrachteten Gene in der Genliste des jeweiligen Merkmals aufzufinden sind, wobei "Nein" angibt, dass dies nicht der Fall ist. Die Zelle "A" gibt somit die Anzahl an Genen an, die in den Genlisten beider Merkmale zu finden sind. In Zelle "B" wird die Anzahl an Genen aufgeführt, welche nur in der Genliste des ersten der beiden Merkmale enthalten sind. Zelle "C" zeigt die Anzahl an Genen, welche nur in der Genliste des zweiten der beiden Merkmale zu finden sind. Zelle "D" führt die Anzahl an Genen auf, welche in keinem der beiden betrachten Genlisten enthalten sind.

		Genliste Merkmal 2	
		Ja	Nein
Genliste	Ja	Α	В
1 1	Nein	С	D

In jeder Kontingenztabelle sind folgende Parameter aufgetragen: 1) Die Anzahl an gemeinsamen Genen zwischen den Genlisten der korrespondierenden Merkmale (entspricht **A** in **Tabelle 3**) 2) Die Anzahl an Genen, welche nur in der Genliste des ersten der beiden Merkmale enthalten sind (entspricht **B** in **Tabelle 3**) 3) Die Anzahl an Genen, welche nur in der Genliste des zweiten der beiden Merkmale enthalten sind (entspricht **C** in **Tabelle 3**) 4) Die Anzahl an Genen, welche auf keiner der Genlisten der beiden Merkmale enthalten sind (entspricht **C** in **Tabelle 3**) 4) Die Anzahl an Genen, welche auf keiner der Genlisten der beiden Merkmale enthalten sind (entspricht **D** in **Tabelle 3**). Die letzte Maßzahl ergibt sich hierbei aus der Gesamtanzahl an betrachteten Genen im entsprechenden Genexpressionsdatensatz, abzüglich der Anzahl an Genen, welche bereits bei Punkt 1 bis 3 genannt wurden.

Um die Kontingenztabelle jedes Merkmalpaares und damit die Überlappung zwischen den Genlisten auf Signifikanz zu prüfen, kam der exakte Test nach Fisher zur Anwendung [84]. Die P-Werte wurden mithilfe der FDR für mehrfaches Testen korrigiert. Zur Visualisierung wurde mithilfe des Programmes Gephi je Gewebe eine Abbildung erstellt, welche die Gemeinsamkeiten zwischen den Genlisten der Merkmale aufzeigt [70]. Zudem wurde zu jedem der beiden Gewebe unter Einsatz der R Erweiterung circlize ein Circosplot generiert, um die Anzahl an sich überschneidenden Genen der einzelnen Merkmale zu jedem der beiden Gewebe grafisch darzustellen [85].

Bezüglich jedes gemeinsamen Gens zwischen den Genlisten wurde die Ursache für die Gemeinsamkeit zwischen den Genlisten untersucht. Dabei wurden zwei Fälle gemeinsamer Gene unterschieden: Einerseits Gene, deren Gemeinsamkeit zwischen den Genlisten durch die Effekte auf die Genexpression durch denselben SNP zu erklären waren und andererseits Gene, bei denen der Einfluss auf die Genexpression durch unterschiedliche SNPs erfolgte. Diese Unterscheidung wurde anhand der zuvor berechneten eQTL und den zugehörigen SNPs sowie Genen durchgeführt. Die zugehörigen SNPs wurden anschließend mithilfe der Webanwendung LDLink auf ihr LD geprüft [74]. **Abbildung 4** zeigt eine Übersicht der genannten Schritte.

Zur weiteren Untersuchung der gemeinsamen Gene erfolgte eine Literaturrecherche, um den gegenwärtigen Wissensstand zu jedem Gen zu eruieren (Stand September 2019). Hierfür erfolgte eine Literaturrecherche mithilfe der PubMed Gene Database [77]. Die dortige Anzahl an mit dem Gen assoziierten PubMed Artikeln wurde als Maßzahl des Wissensstandes zu jedem Gen verwendet. Hierbei gab der Befehl "PubMed Links for Gene (Select x)" der PubMed Advanced Search die mit der Gene Database zum jeweiligen Gen verlinkten PubMed Artikel aus, wobei mit x die NCBI Gene ID des entsprechenden Gens angegeben wurde. Die verlinkten Publikationen wurden anschließend auf Zusammenhänge zwischen den jeweiligen Genen und Phänotypen durchsucht.



Abbildung 4: Arbeitsablauf basierend auf den GTEx Genexpressions- und Genotypdaten.

Ovale Figuren repräsentieren verwendete und erzeugte Datensätze, rechteckige Figuren repräsentieren Analysen oder Berechnungen. Die GTEx Genotypdaten wurden auf die GWAS SNPs und europäische Population gefiltert und hieraus die GRSs sowie deren Korrelationen untereinander berechnet. Mithilfe der auf proteinkodierende Gene gefilterten GTEx Genexpressions- und Genotypdaten wurden die eQTL zu den GWAS SNPs berechnet. Aus den eQTL Ergebnissen wurden nach einer Signifikanzprüfung (p < 1e-05) die eQTL Gene extrahiert und in einem 1 zu 1 Vergleich auf gemeinsame Gene zwischen den Merkmalen untersucht. Die gemeinsamen Gene wurden anschließend analysiert und einer von zwei Gruppen zugeordnet: 1) Gemeinsame Gene, deren Gemeinsamkeit zwischen den Genlisten auf anderen Effekten beruht.

3 Ergebnisse

3.1 Literaturrecherche

Zunächst wurde im Rahmen einer Literaturrecherche nach GWAS-Datensätzen gesucht, um aus diesen summary statistics Daten von SNPs zu entnehmen. Die GWAS sowie die daraus entnommenen SNP-Daten wurden einer strengen Qualitätskontrolle unterzogen, wobei die meisten Studien den angelegten Kriterien nicht entsprachen. In vielen GWAS wurden keine Effektgrößen oder Signifikanzwerte zu den SNP Assoziationen angegeben, oder die Informationen zeigten sich widersprüchlich. In manchen Publikationen existierten zudem keine detaillierten Informationen zu den SNP Allelen: So wurde teilweise das risikoerhöhende Allel nicht eindeutig genannt oder die Allelinformationen fehlten vollständig. Da die Haplotypenstruktur zwischen Ethnien stark schwankt, wurden für diese Arbeit zudem nur Ergebnisse von GWAS verwendet, welche entweder anhand europäischer Populationen oder gemischter Population mit hohem europäischem Anteil an Probanden durchgeführt wurden. Da zum X-Chromosom keine Genotypdaten im GTEx Datensatz vorhanden waren, wurden auf dem X-Chromosom lokalisierte SNPs ebenfalls ausgeschlossen. Weiterhin wurden die SNPs eines Merkmals auf LD geprüft und SNPs, die in LD zueinanderstehen ($r^2 > 0.5$) aus der Studie ausgeschlossen **Tabelle 4** zeigt eine Übersicht zu den für die weitere Analyse verwendeten Merkmalen und SNPs. Eine detaillierte Tabelle mit Quellen, den Gründen für den Ausschluss einzelner SNPs sowie der Anzahl ausgeschlossener SNPs befindet sich im Anhang (Anhangtabelle 1). Insgesamt wurden in den weiteren Analysen 870 SNPs verwendet, wobei die Anzahl pro Merkmal zwischen 4 (Alanin-Aminotransferase, ALT) und 154 (Koronare Herzkrankheit, CAD) schwankte.

Tabelle 4: Übersicht über verwendete Merkmale und SNPs

Die Tabelle zeigt die in der Arbeit untersuchten Erkrankungen/Merkmale inklusive ihrer zugehörigen Gruppe sowie die Anzahl an SNPs, welche bezüglich jeder Erkrankung/Merkmal zur weiteren Analyse verwendet wurden. Um in die weitere Analyse eingeschlossen zu werden, mussten die SNPs folgende Kriterien erfüllen: 1) Vorhandene Angaben zu P-Werten der Assoziationen der jeweiligen SNPs mit dem Merkmal 2) Vorhandene Informationen zu den jeweiligen Effektgrößen und Effektallelen 3) GWS (p < 5e-08) entweder in europäischer Bevölkerung oder in gemischter Bevölkerung mit mehrheitlich europäischen Probanden 4) Nicht auf dem X-Chromosom gelegen 5) Unabhängigkeit bzw. geringes LD in Bezug auf die anderen verwendeten SNPs (r² < 0,5).

Gruppe Erkrankung/Merkmal		Abkürzung	Verwendete SNPs
AMD Altersabhängige		AMD	44
	Makuladegeneration		
Leberenzyme	Alkalische	ALP	14
-	Phosphatase		
	Alanin-	ALT	4
	Aminotransferase		
	Gamma-	GGT	25
	Glutamyltransferase		
	Gesamtleberenzyme	LEP	42
im Plasma			
Kardiovaskuläre	Koronare	CAD	154
Merkmale	Herzkrankheit		
	Diastolischer	DBP	94
	Blutdruck		
	Systolischer Blutdruck	SBP	81
	Blutdruck - Allgemein	GBP	29
	Blutdruckamplitude	PP	42
	Hypertonie	HTN	11
Blutlipide	Lipoproteine niederer	poproteine niederer LDL 76	
	Dichte/ Low Density		
	Lipoprotein		
	Lipoproteine hoher	HDL	99
	Dichte/		
	High Density		
	Lipoprotein		
	Gesamtcholesterin	TC	89
Triglyceride		TG	66

3.2 Genetische Risiko Score (GRS) Berechnung

Unter Einsatz der mit den Merkmalen assoziierten SNPs wurden GRSs berechnet, um potenzielle genetische Korrelationen zwischen den einzelnen Merkmalen zu analysieren. Dazu wurden die Genotypdaten von 635 Individuen des GTEx Projekts (release v7) verwendet [36]. Die verwendeten Daten wurden den nachfolgend beschriebenen Verarbeitungsschritten sowie einer Qualitätskontrolle unterzogen. Hierbei wurden die Allelfrequenzen der SNPs berechnet und sowohl vor als auch nach Ausrichtung auf das risikoerhöhende Allel den in den GWAS publizierten Allelfrequenzen mithilfe von Streudiagrammen gegenübergestellt. Im Rahmen der GRS Berechnung wurden die Risikoallele nach ihrer Effektgröße gewichtet und addiert. Das Ergebnis wurde daraufhin durch die durchschnittliche Effektgröße aller SNPs des jeweiligen GRS dividiert. Eine Erhöhung des GRS um eine Einheit entspricht somit der Risikoerhöhung durch ein Risikoallel von durchschnittlicher Stärke. Als weiterer Kontrollschritt wurden die Verteilungen der GRSs mithilfe von Histogrammen dargestellt. Die einzelnen Plots, welche zur Qualitätskontrolle verwendet wurden, sind am Beispiel von CAD in **Abbildung 5** dargestellt.



Abbildung 5: Grafiken zur Qualitätskontrolle am Beispiel des Merkmals CAD.

a) Vergleich der berechneten Allelfrequenzen der einzelnen SNPs mit den Allelfrequenzen aus den GWAS vor Ausrichtung auf die risikoerhöhenden Allele. Jeder Punkt markiert einen SNP. Die x-Achse gibt die in den jeweiligen GWAS publizierten Allelfrequenzen an, die y-Achse die anhand des GTEx Datensatzes berechneten Allelfrequenzen.

c) Histogramm der GRS Verteilung. Die x-Achse gibt die Höhe der GRSs an, die y-Achse die relative Häufigkeit der verschiedenen GRS Werte. Die gestrichelte rote Linie gibt den Mittelwert des GRSs des jeweiligen Merkmals an.

Die Histogramme der GRS Verteilungen der einzelnen Erkrankungen/Merkmale zeigten hierbei, mit Ausnahme des GRS von ALT, eine Normalverteilung. Da die GRSs einiger Merkmale teilweise auf identischen SNPs basierten und manche Merkmale zudem physiologisch eng verwandt sind, wurden Korrelationen zwischen den GRSs gewisser Merkmale erwartet, wie beispielsweise zwischen den Merkmalen der kardiovaskulären Merkmalsgruppe. Um diese Annahmen zu bestätigen und somit die Qualität der verwendeten SNPs sowie der angewandten Methodik sicherzustellen, wurden die Korrelationen der einzelnen GRSs berechnet und in einer Heatmap (**Abbildung 6**) dargestellt.

b) Vergleich der berechneten Allelfrequenzen der einzelnen SNPs mit den Allelfrequenzen aus den GWAS nach Ausrichtung auf die risikoerhöhenden Allele. Jeder Punkt markiert einen SNP. Die x-Achse gibt die in den jeweiligen GWAS publizierten Allelfrequenzen an, die y-Achse die anhand des GTEx Datensatzes berechneten Allelfrequenzen.



Abbildung 6: Korrelationen zwischen den GRSs der verwendeten Merkmale.

Die GRSs der einzelnen Merkmale wurden auf Korrelationen getestet und die Ergebnisse in einer Heatmap wiedergegeben. Merkmale, deren GRSs Korrelationen aufwiesen, wurden in Gruppen angeordnet. Im rechten unteren Teil der durch die dunkelrote Diagonale getrennten Heatmap finden sich die Korrelationskoeffizienten. Ein blauer Hintergrund verdeutlicht eine negative, ein roter Hintergrund eine positive Korrelation zwischen den entsprechenden Paaren der GRSs. Im linken oberen Teil der Heatmap sind die nach FDR korrigierten P-Werte aufgetragen (*: FDR<0.001, **: FDR<0.0001, ns: nicht signifikant).

Merkmale, deren GRSs Korrelationen aufwiesen, wurden in Gruppen angeordnet. So ist in **Abbildung 6** in der oberen rechten Ecke eine Gruppierung der kardiovaskulären Merkmale allgemeiner Blutdruck (GBP), Bluthochdruck (HTN), systolischer Blutdruck (SBP), diastolischer Blutdruck (DBP), Blutdruckamplitude (PP) und CAD zu erkennen. Zu sehen sind starke Korrelationen zwischen GBP, HTN, SBP und DBP, mit FDR Werten < 0.0001. CAD wies eine schwache Korrelation mit DBP auf ($r^2 = 0.08$, FDR< 0.01), sowie eine Korrelation mit PP ($r^2 = 0.06$, FDR<0.01).

Auch zwischen Blutlipidmerkmalen bildeten sich entsprechende Gruppierungen, wie im mittleren Teil von **Abbildung 6** zu sehen ist. So ließ sich eine starke Korrelation ($r^2 = 0.92$; FDR < 0.0001) zwischen den Blutlipidmerkmalen Lipoproteine niederer Dichte

(LDL) und Gesamtcholesterin (TC) beobachten. Für Triglyceride (TG) und Lipoproteine hoher Dichte (HDL) ergab sich ein negativer Zusammenhang ($r^2 = -0.26$) bei sehr hoher Signifikanz (FDR < 0.0001). Auch TG und LDL ($r^2 = 0.17$) sowie TG und TC (r^2 = 0.2) zeigten hochsignifikante Korrelationen (FDR < 0.0001). Ein negativer Zusammenhang ($r^2 = -0.11$) bei geringer Signifikanz (FDR < 0.01) war zwischen LDL und HDL zu finden.

Zwischen den Gesamtleberenzymen im Plasma (LEP) und Gamma-Glutamyltransferase (GGT) ($r^2 = 1.13$), LEP und Alkalischer Phosphatase (ALP) ($r^2=0.71$) sowie LEP und ALT ($r^2=0.76$) fanden sich stark positive Korrelationen, welche hochsignifikant ausfielen (FDR<0.0001).

Die Blutlipide LDL ($r^2 = 0.16$) sowie TC ($r^2 = 0.14$) zeigten sich mit dem kardiovaskulären Merkmal CAD bei sehr hoher Signifikanz (FDR<0.0001) positiv korreliert. Zwischen TC und dem Leberenzym ALP ließ sich eine negative Korrelation ($r^2 = -0.49$) aufzeigen, ebenso zwischen LDL und ALP ($r^2 = -0.39$), wobei beide Fälle eine sehr hohe Signifikanz aufwiesen (FDR < 0.0001). Das Merkmal AMD zeigte eine positive Korrelation ($r^2 = 0.13$) mit HDL, wobei diese Korrelation das Signifikanzniveau nicht erfüllte.

3.3 Expression Quantitative Trait Loci (eQTL) Berechnung

Unter Einsatz der bei der GRS Berechnung verwendeten SNPs wurde eine eQTL Analyse durchgeführt, um den Einfluss dieser SNPs auf das Genexpressionsniveau aufzuzeigen. Dabei wurden die SNPs mit den Genexpressionsdaten des jeweiligen Gewebes korreliert. Hierdurch sollten Gene identifiziert werden, die als Ursache des Zusammenhangs zwischen SNP und dem jeweiligen assoziierten Merkmal in Betracht kommen. Bei der eQTL Analyse wurden die Genotypdaten der GRS Berechnung sowie die Genexpressionsdaten aus der GTEx Datenbank verwendet. Die Genexpressionsdaten wurden auf proteinkodierende Gene gefiltert, um nur Gene zu untersuchen, deren physiologische Relevanz anhand ihrer Proteine nachvollzogen werden kann. Diese Berechnung wurde für die 15 Erkrankungen oder Merkmale (siehe **Tabelle 4**) anhand der Genexpressionsdaten der Gewebe Leber und Blut durchgeführt. Um nicht signifikante sowie falsch positive Ergebnisse innerhalb der Listen zu adjustieren, wurden alle Ergebnisse mit einem P-Wert größer 1e-03 verworfen. Da die Anzahl an signifikanten Resultaten mit diesem Grenzwert weiterhin hoch ausfiel (Leber: 14436 eQTL Ergebnisse; Blut: 15190 eQTL Ergebnisse), wurde der Grenzwert/die Stringenz auf einen P-Wert von 1e-05 weiter gesenkt. Die deutliche Reduktion an signifikanten eQTL durch die Veränderung der P-Wert Grenzen lässt sich anhand **Tabelle 5** nachvollziehen. Sie zeigt die Anzahl an signifikanten eQTL, aufgetragen nach Gewebeart, P-Wert Schwelle und Merkmal.

Zu einer P-Wert Schwelle von P < 1e-05 betrug die Summe an eQTL Ergebnissen bezüglich Lebergewebe 268, wobei die Werte für die einzelnen Merkmale ein Spektrum von 0 (ALP) bis 84 (AMD) aufwiesen. Anhand des Gewebes Blut wurden insgesamt 194 eQTL identifiziert, mit einer Bandbreite der Einzelergebnisse zwischen 0 (ALP) und 48 (CAD).

Tabelle 5: Anzahl der eQTL Ergebnisse der Gewebe und Merkmale anhand verschiedener P-Wert Grenzen Die Tabelle zeigt die Anzahl der als signifikant identifizierten eQTL, aufgetragen nach untersuchter Gewebeart, P-Wert Schwelle und betrachtetem Merkmal. Zur weiteren Analyse wurden die fett hervorgehobenen eQTL Ergebnisse mit P-Wert < 1e-05 verwendet.

	Leber		Blut	
	P < 1e-03	P < 1e-05	P < 1e-03	P < 1e-05
AMD	1773	84	1065	20
ALP	222	5	172	4
ALT	55	0	37	0
GGT	257	4	804	2
LEP	518	9	1000	6
CAD	3143	41	2618	48
DBP	1599	14	1385	19
SBP	885	10	1260	15
GBP	254	4	284	4
PP	376	12	454	4
HTN	88	2	94	1
LDL	1261	22	1245	14
HDL	1485	14	1562	19
TC	1275	29	1646	19
TG	1232	18	1564	19
Summe	14423	268	15190	194

3.4 Gemeinsamkeiten zwischen den Genlisten der Merkmale

Aus den auf den P-Wert < 1e-05 gefilterten eQTL Ergebnissen wurden die eQTL Gene extrahiert und bezüglich jedes Merkmals und Gewebes in einer Genliste aufgeführt. Die resultierenden Genlisten wurden anschließend miteinander verglichen. Dies geschah in der Absicht, Überschneidungen und somit Gemeinsamkeiten zwischen den Genlisten der Merkmale zu identifizieren und so möglicherweise pleiotrop wirkende Gene aufzudecken. Hierzu wurde zuerst eine Berechnung aller möglichen Merkmalspaare durchgeführt. Diese ergab 105 Kombinationen von jeweils 2 die Merkmalen, wobei Genlisten jeder dieser Merkmalspaare in einer Kontingenztabelle gegenübergestellt wurden. Hierdurch konnte die Anzahl von Genen ermittelt werden, die sich zwischen den jeweiligen Genlisten überschneiden. Nach einer Signifikanzprüfung der Kontingenztabellen beziehungsweise der Überlappung der Genlisten mittels dem exakten Test nach Fisher wurden die Ergebnisse mit weiteren Informationen annotiert und in Tabellen aufgetragen. Die Tabellen inklusive der Anzahl der Gene in der Genliste eines jeden Merkmals, der Anzahl an gemeinsamen Genen zwischen den Genlisten sowie P-Wert und FDR befinden sich zu jedem der beiden untersuchten Gewebe Leber und Blut im Anhang (Anhangtabelle 2 und Anhangtabelle 3). Für das Gewebe Leber ließen sich so 20 Merkmalspaare ermitteln, deren Genlisten mindestens ein gemeinsames Gen aufwiesen. Von den Überschneidungen der Genlisten dieser 20 Merkmalspaare erwiesen sich 16 als statistisch signifikant (FDR < 0.05). Anhand des Gewebes Blut konnten entsprechend 32 Merkmalspaare identifiziert werden, wovon 28 das Signifikanzniveau erfüllten. Abbildung 7 veranschaulicht die genannten Gemeinsamkeiten zwischen den Genlisten der Merkmale zu beiden Geweben.



Abbildung 7: Gemeinsamkeiten zwischen den Genlisten der Merkmale.

Die Abbildung veranschaulicht die Gemeinsamkeiten zwischen den auf den eQTL Ergebnissen beruhenden Genlisten der jeweiligen Merkmale. Hierbei zeigt die Teillabbildung a) die Gemeinsamkeiten bezüglich dem Gewebe Leber und Teilabbildung b) bezüglich dem Gewebe Blut. Zwei Merkmale sind durch eine Linie verbunden, falls die auf den eQTL Ergebnissen basierenden Genlisten der jeweiligen Merkmale mindestens ein gemeinsames Gen aufweisen. Eine gepunktete Linie verdeutlicht gemeinsame Gene zwischen den verbundenen Merkmalen, wobei die Signifikanzschwelle (FDR < 0.05) nicht erfüllt wurde. Verbindungslinien zwischen unterschiedlichen Merkmalsgruppen sind in der Mischfarbe der verbundenen Merkmalsgruppen erstellt. Im Gewebe Leber ließen sich Überschneidungen der Genlisten zwischen den Lipidmerkmalen TC und HDL, TC und LDL sowie LDL und HDL beobachten, wobei sich letztgenannte Beziehung als nicht signifikant erwies. Für HDL ergab sich außerdem eine Überschneidung der zugehörigen Genliste mit AMD. Die kardiovaskulären Merkmalspaare HTN und DBP, HTN und SBP, HTN und GBP, DBP und SBP, DBP und GBP, sowie GBP und SBP wiesen ebenfalls signifikante Gemeinsamkeiten in ihren Genlisten auf. Zwischen den Genlisten von CAD und den Lipidmerkmalen LDL und TC wurden ebenfalls gemeinsame Gene gefunden, sowie zwischen CAD und den Leberenzymen LEP und ALP, wobei die beiden letztgenannten Ergebnisse die Signifikanzschwelle nicht erfüllten. Die Genlisten von TC und LDL auf.

Bezüglich des Gewebes Blut wurden signifikante Überschneidungen der Genlisten der Lipidmerkmale HDL, TC und TG jeweils mit der AMD Genliste festgestellt. Weiterhin konnten signifikante Überschneidungen zwischen den Genlisten der folgenden Merkmalspaare beobachtet werden: 1) HDL jeweils mit TG, TC und LDL 2) TG mit TC und LDL 3) TC mit LDL. Die Genlisten aller 4 Lipidmerkmale überschnitten sich zudem signifikant mit denen der Lebermerkmale LEP und ALP. Bei den kardiovaskulären Merkmalen fanden sich Gemeinsamkeiten in den Genlisten folgender Merkmalspaare: HTN - SBP, HTN -DBP, HTN - GBP, CAD - DBP, CAD - GBP, CAD - PP, CAD - SBP, SBP - DBP, SBP - GBP, sowie GBP - DBP. Diese Überschneidungen erwiesen sich mit Ausnahme der Beziehungen von CAD mit DBP sowie CAD mit SBP als statistisch signifikant.

3.5 Gemeinsame Gene zwischen den Genlisten der Merkmale

Um die genannten Gemeinsamkeiten zwischen den Genlisten weiter zu spezifizieren und insbesondere die Anzahl an sich überschneidenden Genen der einzelnen Merkmale grafisch darzustellen, wurde bezüglich jedes der beiden Gewebe ein Circosplot generiert (siehe **Abbildung 8**).


Abbildung 8: Circosplot der gemeinsamen Gene zwischen den Genlisten der Merkmale.

Die Abbildung veranschaulicht die gemeinsamen Gene zwischen den auf den eQTL Ergebnissen beruhenden Genlisten der jeweiligen Merkmale. Teilabbildung a) zeigt die Ergebnisse bezüglich des Gewebes Leber, Teilabbildung b) bezüglich des Gewebes Blut. Jede Verbindungskurve repräsentiert ein gemeinsames Gen zwischen den Genlisten der Merkmale. Die Segmentgrößen der Merkmale entsprechen dem Anteil der eQTL Ergebnisse jedes Merkmals an der Gesamtzahl der eQTL Ergebnisse des jeweiligen Gewebes. Es wurden alle Gene aufgetragen, die zu mehr als einem Merkmal ein signifikantes eQTL Ergebnis zeigten. Verbindungskurven zwischen unterschiedlichen Merkmalsgruppen sind in der Mischfarbe der verbundenen Merkmalsgruppen erstellt. Anhand des Lebergewebes konnten bei einer Gesamtzahl von 268 eQTL mit den 268 resultierenden eQTL Genen insgesamt 65 Überschneidungen zwischen den eQTL Genlisten der untersuchten Merkmale/Erkrankungen identifiziert werden. Beim Lebergewebe entfiel auf das Merkmal AMD circa ein Drittel aller eQTL Ergebnisse (84 von 268), jedoch nur ein gemeinsames Gen von insgesamt 65. Im Gegensatz hierzu ergaben sich für die Leberenzymmerkmalsgruppe eine Gesamtzahl von 18 eQTL Ergebnissen. Es fanden sich jedoch 9 gemeinsame Gene zwischen den Genlisten der einzelnen Leberenzymmerkmale sowie 10 gemeinsame Gene mit Genlisten der anderen Merkmalsgruppen. Für die kardiovaskulären Merkmale resultierten 83 eQTL. Von insgesamt 30 gemeinsamen Genen der kardiovaskulären Genlisten entfielen 20 auf gemeinsame Gene zwischen den Genlisten der einzelnen kardiovaskulären Merkmale und 10 auf gemeinsame Gene mit Genlisten der anderen Merkmalsgruppen. Insbesondere in der Genliste von CAD überschnitten sich 8 Gene mit Genlisten der Blutlipidmerkmale (4 mit LDL, 4 mit TC). Ein gemeinsames Gen zeigte sich je zwischen CAD und den Leberenzymmerkmalen ALP und LEP. Für die Blutlipidmerkmale ließen sich neben den bereits genannten Gemeinsamkeiten noch 17 gemeinsame Gene zwischen den Genlisten der Lipidmerkmalsgruppe identifizieren (13 zwischen LDL und TC, 3 zwischen HDL und TC, 1 zwischen HDL und LDL). Die Gesamtzahl an eQTL der Lipidmerkmalsgruppe betrug 83.

Anhand des Gewebes Blut konnten insgesamt 194 signifikante eQTL mit zugehörigen Genen nachgewiesen werden, wobei 76 Überschneidungen zwischen den Genlisten gefunden wurden. AMD wies bei 20 eQTL Ergebnissen 6 gemeinsame Gene mit Genlisten der Blutlipidmerkmale auf (je 2 mit TG, TC und HDL). Bezüglich den Leberenzymen wurden 12 signifikante eQTL gefunden. Bei den zugehörigen Genen konnten 14 Überschneidungen mit den Genlisten anderer Merkmale festgestellt werden. Hiervon repräsentierten 6 Verbindungen die Gemeinsamkeiten zwischen den Genlisten innerhalb der einzelnen Leberenzymmerkmale und 8 Verbindungen beruhten auf Gemeinsamkeiten mit den Genlisten von Blutlipidmerkmalen (je 2 von LDL, HDL, TC und TG). Für die kardiovaskulären Merkmale ergab sich eine Gesamtzahl von 91 eQTL und 30 gemeinsamen Genen. 27 dieser Gene waren Gemeinsamkeiten zwischen den Genlisten der kardiovaskulären Merkmals HDL und den Genlisten der Merkmale GBP, SBP und DBP. Aus den Berechnungen der

Merkmalsgruppe der Blutlipide resultierten 71 signifikante eQTL und 43 gemeinsame Gene. 26 dieser Gene waren gemeinsame Gene zwischen den Genlisten der einzelnen Lipidmerkmale, der Rest entfiel auf Überschneidungen mit den Genlisten der anderen Merkmale. Diese Überschneidungen wurden bereits bei der Erläuterung der entsprechenden Merkmale beschrieben.

Im Vergleich der beiden Circosplots zeigten sich einige Auffälligkeiten. So war im Gewebe Leber der Anteil der AMD eQTL Ergebnisse an der gesamten eQTL Anzahl deutlich größer als im Gewebe Blut (Leber: 84 von insgesamt 268 eQTL entspricht einem Anteil von circa 31 %; Blut: 20 von insgesamt 194 eQTL, entspricht einem Anteil von circa 10 %). Die Zahl an gemeinsamen Genen von AMD mit anderen Merkmalen verhielt sich jedoch gegenteilig hierzu (Leber: 1 gemeinsames Gen, Blut: 6 gemeinsame Gene). Für kardiovaskuläre Merkmale wurden anhand des Gewebes Leber 83 eQTL identifiziert. Dies entsprach einem Anteil von circa 30 % an der Gesamtanzahl. Im Gewebe Blut betrug dieser Anteil ungefähr 47 % (91/194). Anhand der Blutlipidmerkmale ließen sich mit circa 31 % (Leber) beziehungsweise 36 % (Blut) ähnliche Anteilsverhältnisse beobachten. Auch bei den Leberenzymen glichen sich die Werte (6 % bei Leber, 6 % bei Blut).

Die gemeinsamen Gene zwischen den Genlisten wurden anschließend hinsichtlich der Ursache für die Gemeinsamkeit zwischen den Genlisten untersucht. Hierbei wurden zwei Fälle unterschieden. Einerseits Gene, deren Gemeinsamkeit zwischen den Genlisten durch die Effekte auf die Genexpression desselben SNPs zu erklären sind und andererseits gemeinsame Gene, deren Effekt auf Genexpression durch unterschiedliche SNPs verursacht wurde. Um diese Unterscheidung zu ermöglichen, wurden hinsichtlich jedes gemeinsamen Gens die zugrundeliegenden eQTL untersucht. Die jeweiligen SNPs der Gene, deren Gemeinsamkeit zwischen den Genlisten nicht durch die Effekte auf die Genexpression desselben SNPs zu erklären sind, wurden anschließend auf ihr LD ($r^2 > 0.5$) geprüft.

Von den 65 Genüberschneidungen bezüglich des Gewebes Leber konnten 48 durch gemeinsame SNPs begründet werden. Die restlichen 17 gemeinsamen Gene basierten nicht auf gemeinsamen SNPs. Die SNPs dieser restlichen 17 gemeinsamen Gene standen bis auf eine Ausnahme (AMD/HDL; $r^2 = 0.0004$) in LD zueinander. Die

r² Werte der in LD stehenden SNPs wiesen hierbei eine Spannweite zwischen 0,83 und 1 auf. **Anhangtabelle 4** zeigt bezüglich des Gewebes Leber die gemeinsamen Gene, deren Gemeinsamkeit zwischen den Genlisten durch die Effekte auf die Genexpression desselben SNPs zu erklären sind. In **Anhangtabelle 5** sind zum Gewebe Leber die Gene, deren Gemeinsamkeit zwischen den Genlisten nicht durch Effekte auf die Genexpression desselben SNPs zu erklären sind, aufgetragen.

Die Anzahl an Genen, deren Gemeinsamkeit zwischen den Genlisten durch die Effekte auf die Genexpression desselben SNPs zu erklären sind, variierte zwischen 13 (Merkmalspaar LDL/TC) und 0 (Merkmalspaare CAD/LDL, CAD/TC, AMD/HDL). Die Anzahl an gemeinsamen Genen des Gewebes Leber, deren Gemeinsamkeit zwischen den Genlisten nicht durch die Effekte auf die Genexpression desselben SNPs zu erklären sind, wies eine Spannweite von 0 (Merkmalspaare LDL/TC, ALP/LEP, DBP/SBP, GGT/LEP, GBP/HTN, HDL/TC, ALP/CAD, HDL/LDL, CAD/LEP) bis 4 Genen auf (Merkmalspaare CAD/LDL, CAD/TC). **Anhangtabelle 6** gibt bezüglich des Gewebes Leber die Merkmalspaare, zu welchen gemeinsame Gene zwischen den Genlisten identifiziert wurden, sowie die Anzahl an gemeinsamen Genen wieder.

Bezüglich des Gewebes Blut basierten von der Gesamtzahl an 76 Genüberschneidungen 60 Überschneidungen auf übereinstimmenden SNPs, während 16 nicht durch gemeinsame SNPs begründbar waren. Alle SNPs dieser 16 Gene standen in LD zueinander, wobei sich die zugehörigen r² Werte zwischen 0,82 und 0,96 erstreckten.

Anhangtabelle 7 zeigt bezüglich des Gewebes Blut die gemeinsamen Gene, deren Gemeinsamkeit zwischen den Genlisten durch die Effekte auf die Genexpression desselben SNPs zu erklären sind. In **Anhangtabelle 8** sind zum Gewebe Blut die Gene, deren Gemeinsamkeit zwischen den Genlisten nicht durch Effekte auf die Genexpression desselben SNPs zu erklären sind, aufgetragen.

Die Anzahl gemeinsamer Gene, deren Gemeinsamkeit zwischen den Genlisten durch die Effekte auf die Genexpression desselben SNPs zu erklären sind, reichte bezüglich des Gewebes Blut von 0 (Merkmalspaare CAD/DBP, CAD/PP, TG/LEP, TC/LEP, HDL/LEP, LDL/LEP, ALP/TG, ALP/TC, ALP/HDL, ALP/LDL, AMD/TG, AMD/TC,

AMD/HDL) bis 14 (Merkmalspaar DBP/SBP). Die Anzahl an gemeinsamen Genen des Gewebes Blut, deren Gemeinsamkeit zwischen den Genlisten nicht durch die Effekte auf die Genexpression desselben SNPs zu erklären sind, variierte zwischen 0 und 2 Genen (Merkmalspaare AMD/HDL, AMD/TC, AMD/TG). Zum Gewebe Blut sind die Merkmalspaare, zu welchen gemeinsame Gene zwischen den Genlisten identifiziert wurden, sowie die Anzahl an gemeinsamen Genen in **Anhangtabelle 9** zu finden.

Insgesamt konnten 65 Gene identifiziert werden, welche sich in den Genlisten von mindestens zwei Merkmalen fanden und somit vermutlich pleiotrop wirken. Von diesen 65 Genen fanden sich 57 in den Genlisten von Merkmalen aus derselben Merkmalsgruppe und 8 in den Genlisten von Merkmalen aus verschiedenen Merkmalsgruppen.

Um den aktuellen Wissensstand zu jedem der 65 vermutlich pleiotropen Gene zu quantifizieren und die bei dieser Arbeit generierten Ergebnisse zu validieren, erfolgte eine Literaturrecherche mithilfe der PubMed Gene Database. Hierbei wurde die Anzahl an assoziierten PubMed Publikationen bezüglich jedes Gens eruiert und nach Zusammenhängen zwischen bekannten Genen und Phänotypen gesucht (Anhangtabelle 10). Die Anzahl an assoziierten PubMed Publikationen zu den einzelnen Genen reichte von 3 bis 416 bei einem Median von 30. Auch die Anzahl an Phänotypen, welche laut Recherche einen Zusammenhang mit den Genen zeigen, wies eine hohe Variabilität auf. Während zu einigen Genen keine bekannten Assoziationen mit Phänotypen gefunden werden konnten, ergaben sich für andere Gene bis zu 30 bekannte Zusammenhänge. Die meisten Veröffentlichungen nutzten der Assoziationen korrelative oder auch zur Ermittlung bioinformatische Studienmodelle, wie beispielsweise GWAS oder eQTL Analysen. Durch die Literaturrecherche konnten zudem einige der bei dieser Arbeit gefundenen Zusammenhänge zwischen Genen und Merkmalen direkt bestätigen werden.

So steht das Gen proline and serine rich coiled-coil 1 (*PSRC1*) laut mehreren Publikationen mit CAD [86–89] sowie LDL [87,90,91] in Relation. Bezüglich des Gens sortilin 1 (*SORT1*) konnten mehrere Publikationen ausfindig gemacht werden, die sowohl eine Assoziation mit dem kardiovaskulären Merkmal CAD [89,92] als auch dem Blutlipidmerkmal LDL [89,93–96] sowie TC [89,97] aufzeigen. Auch für das Gen

cadherin EGF LAG seven-pass G-type receptor 2 (*CELSR2*) ergab die Literaturrecherche Assoziationen mit dem kardiovaskulären Merkmal CAD [98,99] sowie den Blutlipidmerkmalen LDL [93,100,101] und TC [100–102]. Weitere Veröffentlichungen diskutieren den Zusammenhang der Gene ST3 beta-galactoside alpha-2,3-sialyltransferase 4 (*ST3GAL4*) und dynein axonemal heavy chain 11 (*DNAH11*) sowie fatty acid desaturase 2 (*FADS2*) mit den Blutlipidmerkmalen LDL [97,103–105] und TC [97,103,104], wobei *FADS2* zusätzlich mit HDL und TG [95,97,104] in Verbindung gebracht wird. Mehrere Publikationen stellen zudem einen Zusammenhang des Gens unc-51 like kinase 4 (*ULK4*) mit dem Merkmal DBP [106–109] sowie dem allgemeinen Blutdruck und Hypertonie dar [110,111].

4 Diskussion

Die Genetik komplexer Erkrankungen beziehungsweise komplexer Merkmale wird durch eine Vielzahl meist unbekannter Gene beeinflusst, die zumeist mit zahlreichen Umweltfaktoren interagieren [4]. Studien konnten bereits genetische Zusammenhänge zwischen vielen komplexen Erkrankungen aufzeigen, doch die Ursachen dieser Korrelationen sind größtenteils noch ungeklärt [69]. In dieser Arbeit wird eine Vorgehensweise präsentiert, die mithilfe von eQTL Analysen potenziell pleiotrop wirkende Gene identifiziert. Dies kann dazu beitragen, die Lücke zwischen der genetischen Korrelation komplexer Merkmale und den zugrundeliegenden physiologischen Mechanismen zu schließen.

Hierzu wurden im Rahmen einer Literatursuche genetische Varianten ermittelt, die laut GWAS genomweit signifikante Assoziationen zu ausgewählten komplexen Merkmalen aufweisen. Durch eine nachfolgende eQTL Analyse konnten Korrelationen zwischen diesen genetischen Varianten und der Expression von Genen untersucht werden. Dies wiederum erlaubt Rückschlüsse auf Gene, die dem Zusammenhang zwischen den Varianten und den assoziierten komplexen Merkmalen ursächlich sind. Die identifizierten eQTL Gene wurden für jedes untersuchte Merkmal in einer Genliste aufgeführt und die einzelnen Genlisten untereinander verglichen. So konnten möglicherweise pleiotrop wirkende Gene ausfindig gemacht werden, deren Genexpression durch mit komplexen Merkmalen assoziierten genetischen Varianten beeinflusst wird.

Die pleiotropen Gene und die hieraus resultierenden Proteine wurden verschiedenen Signalwegen und physiologischen Vorgängen zugeordnet, um so Rückschlüsse auf Mechanismen zu gewinnen, die den komplexen Erkrankungen oder Merkmalen zugrunde liegen könnten. Auch Pathomechanismen, die mehrere komplexe Erkrankungen oder Merkmale zugleich beeinflussen, sind so identifizierbar. Da dies einen neuen und unerprobten Ansatz darstellt und ein Beweis der Zuverlässigkeit noch aussteht, wurden komplexe Merkmale beziehungsweise Erkrankungen zur Untersuchung ausgewählt, deren Korrelationen untereinander zum Teil bereits bekannt sind. So wurden neben AMD verschiedene Blutlipide, Leberenzyme und kardiovaskuläre Merkmale betrachtet.

Um die Qualität der aus den GWAS entnommenen SNPs sicherzustellen, wurde neben einer strengen Qualitätskontrolle der verwendeten GWAS-Daten zuerst eine Berechnung von GRSs durchgeführt und deren Korrelationen untersucht. Nach der Berechnung wurde mithilfe von Histogrammen die Verteilung der GRSs dargestellt, wobei sich bei allen GRSs mit Ausnahme von ALT eine Normalverteilung zeigte. Hierbei ist anzumerken, dass zu ALT mit 4 SNPs die geringste Zahl von SNPs aller Merkmale zur Berechnung des GRSs genutzt wurde. Die Abweichung von der Normalverteilung lässt sich somit als Konsequenz der niedrigen Anzahl verwendeter SNPs für dieses Merkmal erklären, da dies eine höhere Ungenauigkeit in der Trennschärfe der GRSs vermuten lässt.

Um in dieser Arbeit einen bewährten Ansatz zur GRS Berechnung zu verwenden, wurden die GRSs nach der Methodik von Grassmann et al. (2017) erstellt, wobei die GRSs von dieser Forschungsgruppe zur Untersuchung der genetischen Assoziationen zwischen AMD und zahlreichen komplexen Erkrankungen genutzt wurden. Die Studie von Grassmann et al. (2017) beschreibt ebenfalls einige der in der vorliegenden Arbeit identifizierten GRS Korrelationen [69]. Ein Beispiel ist die GRS Korrelation zwischen CAD aus der Gruppe der kardiovaskulären Merkmale und anderen in dieser Arbeit untersuchten Merkmalen. Zu CAD konnte in der vorliegenden Arbeit bei einem FDR Wert < 0.01 eine leicht positive Korrelation mit DBP sowie mit PP beobachtet werden. Auch mit LDL und TC korrelierte CAD bei einer FDR < 0.0001 positiv. Die genannten Korrelationen von CAD mit DBP, PP, LDL und TC zeigen sich in Übereinstimmung mit bekannten Erkenntnissen zur Pathogenese der CAD, welche eine Beteiligung von Blutdruckmerkmalen und verschiedenen Lipidspiegeln an dieser Erkrankung nahelegen, womit sich die Resultate physiologisch bekräftigen lassen [112–119].

Die Ergebnisse der vorliegenden Arbeit bezüglich CAD werden von den Resultaten von Grassmann et al. (2017) bestätigt, welche ebenfalls GRS Korrelationen von CAD mit DBP, LDL und TC zeigen konnten [69]. Die Forschungsgruppe konnte jedoch auch signifikante Zusammenhänge von Merkmalen finden, die in der vorliegenden Arbeit das Signifikanzniveau nicht erreichten. So zeigten sie beispielsweise GRS Korrelationen von CAD mit SBP, GBP, TG und HDL. Hier sollte jedoch bedacht werden, dass Grassmann et al. (2017) neben der Verwendung anderer SNPs weitaus größere Stichproben mit über 33.000 Individuen zur GRS Berechnung nutzten. Die

dadurch deutlich höhere statistische Teststärke kann somit potenzielle Unterschiede in der Anzahl signifikanter Ergebnisse zwischen der gegenwärtigen Arbeit und der genannten Publikation trotz gleicher GRS Berechnungsmethodik erklären.

Bei Betrachtung der in dieser Arbeit berechneten GRSs fielen zudem einige Merkmale auf, die unerwarteterweise nicht signifikant korrelierten. So ließen sich beispielsweise zwischen den Leberenzymmerkmalen LEP, GGT, ALP und ALT stark positive Korrelationen beobachten, wobei diese jedoch nur zwischen LEP und GGT, LEP und ALP sowie LEP und ALT signifikant (FDR < 0.0001) ausfielen. Da sich die Liste an SNPs des Merkmals LEP aus den SNPs der restlichen Leberenzyme zusammensetzt, war das Ergebnis bezüglich dieses einen Merkmals zu erwarten. Die fehlende Signifikanz der Korrelationen der anderen Leberenzyme ist jedoch überraschend und lässt zu niedrige SNP Zahlen im Rahmen der GRS Berechnung sowie zu geringe Probenzahlen bezüglich dieser Merkmale vermuten.

Insgesamt zeigte sich ein Großteil der komplexen Merkmale, für welche in der Literatur sowie von Grassmann et al. (2017) Zusammenhänge beschrieben werden, in der gegenwärtigen Arbeit ebenfalls korreliert. Dies bestätigt die Korrektheit der in dieser Arbeit verwendeten SNPs sowie der angewandten Methodik und kann somit als zusätzliche Qualitätskontrolle angesehen werden.

Neben der Funktion als Qualitätskontrolle sollten die GRSs ursprünglich dazu verwendet werden, Gene zu identifizieren, deren Genexpression mit der Höhe der GRSs in Zusammenhang steht. Hierzu wurde versucht, die Höhe der GRSs der einzelnen Individuen mit deren Genexpressionsniveau zu korrelieren. Mit diesem Vorgehen konnten jedoch keine signifikanten Ergebnisse (FDR < 0.05) erzielt werden. Der genaue Grund hierfür ist unklar, jedoch lässt sich eine nicht ausreichende Trennschärfe der verwendeten GRSs in Bezug auf die Genexpressionsniveaus vermuten. Um diese Trennschärfe zu erhöhen, wurde anschließend der Ansatz verfolgt, nur Individuen mit relativ hohem beziehungsweise relativ niedrigem GRS zu betrachten und die entsprechenden Individuen einer Hochrisiko- respektive Niedrigrisikogruppe zuzuweisen. Anschließend wurde der Zusammenhang zwischen Gruppenzugehörigkeit und Genexpression betrachtet, um so Gene zu finden, die zwischen den Gruppen einen Unterschied in ihrer Expression aufweisen.

Die Auswahl der Gruppengröße unterlag hierbei jedoch gewissen Einschränkungen. Große Hoch- und Niedrigrisikogruppenanteile zeigten den Vorteil größerer Probenzahlen und damit eine höhere statistische Teststärke, jedoch ging in entsprechendem Maße der beabsichtigte Effekt der größeren Trennschärfe verloren. Wurde der Anteil der Gruppen an der Gesamtanzahl der Individuen verringert, konnten genau gegenteilige Effekte beobachtet werden. Da der Genexpressionsdatensatz des Gewebes Leber 131 Spender umfasst, wurden beispielsweise bei einem Gruppengrößenanteil von 20 % nur 26 Individuen pro Gruppe betrachtet. So lässt sich begründen, wieso trotz Anpassungen der Gruppengröße keine signifikanten Ergebnisse (FDR < 0.05) erzielt werden konnten. Es kann somit festgehalten werden, dass die Korrelation der GRS Höhe mit dem Genexpressionsniveau der Individuen unter Einsatz der in dieser Arbeit verwendeten Datensätze und Methoden nicht erfolgreich schien und dieser Ansatz deswegen nicht weiterverfolgt wurde.

Aus diesem Grund wurde in der endgültigen Vorgehensweise dieser Arbeit anstatt der Korrelation der GRS Höhe mit der Genexpression schließlich eine eQTL Analyse durchgeführt, mithilfe derer die Effekte einzelner SNPs auf das Genexpressionsniveau untersucht wurden. Die eQTL Analyse wurde bezüglich 15 Merkmalen anhand der Genexpressionsdaten der Gewebe Leber und Blut etabliert und die Daten hierzu der GTEx Projekt Datenbank entnommen. Diese Datenbank wurde gewählt, da Genexpressionsdaten unterschiedlicher Gewebedatensätze im Regelfall nur nach vorheriger Normalisierung miteinander vergleichbar sind. Innerhalb der GTEx Datenbank entfällt dieses Problem, da bezüglich aller Gewebedatensätze dieselben Mess- und Prozessprotokolle angewendet wurden. Die eQTL Ergebnisse, die mithilfe der GTEx Genexpressionsdaten gewonnen werden, sind somit direkt untereinander hinsichtlich ihrer Effektgrößen vergleichbar. Die Auswahl der einzelnen Gewebe zur Entnahme der Genexpressionsdaten wurde im Hinblick auf die zu untersuchenden Merkmalsgruppen getroffen. Leber spielt eine zentrale Rolle im Lipidmetabolismus und wurde deshalb als geeignet erachtet, um die Expression von Genen zu beurteilen, welche mit den untersuchten Blutlipiden TC, HDL, LDL und TG in Zusammenhang stehen [120]. Weiterhin enthalten die Zellen des Lebergewebes, die sogenannten Hepatozyten, große Mengen der Enzyme ALP, ALT und GGT, welche deswegen im Rahmen der klinisch-diagnostischen Anwendung auch als Leberenzyme bezeichnet werden [121]. Diese Leberenzyme stellen ebenfalls eine der in dieser Arbeit untersuchten Merkmalsgruppen dar. Lebergewebe kann zudem auch mit AMD in Zusammenhang gebracht werden, da es sich als Hauptexpressionsort von Genen des Komplementsystems und des Lipidmetabolismus präsentiert. Diesen Prozessen wird ein relevanter Zusammenhang mit der AMD Pathogenese zugeschrieben [36,56,58-61,122]. Es wird angenommen, dass Genprodukte des Lipid und Komplementmetabolismus Bereich der Leber im beispielsweise als Komplementfaktoren in die Blutbahn abgegeben werden und so Effekte im Bereich der Retina und somit auf die AMD Pathogenese ausüben könnten [123].

Die Genexpressionsdaten des Gewebes Blut wurden zur eQTL Analyse ausgewählt, da Blut sowohl eine enge örtliche als auch physiologische Beziehung mit den kardiovaskulären Merkmalen, also den Blutdruckmerkmalen HTN, DBP, SBP, GBP und der vaskulären Erkrankung CAD aufweist. Weiterhin ist die Retina beziehungsweise die anliegende und die Retina versorgende Choriokapillaris stark durchblutet, wodurch sich Einflüsse auf den Krankheitsprozess der AMD vermuten lassen [124]. Bei der Choriokapillaris handelt es sich um eine Schicht der Choroidea, wobei Grunwald et al. (2005) beispielsweise eine Assoziation zwischen einer verringerten durchschnittlichen Durchblutung der Choroidea mit einem erhöhten Risiko zur Entwicklung einer feuchten AMD nachweisen konnten [125]. Da die Höhe des arteriellen Blutdrucks nachweislich die choroidale Durchblutung beeinflusst, lassen sich so Korrelationen zwischen den Blutdruckmerkmalen und der AMD Pathogenese vermuten [126]. Da zur eQTL Analyse zudem SNPs verwendet werden, welche sich mit den Blutplasmaspiegeln der Leberenzyme und der Lipidgruppe assoziiert zeigten, können somit mögliche Zusammenhänge zwischen der Genexpression von Blutzellen und der Höhe der erwähnten Blutspiegel untersucht werden.

Hinsichtlich der Auswahl der Gewebe sollten einige ergänzende Aspekte bedacht werden. So kann *a priori* nur unpräzise beurteilt werden, welches der Gewebe das der Erkrankung kausale Gewebe darstellt beziehungsweise in welchen Geweben die Genexpression durch die mit einer bestimmten Erkrankung assoziierten SNPs vorrangig beeinflusst wird. Obwohl circa 50 % der bekannten eQTL mehreren Geweben gemeinsam sind, sind viele eQTL spezifisch bezüglich eines bestimmten Gewebes [123,127]. Weiterhin sind die Gewebe, welche leichthin mit einem Phänotyp in Zusammenhang gebracht werden, nicht zwangsläufig die Gewebe, deren

Genexpression am stärksten beeinflusst wird. So wurde beispielsweise lange vermutet, Fettgewebe wäre bezüglich des Body-Mass-Index (BMI) das relevanteste Gewebe, doch jüngere Studien lassen vermuten, dass vielmehr das Hirngewebe eine vermehrte Expression von BMI assoziierten Genen aufweist [128]. Bei Erkrankungen oder Merkmalen zu denen kein explizit kausales Gewebe existiert, und der Gesamtphänotyp das Resultat einer Vielzahl an Variationen der Genexpression in unterschiedlichen Körpergeweben darstellt, müssten zudem optimalerweise alle am Phänotyp beteiligten Gewebe untersucht werden, um die genetischen Interaktionen in ihrer Gesamtheit und Komplexität erfassen zu können. So kann es beispielsweise auch vorkommen, dass bezüglich eines eQTL in verschiedenen Geweben gegensätzliche Effekte auf die Expression eines Gens ausgeübt werden [129].

Zu beachten ist zudem der grundsätzliche Aufbau der in dieser Arbeit verwendeten Gewebegenexpressionsdaten. Hierbei wurden durch das GTEx Projekt die mRNA Daten von teilweise sehr heterogen aufgebauten Geweben verwendet, ungeachtet der Tatsache, dass sich diese aus zahlreichen einzelnen Zelltypen zusammensetzen. Laut verschiedenen Studien sind 29 % bis 80 % der eQTL spezifisch bezüglich eines gewissen Zelltyps [130–134]. Durch diesen Umstand könnten in dieser Arbeit signifikante eQTL Ergebnisse aufgrund unterschiedlicher Genexpressionen der einzelnen Zellarten verdeckt werden. Werden beispielsweise die Effekte von monozytären und nicht monozytären Blutzellen übergeordnet als Teil des Gesamtblutgewebes im Rahmen einer eQTL Analyse betrachtet, könnten sie sich ausgleichen und zu keinen signifikanten eQTL Resultaten führen, wenngleich solche auf Zellebene faktisch vorhanden wären. Dieser Umstand wird zudem durch die Tatsache verschärft, dass die in dieser Arbeit untersuchten *trans* eQTL erwiesenermaßen zellspezifischer sind als *cis* eQTL [131,135].

Zusammenfassend sind mit der Auswahl der zu untersuchenden Gewebe und der zugehörigen Genexpressionsdaten einige zu bedenkende Problematiken verbunden, die eine zweckmäßige eQTL Analyse erschweren können. Da sich die in der vorliegenden Arbeit untersuchten Gewebe Leber und Blut an zahlreichen Körpervorgängen beteiligt zeigen, welche mit den in dieser Arbeit betrachteten Merkmalsgruppen assoziiert sind, sollte durch die Wahl dieser beiden Gewebe jedoch

eine vertiefte Erforschung der genetischen Korrelationen der untersuchten Merkmalsgruppen möglich sein.

Anstatt sich wie die meisten eQTL Studien auf cis eQTL zu beschränken, wurde die eQTL Analyse dieser Arbeit unter Einbezug von trans eQTL durchgeführt. Im Vergleich zu cis eQTL Effekten sind die Effektgrößen von trans eQTL geringer und die benötigten Probengrößen entsprechend größer, weshalb die Anzahl an bisher bekannten trans eQTL vergleichsweise klein ist [134,136]. Ob die Überzahl an cis gegenüber trans eQTL die physiologische Realität widerspiegelt oder nur der vergleichsweise geringen Stärke der trans Effekte zuzuschreiben ist, wird debattiert [137,138]. Insgesamt konnte jedoch gezeigt werden, dass ein wesentlicher Anteil der Genexpression über trans Effekte beeinflusst wird [136]. Bisherige trans eQTL Studien zeigten, dass sie wertvolle Einsichten in die Pathogenese von Erkrankungen aufzeigen können und dabei helfen, kohärente Gennetzwerke zu identifizieren, welche wahrscheinlich kausal an der Entstehung von Erkrankungen beteiligt sind [37,139] Trotz dieses Potenzials muss bei der Untersuchung von trans eQTL jedoch bedacht werden, dass diese aufgrund ihrer niedrigeren Effektgrößen sensibler auf äußerliche Beeinflussungen reagieren und sich oftmals nicht reproduzieren lassen [140,141]. Dies macht Folgestudien erforderlich, um die erlangten Ergebnisse auf Reproduzierbarkeit zu testen.

Im Rahmen der in dieser Arbeit durchgeführten Berechnung der trans eQTL wird mithilfe linearer Regression jeder einzelne SNP auf Assoziation mit jedem Gen des betrachteten Genexpressionsdatensatzes untersucht. Beim Merkmal ALT, für dessen eQTL Analyse beispielsweise die Korrelation von 4 SNPs mit den Genexpressionsniveaus von 15.939 proteinkodierenden Genen anhand des Gewebes Leber betrachtet wurde, beträgt die Anzahl an durchgeführten Tests für dieses eine Merkmal somit bereits 63.756 (4 x 15.939). Aus diesem Grund wurde eine Signifikanzschwelle von P < 1e-03 angelegt. Die Anzahl an signifikanten Resultaten mit diesem Grenzwert war jedoch weiterhin hoch (Leber: 14.436 eQTL Ergebnisse; Blut: 15.190 eQTL Ergebnisse), sodass der Grenzwert auf einen P-Wert von 1e-05 erniedrigt wurde, um die Anzahl an potenziell falsch positiven Ergebnissen zu reduzieren. Zu der P-Wert Schwelle von 1e-05 betrug die Summe an eQTL Ergebnissen bezüglich Lebergewebe 268, wobei die Werte für die einzelnen Merkmale ein Spektrum von 0 (ALT) bis 84 (AMD) aufweisen. Anhand des Gewebes Blut wurden insgesamt 194 eQTL identifiziert, mit einer Bandbreite der Einzelergebnisse zwischen 0 (ALT) und 48 (CAD). ALT war hierbei das einzige Merkmal, für das anhand der beiden Gewebe keine signifikanten eQTL Ergebnisse erzielt wurden. Da die Anzahl an eQTL Resultaten wie bereits beschrieben von der Menge an untersuchten SNPs abhängt, lassen sich die fehlenden signifikanten ALT eQTL Ergebnisse durch die geringe Anzahl an verwendeten, mit ALT assoziierten SNPs erklären.

Die Überlappung der Genlisten der Merkmalspaare wurde mithilfe des exakten Tests nach Fisher auf Signifikanz geprüft [84]. Dieses Vorgehen wurde beispielsweise von Freund et al. (2018) zur Bestimmung genetischer Gemeinsamkeiten zwischen komplexen Merkmalen und Mendelschen Erkrankungen verwendet [142]. Aufgrund der nachgewiesenen Gültigkeit wurde dieser bewährte Ansatz auch zur Signifikanzprüfung der Überschneidungen zwischen den Genlisten in der hier vorliegenden Arbeit genutzt.

Insgesamt fiel auf, dass vor allem die einzelnen Merkmale innerhalb der Merkmalsgruppen Gemeinsamkeiten in ihren Genlisten aufwiesen. Dies beruhte hauptsächlich auf den mit den Merkmalen assoziierten SNPs, mithilfe derer die eQTL Berechnung erfolgte. So überschnitten sich eng verwandte Merkmale wie beispielsweise SBP, DBP und GBP teilweise in ihren assoziierten genetischen Varianten. Hierbei überlappten sich beispielsweise zwischen SBP und DBP 45 SNPs. Dieser Aspekt führte ebenfalls zu den bereits beschriebenen Korrelationen der GRSs dieser Merkmale. Die zu beobachtenden Ähnlichkeiten der GRS Korrelationen mit den Gemeinsamkeiten zwischen den Genlisten war zu erwarten. Dies ist der Fall, da übereinstimmende SNPs einerseits eine genetische Korrelation bei der GRS Berechnung bedingen, jedoch andererseits auch zu einer Beeinflussung der Genexpression derselben Gene führen und so folglich die Überschneidungen in den Genlisten nach sich ziehen.

Die Anzahl an sich überschneidenden Genen der einzelnen Merkmale sowie die Anzahl der zugehörigen eQTL Ergebnisse zu den beiden untersuchten Geweben wurde grafisch in Circosplots dargestellt. Es fiel auf, dass zum Gewebe Leber der Anteil an AMD eQTL Ergebnissen an der Gesamtzahl an eQTL Ergebnissen deutlich größer war als bezüglich des Gewebes Blut (Leber: 84 von insgesamt 268 eQTL (31 %); Blut: 20 von insgesamt 194 eQTL (10 %)). Diese Beobachtungen erscheinen stimmig, da Lebergewebe Hauptort der Synthese von Komplementfaktoren ist und eines der zentralen Organe des Fett- und Cholesterinmetabolismus darstellt [120,143–145]. Diese Körperprozesse zeigten sich bereits in vergangenen Studien mit der AMD Pathogenese assoziiert [53,56,57]. Hinsichtlich des geringen Anteils an AMD eQTL Resultaten bezüglich des Gewebes Blut muss bedacht werden, dass Blut ein stark heterogenes Gewebe darstellt und sich folglich zelltypspezifische eQTL in ihrem Vorhandensein oder ihrer Ausrichtung widersprechen könnten. Dies würde nachfolgend das Signal von vorhandenen eQTL überdecken und könnte eine der Ursachen für die oben genannten Differenzen in der eQTL Verteilung zwischen den beiden untersuchten Geweben sein.

Im Gegensatz zu den eQTL Ergebnissen zeigte sich für die Zahl an gemeinsamen Genen zwischen der AMD Genliste und den Genlisten anderer Merkmale ein genau gegensätzliches Muster. So wurde anhand Lebergewebe nur ein gemeinsames Gen (INTS12) der AMD Genliste mit anderen Merkmalen festgestellt (mit HDL). Bezüglich des Gewebes Blut wurden dagegen 6 gemeinsame Gene der AMD Genliste identifiziert, jeweils 2 mit den einzelnen Lipidmerkmalen TG, TC und HDL. Dieses Ergebnis muss jedoch differenziert betrachtet werden. Der Blick in Anhangtabelle 8 offenbart, dass es sich bei den 2 Genen zwischen den Genlisten von AMD und TG, TC und HDL jeweils um die gleichen Gene handelt (HIST1H2BF und SLFN13), welche noch dazu durch dieselben SNPs (rs2043085 und rs1532085) in ihrer Expression beeinflusst werden. Die Überschneidung an GWAS SNPs zwischen den Merkmalen TG, TC und HDL hatte somit zu diesem Ergebnis geführt, womit sich bezüglich des Gewebes Blut real nur 2 Gene zwischen den Genlisten der AMD und den genannten Blutlipiden finden. Dieses Beispiel demonstriert, dass dem rein quantitativen Vergleich an Überschneidungen zwischen Genlisten der beiden Gewebe somit immer auch ein prüfender Blick auf die zugrundliegenden eQTL Gene und eQTL SNPs folgen muss, um die relativierenden Effekte durch gehäuft vorkommende eQTL Gene und eQTL SNPs in der Bewertung nicht außer Acht zu lassen.

Beim Vergleich der kardiovaskulären Merkmalsgruppe fielen ebenfalls Unterschiede auf. So wurden zu den kardiovaskulären Merkmalen anhand des Gewebes Leber 83 eQTL identifiziert. Dies entsprach einem Anteil von circa 30 % an der Gesamtanzahl an eQTL in diesem Gewebe. Im Gewebe Blut betrug dieser Anteil ungefähr 47 % (91 eQTL zu kardiovaskulären Merkmalen bei insgesamt 194 eQTL bezüglich des Gewebes Blut). Eine mögliche Erklärung für diese Dominanz im Gewebe Blut wäre das enge physiologische Verhältnis zwischen dem Blutfluss und den kardiovaskulären Merkmalen wie beispielsweise CAD. So sind viele der im Blut enthaltenen Zelltypen, wie beispielsweise Lymphozyten und Makrophagen, eng an der Pathogenese von CAD beteiligt [146,147]. Diese Tatsache kann erklären, wieso die CAD Risikovarianten die Expression besonders vieler Gene in diesem Gewebe signifikant beeinflussten. Äquivalent hierzu können so auch die eQTL Effekte durch die Risikovarianten der anderen kardiovaskulären Merkmale im Gewebe Blut erklärt werden.

Die gemeinsamen Gene zwischen den Genlisten wurden zudem hinsichtlich der Ursachen für deren Gemeinsamkeiten geprüft. Hierbei wurden zwei Fälle unterschieden: Einerseits Gene, deren Gemeinsamkeit zwischen den Genlisten durch die Effekte desselben SNPs zu erklären sind und andererseits gemeinsame Gene, deren Expression durch zwei verschiedene SNPs beeinflusst wird. Bei der Untersuchung des LD der SNPs der Gene, die durch zwei verschiedene SNPs beeinflusst werden, konnte interessanterweise festgestellt werden, dass alle SNP Paare der entsprechenden Gene mit nur einer Ausnahme in relativ hohem LD ($r^2 > 0,5$) zueinanderstehen (siehe **Anhangtabelle 5** und **Anhangtabelle 8**). Die SNPs, welche zueinander in LD stehen, weisen in ihrer Allelvererbung Assoziationen auf und sind folglich nicht unabhängig voneinander. Aus diesem Grund sollten die von diesen SNPs in ihrer Expression beeinflussten Gene somit ebenfalls als Gene betrachtet werden, deren Gemeinsamkeit zwischen den Genlisten auf gemeinsamen SNPS beruhen, auch wenn die SNPs formal verschieden sind.

Die erwähnte Ausnahme stellte hierbei das Gen *INTS12* dar, welches in den Genlisten der Merkmale AMD und HDL identifiziert wurde. Dieses Gen zeigte sich in der eQTL Analyse von zwei unterschiedlichen SNPs beeinflusst, welche sich zudem nicht in LD zueinander befanden. Bei dem mit AMD assoziierten SNP handelte es sich um *rs943080*, während *rs4650994* mit HDL in Zusammenhang stand, wobei beide SNPs signifikant die Genexpression des Gens *INTS12* beeinflussten. Der SNP *rs943080* ist auf Chromosom 6 und der SNP *rs4650994* ist auf Chromosom 1 lokalisiert. Das Gen *INTS12* wiederum ist auf Chromosom 4 positioniert [72]. Die Effekte der beiden

genannten SNPs auf *INTS12* sind somit chromosomenübergreifend und folglich indirekt, eine typische Charakteristik von *trans* eQTL.

Insgesamt konnten in der vorliegenden Arbeit 65 verschiedene Gene identifiziert werden, welche vermutlich pleiotrop wirken. Von diesen 65 Genen fanden sich 57 in den Genlisten von Merkmalen aus derselben Merkmalsgruppe. Nachfolgend soll der Fokus auf die 8 Gene gerichtet werden, die sich in den Genlisten der Merkmale aus unterschiedlichen Merkmalsgruppen fanden. Durch die Betrachtung der Genüberschneidungen unterschiedlicher Merkmalsgruppen können die gemeinsamen Gene und physiologischen Zusammenhänge weniger eng verwandter Merkmale gezielt diskutiert werden. Durch diesen Ansatz wird sowohl eine höhere Aussagekraft der Ergebnisse erwartet als auch ein Schwerpunkt auf Aspekte der Pleiotropie der sich daraus ergebenden Gene gelegt.

Die folgenden 5 Gene fanden sich sowohl in Genlisten von Merkmalen aus der kardiovaskulären Merkmalsgruppe als auch in Genlisten von Merkmalen aus der Blutlipidmerkmalsgruppe und sollen im Folgenden eingehender diskutiert werden: 1) Williams-Beuren syndrome chromosomal region 28 (*WBSCR28*); 2) Signal peptide, CUB domain and EGF like domain containing 2 (*SCUBE2*); 3) Cadherin EGF LAG seven-pass G-type receptor 2 (*CELSR2*); 4) Proline and serine rich coiled-coil 1 (*PSRC1*); 5) Sortilin 1 (*SORT1*).

Das Gen *WBSCR28* wurde in den Genlisten der Blutlipidmerkmale LDL und TC, des Leberenzyms ALP sowie des kardiovaskulären Merkmals CAD identifiziert und könnte somit pleiotrope Effekte in Bezug auf diese Merkmale beziehungsweise Erkrankungen aufweisen. Zu diesem Gen ist nur wenig bekannt. Micale et al. gelang es 2008 erstmals im Rahmen einer DNA-Sequenzanalyse das *WBSCR28* Gen auf Chromosom 7q11.23 innerhalb des bei Williams-Beuren Syndrom (WBS) deletierten Bereiches zu lokalisieren [148]. Das WBS hat seine Ursache in einer heterozygoten Deletion des langen Armes des Chromosoms 7 mit nachfolgender Haploinsuffizienz der Gene des deletierten Bereiches [149]. Neben *WBSCR28* kommt es hierbei zu einer Deletion von bis zu 27 weiteren Genen [150,151]. Aus der heterozygoten Deletion resultiert eine reduzierte Genexpression der betroffenen Gene [151]. Kinder mit diesem Krankheitsbild fallen meist durch ihre typischen Gesichtsmerkmale ("Elfengesicht")

sowie geistige und körperliche Entwicklungsstörungen auf [152]. Die Zusammenhänge zwischen WBS und LDL beziehungsweise TC werden durch mehrere Studien untermauert, auch wenn die Forschungslage in Bezug auf die Richtung der Beeinflussung der beiden Lipidwerte umstritten ist [153,154]. Anhand der eQTL Resultate der vorliegenden Arbeit lässt sich zeigen, dass die mit LDL und TC in Zusammenhang stehenden Risikoallele der betreffenden SNPs jeweils mit einer Herabregulierung des *WBSCR28* Gens assoziiert sind. Da die Risikoallele der SNPs mit einer verminderten Expression des *WBSCR28* Gens assoziiert sind, wie sie auch bei einer heterozygoten Gendeletion auftritt, ist folglich der Zusammenhang zwischen den Phänotypen bei WBS Patienten und den in dieser Arbeit generierten Ergebnissen zu *WBSCR28* plausibel erklärbar.

WBSCR28 wurde darüber hinaus in der Genliste des Merkmals CAD identifiziert. Circa 80 % aller von WBS betroffenen Personen zeigen kardiovaskuläre Defekte wie beispielsweise Aortenstenosen [155,156]. Zudem manifestieren bei WBS gehäuft Arteriopathien in Koronararterien, wobei es hier zu diffusen Stenosen der Koronarostien mit der Folge einer erhöhten myokardialen Ischämieneigung kommt, ähnlich dem Befundbild im Rahmen einer CAD [157,158]. Auch hier führen laut den Ergebnissen der vorliegenden Arbeit die mit CAD in Zusammenhang stehenden Risikoallele der betreffenden SNPs zu einer Herabregulation der Genexpression von *WBSCR28*. Somit könnte das vermehrte Auftreten von CAD durch die Deletion von WBSCR28 ausgelöst oder zumindest beeinflusst werden.

Laut dieser Arbeit ist die Höhe der Expression von *WBSCR28* weiterhin mit SNPs assoziiert, welche sich mit dem Serumspiegeln des Leberenzyms ALP assoziiert zeigten. Die Gruppe um Palmieri et al. (2018) konnten ähnliche Resultate erzielen, indem sie bei einer Studie mit 29 WBS Patienten signifikant erniedrigte ALP Spiegel feststellten [159]. Das Risikoallel des mit ALP assoziierten SNP *rs579459* steht laut den Ergebnissen der vorliegenden Arbeit mit einer erhöhten Expression des *WBSCR28* Gens in Korrelation. Da das risikosenkende Allel, welches mit niedrigen ALP Spiegeln assoziiert ist, somit die Genexpression von WBSCR28 senkt, stimmt dies mit den Ergebnissen von Palmieri et al. (2018) überein, und lässt somit ebenfalls ein Mitwirken von *WBSCR28* an den verminderten ALP Spiegeln von WBS Patienten als Schluss zu.

Bei Betrachtung der genannten Zusammenhänge sollte beachtet werden, dass im Rahmen der heterozygoten Deletion des WBS neben *WBSCR28* noch mindestens 27 weitere Gene betroffen sind und deren Deletion ebenfalls als Ursache der beschriebenen Phänotypen in Frage kommt. Die relevantesten Kandidaten hierfür sind das MLX interacting protein like (*MLXIPL*) Gen sowie das Elastin (*ELN*) Gen, wobei beide nachgewiesenermaßen an phänotypischen Charakteristika des WBS beteiligt sind [154,160–165]. Wenngleich die Ergebnisse der vorliegenden Arbeit plausibel anhand der Literatur bestätigt werden können, erfordert die Komplexität der Zusammenhänge zwischen *WBSCR28*, WBS sowie der Merkmale ALP, LDL, TC und CAD weitere Forschungsarbeit, um so die Art und Kausalität dieser Verbindungen weiter zu beleuchten.

Ein weiteres Gen, welches in dieser Arbeit als möglicherweise pleiotrop identifiziert wurde, ist SCUBE2. Es wurde als gemeinsames Gen der Genlisten der kardiovaskulären Merkmale SBP, DBP und GBP sowie des Blutlipidmerkmals HDL ermittelt. Dieses Gen kodiert für das Protein SCUBE2. Das Protein agiert als Tumorsuppressor und beeinflusst das Wachstum Blutgefäßen von aus vorbestehenden Gefäßen, die sogenannte Angiogenese [166]. Hierbei fördert SCUBE2 im Rahmen der Angiogenese die Interaktion des endothelialen Vascular Endothelial Growth Factor Receptor 2 (VEGFR2) mit dem Vascular Endothelial Growth Factor (VEGF) [167]. Der VEGF Signalweg zeigt sich wiederum an der Regulation des Blutdrucks beteiligt [168]. Dementsprechend verursachen VEGF Inhibitoren, wie sie beispielsweise bei onkologischen Erkrankungen oder bei der Therapie der feuchten Form der Makuladegeneration angewendet werden, bei bis zu 80 % der behandelten Patienten eine Erhöhung des arteriellen Blutdrucks [169]. Dies demonstriert anschaulich einen zugrundeliegenden physiologischen Mechanismus bezüglich des Zusammenhangs von SCUBE2 mit den Blutdruckmerkmalen SBP, DBP und GBP.

Ein weiterer Signalweg, mit welchem sich das SCUBE2 Protein assoziiert zeigt, ist der Hedgehog Signalweg, benannt nach seinem Liganden, dem Hedgehog Signalprotein. Hierbei konnte gezeigt werden, dass die SCUBE Proteinfamilie und insbesondere SCUBE2 die Sekretion des Hedgehog Liganden vermitteln [170,171]. Der Hedgehog Signalweg stellt einen starken Regulator des Lipidmetabolismus in der Leber dar [172]. Die Forschungsgruppe um Lever et al. (2007) verabreichte Mäusen einen Antikörper gegen den Hedgehog Liganden und konnte hierdurch einen signifikanten Anstieg atherosklerotischer Veränderungen in den Aorten der Mäuse induzieren, wobei sich die TC Werte der Mäuse signifikant erniedrigt zeigten [173]. In der gegenwärtigen Arbeit wurde zwar nur eine Assoziation von SCUBE2 mit HDL identifiziert, jedoch steht HDL als Teil des TC Spiegels mit diesem in engem Zusammenhang. Obwohl Lever et al. (2007) eindeutig Assoziationen des Hedgehog Signalweges und damit auch SCUBE2 mit atherogenen Prozessen zeigen konnten, wie sie auch bei der Pathogenese der CAD auftreten, wurde in der vorliegenden Arbeit zu diesem Merkmal kein signifikanter Zusammenhang gefunden. Ali et al. (2019) konnten zudem eine signifikant vermehrte Expression von SCUBE2 bei Patienten mit Diabetes Mellitus und veränderten Blutlipidwerten im Vergleich zur Kontrollgruppe nachweisen, insbesondere bei Patienten mit erhöhten LDL, TC und TG Werten sowie erniedrigten HDL Werten. Interessanterweise wurden bei der Patientengruppe eine signifikant erhöhte Endothelin-1 (EDN1) Expression nachgewiesen, wobei diese laut mehrerer Publikationen mit erhöhter VEGF Expression assoziiert ist [174-176]. Diese Ergebnisse bekräftigen sowohl die in dieser Arbeit detektierte Assoziation von SCUBE2 mit HDL Werten als auch den bereits diskutierten Zusammenhang mit den Blutdruckmerkmalen SBP, DBP und GBP über den VEGF Signalweg.

Zusammenfassend implizieren die Ergebnisse der vorliegenden Arbeit pleiotrope Zusammenhänge zwischen den Blutdruckmerkmalen SBP, DBP und GBP sowie dem Lipidmerkmal HDL. Anhand aktueller Publikationen können eindeutige Zusammenhänge zwischen *SCUBE2* und kardiovaskulären Merkmalen sowie dem Lipidmetabolismus gezeigt werden. Der VEGF sowie der Hedgehog Signalweg könnten hier laut Literatur möglicherweise eine Schlüsselrolle spielen und die Verbindung von *SCUBE2* mit den genannten Merkmalen erklären.

Weitere in dieser Arbeit als möglicherweise pleiotrop wirkend identifizierte Gene sind *PSRC1, CELSR2* und *SORT1.* Diese Gene wurden in den Genlisten des kardiovaskulären Merkmals CAD sowie in den Genlisten der Blutlipidmerkmale LDL und TC identifiziert und sind Teil des 1p13.3 Genlokus [89,177]. Arvind et al. (2014) konnten im Rahmen einer Bionetzwerkanalyse Interaktionen der 3 Gene untereinander aufzeigen, womit sie als Teil eines sich gegenseitig beeinflussenden Netzwerkes betrachtet werden sollten. Zudem konnte die Forschungsgruppe signifikant protektive

Assoziationen einer erhöhten *PSRC1* Genexpression in Bezug auf CAD feststellen, sowie Korrelationen zwischen der *PSRC1* Genexpression und TC, LDL und HDL Spiegeln. Auch zeigten sie Zusammenhänge zwischen der *CELSR2* Genexpression und TC, LDL und TG Spiegeln [89]. Andaleon et al. (2019) wiesen in Ergänzung hierzu eine Assoziation von erhöhten *PSRC1* Genexpressionsniveaus mit erniedrigten TC und LDL Spiegeln in verschiedenen Ethnien und Geweben, unter anderem Lebergewebe, nach [178]. Es konnten zudem Zusammenhänge von *CELSR2* und *PSRC1* mit CAD über die Lipoprotein-assoziierten Phospholipase A2 (LP-PLA2) gezeigt werden [179,180]. Das Enzym LP-PLA2 findet sich in hoher Menge beispielsweise in atherosklerotischen Plaques und spielt eine wichtige Rolle in der Pathogenese der Atherosklerose und CAD [181–183]. Mit Darapladib befindet sich bereits ein Medikament in Entwicklung, um in seiner Funktion als LP-PLA2 Inhibitor als Therapiemöglichkeit der Atherosklerose zu fungieren [184]. Diese Ergebnisse bestätigen somit anschaulich die in dieser Arbeit identifizierten Zusammenhänge der Gene mit den jeweiligen Phänotypen.

Das dritte Gen des 1p13.3 Lokus, SORT1, wird ebenfalls mit Atherosklerose und Lipidspiegeln in Verbindung gebracht [96,185,186]. Das SORT1 Protein zeigt weiterhin einen engen Zusammenhang mit dem Serumspiegel des Proteins Progranulin, welches wiederum mit Diabetes Typ 2 und Atherosklerose korreliert [187-189]. Außerdem ist das SORT1 Protein an der Sekretion der Serinprotease PCSK9 beteiligt [190]. Das PCSK9 Protein sowie das zugehörige Gen konnten wiederum mit Atherosklerose sowie dem Cholesterolstoffwechsel assoziiert werden [191–193]. Die seit 2015 zugelassenen monoklonalen PCSK9 Antikörper Alirocumab und Evolocumab werden dementsprechend Behandlungsverfahren als der Hypercholesterinämie, also erhöhter Blutcholesterinspiegel, eingesetzt [194]. 2020 wurde für dieses Erkrankungsbild in Europa zudem das Medikament Inclisiran zugelassen, welches als einer der ersten Vertreter der sogenannten RNAi-Therapeutika die Translation von PCSK9 inhibiert [195].

Die Richtung der genannten Zusammenhänge in Bezug auf *SORT1* und hiermit assoziierter Prozesse befindet sich jedoch nicht in Übereinstimmung mit den Ergebnissen dieser Arbeit. Im Gegensatz zu obigen Studien implizieren die vorliegenden Ergebnisse eine Korrelation niedriger *SORT1* Expression mit dem

Auftreten von CAD sowie erhöhter LDL und TC Spiegel. Einige Studien konnten jedoch auch die Ergebnisse der gegenwärtigen Arbeit bestätigen [196–198]. Die Beziehung von *SORT1* mit den Phänotypen CAD, LDL und TC scheint somit komplex und noch nicht endgültig geklärt zu sein [199]. Ein Aspekt, der in die Betrachtung mit einbezogen werden sollte, ist die Tatsache, dass regulatorische Mechanismen auf allen Stufen entlang der Genexpressionskaskade zwischen DNA und Protein wirken [200]. Da die in der vorliegenden Arbeit verwendete Analyse mittels eQTL nur Prozesse bis zur Ebene der Transkription erfasst, sind regulatorische Mechanismen sowie andere Einflussfaktoren, die nach dieser Ebene wirken, den in dieser Arbeit verwendeten Methoden nicht zugänglich. Diese Tatsache könnte eine Ursache für die Differenzen zwischen den Resultaten darstellen.

Zusammenfassend zeigen sich die Studien zu den Genen CELSR2 und PSRC1 in Übereinstimmung Ergebnissen vorliegenden Der mit den der Arbeit. Wirkmechanismus über das Enzym LP-PLA2 scheint hierbei ein aussichtsreicher Kandidat zu sein, welcher der Verbindungen der genannten Gene mit CAD zugrunde liegen könnte. Im Gegensatz hierzu ist die Studienlage zu SORT1 komplex und teilweise umstritten. So kommen die Arbeiten in Bezug auf die Korrelation der SORT1 Expression mit CAD, LDL und TC zu widersprüchlichen Ergebnissen und lassen somit keine endgültige Aussage in Bezug auf die in dieser Arbeit generierten Resultate zu. In Abbildung 9 sind die erläuterten Zusammenhänge bezüglich der Gene WBSCR28, SCUBE2 sowie CELSR2-PSRC1-SORT1 grafisch dargestellt.

Diskussion



Abbildung 9: Übersicht zu den Zusammenhängen der Gene WBSCR28, SCUBE2 und CELSR2-PSRC1-SORT1.

Bezüglich WBSCR28 wurde zu Zusammenhängen entsprechend den Ergebnissen dieser Arbeit mit CAD, LDL, TC und ALP recherchiert, bezüglich SCUBE2 zu Zusammenhängen mit SBP, DBP, GBP und HDL, und bezüglich CELSR2-PSRC1-SORT1 zu Zusammenhängen mit CAD, LDL und TC. Blaue Kästen repräsentieren die Ausgangsgene, zu deren Zusammenhängen mit den entsprechenden Merkmalen beziehungsweise Ergebnissen dieser Arbeit gezielt recherchiert wurde. Weiße Figuren repräsentieren die Merkmale, welche über die in grau dargestellten Mechanismen oder Proteine mit den Ausgangsgenen in Zusammenhang stehen. Die Verbindungslinien zwischen den Figuren geben Assoziationen an. Zu den Verbindungslinien sind die Publikationen angegeben, die die Assoziation beschreiben. Die folgenden 3 Gene konnten in dieser Arbeit sowohl in Genlisten von AMD als auch in Genlisten von Merkmalen der Blutlipidmerkmalsgruppe identifiziert werden und scheinen folglich pleiotrop in Bezug auf diese Merkmalsgruppen zu wirken: *1*) Schlafen family member 13 (*SLFN13*); *2*) Histone cluster 1 H2B family member F (*HIST1H2BF*); *3*) Integrator complex subunit 12 (*INTS12*).

In den Genlisten des Merkmals AMD und der Blutlipidmerkmale HDL, TC und TG wurde das Gen *SLFN13* identifiziert. *SLFN13* kodiert für eine Endoribonuklease, welche Funktionen im Rahmen der Translationskontrolle wahrnimmt [201]. Dieses Protein gehört zur kleinen Gruppe der Schlafen (SLFN) Proteinfamilie, deren Mitglieder beispielsweise an der Regulation der Zellproliferation sowie Induktion von Immunantworten beteiligt sind [202–204]. Weitere Informationen konnten zum *SLFN13* Gen und dem dazugehörigen Protein im Rahmen der Literaturrecherche nicht gefunden werden.

Das Gen HIST1H2BF wurde in dieser Arbeit in den Genlisten der Merkmale AMD sowie der Blutlipidmerkmale HDL, TC und TG identifiziert. Dieses Gen kodiert für die Isoform F des Histons H2B [205]. Histon H2B ist eines der 4 Haupthistone des Menschen und am Aufbau der Chromatinstruktur beteiligt [206]. Zu den Isoformen und insbesondere Isoform F ist wenig bekannt, jedoch ist es nach Molden et al. (2015) sehr wahrscheinlich, dass die einzelnen Isoformen unterschiedliche Funktionen ausüben und die Art der beeinflussten Prozesse zudem je nach Gewebe variiert [207]. Mithilfe von Mausmodellen konnten Santoro et al. (2012) beispielsweise die Funktion der replikationsunabhängigen Isoform H2BE beschreiben, welche ausschließlich durch olfaktorische Neuronen des olfaktorischen Epithels exprimiert wird. So hatten H2BE Knockout Mäuse eine signifikant verringerte Genexpression sowie Dichte ihrer olfaktorischen Rezeptoren und infolgedessen eine Beeinträchtigung ihrer Geruchssinnesleistung [208]. Die Isoform H2BF könnte entsprechend an bisher unbekannten physiologischen Prozessen im Rahmen der AMD Pathogenese sowie der Regulation von Blutlipiden mitwirken.

Das Gen *INTS12* fand sich in dieser Arbeit in den Genlisten der Merkmale AMD und HDL. Es wurden Assoziationen dieses Gens und des zugehörigen Proteins mit der Lungenfunktion aufgezeigt [209]. Kheirallah et al. (2017) konnten zudem

demonstrieren, dass *INTS12* an der Regulierung der Expression von Genen beteiligt ist, welche wiederum eng mit Signalwegen der Proteinsynthese assoziiert sind [210]. Das Gen *INTS12* kodiert für die Untereinheit 12 des Integratorkomplexes, welchem entscheidende Bedeutung im Rahmen der Transkription zukommt. So begleitet er das C-terminale Ende der RNA Polymerase II (POLII), wobei diese bei Eukaryoten für die Transkription verantwortlich ist [211,212]. Der Integratorkomplex ist zudem an der Biogenese und Prozessierung der small nuclear RNA (snRNA) beteiligt, welche als Bestandteil des Spleißosoms das Spleißen der Introns der im Zellkern enthaltenen prämRNA im Rahmen der Transkription katalysieren [213,214].

Aufgrund der Beteiligung von *SLFN13*, *HIST1H2BF* und *INTS12* an zahlreichen regulatorischen und ubiquitär vorkommenden Prozessen ist ein Zusammenhang mit den jeweiligen komplexen Merkmalen durchaus denk- und begründbar. Die Tatsache, dass die Gene bisher mit den genannten Merkmalen noch nicht in Verbindung gebracht werden konnten, lässt sich durch die vermutlich relativ geringe Wirkung beziehungsweise Beteiligung der Gene an den betrachteten komplexen Merkmalen erklären. In **Abbildung 10** sind die erläuterten Zusammenhänge bezüglich der Gene *SLFN13*, *HIST1H2BF* sowie *INTS12* grafisch dargestellt.



Abbildung 10: Übersicht zu den Zusammenhängen der Gene SLFN13, HIST1H2BF und INTS12.

Bezüglich SLFN13 wurde zu Zusammenhängen entsprechend den Ergebnissen dieser Arbeit mit AMD, HDL, TC und TG recherchiert, bezüglich HIST1H2BF zu Zusammenhängen mit AMD, HDL, TC und TG, und bezüglich INTS12 zu Zusammenhängen mit AMD und HDL. Blaue Kästen repräsentieren die Ausgangsgene, zu deren Zusammenhängen mit den entsprechenden Merkmalen beziehungsweise Ergebnissen dieser Arbeit gezielt recherchiert wurde. Weiße Figuren repräsentieren die Merkmale, welche über die in grau dargestellten Mechanismen oder Proteine mit den Ausgangsgenen in Zusammenhang stehen. Die Verbindungslinien zwischen den Figuren geben Assoziationen an. Zu den Verbindungslinien sind die Publikationen angegeben, die die Assoziation beschreiben.

Neben diesen 8 detailliert beschriebenen Genen wurden noch 57 weitere, vermutlich pleiotrop wirkende Gene identifiziert. All diese Gene beruhen auf den Ergebnissen einer mit GWAS SNPs durchgeführten eQTL Analyse. Es sollte angemerkt werden, dass diese Vorgehensweise gewisse Limitierungen aufweist: So muss in Hinblick auf eQTL bedacht werden, dass die Expression eines Gens nur teilweise mit der Expression des zugehörigen Proteins korreliert [215]. Weiterhin wirken

Diskussion

merkmalsassoziierte Varianten auf die Ausprägung des Phänotyps durch verschiedenste Prozesse, welche in ihrer Komplexität durch eQTL nicht vollständig erfasst werden können. Um solche Wirkungsmechanismen zu analysieren, sind somit ergänzende Methodiken jenseits von eQTL Studien erforderlich. Hierzu gibt es zahlreiche weitere Arten von Quantitative Trait Loci (QTL). Ansätze wie splicing Quantitative Trait Loci (sQTL) Analysen können beispielsweise den Einfluss genetischer Varianten auf Splicingprozesse beschreiben [216]. Im Rahmen von protein Quantitative Trait Loci (pQTL) wird die Korrelation zwischen genetischer Variation und der Proteinexpression untersucht [217]. Auch das Epigenom kann durch die Untersuchung von methylation Quantitative Trait Loci (mQTL) oder histone acetylation Quantitative Trait Loci (haQTL) miteinbezogen werden. Hierbei wird der Einfluss genetischer Varianten auf die DNA Methylierung respektive Histonmodifikation betrachtet [218,219]. Die in dieser Arbeit verwendeten eQTL Daten sind somit nur fähig, einen kleinen Teil der komplexen molekularen Interaktionen zwischen genetischer Variation und den assoziierten Phänotypen zu beschreiben. Neben diesen Aspekten impliziert das Übereinstimmen zweier Korrelationen, im in dieser Studie vorliegenden Fall die Korrelation des SNPs mit der Krankheit sowie des SNPs mit der Genexpression, nicht zwingend Kausalität. Stattdessen wäre es auch möglich, dass der SNP, welcher Assoziationen mit der Krankheit zeigt, die Genexpression unabhängig von der Erkrankung beeinflusst. Die Genexpression wäre in diesem Fall nicht ursächlich für das Auftreten des untersuchten Phänotyps. Zudem kann die krankheitsassoziierte kausale Variante in starkem LD mit einer weiteren Variante stehen, welche wiederum für Modulation der Genexpression ursächlich ist. Hierbei wären das entsprechende Gen und seine Expression nicht kausal mit der untersuchten Erkrankung beziehungsweise dem Merkmal assoziiert. Eine Möglichkeit um solche Konstellationen miteinzubeziehen und kausale Zusammenhänge zu bestimmen, stellen zum Beispiel weiterführende in silico Methodiken wie Kolokalisationsanalysen und Mendelsche Randomisierung dar [27,217,220–222]. Insgesamt muss festgehalten werden, dass die in der vorliegenden Arbeit angewandte bioinformatische Vorgehensweise aufgrund der genannten Aspekte vorrangig dazu verwendet werden sollte, neue Hypothesen bezüglich potenziell kausaler Verbindungen zwischen Varianten, Genen und Merkmalen zu generieren. Dies kann die Anzahl der weiter zu untersuchenden Zusammenhänge und physiologischen Mechanismen eingrenzen.

Für den endgültigen Nachweis der implizierten Verbindungen sind jedoch Folgestudien nötig [220,223].

Einzelne Varianten, welche dieser Arbeit zufolge möglicherweise pleiotrope Effekte aufweisen, könnten beispielsweise mit Techniken der Genom-Editierung gezielt untersucht werden, die phänotypischen Effekte der allelspezifischen um Veränderungen der Genexpression weiter zu beleuchten. Die im Rahmen der vorliegenden Arbeit identifizierten und vermutlich pleiotrop wirkenden Gene wären zudem Kandidaten, um im Rahmen von Knockout, Knockdown oder Überexpressionsstudien mit Modellorganismen vertieft untersucht zu werden. Der Einbezug der in dieser Arbeit generierten Informationen, zu welchen der Merkmale das jeweilige Gen pleiotrop wirkt, könnte hierbei genutzt werden, den Fokus gezielt auf Eigenschaften der untersuchten Organismen zu richten, welche mit den beeinflussten Phänotypen solche weiterführenden in Zusammenhang stehen. Falls Forschungsansätze erfolgreich sind, könnten sich beispielsweise neue pharmakologische Behandlungsansätze ableiten lassen.

Neben dieser therapeutischen Nutzung käme auch die Anwendung im Bereich der Diagnostik in Betracht. So könnte die Genexpression der in dieser Arbeit identifizierten Gene als Biomarker genutzt werden. Da das Expressionsniveau dieser eQTL Gene mit verschiedenen Erkrankungen assoziiert ist, scheint dieser Parameter *a priori* aussichtsreich für eine solche Anwendung. Hierdurch könnten Risikovoraussagen gegenüber komplexen Erkrankungen einzelner Individuen getroffen werden. Auch Klassifizierungen von Erkrankungssubtypen unter den Betroffenen könnten so ermöglicht werden, mit dem Ziel einer individuellen Behandlung dieser Subtypen [223].

Um mehr *trans* eQTL mit den einhergehenden kleineren Effektgrößen aufzudecken, sollte der vorliegende Ansatz zukünftig hochskaliert und auf größere Stichproben angewendet werden, ähnlich wie dies bereits bei GWAS durchgeführt wurde. Hierdurch könnten eQTL mit kleineren Effektgrößen identifiziert werden und so keine relevanten eQTL mehr verfehlt werden [200]. Viele der in neueren Studien identifizierten eQTL waren jenseits der Detektionsschwellen älterer Analysen, und es ist zu vermuten, dass noch viele tausend eQTL und deren Gene jenseits unserer aktuellen statistischen Reichweite unentdeckt geblieben sind.

Ein Aspekt, dem bei der Anwendung der vorliegenden Methodik zukünftig Aufmerksamkeit zuteilwerden sollte, ist die Auswahl der verwendeten Genotypdaten. Die Daten dieser Arbeit beziehen sich auf Probanden vorrangig europäischer Herkunft. Diese Auswahl wurde aus pragmatischen Gründen vollzogen. Bezüglich des GWAS Catalog, welcher dieser Arbeit neben PubMed als Quelle diente, sind 79 % der untersuchten Individuen europäischer Abstammung, trotz des deutlichen niedrigeren Anteils dieser Gruppe an der Weltbevölkerung [224]. Durch die unterschiedliche Varianten- und Haplotypenverteilungen zwischen Ethnien können die Ergebnisse der vorliegenden Arbeit folglich nicht auf eine Population anderer Abstammung übertragen werden. Die Nutzung von Genotypdaten verschiedener Bevölkerungsgruppen könnte zukünftig dazu beitragen, die erlangten Ergebnisse auch auf weitere Ethnien anwendbar zu machen.

Auch die Auswahl der zu nutzenden Genexpressionsdaten ist von entscheidender Bedeutung. Um die Effekte von regulatorischen Varianten auf möglichst viele Gene erfassen zu können, ist es sinnvoll, die Anzahl untersuchter Gewebe nach Möglichkeit auszuweiten. Hierdurch können die gewebespezifischen Effekte genetischer Varianten gezielt erfasst werden. Dies ist notwendig, um zu verstehen, wieso ein gewisses Gewebe das kausale "Krankheitsgewebe" darstellt und nicht ein anderes Gewebe. Neben GTEx könnten zur Gewinnung von Genexpressionsdaten weitere Quellen wie beispielsweise die Multiple Tissue Human Expression Resource (MuTHER) Datenbank zum Einsatz kommen [133]. Ein weiterer Schritt wäre die Verwendung zellspezifischer Genexpressionsdaten zur eQTL Berechnung, um so die Effekte genetischer Varianten auf einzelne Zelltypen zu erfassen. Die so erlangten Informationen könnten genutzt werden, um Rückschlüsse auf die bei einer Erkrankung betroffenen Zellarten zu ermöglichen [225,226]. Dies ist von besonderer Bedeutung, da die in dieser Arbeit untersuchten trans eQTL erwiesenermaßen zellspezifischer sind als *cis* eQTL [131,135]. Es sind bereits Bestrebungen im Gange, eQTL in einzelnen Zelltypen, wie beispielsweise bestimmten Immunzellen, zu identifizieren [131,225,227]. Zelltyp-spezifische Expressionsdaten könnten für zukünftige Studien beispielsweise der Reference Expression (RefEx) Datenbank entnommen werden [228]. Auch bezüglich des GTEx Projekts wurde inzwischen begonnen, Datensätze einzelner Zelltypen zu implementieren [130]. In Ergänzung zu den genannten zellspezifischen Effekten zeigt sich der Einfluss der genetischen Variabilität auf die Genexpression vielfach auch erst unter bestimmten Umwelteinflüssen, wie exogener

Zellstimulation oder nach Krankheitsbeginn, infolge der hierdurch aktivierten zellulären Signalprozesse [225]. Eine eQTL Studie an unstimulierten sowie mit viralen oder bakteriellen Bestandteilen stimulierten Monozyten konnte aufzeigen, dass bis zu 18 % aller eQTL stimulusspezifisch waren [229]. Die Verwendung kontextspezifischer Genexpressionsdaten könnte somit ebenfalls dazu beitragen, weitere bisher unentdeckte eQTL zu identifizieren. Dies ist insofern von besonderer Bedeutung, da die große Mehrheit der bisher durchgeführten eQTL Studien anhand gesunder und unbeeinflusster Gewebe durchgeführt wurde.

Zusammenfassend wird in dieser Arbeit eine Methode präsentiert, die unter Einbezug von Daten aus unabhängigen GWAS mithilfe von eQTL Analysen Gene identifiziert, deren Genexpressionen sich mit verschiedenen komplexen Merkmalen assoziiert zeigen. Hierdurch konnten 65 Gene bestimmt werden, welche vermutlich pleiotrop in Bezug auf insgesamt 15 verschiedene Erkrankungen und Merkmale wirken. Die vorliegende Arbeit ist hierbei die erste, welche neben einer *trans* eQTL Analyse den Aspekt der pleiotropen Wirkung der identifizierten *trans* eQTL auf verschiedene Merkmale und Krankheitsbilder miteinbezieht und so die Integration von Genotyp und Genexpressionsdaten und deren pleiotropen Zusammenhängen erlaubt. Durch Einbezug pleiotroper Aspekte ermöglicht die Methodik einen breiteren Blick auf die Interaktionen von Genen mit verschiedenen Erkrankungen sowie auf die Zusammenhänge zwischen den Erkrankungen selbst.

Zu einigen der in dieser Arbeit als möglicherweise pleiotrop identifizierten Gene wurden die aufgezeigten Korrelationen bereits in der Literatur beschrieben, wie beispielsweise die Assoziation von *CELSR2* oder *PSRC1* mit LDL und CAD. Im Gegensatz hierzu konnten jedoch auch pleiotrope Effekte bei Genen gezeigt werden, zu denen diese Korrelationen zuvor unbekannt waren und zu welchen die Zusammenhänge mithilfe bekannter Literaturdaten nicht ausreichend erklärt werden können. Dies ist durch die typischerweise kleine Beteiligung solcher Gene an den komplexen Merkmalen begründbar und war folglich zu erwarten. Nichtsdestotrotz wirkt jeder einzelne und noch so kleine genetische Beitrag letztendlich an der Entstehung dieser Merkmale mit und muss somit erforscht werden, um die genetischen Hintergründe komplexer Merkmale in ihrer Gänze erfassen zu können. Das Wissen um die Existenz solcher Verbindungen kann zudem wertvolle Hinweise liefern, zu

welchen Genen und Mechanismen gezielte Untersuchungen lohnend sein könnten, um so die den Zusammenhängen zugrundeliegenden molekularen Prozesse näher zu beleuchten.

Die Untersuchung der in der hier vorliegenden Arbeit identifizierten pleiotropen Gene konnte weiterhin aufzeigen, dass einige dieser Gene an regulatorischen und ubiquitären Prozessen wie beispielsweise der Translation oder Aufbau der DNA beteiligt sind und sich somit nicht konkreten Phänotypen zuordnen lassen. Folglich muss auch solchen Genen bei der zukünftigen Erforschung der Genetik komplexer Merkmale besondere Beachtung geschenkt werden.

Ergebnisse zeigen, wie die Integration von Genotyp Die erlangten und Genexpressionsdaten in Verbindung mit dem Aspekt der Pleiotropie helfen kann, die Zusammenhänge komplexer Erkrankungen zu entschlüsseln. Da die verschiedenen genomischen Datensätze auch zukünftig weiter anwachsen werden, ist diese Integration von zunehmender Bedeutung. Die Menge an identifizierten pleiotrop wirkenden Genen, welche zudem in die verschiedensten molekularen Prozesse involviert sind, demonstriert zudem die breite genetische Basis komplexer Erkrankungen. Dies unterstreicht die Notwendigkeit der weiteren Untersuchung der Erkrankungen Zusammenhänge komplexer und den zugrundeliegenden Mechanismen, um mithilfe dieser Einsichten zukünftige Forschungsschwerpunkte sowie mögliche Therapieziele abzuleiten.

Abkürzung	Erklärung
AMD	Altersabhängige Makuladegeneration
ALP	Alkalische Phosphatase
ALT	Alanin-Aminoransferase
BMI	Body-Mass-Index
CAD	Coronary Artery Disease / Koronare Herzkrankheit
CELSR2	Cadherin EGF LAG seven-pass G-Type receptor 2
CFH	Complement factor H
CFTR	Cystic fibrosis transmembrane conductance regulator
DBP	Diastolischer Blutdruck
DNAH11	Dynein axonemal heavy chain 11
ELN	Elastin
eQTL	Expression Quantitative Trait Locus
FADS2	Fatty acid desaturase 2
FDR	False Discovery Rate/ Falscherkennungsrate
GBP	Blutdruck- Allgemein
GGT	Gamma-Glutamyltransferase
GRS	Genetischer Risiko Score
GTEx	Genotype-Tissue-Expression
GWAS	Genomweite Assoziationsstudie
HDL	High Density Lipoprotein / Lipoproteine höherer Dichte
HIST1H2BF	Histone cluster 1 H2B family member F
HTN	Hypertonie
INTS12	Integrator complex subunit 12
LD	Linkage disequilibrium / Kopplungsungleichgewicht
LDL	Low Density Lipoprotein / Lipoproteine geringerer Dichte
LEP	Leberenzyme im Plasma
Lp-PLA2	Lipoprotein-assoziierte Phospholipase A2
Mb	Megabase
MLXIPL	MLX interacting protein like

Abkürzungsverzeichnis

mQTL	Methylation Quantitative Trait Locus
MuTHER	Multiple Tissue Human Expression Resource
NAFLD	Nichtalkoholische Fettlebererkrankung
NASH	Nichtalkoholische Steatohepatitis
NHGRI	National Human Genome Research Institutes
OR	Odds Ratio
POLII	RNA Polymerase II
pQTL	Protein Quantitative Trait Locus
PP	Blutdruckamplitude
PSRC1	Proline and serine rich coiled-coil 1
QTL	Quantitative Trait Locus
RPE	Retinales Pigmentepithel
SBP	Systolischer Blutdruck
SCUBE2	Signal peptide, CUB domain and EGF like domain containing 2
SLFN	Schlafen
SLFN13	Schlafen family member 13
SNP	Single Nucleotide Polymorphism
snRNA	Small nuclear RNA
sQTL	Splicing Quantitative Trait Locus
ST3GAL4	ST3 Beta-Galactoside Alpha-2,3-Sialyltransferase 4
SORT1	Sortilin 1
ТС	Gesamtcholesterin
TG	Triglyceride
ULK4	Unc-51 like kinase 4
VEGF	Vascular endothelial growth factor
VLDL	Very Low-Density Lipoprotein
WBS	Williams-Beuren Syndrom
WBSCR14	Williams-Beuren syndrome chromosome region 14
WBSCR28	Williams-Beuren syndrome chromosomal region 28

Abbildungsverzeichnis

Abbildung 1: eQTL und ihre Wirkungsweisen 11
Abbildung 2: Manhattan Plot der mit AMD assoziierten Loci
Abbildung 3: Beziehungen von Merkmalen, SNPs und der Expression
pleiotroper Gene
Abbildung 4: Arbeitsablauf basierend auf den GTEx Genexpressions- und
Genotypdaten
Abbildung 5: Grafiken zur Qualitätskontrolle am Beispiel des Merkmals CAD. 28
Abbildung 6: Korrelationen zwischen den GRSs der verwendeten Merkmale 29
Abbildung 7: Gemeinsamkeiten zwischen den Genlisten der Merkmale
Abbildung 8: Circosplot der gemeinsamen Gene zwischen den Genlisten der
Merkmale
Abbildung 9: Übersicht zu den Zusammenhängen der Gene WBSCR28, SCUBE2
und CELSR2-PSRC1-SORT1
Abbildung 10: Übersicht zu den Zusammenhängen der Gene SLFN13,
HIST1H2BF und INTS12

Tabellenverzeichnis

Tabelle 1: Programme	19
Tabelle 2: Datenbanken	19
Tabelle 3: Musterkontingenztabelle zu den Überschneidungen der Genlisten.	23
Tabelle 4: Übersicht über verwendete Merkmale und SNPs	27
Tabelle 5: Anzahl der eQTL Ergebnisse der Gewebe und Merkmale anha	and
verschiedener P-Wert Grenzen	31

Anhangtabellenverzeichnis

Anhangtabelle 1: Übersicht verwendete Merkmale und SNPs 111
Anhangtabelle 2: Merkmalspaare Lebergewebe 112
Anhangtabelle 3: Merkmalspaare Blutgewebe 113
Anhangtabelle 4: Gemeinsame Gene aufgrund gemeinsamer SNPs in Leber 114
Anhangtabelle 5: Gemeinsame Gene aufgrund anderer Effekte in Leberg 116
Anhangtabelle 6: Merkmalspaare Lebergewebe 117
Anhangtabelle 7: Gemeinsame Gene aufgrund gemeinsamer SNPs in Blutg. 118
Anhangtabelle 8: Gemeinsame Gene aufgrund anderer Effekte in Blutg 120
Anhangtabelle 9: Merkmalspaare Blutgewebe 121
Anhangtabelle 10: Pleiotrope Gene mit Übersicht zu aktuellem Wissensstand
und bekannten Assoziationen
Literaturverzeichnis

- Rommens J, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou J, Et A (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science (80-) 245:1066–1073
- Beaudet A, Bowcock A, Buchwald M, Cavalli-Sforza L, Farrall M, King MC, Klinger K, Lalouel JM, Lathrop G, Naylor S (1986) Linkage of cystic fibrosis to two tightly linked DNA markers: Joint report from a collaborative study. Am J Hum Genet 39:681–693
- 3. Pulst SM (1999) Genetic Linkage Analysis. Arch Neurol 56:667
- 4. Weeks DE, Lathrop GM (1995) Polygenic disease: methods for mapping complex disease traits. Trends Genet 11:513–519
- 5. International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921
- 6. The International HapMap Consortium (2005) A haplotype map of the human genome. Nature 437:1299–1320
- Donaldson P, Daly A, Ermini L, Bevitt D (2016) Genetics of complex disease. Garland Science
- 8. Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95–108
- 9. Reich DE, Lander ES (2001) On the allelic spectrum of human disease. Trends Genet 17:502–510
- Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN (2003) Metaanalysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 33:177–182
- Rao DC (2008) An Overview of the Genetic Dissection of Complex Traits. Adv Genet 60:3–34
- Bush WS, Moore JH (2012) Chapter 11: Genome-Wide Association Studies.
 PLoS Comput Biol. https://doi.org/10.1371/journal.pcbi.1002822
- Pearson TA, Manolio TA (2008) How to Interpret a Genome-wide Association Study. JAMA 299:1335
- Marian AJ (2012) Molecular Genetic Studies of Complex Phenotypes. Transl Res 159:64–79
- 15. Buniello A, Macarthur JAL, Cerezo M, et al (2019) The NHGRI-EBI GWAS

Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res 47:D1005–D1012

- Klein RJ, Zeiss C, Chew EY, Henning AK, Sangiovanni JP, Mane SM, Ott J, Barnstable C, Hoh J (2005) Complement Factor H Polymorphism in Age-Related Macular Degeneration. Science (80-) 308:385–390
- 17. Durbin RM, Altshuler DL, Durbin RM, et al (2010) A map of human genome variation from population-scale sequencing. Nature 467:1061–1073
- Gibbs RA, Boerwinkle E, Doddapaneni H, et al (2015) A global reference for human genetic variation. Nature 526:68–74
- 19. Ding K, Kullo IJ (2007) Methods for the selection of tagging SNPs: a comparison of tagging efficiency and performance. Eur J Hum Genet 15:228–236
- Meng Z, Zaykin D V., Xu C-F, Wagner M, Ehm MG (2003) Selection of Genetic Markers for Association Analyses, Using Linkage Disequilibrium and Haplotypes. Am J Hum Genet 73:115–130
- 21. Reich DE, Cargill M, Bolk S, et al (2001) Linkage disequilibrium in the human genome. Nature 411:199–204
- 22. Slatkin M (2008) Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. Nat Rev Genet 9:477–85
- Kriebel J, Illig T, Grallert H (2012) Genomweite Assoziationsstudien (GWAS) -Möglichkeiten und Grenzen. BioSpektrum 18:508–510
- Maurano MT, Humbert R, Rynes E, et al (2012) Systematic localization of common disease-associated variation in regulatory DNA. Science (80-) 337:1190–1195
- 25. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 106:9362–7
- 26. Brodie A, Azaria JR, Ofran Y (2016) How far from the SNP may the causative genes be? Nucleic Acids Res 44:6046–6054
- 27. Zhu Z, Zhang F, Hu H, et al (2016) Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 48:481–487
- 28. Smemo S, Tena JJ, Kim K-H, et al (2014) Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature 507:371–375
- 29. Lewis CM, Ravindrarajah R, Munroe PB, et al (2007) Genome-wide association

study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661–678

- 30. Gibson G (2012) Rare and common variants: twenty arguments. Nat Rev Genet 13:135–145
- 31. Petretto E, Mangion J, Dickens NJ, et al (2006) Heritability and Tissue Specificity of Expression Quantitative Trait Loci. PLoS Genet 2:e172
- 32. Jansen RC, Nap J-P (2001) Genetical genomics: the added value from segregation. Trends Genet 17:388–391
- 33. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M (2009) Mapping complex disease traits with global gene expression. Nat Rev Genet 10:184–194
- 34. Folkersen L, Hooft F van't, Chernogubova E, et al (2010) Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease. Circ Cardiovasc Genet 3:365–373
- 35. Melé M, Ferreira PG, Reverter F, et al (2015) The human transcriptome across tissues and individuals. Science (80-) 348:660–665
- 36. The GTEx Consortium, Lonsdale J, Thomas J, et al (2013) The Genotype-Tissue Expression (GTEx) Project. Nat Genet 45:580–585
- 37. Westra HJ, Franke L (2014) From genome to function by studying eQTLs.Biochim Biophys Acta Mol Basis Dis 1842:1896–1902
- 38. Wong WL, Su X, Li X, Cheung CMG, Klein R, Cheng C-YY, Wong TY (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–e116
- 39. DeAngelis MM, Owen LA, Morrison MA, et al (2017) Genetics of age-related macular degeneration (AMD). Hum Mol Genet 26:R45–R50
- 40. Ferris F, Davis M (2005) A Simplified Severity Scale for Age-Related Macular Degeneration. Arch Ophthalmol 123:1570
- Fritsche LG, Fariss RN, Stambolian D, Abecasis GR, Curcio CA, Swaroop A (2014) Age-related macular degeneration: genetics and biology coming together. Annu Rev Genomics Hum Genet 15:151–71
- 42. Fritsche LG, Igl W, Bailey JNC, et al (2016) A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. Nat Genet 48:134–143
- 43. McCarty CA, Mukesh BN, Fu CL, Mitchell P, Wang JJ, Taylor HR (2001) Risk

Factors for Age-Related Maculopathy. Arch Ophthalmol 119:1455

- Age-Related Eye Disease Study Research Group A-REDSR (2000) Risk factors associated with age-related macular degeneration. A case-control study in the age-related eye disease study: Age-Related Eye Disease Study Report Number 3. Ophthalmology 107:2224–32
- 45. Bressler SB, Muñoz B, Solomon SD, West SK (2008) Racial Differences in the Prevalence of Age-Related Macular Degeneration. Arch Ophthalmol 126:241
- Ambati J, Ambati BK, Yoo SH, Ianchulev S, Adamis AP (2003) Age-Related Macular Degeneration: Etiology, Pathogenesis, and Therapeutic Strategies. Surv Ophthalmol 48:257–293
- 47. Delcourt C, Diaz J-L, Ponton-Sanchez A, Papoz L (1998) Smoking and Agerelated Macular Degeneration. Arch Ophthalmol 116:1031
- Ding X, Patel M, Chan CC (2009) Molecular pathology of age-related macular degeneration. Prog Retin Eye Res 28:1–18
- Beatty S, Koh H-H, Phil M, Henson D, Boulton M (2000) The Role of Oxidative Stress in the Pathogenesis of Age-Related Macular Degeneration. Surv Ophthalmol 45:115–134
- Klein R, Knudtson MD, Klein BEK, et al (2008) Inflammation, Complement Factor
 H, and Age-Related Macular Degeneration: The Multi-Ethnic Study of
 Atherosclerosis. Ophthalmology 115:1742–1749
- 51. Hammomnd CJ, Webster AR, Sneider H, Bird AC, Gilbert CE, Spector TD, Sharma S (2002) Genetic influence on early age-related maculophaty: A Twin Study. Evidence-Based Eye Care 3:162–163
- 52. Shahid H, Khan JC, Cipriani V, et al (2012) Age-related macular degeneration: the importance of family history as a risk factor. Br J Ophthalmol 96:427–31
- 53. Haines JL (2005) Complement Factor H Variant Increases the Risk of Age-Related Macular Degeneration. Science (80-) 308:419–421
- 54. Rivera A, Fisher SA, Fritsche LG, Keilhauer CN, Lichtner P, Meitinger T, Weber BHF (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
- 55. Yang Z, Camp NJ, Sun H, et al (2006) A variant of the HTRA1 gene increases susceptibility to age-related macular degeneration. Science 314:992–3
- 56. Johnson L V, Leitner WP, Staples MK, Anderson DH (2001) Complement

Activation and Inflammatory Processes in Drusen Formation and Age Related Macular Degeneration. Exp Eye Res 73:887–896

- 57. Mullins RF, Aptsiauri N, Hageman GS (2001) Structure and composition of drusen associated with glomerulonephritis: Implications for the role of complement activation in drusen biogenesis. Eye 15:390–395
- 58. Paun CC, Ersoy L, Schick T, Groenewoud JMM, Lechanteur YT, Fauser S, Hoyng CB, de Jong EK, den Hollander AI (2015) Genetic Variants and Systemic Complement Activation Levels Are Associated With Serum Lipoprotein Levels in Age-Related Macular Degeneration. Investig Opthalmology Vis Sci 56:7766
- 59. Wang Y, Wang M, Zhang X, Zhang Q, Nie J, Zhang M, Liu X, Ma L (2016) The association between the lipids levels in blood and risk of age-related macular degeneration. Nutrients 8:1–15
- Cougnard-Grégoire A, Delyfer M-N, Korobelnik J-F, Rougier M-B, Le Goff M, Dartigues J-F, Barberger-Gateau P, Delcourt C (2014) Elevated High-Density Lipoprotein Cholesterol and Age-Related Macular Degeneration: The Alienor Study. PLoS One 9:e90973
- Tomany SC, Wang JJ, van Leeuwen R, Klein R, Mitchell P, Vingerling JR, Klein BEK, Smith W, de Jong PTVM (2004) Risk factors for incident age-related macular degeneration: Pooled findings from 3 continents. Ophthalmology 111:1280–1287
- Cuvillier O (2002) Sphingosine in apoptosis signaling. Biochim Biophys Acta -Mol Cell Biol Lipids 1585:153–162
- 63. Downes CP, Gray A, Lucocq JM (2005) Probing phosphoinositide functions in signaling and membrane trafficking. Trends Cell Biol 15:259–268
- Kiel C, Lastrucci C, Luthert PJ, Serrano L (2017) Simple and complex retinal dystrophies are associated with profoundly different disease networks. Sci Rep 7:1–10
- 65. Stearns FW (2010) One hundred years of pleiotropy: a retrospective. Genetics 186:767–73
- 66. Wagner GP, Zhang J (2011) The pleiotropic structure of the genotype–
 phenotype map: the evolvability of complex organisms. Nat Rev Genet 12:204–
 213
- 67. Sivakumaran S, Agakov F, Theodoratou E, Prendergast JG, Zgaga L, Manolio T, Rudan I, McKeigue P, Wilson JF, Campbell H (2011) Abundant Pleiotropy in

Human Complex Diseases and Traits. Am J Hum Genet 89:607-618

- 68. Cotsapas C, Voight BF, Rossin E, et al (2011) Pervasive Sharing of Genetic Effects in Autoimmune Disease. PLoS Genet 7:e1002254
- Grassmann F, Kiel C, Zimmermann ME, et al (2017) Genetic pleiotropy between age-related macular degeneration and 16 complex diseases and traits. Genome Med 9:1–13
- Bastian M, Heymann S, Jacomy M (2009) Gephi: An Open Source Software for Exploring and Manipulating Networks. Icwsm 361–362
- R Core Development Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://doi.org/10.1007/978-3-540-74686-7
- 72. Zerbino DR, Achuthan P, Akanni W, et al (2018) Ensembl 2018. Nucleic Acids Res 46:D754–D761
- MacArthur J, Bowler E, Cerezo M, et al (2017) The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res 45:D896–D901
- 74. Machiela MJ, Chanock SJ (2015) LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics 31:3555–3557
- 75. Blake JA, Richardson JE, Bult CJ, et al (2003) MGD: The mouse genome database. Nucleic Acids Res 31:193–195
- 76. Hamosh A, Scott AF, Amberger J, Valle D, McKusick VA (2000) Online Mendelian Inheritance in Man (OMIM). Hum Mutat 15:57–61
- 77. Brown GR, Hem V, Katz KS, et al (2015) Gene: a gene-centered information resource at NCBI. Nucleic Acids Res 43:D36–D42
- Consortium TU (2019) UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 47:D506–D515
- 79. Grassmann F, Fritsche LG, Keilhauer CN, Heid IM, Weber BHF (2012) Modelling the genetic risk in age-related macular degeneration. PLoS One 7:e37979
- Benjamini Y, Hochberg Y, Benjamini, Yoav HY (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 57:289–300
- 81. Warnes AGR, Bolker B, Bonebakker L, Huber W, Liaw A, Lumley T, Magnusson A, Moeller S, Schwartz M (2019) Package gplots: Various R programming tools

for plotting data.

- 82. Neuwirth E (2014) ColorBrewer Palettes. Packag RColorBrewer 1.1-2:1–5
- 83. Shabalin AA (2012) Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. Bioinformatics 28:1353–1358
- 84. Fisher RA (1922) On the Interpretation of χ 2 from Contingency Tables, and the Calculation of P. J R Stat Soc 85:87
- 85. Gu Z, Gu L, Eils R, Schlesner M, Brors B (2014) Circlize implements and enhances circular visualization in R. Bioinformatics 30:2811–2812
- Ripatti S, Tikkanen E, Orho-Melander M, et al (2010) A multilocus genetic risk score for coronary heart disease: Case-control and prospective cohort analyses. Lancet 376:1393–1400
- Zhou YJ, Hong SC, Yang Q, Yin RX, Cao XL, Chen WX (2015) Association of variants in CELSR2-PSRC1-SORT1 with risk of serum lipid traits, coronary artery disease and ischemic stroke. Int J Clin Exp Pathol 8:9543–9551
- Samani NJ, Sci FM, Erdmann J, et al (2007) UKPMC Funders Group Author Manuscript Genomewide Association Analysis of Coronary Artery Disease. English J 357:443–453
- 89. Arvind P, Nair J, Jambunathan S, Kakkar V V., Shanker J (2014) CELSR2-PSRC1-SORT1 gene expression and association with coronary artery disease and plasma lipid levels in an Asian Indian cohort. J Cardiol 64:339–346
- Lanktree MB, Anand SS, Yusuf S, Hegele RA (2009) Replication of genetic associations with plasma lipoprotein traits in a multiethnic sample. J Lipid Res 50:1487–1496
- 91. Jiménez-Casas A (2008) Invariant regions and global existence for a phase field model. Discret Contin Dyn Syst Ser S 1:273–281
- 92. Dichgans M, Malik R, König IR, Rosand J, Clarke R, Gretarsdottir S, Thorleifsson G, Mitchell BD, Assimes TL, Levi C (2015) Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. 45:24–36
- 93. Surakka I, Horikoshi M, Mägi R, et al (2015) The impact of low-frequency and rare variants on lipid levels. Nat Genet 47:589–597
- 94. Lettre G, Palmer CD, Young T, et al (2011) Genome-Wide association study of coronary heart disease and its risk factors in 8,090 african americans: The nhlbi CARe project. PLoS Genet. https://doi.org/10.1371/journal.pgen.1001300

- 95. Kathiresan S, Willer CJ, Peloso G, et al (2010) Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet 41:56–65
- 96. Breiderhoff T, Kjolby M, Madsen P, et al (2010) Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export. Cell Metab 12:213–223
- 97. Teslovich TM, Musunuru K, Smith A V., et al (2010) Biological, clinical and population relevance of 95 loci for blood lipids. Nature 466:707–713
- 98. Peden JF, Hopewell JC, Saleheen D, et al (2011) A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Nat Genet 43:339–346
- 99. Dawn M Waterworth, Sally L Ricketts, , Kijoung Song(2015) Joint effects of common genetic variants from multiple genes and pathways on the risk of premature coronary artery disease. 85:1–27
- 100. Ligthart S, Vaez A, Hsu YH, Stolk R, Uitterlinden AG, Hofman A, Alizadeh BZ, Franco OH, Dehghan A (2016) Bivariate genome-wide association study identifies novel pleiotropic loci for lipids and inflammation. BMC Genomics 17:1– 10
- 101. Below JE, Parra EJ, Gamazon ER, et al (2016) Meta-analysis of lipid-traits in Hispanics identifies novel loci, population-specific effects, and tissue-specific enrichment of eQTLs. Sci Rep 6:1–13
- 102. Nagy R, Boutin TS, Marten J, et al (2017) Exploration of haplotype research consortium imputation for genome-wide association studies in 20,032 Generation Scotland participants. Genome Med 9:1
- 103. Liu C-W, Sun J-Q, Wu J-Z, Lin Q-Z, Pan S-L, Wu J, Guo T, Yin R-X, Shen S-W, Shi G-Y (2014) Association of the ST3GAL4 rs11220462 polymorphism and serum lipid levels in the Mulao and Han populations. Lipids Health Dis 13:123
- Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Iris M (2009) Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. 41:47–55
- 105. Chiara Sabatti1, 2, 28, Susan K Service3, 28, Anna-Liisa Hartikainen4, Anneli Pouta5, Samuli Ripatti6, Jae Brodsky2, Chris G Jones3, 7, Noah A Zaitlen7, Teppo Varilo8, 9, Marika Kaakinen10, Ulla Sovio11, Aimo Ruokonen12, Jaana Laitinen13, Eveliina Jakkula and LP (2009) Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. Nat Genet 41:35–

46

- 106. Liang J, Le TH, Edwards DRV, et al (2017) Single-trait and multi-trait genomewide association analyses identify novel loci for blood pressure in Africanancestry populations. PLoS Genet. https://doi.org/10.1371/journal.pgen.1006728
- 107. Levy D, Ehret GB, Rice K, et al (2009) Genome-wide association study of blood pressure and hypertension. Nat Genet 41:677–687
- 108. Franceschini N, Fox E, Zhang Z, et al (2013) Genome-wide association analysis of blood-pressure traits in african-ancestry individuals reveals common associated genes in African and Non-African populations. Am J Hum Genet 93:545–554
- 109. Georg B Ehret, Teresa Ferreira DIC (2016) The genetics of blood pressure regulation and its target organs from association studies in 342,415 individuals. Nat Genet 48:1171–1184
- 110. Natekar A, Olds RL, Lau MW, Min K, Imoto K, Slavin TP (2014) Elevated blood pressure: Our family's fault? The genetics of essential hypertension. World J Cardiol 6:327
- 111. Shen X, Espin-Garcia O, Qiu X, Brhane Y, Liu G, Xu W (2014) Haplotype approach for association analysis on hypertension. BMC Proc 8:1–5
- 112. Wilson PWF (1990) High-density lipoprotein, low-density lipoprotein and coronary artery disease. Am J Cardiol 66:7–10
- 113. Rapsomaniki E, Timmis A, George J, et al (2014) Blood pressure and incidence of twelve cardiovascular diseases: Lifetime risks, healthy life-years lost, and agespecific associations in 1.25 million people. Lancet 383:1899–1911
- 114. Gill D, Georgakis MK, Zuber V, Karhunen V, Burgess S, Malik R, Dichgans M (2020) Genetically predicted midlife blood pressure and coronary artery disease risk: Mendelian randomization analysis. J Am Heart Assoc. https://doi.org/10.1161/JAHA.120.016773
- 115. Van Der Wal AC, Becker AE, van der Loos CM, Das PK (1994) Site of intimal rupture or erosion of thrombosed coronary atherosclerotic plaques is characterized by an inflammatory process irrespective of the dominant plaque morphology. Circulation 89:36–44
- 116. Badimon L, Vilahur G (2012) LDL-cholesterol versus HDL-cholesterol in the atherosclerotic plaque: inflammatory resolution versus thrombotic chaos. Ann N

Y Acad Sci 1254:18–32

- Chistiakov DA, Melnichenko AA, Myasoedova VA, Grechko A V., Orekhov AN (2017) Mechanisms of foam cell formation in atherosclerosis. J Mol Med 95:1153–1165
- 118. Fontbonne A, Eschwge E, Cambien F, Richard J-L, Ducimetire P, Thibult N, Warnet J-M, Claude J-R, Rosselin G-E (1989) Hypertriglyceridaemia as a risk factor of coronary heart disease mortality in subjects with impaired glucose tolerance or diabetes. Diabetologia 32:300–304
- 119. Sarwar N, Danesh J, Eiriksdottir G, Sigurdsson G, Wareham N, Bingham S, Boekholdt SM, Khaw K-T, Gudnason V (2006) Triglycerides and the Risk of Coronary Heart Disease. Circulation 115:450–458
- Dumon H, Diez M, Nguyen P, Leray V, Bloc'h J Le, Serisier S, Siliart B (2008)
 Liver lipid metabolism. J Anim Physiol Anim Nutr (Berl) 92:272–283
- 121. Broe ME DE, Borgers M, Wieme RJ (1975) The Serparation And Characterization of Liver Plasma Membrane Fragments circulating in the Blood of Patients with Cholestasis. Clin Chim Acta 59:369–372
- 122. Kubes P, Jenne C (2018) Immune Responses in the Liver. Annu Rev Immunol 36:247–277
- 123. Strunz T, Nahkuri S, Maugeais C, Nogoceke E, Grassmann F, Fauser S, Weber BHF, Souza-Costa D, Strunz T, Gayán J (2018) A mega-analysis of expression quantitative trait loci (eQTL) provides insight into the regulatory architecture of gene expression variation in liver. Sci Rep 8:5865
- 124. McLeod DS, Grebe R, Bhutto I, Merges C, Baba T, Lutty GA (2009) Relationship between RPE and Choriocapillaris in Age-Related Macular Degeneration. Investig Opthalmology Vis Sci 50:4982
- 125. Grunwald JE, Metelitsina TI, DuPont JC, Ying G-S, Maguire MG (2005) Reduced Foveolar Choroidal Blood Flow in Eyes with Increasing AMD Severity. Investig Opthalmology Vis Sci 46:1033
- 126. Polak K, Polska E, Luksch A, Dorner G, Fuchsjäger-Mayrl G, Findl O, Eichler H-G, Wolzt M, Schmetterer L (2003) Choroidal blood flow and arterial blood pressure. Eye 17:84–88
- 127. The GTEx Consortium (2015) The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science (80-) 348:648–660
- 128. Locke AE, Kahali B, Berndt SI, et al (2015) Genetic studies of body mass index

yield new insights for obesity biology. Nature 518:197–206

- 129. Mizuno A, Okada Y (2019) Biological characterization of expression quantitative trait loci (eQTLs) showing tissue-specific opposite directional effects. Eur J Hum Genet 1–12
- 130. The GTEx Consortium (2020) The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science (80-) 1–32
- 131. Zhang T, Choi J, Kovacs MA, et al (2018) Cell-type-specific eQTL of primary melanocytes facilitates identification of melanoma susceptibility genes. Genome Res 28:1621–1635
- Dimas AS, Deutsch S, Stranger BE, et al (2009) Common Regulatory Variation Impacts Gene Expression in a Cell Type–Dependent Manner. Science (80-) 325:1246–1250
- 133. Nica AC, Parts L, Glass D, et al (2011) The Architecture of Gene Regulatory Variation across Multiple Human Tissues: The MuTHER Study. PLoS Genet 7:e1002003
- 134. Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A, Ellis P, Langford C, Vannberg FO, Knight JC (2012) Genetics of gene expression in primary immune cells identifies cell type–specific master regulators and roles of HLA alleles. Nat Genet 44:502–510
- GTEx Consortium (2017) Genetic effects on gene expression across human tissues. Nature 550:204–213
- 136. Grundberg E, Small KS, Hedman ÅK, et al (2012) Mapping cis- and transregulatory effects across multiple tissues in twins. Nat Genet 44:1084–1089
- 137. Wittkopp PJ, Haerum BK, Clark AG (2008) Regulatory changes underlying expression differences within and between Drosophila species. Nat Genet 40:346–350
- Wray GA (2007) The evolutionary significance of cis-regulatory mutations. Nat Rev Genet 8:206–216
- 139. Small KS, Hedman ÅK, Grundberg E, et al (2011) Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. Nat Genet 43:561–564
- 140. Kang HM, Ye C, Eskin E (2008) Accurate Discovery of Expression Quantitative Trait Loci Under Confounding From Spurious and Genuine Regulatory Hotspots Hyun. Genetics. https://doi.org/10.1534/genetics.108.094201

- 141. Breitling R, Li Y, Tesson BM, et al (2008) Genetical Genomics: Spotlight on QTL Hotspots. PLoS Genet 4:e1000232
- 142. Freund MK, Burch KS, Shi H, et al (2018) Phenotype-Specific Enrichment of Mendelian Disorder Genes near GWAS Regions across 62 Complex Traits. Am J Hum Genet 103:535–552
- 143. Alper hesCter A, Johnson AM, Birtch AG, Moore FD (1968) Human C'3:
 Evidence for the Liver as the Primary Site of Synthesis. Science (80-) 163:286–288
- 144. Morris KM, Aden DP, Knowles BB, Colten HR (1982) Complement biosynthesis by the human hepatoma-derived cell line HepG2. J Clin Invest 70:906–13
- 145. MORGAN BP, GASQUE P (1997) Extrahepatic complement biosynthesis: where, when and why? Clin Exp Immunol 107:1–7
- 146. Galkina E, Ley K (2007) Leukocyte influx in atherosclerosis. Curr Drug Targets8:1239–48
- 147. Weber C, Zernecke A, Libby P (2008) The multifaceted contributions of leukocyte subsets to atherosclerosis: lessons from mouse models. Nat Rev Immunol 8:802–815
- Micale L, Fusco C, Augello B, Napolitano LMR, Dermitzakis ET, Meroni G, Merla G, Reymond A (2008) Williams-Beuren syndrome TRIM50 encodes an E3 ubiquitin ligase. Eur J Hum Genet 16:1038–1049
- 149. Peoples R, Franke Y, Wang YK, Pérez-Jurado L, Paperna T, Cisco M, Francke U (2000) A physical map, including a BAC/PAC clone contig, of the Williams-Beuren syndrome--deletion region at 7q11.23. Am J Hum Genet 66:47–68
- 150. Schubert C, Laccone F (2006) Williams-Beuren syndrome: Determination of deletion size using quantitative real-time PCR. Int J Mol Med 18:799–806
- 151. Merla G, Howald C, Henrichsen CN, Lyle R, Wyss C, Zabot M-T, Antonarakis SE, Reymond A (2006) Submicroscopic deletion in patients with Williams-Beuren syndrome influences expression levels of the nonhemizygous flanking genes. Am J Hum Genet 79:332–41
- Carrasco X, Castillo S, Aravena T, Rothhammer P, Aboitiz F (2005) Williams syndrome: Pediatric, neurologic, and cognitive development. Pediatr Neurol 32:166–172
- 153. Takeuchi D, Furutani M, Harada Y, Furutani Y, Inai K, Nakanishi T, Matsuoka R (2015) High prevalence of cardiovascular risk factors in children and adolescents

with Williams-Beuren syndrome. BMC Pediatr 15:126

- 154. Palacios-Verdú MG, Segura-Puimedon M, Borralleras C, Flores R, Campo M Del, Campuzano V, Pérez-Jurado LA (2015) Metabolic abnormalities in Williams-Beuren syndrome. J Med Genet 52:248–255
- 155. Zalzstein E, Moes CAF, Musewe NN, Freedom RM (1991) Spectrum of cardiovascular anomalies in Williams-Beuren syndrome. Pediatr Cardiol 12:219–223
- 156. Collins RT, Kaplan P, Somes GW, Rome JJ (2010) Long-Term Outcomes of Patients With Cardiovascular Abnormalities and Williams Syndrome. Am J Cardiol 105:874–878
- 157. Rein AJJT, Preminger TJ, Perry SB, Lock JE, Sanders SP (1993) Generalized arteriopathy in williams syndrome: An intravascular ultrasound study. J Am Coll Cardiol 21:1727–1730
- 158. Ergul Y, Nisli K, Kayserili H, Karaman B, Basaran S, Dursun M, Yilmaz E, Ergul N, Unal SN, Dindar A (2012) Evaluation of coronary artery abnormalities in williams syndrome patients using myocardial perfusion scintigraphy and CT angiography. Cardiol J 19:301–308
- 159. Palmieri S, Orsi E, Eller-Vainicher C, et al (2018) Bone involvement and mineral metabolism in Williams' syndrome. J Endocrinol Invest 42:337–344
- 160. Cairo S, Merla G, Urbinati F, Ballabio A, Reymond A (2001) WBSCR14, a gene mapping to the Williams-Beuren syndrome deleted region, is a new member of the Mlx transcription factor network. Hum Mol Genet 10:617–627
- 161. de Luis O, Valero MC, Pérez Jurado LA (2000) WBSCR14, a putative transcription factor gene deleted in Williams-Beuren syndrome: complete characterisation of the human gene and the mouse ortholog. Eur J Hum Genet 8:215–222
- 162. Dentin R, Benhamed F, Hainault I, Fauveau V, Foufelle F, Dyck JRB, Girard J, Postic C (2006) Liver-Specific Inhibition of ChREBP Improves Hepatic Steatosis and Insulin Resistance in ob/ob Mice. Diabetes 55:2159–2170
- 163. Lowery MC, Morris CA, Ewart A, Brothman LJ, Zhu XL, Leonard CO, Carey JC, Keating M, Brothman AR (1995) Strong correlation of elastin deletions, detected by FISH, with Williams syndrome: evaluation of 235 patients. Am J Hum Genet 57:49–53
- 164. Wagenseil JE, Ciliberto CH, Knutsen RH, Levy MA, Kovacs A, Mecham RP

(2009) Reduced vessel elasticity alters cardiovascular structure and function in newborn mice. Circ Res 104:1217–1224

- 165. Yamashina A, Tomiyama H et al. (2003) Brachial-Ankle Pulse Wave Velocity as a Marker of Atherosclerotic Vascular Damage and Cardiovascular Risk. Hypertens Res 26:615–622
- 166. Lin YC, Chen CC, Cheng CJ, Yang RB (2011) Domain and functional analysis of a novel breast tumor suppressor protein, SCUBE2. J Biol Chem 286:27039– 27047
- 167. Lin YC, Chao TY, Yeh CT, Roffler SR, Kannagi R, Yang RB (2017) Endothelial SCUBE2 interacts with VEGFR2 and regulates VEGF-induced angiogenesis. Arterioscler Thromb Vasc Biol 37:144–155
- 168. Facemire CS, Nixon AB, Griffiths R, Hurwitz H, Thomas M (2010) Vascular Endothelial Growth Factor Receptor 2 (Vegfr2) Controls Blood Pressure by Regulating Nitric Oxide Synthase Expression. 54:652–658
- 169. Robinson ES, Matulonis UA, Ivy P, Berlin ST, Tyburski K, Penson RT, Humphreys BD (2010) Rapid development of hypertension and proteinuria with cediranib, an oral vascular endothelial growth factor receptor inhibitor. Clin J Am Soc Nephrol 5:477–83
- 170. Tukachinsky H, Salic A, Liu J, Kuzmickas RP, Jao CY (2012) Dispatched and Scube Mediate the Efficient Secretion of the Cholesterol-Modified Hedgehog Ligand. Cell Rep 2:308–320
- 171. Johnson JLFA, Hall TE, Dyson JM, Sonntag C, Ayers K, Berger S, Gautier P, Mitchell C, Hollway GE, Currie PD (2012) Scube activity is necessary for Hedgehog signal transduction in vivo. Dev Biol 368:193–202
- 172. Schmidt-Heck W, Boettger J, Klöting N, et al (2016) Hedgehog signaling is a potent regulator of liver lipid metabolism and reveals a GLI-code associated with steatosis. Elife 5:1–28
- 173. Lever E, Sheer D (2007) Disruption of hedgehog signalling in ApoE-/- mice reduces plasma lipid levels, but increases atherosclerosis due to enhanced lipid uptake by macrophages. J Pathol 212:420–428
- 174. Ali H, Aprilia D, Arizal C, Bakti I, Rahmadea P (2019) Upregulation of SCUBE2 expression in dyslipidemic type 2 diabetes mellitus is associated with endothelin-1. 13:2869–2872
- 175. Wu M, Huang C, Lin J, Wang S, Peng C, Cheng H, Tang C (2014) Endothelin-1

promotes vascular endothelial growth factor-dependent angiogenesis in human chondrosarcoma cells. Oncogene 33:1725–1735

- 176. Lee EJ, Hwang I, Kim G, Moon D, Kang SY, Hwang I, Lee S, Marie PJ, Kim H (2019) Endothelin-1 Augments Therapeutic Potency of Human Mesenchymal Stem Cells via CDH2 and VEGF Signaling. Mol Ther Methods Clin Dev 13:503– 511
- 177. Kjolby M, Nielsen MS, Petersen CM (2015) Sortilin, Encoded by the Cardiovascular Risk Gene SORT1, and Its Suggested Functions in Cardiovascular Disease. Curr Atheroscler Rep 17:1–9
- 178. Andaleon A, Mogil LS, Wheeler HE (2019) Genetically regulated gene expression underlies lipid traits in Hispanic cohorts. PLoS One. https://doi.org/10.1371/JOURNAL.PONE.0220827
- 179. Fontes JD, Benjamin EJ, Stewart AFR, et al (2011) Eight genetic loci associated with variation in lipoprotein-associated phospholipase A2 mass and activity and coronary heart disease: meta-analysis of genome-wide association studies from five community-based studies. Eur Heart J 33:238–251
- 180. Suchindran S, Rivedal D, Guyton JR, Milledge T, Gao X, Benjamin A, Rowell J, Ginsburg GS, McCarthy JJ (2010) Genome-wide association study of Lp-PLA2 activity and mass in the framingham heart study. PLoS Genet. https://doi.org/10.1371/journal.pgen.1000928
- 181. Takahara N, Kashiwagi A, Maegawa H, Shigeta Y (1996) Lysophosphatidylcholine stimulates the expression and production of MCP-1 by human vascular endothelial cells. Metabolism 45:559–564
- 182. Kume N, Cybulsky MI, Gimbrone MA (1992) Lysophosphatidylcholine, a component of atherogenic lipoproteins, iLysophosphatidylcholine, a component of atherogenic lipoproteins, induces mononuclear leukocyte adhesion molecules in cultured human and rabbit arterial endothelial cells. J Clin Invest 90:1138–44
- 183. Zalewski A, Macphee C (2005) Role of Lipoprotein-Associated Phospholipase A
 2 in Atherosclerosis. Arterioscler Thromb Vasc Biol 25:923–931
- 184. Mohler ER, Ballantyne CM, Davidson MH, Hanefeld M, Ruilope LM, Johnson JL, Zalewski A (2008) The Effect of Darapladib on Plasma Lipoprotein-Associated Phospholipase A2 Activity and Cardiovascular Biomarkers in Patients With Stable Coronary Heart Disease or Coronary Heart Disease Risk Equivalent: The Results of a Multicenter, Randomized, Double-BI. J Am Coll Cardiol 51:1632–

1641

- Patel KM, Strong A, Tohyama J, Jin X, Morales CR, Billheimer J, Millar J, Kruth H, Rader DJ (2015) Macrophage sortilin promotes LDL uptake, foam cell formation, and atherosclerosis. Circ Res 116:789–796
- 186. Mortensen MB, Kjolby M, Gunnersen S, Larsen J V., Palmfeldt J, Falk E, Nykjaer A, Bentzon JF (2014) Targeting sortilin in immune cells reduces proinflammatory cytokines and atherosclerosis. J Clin Invest 124:5317–5322
- 187. Hu F, Padukkavidana T, Vægter CB, Brady OA, Zheng Y, Mackenzie IR, Feldman HH, Nykjaer A, Strittmatter SM (2010) Sortilin-mediated endocytosis determines levels of the frontotemporal dementia protein, progranulin. Neuron 68:654–667
- 188. Lee N, Bluher M, Pi K-B, et al (2008) Serum Progranulin Concentrations May Be Associated With Macrophage Infiltration Into Omental Adipose Tissue. Diabetes 58:627–636
- 189. Kawase R, Ohama T, Matsuyama A, et al (2013) Deletion of progranulin exacerbates atherosclerosis in ApoE knockout mice. Cardiovasc Res 100:125– 133
- 190. Gustafsen C, Kjolby M, Nyegaard M, et al (2014) The hypercholesterolemia-risk gene SORT1 facilitates PCSK9 secretion. Cell Metab 19:310–318
- 191. Denis M, Marcinkiewicz J, Zaid A, Gauthier D, Poirier S, Lazure C, Seidah NG, Prat A (2012) Gene inactivation of proprotein convertase subtilisin/kexin type 9 reduces atherosclerosis in mice. Circulation 125:894–901
- 192. Abifadel M, Rabès JP, Devillers M, Munnich A, Erlich D, Junien C, Varret M, Boileau C (2009) Mutations and polymorphisms in the proprotein convertase subtilisin kexin 9(PCSK9) gene in cholesterol metabolism and disease. Hum Mutat 30:520–529
- 193. Stein EA, Mellis S, Yancopoulos GD, et al (2012) Effect of a Monoclonal Antibody to PCSK9 on Low-Density Lipoprotein Cholesterol. Am Med Assoc 366:1108– 1118
- 194. Raedler LA (2016) Praluent (Alirocumab): First PCSK9 Inhibitor Approved by the FDA for Hypercholesterolemia. Am Heal drug benefits 9:123–6
- 195. Lamb YN (2021) Inclisiran : First Approval. Drugs 81:389–395
- 196. Strong A, Ding Q, Edmondson AC, et al (2012) Hepatic sortilin regulates both apolipoprotein B secretion and LDL catabolism. J Clin Invest 122:2807–2816

- Musunuru K, Strong A, Frank-Kamenetsky M, et al (2010) From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature 466:714– 719
- 198. Conlon DM (2019) Role of sortilin in lipid metabolism. Curr Opin Lipidol 30:198–204
- Strong A, Patel K, Rader DJ (2014) Sortilin and lipoprotein metabolism: making sense out of complexity. Curr Opin Lipidol 25:350–7
- 200. Albert FW, Kruglyak L (2015) The role of regulatory variation in complex traits and disease. Nat Rev Genet 16:197–212
- 201. Li Y-P, Gao S, Song L-B, et al (2018) Structure of Schlafen13 reveals a new class of tRNA/rRNA- targeting RNase engaged in translational control. Nat Commun. https://doi.org/10.1038/s41467-018-03544-x
- 202. Perez-Lamigueiro MA, de la Casa-Esperón E, Naik S, Ayers G, Bustos O, Casola C, Chippindale PT, Pritham EJ (2009) Evolution of the Schlafen genes, a gene family associated with embryonic lethality, meiotic drive, immune processes and orthopoxvirus virulence. Gene 447:1–11
- 203. Puck A, Aigner R, Modak M, Cejka P, Blaas D, Stöckl J (2015) Expression and regulation of Schlafen (SLFN) family members in primary human monocytes, monocyte-derived dendritic cells and T cells. Results Immunol 5:23–32
- 204. Katsoulidis E, Carayol N, Woodard J, Konieczna I, Majchrzak-Kita B, Jordan A, Sassano A, Eklund EA, Fish EN, Platanias LC (2009) Role of schlafen 2 (SLFN2) in the generation of interferon α-induced growth inhibitory responses. J Biol Chem 284:25051–25064
- 205. Bonenfant D, Towbin H, Schindler P, Oostrum J Van (2006) Characterization of Histone H2A and H2B Variants and Their Post-translational Modifications by Mass Spectrometry *. Mol Cell Proteomics 53 541–552
- 206. Bhasin M, Reinherz EL, Reche PA (2006) Recognition and Classification of Histones Using Support Vector Machine. J Comput Biol 13:102–112
- 207. Molden RC, Bhanu N V., LeRoy G, Arnaudo AM, Garcia BA (2015) Multi-faceted quantitative proteomics analysis of histone H2B isoforms and their modifications. Epigenetics and Chromatin 8:1–17
- 208. Santoro SW, Dulac C (2012) The activity-dependent histone variant H2BE modulates the life span of olfactory neurons. Elife 1:1–32
- 209. Obeidat M, Miller S, Probert K, et al (2013) GSTCD and INTS12 Regulation and

Expression in the Human Lung. PLoS One 8:1-16

- 210. Kheirallah AK, Moor CH De, Faiz A, Sayers I, Hall IP (2017) Lung function associated gene Integrator Complex subunit 12 regulates protein synthesis pathways. 1–20
- 211. Stadelmayer B, Micas G, Gamot A, et al (2014) Integrator complex regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. Nat Commun 5:5531
- 212. Gardini A, Baillat D (2014) Integrator Regulates Transcriptional Initiation and Pause Release Following Activation. Mol Cell 56:128–139
- 213. Ezzeddine N, Chen J, Waltenspiel B, Burch B, Albrecht T, Zhuo M, Warren WD, Marzluff WF, Wagner EJ (2011) A subset of Drosophila integrator proteins is essential for efficient U7 snRNA and spliceosomal snRNA 3'-end formation. Mol Cell Biol 31:328–41
- 214. Cooch N, Näär AM, Baillat D, Hakimi M-A, Shiekhattar R, Shilatifard A (2005) Integrator, a Multiprotein Mediator of Small Nuclear RNA Processing, Associates with the C-Terminal Repeat of RNA Polymerase II. Cell 123:265–276
- 215. Battle A, Khan Z, Wang SH, Mitrano A, Ford MJ, Pritchard JK, Gilad Y (2015) Impact of regulatory variation from RNA to protein. Science (80-) 347:664–667
- Zhang X, Joehanes R, Chen BH, Huan T, Ying S, Munson PJ, Johnson AD, Levy
 D, O'Donnell CJ (2015) Identification of common genetic variants controlling transcript isoform variation in human whole blood. Nat Genet 47:345–352
- 217. He B, Shi J, Wang X, Jiang H, Zhu HJ (2020) Genome-wide pQTL analysis of protein expression regulatory networks in the human liver. BMC Biol 18:1–16
- 218. Gaunt TR, Shihab HA, Hemani G, et al (2016) Systematic identification of genetic influences on methylation across the human life course. Genome Biol 17:61
- 219. Pelikan RC, Kelly JA, Fu Y, et al (2018) Enhancer histone-QTLs are enriched on autoimmune risk haplotypes and influence gene expression within chromatin networks. Nat Commun. https://doi.org/10.1038/s41467-018-05328-9
- 220. Nica AC, Montgomery SB, Dimas AS, Stranger BE, Beazley C, Barroso I, Dermitzakis ET (2010) Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. PLoS Genet 6:e1000895
- 221. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C,

Plagnol V (2014) Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. PLoS Genet. https://doi.org/10.1371/journal.pgen.1004383

- 222. Richardson TG, Hemani G, Gaunt TR, Relton CL, Davey Smith G (2020) A transcriptome-wide Mendelian randomization study to uncover tissue-dependent regulatory mechanisms across the human phenome. Nat Commun 11:1–11
- 223. Gibson G, Powell JE, Marigorta UM (2015) Expression quantitative trait locus analysis for translational medicine. Genome Med 7:60
- 224. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ (2019) Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet 51:584–591
- 225. Fairfax BP, Humburg P, Makino S, et al (2014) Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. Science 343:1246949
- 226. Ye CJ, Feng T, Kwon H-K, et al (2014) Intersection of population variation and autoimmunity genetics in human T cell activation. Science (80-) 345:1254665– 1254665
- 227. Rotival M, Zeller T, Wild PS, et al (2011) Integrating Genome-Wide Genetic Variations and Monocyte Expression Data Reveals Trans-Regulated Gene Modules in Humans. PLoS Genet 7:e1002367
- 228. Ono H, Ogasawara O, Okubo K, Bono H (2017) RefEx, a reference gene expression dataset as a web tool for the functional analysis of genes. Sci Data. https://doi.org/10.1038/sdata.2017.105
- 229. Kim-Hellmuth S, Bechheim M, Pütz B, et al (2017) Genetic regulatory effects modified by immune activation contribute to autoimmune disease associations. Nat Commun 8:266
- 230. Chambers JC, Zhang W, Sehmi J, et al (2011) Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet 43:1131–1138
- 231. Deloukas P, Kanoni S, Willenborg C, et al (2013) Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 45:25–33
- 232. Nelson CP, Goel A, Butterworth AS, et al (2017) Association analyses based on false discovery rate implicate new loci for coronary artery disease. Nat Genet 49:1385–1391

- 233. Howson JMM, Zhao W, Barnes DR, et al (2017) Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. Nat Genet 49:1113–1119
- 234. Klarin D, Zhu QM, Emdin CA, et al (2017) Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. Nat Genet 49:1392–1397
- 235. Pim van der Harst NV, Van Der Harst P, Verweij N, Pim van der Harst NV (2017) Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. Circ Res. https://doi.org/10.17632/2zdd47c94h.1
- 236. Verweij N, Eppinga RN, Hagemeijer Y, Van Der Harst P (2017) Identification of 15 novel risk loci for coronary artery disease and genetic risk of recurrent events, atrial fibrillation and heart failure. Sci Rep. https://doi.org/10.1038/s41598-017-03062-8
- 237. Anand SS, Ripatti S, Roberts R, et al (2015) A comprehensive 1000 Genomes– based genome-wide association meta-analysis of coronary artery disease. Nat Genet 47:1121–1130
- 238. Ehret GB, Munroe PB, Rice KM, et al (2011) Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 478:103–109
- 239. The UK Biobank Cardio-metabolic Traits Consortium Blood Pressure Working Group, Warren HR, AI EE et, et al (2018) Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. Nat Genet 49:403–415
- 240. Kato N, Loh M, Takeuchi F, et al (2015) Trans-ancestry genome-wide association study identifies 12 genetic loci influencing blood pressure and implicates a role for DNA methylation. Nat Genet 47:1282–1293
- 241. Willer CJ, Schmidt EM, Sengupta S, et al (2013) Discovery and refinement of loci associated with lipid levels. Nat Genet 45:1274–1285
- 242. Hoffmann TJ, Theusch E, Haldar T, et al (2018) A large electronic-health-recordbased genome-wide study of serum lipids. Nat Genet 50:401–413
- 243. Levin C, Koren A, Pretorius E, et al (2015) Deleterious mutation in the FYB gene is associated with congenital autosomal recessive small-platelet thrombocytopenia. J Thromb Haemost 13:1285–1292
- 244. Li C, Jiao S, Wang G, et al (2015) The Immune Adaptor ADAP Regulates

Reciprocal TGF-β1-Integrin Crosstalk to Protect from Influenza Virus Infection. PLoS Pathog. https://doi.org/10.1371/journal.ppat.1004824

- 245. Yuan X, Waterworth D, Perry JRB, et al (2008) Population-Based Genome-wide Association Studies Reveal Six Loci Influencing Plasma Levels of Liver Enzymes. Am J Hum Genet 83:520–528
- 246. Keith L. Keene, Wei-Min Chen, Fang Chen, (2014) Genetic associations with plasma B12, B6, and folate levels in an ischemic stroke population from the Vitamin Intervention for Stroke Prevention (VISP) trial. Zootaxa. https://doi.org/10.3389/fpubh.2014.00112
- 247. Vandepoele K, Van Roy N, Staes K, Speleman F, Van Roy F (2005) A novel gene family NBPF: Intricate structure generated by gene duplications during primate evolution. Mol Biol Evol 22:2265–2274
- 248. Dong-Hui Chen, 1,* Alipi Naydenov, 2 Jacqueline L. Blankman (2013) Two Novel Mutations in ABHD12: Expansion of the Mutation Spectrum in PHARC and Assessment of their Functional Effects. Hum Mutat
- 249. Fiskerstrand T, H'Mida-Ben Brahim D, Johansson S, et al (2010) Mutations in ABHD12 cause the neurodegenerative disease PHARC: An inborn error of endocannabinoid metabolism. Am J Hum Genet 87:410–417
- 250. Spannbauer MM, Trautwein C (2009) Frequent in-frame somatic deletions activate gp130 in inflammatory hepatocellular tumors. Hepatology 49:1387– 1389
- 251. Heim D, Gil-Ibanez I, Herden J, Parplys AC, Borgmann K, Schmidt-Arras D, Lohse AW, Rose-John S, Wege H (2016) Constitutive gp130 activation rapidly accelerates the transformation of human hepatocytes via an impaired oxidative stress response. Oncotarget. https://doi.org/10.18632/oncotarget.10956
- 252. Korotaeva AA, Samoilova E V., Chepurnova DA, Zhitareva I V., Shuvalova YA, Prokazova N V. (2018) Soluble glycoprotein 130 is inversely related to severity of coronary atherosclerosis. Biomarkers 23:527–532
- 253. Chen H, Zhang X, Liao N, Wen F (2016) Increased levels of IL-6, sIL-6R, and sgp130 in the aqueous humor and serum of patients with diabetic retinopathy. 1005–1014
- 254. Kasperska-Zajac A, Grzanka A, Damasiewicz-Bodzek A (2015) IL-6 transsignaling in patients with chronic spontaneous urticaria. PLoS One 10:1–9
- 255. Xu H, Radabaugh T, Lu Z, Galligan M, Billheimer D, Vercelli D, Wright AL, Monks

TJ, Halonen M, Lau SS (2016) Exploration of early-life candidate biomarkers for childhood asthma using antibody arrays. Pediatr Allergy Immunol 27:696–701

- 256. Hirota H, Izumi M, Hamaguchi T, Sugiyama S, Murakami E, Kunisada K, Fujio Y, Oshima Y, Nakaoka Y (2004) Circulating interleukin-6 family cytokines and their receptors in patients with congestive heart failure. 237–241
- 257. Pacana T, Park C-B, Choi A, et al (2015) Activation of the GP130-STAT3 axis and its potential implications in nonalcoholic fatty liver disease. Am J Physiol Liver Physiol 308:G794–G803
- 258. Dechow T, Steidle S, Götze KS, et al (2014) GP130 activation induces myeloma and collaborates with MYC. 124:11–15
- 259. Giovanni Zuliani, MD PhD, Matteo Galvani, MD, Marcello Maggio (2011) Plasma Soluble SGP130 Levels Are Increased In Older Subjects With Metabolic Syndrome. The Role of Imsulin Resistance. 213:319–324
- 260. Stahl EA, Raychaudhuri S, Remmers EF, et al (2010) Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 42:508–514
- 261. Lin YS, Lin LC, Huang MH, Huang AJ, Huang YT (2009) Down-regulation of gp130 in nasopharyngeal carcinoma. Am J Rhinol Allergy 23:28–32
- 262. Inta I, Weber D, Grundt C, Veltkamp R, Winteroll S, Auffarth GU, Bettendorf M, Lemmer B, Schwaninger M (2009) Correlation of soluble gp130 serum concentrations with arterial blood pressure. J Hypertens 27:527–534
- 263. Lei, Yu, Schneider JA, Meissner A, Jager PL De, Bennett DA (2016) NIH Public Access. 72:15–24
- 264. Lambert J-C (2014) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 45:1452–1458
- 265. Wu L, Guo S, Yang D, et al (2014) Copy number variations of HLA-DRB5 is associated with systemic lupus erythematosus risk in Chinese Han population. Acta Biochim Biophys Sin (Shanghai) 46:155–160
- 266. Quandt JA, Huh J, Baig M, et al (2012) Myelin Basic Protein-Specific TCR/HLA-DRB5*01:01 Transgenic Mice Support the Etiologic Role of DRB5*01:01 in Multiple Sclerosis. J Immunol 189:2897–2908
- 267. Caillier SJ, Briggs F, Cree BAC, et al (2008) Uncoupling the Roles of HLA-DRB1 and HLA-DRB5 Genes in Multiple Sclerosis. J Immunol 181:5473–5480
- 268. Shih B, Bayat A (2012) Comparative genomic hybridisation analysis of keloid

tissue in Caucasians suggests possible involvement of HLA-DRB5 in disease pathogenesis. Arch Dermatol Res 304:241–249

- 269. Kudat H, Telci G, Sozen AB, Oguz F, Akkaya V, Ozcan M, Atilgan D, Carin M, Guven O (2006) The role of HLA molecules in susceptibility to chronic rheumatic heart disease. Int J Immunogenet 33:41–44
- 270. Ferrari R, Hernandez DG, Nalls MA, Rohrer JD, Ramasamy A KJ, Dobson-Stone C, Brooks WS, Schofield PR, Halliday GM, Hodges JR PO, Bartley L, Thompson E, Haan E, Hernández I, Ruiz A, Boada M, Borroni B P, et al (2014) Frontotemporal dementia and its subtypes: a genome-wide association study. Lancet Neurol 13:686–99
- 271. Zhu N, Luo F, Chen Q, Li N, Xiong H, Feng Y, Yang Z, Hou W (2014) Influence of HLA-DRB alleles on haemorrhagic fever with renal syndrome in a Chinese Han population in Hubei Province, China. Eur J Clin Microbiol Infect Dis 34:187– 195
- 272. Zhao LP, Alshiekh S, Zhao M, et al (2016) Next-generation sequencing reveals that HLA-DRB3, -DRB4, and -DRB5 may be associated with islet autoantibodies and risk for childhood type 1 diabetes. Diabetes 65:710–718
- 273. Sato H, Miyamoto T, Yogev L, et al (2006) Polymorphic alleles of the human MEI1 gene are associated with human azoospermia by meiotic arrest. J Hum Genet 51:533–540
- 274. Schessl J, Bach E, Rost S, Feldkirchner S, Kubny C, Müller S, Hanisch FG, Kress W, Schoser B (2014) Novel recessive myotilin mutation causes severe myofibrillar myopathy. Neurogenetics 15:151–156
- 275. Liu G-H, Guan T, Datta K, Coppinger J, Yates J, Gerace L (2009) Regulation of Myoblast Differentiation by the Nuclear Envelope Protein NET39. Mol Cell Biol 29:5800–5812
- 276. Huang ZP, Kataoka M, Chen J, et al (2015) Cardiomyocyte-enriched protein CIP protects against pathophysiological stresses and regulates cardiac homeostasis. J Clin Invest 125:4122–4134
- 277. Rosendahl J, Kirsten H, Hegyi E, et al (2018) Genome-wide association study identifies inversion in the CTRB1-CTRB2 locus to modify risk for alcoholic and non-alcoholic chronic pancreatitis. Gut 1855–1863
- 278. Kyo K, Muto T, Nagawa H, Lathrop GM, Nakamura Y (2001) Associations of distinct variants of the intestinal mucin gene MUC3A with ulcerative colitis and

Crohn's disease. J Hum Genet 46:5-20

- 279. Niu T, Liu Y, Zhang Y, Fu Q, Liu Z, Wang Z, Fu H, Xu J, Liu K (2016) Increased expression of MUC3A is associated with poor prognosis in localized clear-cell renal cell carcinoma. Oncotarget. https://doi.org/10.18632/oncotarget.10312
- 280. Kanai M, Akiyama M, Takahashi A, et al (2018) Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. Nat Genet 50:390–400
- 281. Takane K, Midorikawa Y, Yagi K, Sakai A, Aburatani H, Takayama T, Kaneda A (2014) Aberrant promoter methylation of PPP1R3C and EFHD1 in plasma of colorectal cancer patients. Cancer Med 3:1235–1245
- 282. Fox CS, Liu Y, White CC, et al (2012) Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. PLoS Genet. https://doi.org/10.1371/journal.pgen.1002695
- 283. Goettsch C, Hutcheson JD, Aikawa M, et al (2016) Sortilin mediates vascular calcification via its recruitment into extracellular vesicles. J Clin Invest 126:1323– 1336
- 284. Burgess A, Shah K, Hough O, Hynynen K (2016) Serum sortilin associates with aortic calcification and cardiovascular risk in men. 15:477–491
- 285. Köttgen A, Pattaro C, Böger CA, et al (2010) Multiple loci associated with indices of renal function and chronic kidney disease. Nat Genet 42:376–384
- 286. Roselli S, Pundavela J, Demont Y, Faulkner S, Keene S, Attia J, Jiang CC, Zhang XD, Walker MM, Hondermarck H (2015) Sortilin is associated with breast cancer aggressiveness and contributes to tumor cell adhesion and invasion. Oncotarget. https://doi.org/10.18632/oncotarget.3401
- 287. Andreasen N, Zetterberg H, Blennow K, Hansson O, Skoog I, Minthon L, Kettunen P, Nilsson S, Andersson C-H, Wallin A (2016) A Genetic Variant of the Sortilin 1 Gene is Associated with Reduced Risk of Alzheimer's Disease. J Alzheimer's Dis 53:1353–1363
- 288. Buttenschøn HN, Demontis D, Kaas M, et al (2015) Increased serum levels of sortilin are associated with depression and correlated with BDNF and VEGF. Transl Psychiatry. https://doi.org/10.1038/tp.2015.167
- 289. Jones GT, Tromp G, Kuivaniemi H, et al (2017) Meta-Analysis of Genome-Wide Association Studies for Abdominal Aortic Aneurysm Identifies Four New Disease-Specific Risk Loci. Circ Res 120:341–353

- 290. Einarsdottir E, Grauers A, Wang J, et al (2017) CELSR2 is a candidate susceptibility gene in idiopathic scoliosis. PLoS One 12:1–14
- 291. Huang CH, Reid ME, Chen Y (1995) Identification of a partial internal deletion in the RH locus causing the human erythrocyte D--phenotype. Blood 86:784–790
- 292. Mouro I, Colin Y, Cartron BCJ, Cabanel A (1993) Molecular genetic basis of the system. 5:3–6
- 293. Polin H, Brisner M, Reiter A, Danzer M (2018) Identification of four novel RHD alleles by weakened antigen D expression. Transfusion 58:267–268
- 294. Garcia F, Rodriguez MA, Goldman M, et al (2015) New RHD variant alleles. Transfusion 55:427–429
- 295. Chambers JC, Zhang W, Sehmi J, et al (2012) Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet 43:1131–1138
- 296. Song J, Xue C, Preisser JS, Cramer DW, Houck KL, Liu G, Folsom AR, Couper D, Yu F, Dong JF (2016) Association of single nucleotide polymorphisms in the ST3GAL4 Gene with VWF antigen and factor VIII activity. PLoS One 11:1–14
- 297. Monteerarat Y, Suptawiwat O, Boonarkart C, Uiprasertkul M, Auewarakul P, Viprakasit V (2010) Inhibition of H5N1 highly pathogenic influenza virus by suppressing a specific sialyltransferase. Arch Virol 155:889–893
- 298. Anne Louise Sørensen, Viktoria Rumjantseva, Sara Nayeb-Hashemi, Henrik Clausen, John H. Hartwig, Hans H. Wandall, and Karin M. Hoffmeister (2009) Role of sialic acid for platelet life span: exposure of -galactose results in the\nrapid clearance of platelets from the circulation by asialoglycoprotein\nreceptor–expressing liver macrophages and hepatocytes. Blood 114:1645–1654
- 299. Inafuku S, Noda K, Amano M, Ohashi T, Yoshizawa C, Saito W, Murata M, Kanda A, Nishimura SI, Ishida S (2015) Alteration of N-glycan profiles in diabetic retinopathy. Investig Ophthalmol Vis Sci 56:5316–5322
- 300. Jian Y, Weixia Z, Yanying X, et al (2010) Altered mRNA expressions of sialyltransferases in human gastric cancer tissues. Med Oncol 29:84–90
- 301. Gomes C, Osório H, Pinto MT, Campos D, Oliveira MJ, Reis CA (2013) Expression of ST3GAL4 Leads to SLexExpression and Induces c-Met Activation and an Invasive Phenotype in Gastric Carcinoma Cells. PLoS One 8:1–13
- 302. Gallagher PG, Tse WT, Costa F, Scarpa A, Boivin P, Delaunay J, Forget BG

(1991) A splice site mutation of the β -spectrin gene causing exon skipping in hereditary elliptocytosis associated with a truncated β -spectrin chain. J Biol Chem 266:15154–15159

- 303. Park J, Jeong DC, Yoo J, et al (2016) Mutational characteristics of ANK1 and SPTB genes in hereditary spherocytosis. Clin Genet 90:69–78
- 304. Ding H, Xu Y, Bao X, Wang X, Cui G, Wang W, Hui R, Wang DW (2010) Confirmation of genomewide association signals in chinese han population reveals risk loci for ischemic stroke. Stroke 41:177–180
- 305. Chesi A, Mitchell JA, Kalkwarf HJ, et al (2017) A Genomewide Association Study Identifies Two Sex-Specific Loci, at SPTB and IZUMO3, Influencing Pediatric Bone Mineral Density at Multiple Skeletal Sites. J Bone Miner Res 32:1274–1281
- 306. Li A, Davila S, Furu L, Qian Q, Tian X, Kamath PS, King BF, Torres VE, Somlo S (2003) Mutations in PRKCSH cause isolated autosomal dominant polycystic liver disease. Am J Hum Genet 72:691–703
- 307. Drenth JPH, Tahvanainen E, Te Morsche RHM, Tahvanainen P, Kääriäinen H, Höckerstedt K, Van De Kamp JM, Breuning MH, Jansen JBMJ (2004) Abnormal Hepatocystin Caused by Truncating PRKCSH Mutations Leads to Autosomal Liver Disease. Hepatology 39:924–931
- 308. Rimol LM, Agartz I, Djurovic S, et al (2010) Sex-dependent association of common variants of microcephaly genes with brain structure. Proc Natl Acad Sci 107:384–388
- 309. Jiang LLL, Yin J, Ye L, et al (2014) Novel risk loci for rheumatoid arthritis in han chinese and congruence with risk variants in europeans. Arthritis Rheumatol 66:1121–1132
- 310. Schwabe G (2008) Primary Ciliary Dyskinesia Associated With Normal Axoneme Ultrastructure Is Caused by DNAH11 Mutations. Hum Mutat 29:1–8
- 311. Bartoloni L, Blouin J-L, Pan Y, et al (2002) Mutations in the DNAH11 (axonemal heavy chain dynein type 11) gene cause one form of situs inversus totalis and most likely primary ciliary dyskinesia. Proc Natl Acad Sci 99:10282–10286
- 312. Lai M, Pifferi M, Bush A, et al (2016) Gene editing of DNAH11 restores normal cilia motility in primary ciliary dyskinesia. J Med Genet 53:242–249
- 313. Yang Q, Yin RX, Zhou YJ, Cao XL, Guo T, Chen WX (2015) Association of polymorphisms in the MAFB gene and the risk of coronary artery disease and ischemic stroke: A case-control study. Lipids Health Dis 14:7318–7331

- 314. Lemaitre RN, Tanaka T, Tang W, et al (2011) Genetic loci associated with plasma phospholipid N-3 fatty acids: A Meta-Analysis of Genome-Wide association studies from the charge consortium. PLoS Genet. https://doi.org/10.1371/journal.pgen.1002193
- 315. Xie L, Innis S (2008) Genetic Variants of the FADS1 FADS2 Gene Cluster Are Associated with Altered (n-6) and (n-3) Essential Fatty Acids in Plasma and Erythrocyte Phospholipids in Women during Pregnancy and in Breast Milk during Lactation. https://doi.org/10.3945/jn.108.096156.including
- 316. Naglik JR, Moyes D, Makwana J, et al (2009) NIH Public Access. J Allergy Clin Immunol 154:3266–3280
- 317. Yanhong Liu RKM (2011) Elevated Delta-6 Desaturase (FADS2) Gene
 Expression in the Prefrontal Cortex of Patients with Bipolar Disorder. 45:269–
 272
- 318. Wang Z, Han G, Liu Q, Zhang W, Wang J (2018) Silencing of PYGB suppresses growth and promotes the apoptosis of prostate cancer cells via the NF-κB/Nrf2 signaling pathway. Mol Med Rep 18:3800–3808
- 319. Zhang S, Zhou Y, Zha Y, Yang Y, Wang L, Li J, Jin W (2018) PYGB siRNA inhibits the cell proliferation of human osteosarcoma cell lines. Mol Med Rep 18:715–722
- 320. Cubranic Z, Madzar Z, Matijevic S, Dvornik S, Fisic E, Tomulic V, Kunisek J, Laskarin G, Kardum I, Zaputovic L (2012) Diagnostic accuracy of heart fatty acid binding protein (H-FABP) and glycogen phosphorylase isoenzyme BB (GPBB) in diagnosis of acute myocardial infarction in patients with acute coronary syndrome. Biochem medica 22:225–236
- 321. Fister S, Günthert AR, Emons G, Gründker C (2007) Gonadotropin-releasing hormone type II antagonists induce apoptotic cell death in human endometrial and ovarian cancer cells in vitro and in vivo. Cancer Res 67:1750–1756
- 322. Emons CGARGRPMG (2002) Expression of Gonadotropin-Releasing Hormone II (GnRH-II) Receptor in Human Endometrial and Ovarian Cancer Cells and Effects of GnRH-II on Tumor Cell Proliferation. 143:4143–4146
- 323. Guo D chuan, Grove MLL, Prakash SKK, et al (2016) Genetic Variants in LRP1 and ULK4 Are Associated with Acute Aortic Dissections. Am J Hum Genet 99:762–769
- 324. Broderick P, Chubb D, Johnson DC, et al (2016) Common variation at 3p22.1

and 7p15.3 influences multiple myeloma risk. 44:58–61

- 325. Lang B, Pu J, Hunter I, et al (2014) Recurrent deletions of ULK4 in schizophrenia: a gene crucial for neuritogenesis and neuronal motility. J Cell Sci 127:630–640
- 326. Guo E, Liu H, Liu X (2017) Overexpression of SCUBE2 Inhibits Proliferation, Migration, and Invasion in Glioma Cells. Oncol Res Featur Preclin Clin Cancer Ther 25:437–444
- 327. Cheng CJ, Lin YC, Tsai MT, Chen CS, Hsieh MC, Chen CL, Yang RB (2009) SCUBE2 suppresses breast tumor cell proliferation and confers a favorable prognosis in invasive breast cancer. Cancer Res 69:3634–3641
- 328. Margolin DH, Kousi M, Chan Y-M, et al (2013) Ataxia, Dementia, and Hypogonadotropism Caused by Disordered Ubiquitination. N Engl J Med 368:1992–2003
- 329. Wang H, Wang Y, Qian L, Wang X, Gu H, Dong X (2016) RNF216 contributes to proliferation and migration of colorectal cancer via suppressing BECN1dependent autophagy. Oncotarget 7:1–10
- 330. Goyenechea E, Crujeiras AB, Abete I, Martínez JA (2009) Expression of two inflammation-related genes (RIPK3 and RNF216) in mononuclear cells is associated with weight-loss regain in obese subjects. J Nutrigenet Nutrigenomics 2:78–84
- 331. Parsa A, Chang YPC, Kelly RJ, Corretti MC, Ryan KA, Robinson SW, Gottlieb SS, Kardia SLR, Shuldiner AR, Liggett SB (2011) Hypertrophy-associated polymorphisms ascertained in a founder cohort applied to heart failure risk and mortality. Clin Transl Sci 4:17–23
- 332. Gonelle-gispert C, Halban PA, Niemann H, et al (1999) SNAP-25a and -25b isoforms are both expressed in insulin-secreting cells and can function in insulin secretion. Society 165:159–165
- 333. Jamain S, Etain B, Kahn J-P, Dumaine A, Chevalier F, Bellivier F, Deshommes J, Mathieu F, Leboyer M, Henry C (2009) A SNAP25 promoter variant is associated with early-onset bipolar disorder and a high expression level in brain. Mol Psychiatry 15:748–755
- 334. Brophy K, Hawi Z, Kirley A, Fitzgerald M, Gill M (2002) Synaptosomal-associated protein 25 (SNAP-25) and attention deficit hyperactivity disorder (ADHD): Evidence of linkage and association in the Irish population. Mol Psychiatry 7:913–917

- 335. Carroll LS, Kendall K, O'Donovan MC, Owen MJ, Williams NM (2009) Evidence that putative ADHD low risk alleles at SNAP25 may increase the risk of schizophrenia. Am J Med Genet Part B Neuropsychiatr Genet 150:893–899
- 336. Wang Q, Wang Y, Ji W, et al (2015) SNAP25 is associated with schizophrenia and major depressive disorder in the Han Chinese population. J Clin Psychiatry 76:e76–e82
- Barakauskas, E AM (2017) Quantitative mass spectrometry reveals changes in SNAP-25 isoforms in schizophrenia. 177:44–51
- 338. Safari M reza, Omrani MD, Noroozi R, Sayad A, Sarrafzadeh S, Komaki A, Manjili FA, Mazdeh M, Ghaleiha A, Taheri M (2017) Synaptosome-Associated Protein 25 (SNAP25) Gene Association Analysis Revealed Risk Variants for ASD, in Iranian Population. J Mol Neurosci 61:305–311
- 339. Al-Daghri NM, Costa AS, Alokail MS, Zanzottera M, Alenad AM, Mohammed AK, Clerici M, Guerini FR (2016) Synaptosomal protein of 25 kDa (Snap25) polymorphisms associated with glycemic parameters in type 2 diabetes patients. J Diabetes Res. https://doi.org/10.1155/2016/8943092
- 340. Kim JW, Biederman J, Arbeitman L, et al (2007) Investigation of variation in SNAP-25 and ADHD and relationship to co-morbid major depressive disorder. Am J Med Genet Part B Neuropsychiatr Genet 144:781–790
- 341. Oner CBO (2011) Synaptosomal-Associated Protein 25 Gene Polymorphisms and Antisocial Personality Disorder: Association With Temperament and Psychopathy. 1:233–245
- 342. Gosso MF, De Geus EJC, Van Belzen MJ, Polderman TJC, Heutink P, Boomsma DI, Posthuma D (2006) The SNAP-25 gene is associated with cognitive ability: Evidence from a family-based study in two independent Dutch cohorts. Mol Psychiatry 11:878–886
- 343. Ikeda O, Sekine Y, Mizushima A, et al (2010) Interactions of STAP-2 with Brk and STAT3 participate in cell growth of human breast cancer cells. J Biol Chem 285:38093–38103
- 344. Kitai Y, Iwakami M, Saitoh K, et al (2017) STAP-2 protein promotes prostate cancer growth by enhancing epidermal growth factor receptor stabilization. J Biol Chem 292:19392–19399
- 345. Cao B, Yang L, Rong W, Feng L, Han N, Zhang K, Cheng S, Wu J, Xiao T, Gao Y (2015) Latent transforming growth factor-beta binding protein-1 in circulating

plasma as a novel biomarker for early detection of hepatocellular carcinoma. Int J Clin Exp Pathol 8:16046–16054

- 346. Öklü R, Hesketh R, Wicky S, Metcalfe JC (2011) Expression of mRNA isoforms of latent transforming growth factor-β binding protein-1 in coronary atherosclerosis and human tissues. Biochem Genet 49:213–225
- 347. Beaufort N, Scharrer E, Kremmer E, Lux V, Ehrmann M, Huber R, Houlden H, Werring D, Haffner C, Dichgans M (2014) Cerebral small vessel disease-related protease HtrA1 processes latent TGF-β binding protein 1 and facilitates TGF-β signaling. Proc Natl Acad Sci 111:16496–16501
- 348. Pei YF, Zhang L, Yang TL, et al (2012) Genome-wide association study of copy number variants suggests ltbp1 and fgd4 are important for alcohol drinking. PLoS One 7:3–8
- 349. Tritschler I, Gramatzki D, Capper D, Mittelbronn M, Meyermann R, Saharinen J, Wick W, Keski-Oja J, Weller M (2009) Modulation of TGF-β activity by latent TGF-β-binding protein 1 in human malignant glioma cells. Int J Cancer 125:530– 540
- 350. Higashi T, Sasagawa T, Inoue M, Oka R, Shuangying L, Saijoh K (2001) Overexpression of latent transforming growth factor-β1 (TGF-β1) binding protein 1 (LTBP-1) in association with TGF-β1 in ovarian carcinoma. Japanese J Cancer Res 92:506–515
- 351. Inouye M, Ripatti S, Kettunen J, et al (2012) Novel Loci for Metabolic Networks and Multi-Tissue Expression Studies Reveal Genes for Atherosclerosis. PLoS Genet. https://doi.org/10.1371/journal.pgen.1002907
- 352. Maneerat Y, Prasongsukarn K, Benjathummarak S, Dechkhajorn W (2017) PPBP and DEFA1/DEFA3 genes in hyperlipidaemia as feasible synergistic inflammatory biomarkers for coronary heart disease. Lipids Health Dis 16:1–12
- 353. Kinouchi T, Uemura M, Wang C, et al (2017) Expression level of CXCL7 in peripheral blood cells is a potential biomarker for the diagnosis of renal cell carcinoma. Cancer Sci 108:2495–2502
- 354. Shusterman A, Munz M, Richter G, et al (2017) The PF4/PPBP/CXCL5 Gene Cluster Is Associated with Periodontitis. J Dent Res 96:945–952
- 355. Böckelmann D, Naz A, Siddiqi MYJ, Lerner E, Sandrock-Lang K, Shamsi TS, Zieger B (2018) Bernard-Soulier syndrome in Pakistan: Biochemical and molecular analyses leading to identification of a novel mutation in GP1BA.

Haemophilia 24:e18–e22

- 356. Afshar-Kharghan V, Matijevic-Aleksic N, Ahn C, Boerwinkle E, Wu KK, López JA (2004) The variable number of tandem repeat polymorphism of platelet glycoprotein lbα and risk of coronary heart disease. Blood 103:963–965
- 357. Meisel C, Afshar-Kharghan V, Cascorbi I, Laule M, Stangl V, Felix SB, Baumann G, López JA, Roots I, Stangl K (2001) Role of kozak sequence polymorphism of platelet glycoprotein Ibα as a risk factor for coronary artery disease and catheter interventions. J Am Coll Cardiol 38:1023–1027
- 358. Douglas H, Michaelides K, Gorog DA, Durante-Mangoni E, Ahmed N, Davies GJ, Tuddenham EG (2002) Platelet membrane glycoprotein Ibalpha gene -5T/C Kozak sequence polymorphism as an independent risk factor for the occurrence of coronary thrombosis. Heart 87:70–74
- 359. Yoshida T, Yajima K, Hibino T, et al (2007) Association of gene polymorphisms with myocardial infarction in individuals with different lipid profiles. Int J Mol Med 20:581–590
- 360. Ozelo MC, Origa AF, Aranha FJ, Mansur AP, Annichino-Bizzacchi JM, Costa FF, Pollak ES, Arruda VR (2004) Platelet glycoprotein lbα polymorphisms modulate the risk for myocardial infarction. Thromb Haemost 92:384–386
- 361. Baker RI, Eikelboom J, Lofthouse E, Staples N, Afshar-Kharghan V, López JA, Shen Y, Berndt MC, Hankey G (2001) Platelet glycoprotein Ibα Kozak polymorphism is associated with an increased risk of ischemic stroke. Blood 98:36–40
- 362. Kanaji T, Ware J, Okamura T, Newman PJ (2012) GPIbα regulates platelet size by controlling the subcellular localization of filamin. Blood 119:2906–2913
- 363. Vairaktaris E, Serefoglou ZC, Yapijakis C, Agapi C, Vassiliou S, Nkenke E, Antonis V, Sofia S, Neukam FW, Patsouris E (2007) The Platelet Glycoprotein Ib· VNTR Polymorphism is Associated with Risk for Oral Cancer. Anticancer Res 27:4011–4014
- 364. Jin YY, Yu GZ, Wang Y, Cui LY, Xin XM (2009) Variable number of tandem repeats polymorphism of platelet glycoprotein lb α in Chinese people and CC genotype with aspirin sensitivity in patients with cerebral infarction. J Clin Pharm Ther 34:239–243
- 365. Soylu A, Tokaç M, Çora T, Düzenli MA, Acar H (2009) Platelet glycoprotein Ibα gene polymorphism and massive or submassive pulmonary embolism. J Thromb

Thrombolysis 27:259–266

- 366. Jain S, Zuka M, Liu J, et al (2007) Platelet glycoprotein lb supports experimental lung metastasis. Proc Natl Acad Sci 104:9024–9028
- 367. Edvardson S, Erlich Y, Elpeleg O, Hodges E, Hannon G, Dor T, Shaag A (2011) Exome sequencing and disease prediction implicate a mutation in KIF1A as a cause of hereditary spastic paraparesis type 30. Eur J Paediatr Neurol 15:S19
- 368. Hotchkiss L, Donkervoort S, Leach ME, Mohassel P, Bharucha-Goebel DX, Bradley N, Nguyen D, Hu Y, Gurgel-Giannetti J, Bönnemann CG (2016) Novel de Novo Mutations in KIF1A as a Cause of Hereditary Spastic Paraplegia with Progressive Central Nervous System Involvement. J Child Neurol 31:1114–1119
- 369. Krenn M, Zulehner G, Hotzy C, Rath J, Stogmann E, Wagner M, Haack TB, Strom TM, Zimprich A, Zimprich F (2017) Hereditary spastic paraplegia caused by compound heterozygous mutations outside the motor domain of the KIF1A gene. Eur J Neurol 24:741–747
- 370. Citterio A, Arnoldi A, Panzeri E, et al (2015) Variants in KIF1A gene in dominant and sporadic forms of hereditary spastic paraparesis. J Neurol 262:2684–2690
- 371. Langlois S, Tarailo-Graovac M, Sayson B, Drögemöller B, Swenerton A, Ross CJD, Wasserman WW, Van Karnebeek CDM (2016) De novo dominant variants affecting the motor domain of KIF1A are a cause of PEHO syndrome. Eur J Hum Genet 24:949–953
- 372. Guerrero-Preston R, Hadar T, Ostrow KL, et al (2014) Differential promoter methylation of kinesin family member 1a in plasma is associated with breast cancer and DNA repair capacity. Oncol Rep 32:505–512
- 373. Demokan S, Chang X, Chuang A, et al (2010) KIF1A and EDNRB are differentially methylated in primary HNSCC and salivary rinses. Int J Cancer 127:2351–2359
- 374. Rivire JB, Ramalingam S, Lavastre V, et al (2011) KIF1A, an axonal transporter of synaptic vesicles, is mutated in hereditary sensory and autonomic neuropathy type 2. Am J Hum Genet 89:219–301
- 375. Yoon CH, Rho SB, Kim ST, Kho S, Park J, Jang IS, Woo S, Kim SS, Lee JH, Lee SH (2012) Crucial role of TSC-22 in preventing the proteasomal degradation of p53 in cervical cancer. PLoS One 7:1–13
- 376. Nakashiro KI, Kawamata H, Hino S, Uchida D, Miwa Y, Hamano H, Omotehara F, Yoshida H, Sato M (1998) Down-regulation of TSC-22 (transforming growth

factor β -stimulated clone 22) markedly enhances the growth of a human salivary gland cancer cell line in vitro and in vivo. Cancer Res 58:549–555

- 377. Buraczynska M, Baranowicz-Gaszczyk I, Borowicz E, Ksiazek A (2007) TGF-β1 and TSC-22 Gene Polymorphisms and Susceptibility to Microvascular Complications in Type 2 Diabetes. Nephron Physiol 106:p69–p75
- 378. Rentsch CA, Cecchini MG, Schwaninger R, Germann M, Markwalder R, Heller M, Van Der Pluijm G, Thalmann GN, Wetterwald A (2006) Differential expression of TGFβ-stimulated clone 22 in normal prostate and prostate cancer. Int J Cancer 118:899–906
- 379. Hashimoto; R, Ikeda M (2013) Genome-wide association study of cognitive decline in schizophrenia. Am J Psychiatry 170:683–684.
- 380. Aouizera B, Vittinghoff E (2011) GWAS for discovery and replication of genetic loci associated with sudden cardiac arrest in patients with coronary artery disease. BMC Cardiovasc Disord. https://doi.org/10.1186/1471-2261-11-29
- 381. Shioji G, Ezura Y, Nakajima T, et al (2005) Nucleotide variations in genes encoding plasminogen activator inhibitor-2 and serine proteinase inhibitor B10 associated with prostate cancer. J Hum Genet 50:507–515
- 382. Ma Q, Ozel AB, Ramdas S, et al (2014) Genetic variants in PLG, LPA, and SIGLEC 14 as well as smoking contribute to plasma plasminogen levels. Blood 124:3155–3164
- Angata T, Ishii T, Motegi T, et al (2013) Loss of Siglec-14 reduces the risk of chronic obstructive pulmonary disease exacerbation. Cell Mol Life Sci 70:3199– 3210
- Bolling MC, Jongbloed JDH, Boven LG, Diercks GFH, Smith FJD, McLean WHI, Jonkman MF (2014) Plectin mutations underlie epidermolysis bullosa simplex in 8% of patients. J Invest Dermatol 134:273–276
- 385. Charlesworth A, Chiaverini C, Chevrant-Breton J, et al (2013) Epidermolysis bullosa simplex with PLEC mutations: New phenotypes and new mutations. Br J Dermatol 168:808–814
- 386. Pfendner E, Uitto J (2005) Plectin gene mutations can cause epidermolysis bullosa with pyloric atresia. J Invest Dermatol 124:111–115
- 387. Thorolfsdottir RB, Sveinbjornsson G, Sulem P, et al (2017) A Missense Variant in PLEC Increases Risk of Atrial Fibrillation. J Am Coll Cardiol 70:2157–2168
- 388. Kelly; KA, Bardeesy N, Anbazhagan R, Gurumurthy S, Berger J, Alencar H

(2008) Targeted Nanoparticles for Imaging Incipient Pancreatic Ductal Adenocarcinoma. PLOS MEd 5:77–86

- 389. Bausch D, Thomas S, Mino-Kenudson M, Fernández-del Castillo C, Bauer TW, Williams M, Warshaw AL, Thayer SP, Kelly KA (2011) Plectin-1 as a novel biomarker for pancreatic cancer. Clin Cancer Res 17:302–309
- 390. Paumard-Hernández B, Calvete O, Inglada Pérez L, et al (2018) Whole exome sequencing identifies PLEC, EXO5 and DNAH7 as novel susceptibility genes in testicular cancer. Int J Cancer 143:1954–1962
- 391. Zhong J, Chen G, Dang Y, Liao H, Zhang J, Lan D (2017) Novel compound heterozygous PLEC mutations lead to early-onset limb-girdle muscular dystrophy 2Q. Mol Med Rep 15:2760–2764
- 392. Natsuga K, Nishie W, Shinkuma S, Arita K, Nakamura H, Ohyama M, Osaka H, Kambara T, Hirako Y, Shimizu H (2010) Plectin Deficiency Leads to Both Muscular Dystrophy and Pyloric Atresia in Epidermolysis Bullosa Simplex. Hum Mutat 31:1687–1698
- 393. Maselli R, Arredondo J, Cagney O, et al (2011) Congenital myasthenic syndrome associated with epidermolysis bullosa caused by homozygous mutations in PLEC1 and CHRNE. Clin Genet 80:444–451
- 394. Peyser PA, Polasek O, Zoledziewska M, et al (2015) A meta-analysis of 120 246 individuals identifies 18 new loci for fibrinogen concentration. Hum Mol Genet 25:358–370
- 395. Eppinga RN, Hagemeijer Y, Burgess S, Hinds DA, Stefansson K, Gudbjartsson DF, Van Veldhuisen DJ, Munroe PB, Verweij N, Van Der Harst P (2016) Identification of genomic loci associated with resting heart rate and shared genetic predictors with all-cause mortality. Nat Genet 48:1557–1563
- 396. Youfang Liu, Michelle S. Yau, Laura M. Yerges-Armstrong DJD (2017) Genetic Determinants of Radiographic Knee Osteoarthritis in African Americans. Scand J Gastroenterol 44:1652–1658
- 397. Pawar H, Kashyap MK, Sahasrabuddhe NA, et al (2011) Quantitative tissue proteomics of esophageal squamous cell carcinoma for novel biomarker discovery. Cancer Biol Ther 12:510–522
- 398. Saito T, Okada S, Yamada E, Ohshima K, Shimizu H, Shimomura K, Sato M, Pessin JE, Mori M (2003) Syntaxin 4 and Synip (Syntaxin 4 Interacting Protein) Regulate Insulin Secretion in the Pancreatic □ HC-9 Cell *. 278:36718–36725

- 399. Min J, Okada S, Kanzaki M, Elmendorf JS, Coker KJ, Ceresa BP, Syu L, Noda Y, Arbor A, Arbor A (1999) Synip: A Novel Insulin-Regulated Syntaxin 4 Binding Protein Mediating GLUT4 Translocation in Adipocytes 2800 Plymouth Road. 3:751–760
- 400. Yamada E, Saito T, Okada S, Takahashi H, Ohshima K (2014) Synip phosphorylation is required for insulin-stimulated Glut4 translocation and glucose uptake in podocyte. 61:523–527
- 401. Ban HJ, Heo JY, Oh KS, Park KJ (2010) Identification of Type 2 Diabetesassociated combination of SNPs using Support Vector Machine. BMC Genet. https://doi.org/10.1186/1471-2156-11-26
- 402. Lu AT GWAS of epigenetic aging rates in blood reveals a critical role for TERT. https://doi.org/10.1038/s41467-017-02697-5
- 403. Rokudai S, Li Y, Otaka Y, Fujieda M, Owens DM, Christiano AM (2018) STXBP4 regulates APC/C-mediated p63 turnover and drives squamous cell carcinogenesis. https://doi.org/10.1073/pnas.1718546115
- 404. Masoodi TA, Banaganapalli B, Vaidyanathan V, Talluri VR, Shaik NA (2017) Computational Analysis of Breast Cancer GWAS Loci Identifies the Putative Deleterious Effect of STXBP4 and ZNF404 Gene Variants. 4307:4296–4307
- 405. Agrawal PB, Pierson CR, Joshi M, et al (2014) REPORT SPEG Interacts with Myotubularin , and Its Deficiency Causes Centronuclear Myopathy with Dilated Cardiomyopathy. 218–226
- 406. Quick AP, Wang Q, Philippen LE, et al (2016) Cellular Biology SPEG (Striated Muscle Preferentially Expressed Protein Kinase) Is Essential for Cardiac Function by Regulating Junctional Membrane Complex Activity. 110–119
- 407. Arvanitis DA, Flouris GA, Spandidos DA (2005) Genomic rearrangements on VCAM1 , SELE , APEG1 and AIF1 loci in atherosclerosis. 9:153–159
- 408. Kochi Y, Yamada R, Kobayashi K, et al (2004) Analysis of Single-Nucleotide Polymorphisms in Japanese Rheumatoid Arthritis Patients Shows Additional Susceptibility Markers Besides the Classic Shared Epitope Susceptibility Sequences. Arthritis Rheum 50:63–71
- 409. Chang SW, Fann CSJ, Su WH, et al (2014) A genome-wide association study on chronic HBV infection and its clinical progression in male Han-Taiwanese. PLoS One. https://doi.org/10.1371/journal.pone.0099724
- 410. Mbarek H, Ochi H, Urabe Y, et al (2011) A genome-wide association study of

chronic hepatitis B identified novel risk locus in a Japanese population. Hum Mol Genet 20:3884–3892

- 411. Vijai J, Kirchhoff T, Schrader KA, et al (2013) Susceptibility Loci Associated with Specific and Shared Subtypes of Lymphoid Malignancies. https://doi.org/10.1371/journal.pgen.1003220
- 412. Onouchi Y, Ozaki K, Burns JC, et al (2012) A genome-wide association study identifies three new risk loci for Kawasaki disease. Nat Genet 44:517–521
- 413. Mariani E, Frabetti F, Tarozzi A, Pelleri MC, Pizzetti F, Casadei R (2016) Meta-Analysis of Parkinson 's Disease Transcriptome Data Using TRAM Software : Whole Substantia Nigra Tissue and Single Dopamine Neuron Differential Gene Expression. 1–21
- 414. Thonberg H, Chiang H, Lilius L, et al (2017) Identification and description of three families with familial Alzheimer disease that segregate variants in the SORL1 gene. 1–14
- 415. Kurochkin I V., Yonemitsu N, Funahashi S ichi, Nomura H (2001) ALEX1, a novel human armadillo repeat protein that is expressed differentially in normal tissues and carcinomas. Biochem Biophys Res Commun 280:340–347
- 416. Fan Z, Li C, Pang L, et al (2017) ALEX1, a novel tumor suppressor gene, inhibits gastric cancer metastasis via the PAR-1/Rho GTPase signaling pathway. J Gastroenterol 53:71–83
- 417. Zeng F, Liao K, Wu J, et al (2015) ALEX1 may be a novel biomarker for human cervical squamous cell carcinoma. Int J Clin Exp Pathol 8:9434–9439
- 418. Zhang H-T, Song F-Z, Liu G-L, Wu J-Y, Gao Y, Zeng F, Yun H (2015) ALEX1 Regulates Proliferation and Apoptosis in Breast Cancer Cells. Asian Pacific J Cancer Prev 16:3293–3299
- 419. Hasan A, Al-ghimlas F, Warsame S, Al-hubail A, Ahmad R, Bennakhi A, Al-arouj M, Behbehani K, Dehbi M, Dermime S (2014) IL-33 is negatively associated with the BMI and confers a protective lipid / metabolic profile in non-diabetic but not diabetic subjects. BMC Immunol 1–9
- 420. Moffatt MF, Strachan DP, Farrall M, Heath S, Gut IG, Demenais F, Lathrop M, Bouzigon E, von Mutius E, Cookson WOCM (2010) A Large-Scale, Consortium-Based Genomewide Association Study of Asthma. N Engl J Med 363:1211–1221
- 421. Sedigheh BM, Masoud M, Zahra A (2016) Serum IL-33 Is Elevated in Children with Asthma and Is Associated with. 14194:193–196
- 422. Moffatt MF, Phil D, Gut IG, Ph D, Demenais F, Strachan DP (2010) Europe PMC Funders Group Europe PMC Funders Author Manuscripts A Large-Scale , Consortium-Based Genomewide Association Study of Asthma. N Engl J Med 363:1211–1221
- 423. Hinds DA, McMahon G, Kiefer AK, et al (2013) A genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci. Nat Genet 45:907–911
- 424. Ferreira; MAR., Matheson MC., Hopper JL (2015) Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. J Allerg Clin Immunol 133:1564–1571
- 425. Albertsen HM, Chettier R, Farrington P, Ward K (2013) Genome-Wide Association Study Link Novel Loci to Endometriosis. https://doi.org/10.1371/journal.pone.0058257
- 426. Brestoff JR, Kim BS, Saenz SA, et al (2015) Group 2 innate lymphoid cells promote beiging of white adipose tissue and limit obesity. Nature 519:242–246
- 427. Og EA, Just MJ (2018) The Association between Inflammatory Markers (iNOS , HO-1 , IL-33 , MIP-1 β) and Depression with and without Posttraumatic Stress Disorder Pharmacological Reports The association between in fl ammatory markers (iNOS , HO-1 , IL-33 , MIP-1 β) and depre. Proc Natl Acad Sci U S A. https://doi.org/10.1016/j.pharep.2018.06.001
- 428. Kudinova A, Brandon G (2016) Cross-Species Evidence for the Role of Interleukin-33 in Depression Risk. J Abnorm Psychol 125:482–494
- 429. Li Y, Shi J, Qi S, Zhang J, Peng D, Chen Z, Wang G, Wang Z, Wang L (2018)
 IL-33 facilitates proliferation of colorectal cancer dependent on COX2 / PGE 2.
 1–12
- 430. Wu C., Wu Y., Cheng C., · Hong Z., Shi Z., Lin S., · Li J., He X., Zhu A. (2018) Interleukin-33 Predicts Poor Prognosis and Promotes Renal Cell Carcinoma Cell Growth Through i ts Receptor ST2 and the JNK Signaling Pathway. Cell Physiol Cell Physiol Biochem 47:191–200
- 431. Sawada R, Ku Y, Akita M, Otani K, Fujikura K, Itoh T, Ajiki T, Fukumoto T, Kakeji Y, Zen Y (2018) Interleukin-33 overexpression reflects less aggressive tumour features in large-duct type cholangiocarcinomas. 259–272
- 432. Wei S, Chunlei W, Sen L (2017) IL-33 Expression in Chronic Rhinosinusitis with Nasal Polyps and Its Relationship with Clinical Severity. 323–330

- 433. Kim D, Jin HR, Eun KM, Mo J, Cho SH, Oh S, Cho D, Kim DW (2017) The role of interleukin-33 in chronic rhinosinusitis. 635–645
- 434. Still AA, Han JH, Suh C, et al (2017) Serum Levels of Interleukin 33 and Soluble ST2 Are Associated with the Extent of Disease Activity and Cutaneous Manifestations in Patients with Serum Levels of Interleukin 33 and Soluble ST2 Are Associated with the Extent of Disease Activity and Cutaneous. https://doi.org/10.3899/jrheum.170020
- 435. Zeng X, Zhang Z, Gao Q, Wang Y, Yu X, Zhou B, Xi M (2016) Clinical Significance of Serum Interleukin-31 and Interleukin-33 Levels in Patients of Endometrial Cancer : A Case Control Study. Dis. Markers 2016:
- 436. Voloshyna I, Mucci T, Sher J, Fonacier LS, Littlefield MJ, Carsons S, Reiss AB (2015) Plasma IL-33 in atopic patients correlates with pro-inflammatory cytokines and changes cholesterol transport protein expression: a surprising neutral overall impact on atherogenicity Experimental Allergy. Clin Exp Allergy 1554–1565
- 437. Demircan MKK, Demirci ETT, Sahin MSAM (2016) Clinical significance of ADAMTS1, ADAMTS5, ADAMTS9 aggrecanases and IL 17A, IL 23, IL 33 cytokines in polycystic ovary syndrome. J Endocrinol Invest 39:1269–1275
- 438. Yue J, Tong Y, Xie L, Ma T, Yang J (2016) Genetic variant in IL-33 is associated with idiopathic recurrent miscarriage in Chinese Han population. Nat Publ Gr 1–
 7
- 439. Saadah OI, Al-harthi SE, Al-mughales JA, Bin-taleb YY, Baeshen RS (2015)
 Serum Interleukin-33 level in Saudi children with inflammatory bowel disease.
 8:16000–16006
- 440. Koca SS, Pehlivan Y, Kara M, et al (2016) The IL-33 gene is related to increased susceptibility to systemic sclerosis. Rheumatol Int 36:579–584
- 441. Corrales A, González-juanatey C, Miranda-filloy JA, et al (2015) Protective Role of the Interleukin 33 rs3939286 Gene Polymorphism in the Development of Subclinical Atherosclerosis in Rheumatoid Arthritis Patients. PLoS One 10:1–8
- 442. Judd LM, Heine RG, Menheniott TR, et al (2019) Elevated IL-33 expression is associated with pediatric eosinophilic esophagitis , and exogenous IL-33 promotes eosinophilic esophagitis development in mice. 13–25
- 443. Yang J hua, Wu F qin, Wen Q, Zhang W cai, Wang Y e., Xiong X, Su Y wen, Cheng L xian (2015) Association of IL33/ST2 signal pathway gene

polymorphisms with myocardial infarction in a Chinese Han population. J Huazhong Univ Sci Technol - Med Sci 35:16–20

- 444. Xia J, Zhao J, Shang J, Li M, Zeng Z, Zhao J, Wang J, Xu Y, Xie J (2019)
 Biomarkers in Lung Diseases: from Pathogenesis to Prediction to New Therapies Increased IL-33 expression in chronic obstructive pulmonary disease.
 i:619–627
- 445. Kim JY, Lim S, Kim G, Yun HJ, Ahn S, Choi HS (2014) Interleukin-33 / ST2 axis promotes epithelial cell transformation and breast tumorigenesis via upregulation of COT activity. Oncogene 34:4928–4938
- 446. Wu F, He M, Wen Q, Zhang W, Yang J, Zhang X (2014) Associations between Variants in IL-33 / ST2 Signaling Pathway Genes and Coronary Heart Disease Risk. 23227–23239
- 447. Liu C, Shen D, Zhu K, Tang J, Wang X, Zhang L, Zhang J (2014) Characterization of interleukin-33 and matrix metalloproteinase-28 in serum and their association with disease severity in patients with coronary heart disease. 25:498–504
- 448. González-Gay MA, Diamantopoulos AP, Beretta L, et al (2014) A Candidate Gene Approach Identifies an IL33 Genetic Variant as a Novel Genetic Risk Factor for GCA. PLoS One 9:e113476
- 449. Sun Y, Zhang J, Lv S, et al (2014) Interleukin-33 Promotes Disease Progression in Patients with Primary Biliary Cirrhosis. 255–261
- 450. Shao D, Perros F, Caramori G, et al (2014) Nuclear IL-33 regulates soluble ST2 receptor and IL-6 expression in primary human arterial endothelial cells and is decreased in idiopathic pulmonary arterial hypertension. Biochem. Biophys. Res. Commun. 451:
- 451. Li C, Mu R, Guo J, Wu X, Tu X, Liu X, Hu F, Guo S, Zhu J (2014) Genetic variant in IL33 is associated with susceptibility to rheumatoid arthritis. Arthritis Res Ther 1–10
- 452. Hu L, Fu Y, Zhang D, Zhang J (2013) Serum IL-33 as a Diagnostic and Prognostic Marker in Non- small Cell Lung Cancer. Clin Exp Rheumatol 14:2563–2566
- 453. Marvie P, Lisbonne M, Helgoualc AL, Rauch M, Turlin B, Piquet-pellorce C, Samson M (2010) Interleukin-33 overexpression is associated with liver fibrosis in mice and humans. 14:1726–1739

- 454. Second T, Hospital X (2013) Predictive Value of Serum Bone Sialoprotein and Prostate-specific Antigen Doubling Time in Patients with Bone Metastasis of Prostate. 33:559–562
- 455. Wei RJ, Li TY, Yang XC, Jia N, Yang XL, Song HB (2015) Serum levels of PSA , ALP , ICTP , and BSP in prostate cancer patients and the significance of ROC curve in the diagnosis of prostate cancer bone metastases. 15:
- 456. Wang J, Wang L, Xia B, Yang C, Lai H, Chen X (2013) BSP Gene Silencing Inhibits Migration , Invasion , and Bone Metastasis of MDA-MB-231BO Human Breast Cancer Cells. 8:1–8
- 457. Zhang Y, Liu H, Zhang C, Zhang T, Zhang B, Li L, Chen G, Fu D, Wang K (2015) Endochondral ossification pathway genes and postmenopausal osteoporosis : Association and specific allele related serum bone sialoprotein levels in Han Chinese. 1–8
- 458. Styrkarsdottir U, Halldorsson B V, Gretarsdottir S, et al (2010) New sequence variants associated with bone mineral density. Nat Genet 41:2008–2010
- 459. Pesesse L, Sanchez C, Walsh DA, Delcour J, Baudouin C, Msika P, Henrotin Y (2014) Bone sialoprotein as a potential key factor implicated in the pathophysiology of osteoarthritis. Osteoarthr Cartil 22:547–556
- 460. Xu T, Qin R, Zhou J, et al (2012) High Bone Sialoprotein (BSP) Expression Correlates with Increased Tumor Grade and Predicts a Poorer Prognosis of High-Grade Glioma Patients. https://doi.org/10.1371/journal.pone.0048415
- 461. Shaffer JR, Feingold E, Wang X, Lee M (2013) GWAS of Dental Caries Patterns in the Permanent Dentition. J Dent Res 38–44

Anhang

Anhangtabelle 1: Übersicht verwendete Merkmale und SNPs

Übersicht zu den für die eQTL Analyse verwendeten Merkmalen und SNPs. Die Spalten Gruppe, Merkmal und Abk. geben die jeweilige Merkmalsgruppe, die darin enthaltenen Merkmale und deren verwendeten Abkürzungen an. Die Spalte SNPs in GWAS Studien führt die Anzahl der in den Quellen angegebenen SNPs auf, die Spalte Zur Analyse verwendete SNPs die Anzahl der in dieser Arbeit verwendeten SNPs sowie in Klammern die Anzahl der ausgeschlossenen SNPs, wobei der Kleinbuchstabe den Grund für den Ausschluss der jeweiligen SNPs angibt. Fettgedruckte Zahlen geben die Gesamtanzahl der originalen bzw. nicht verwendeten SNPs zum jeweiligen Merkmal an

Gruppe	Merkmal	Abk.	Quelle	SNPs in GWAS Studien	Zur Analyse verwendete SNPs
AMD	Altersabhängige Makuladegeneration	AMD	[42]	52	44 g(8)
Leberenzyme	Alkalische Phosphatase	ALP	[230]	43	14 b(28) e(1)
	Alanin-Aminoransferase	ALT	[230]	43	4 b(38), e(1)
	Gamma-Glutamyltransferase	GGT	[230]	43	25 b(17) e(1)
	Leberenzyme im Plasma	LEP	[230]	43	42 e(1)
Kardiovaskuläre	Koronare Herzkrankheit	CAD	[231]	51	50 a(1)
Merkmale			[232]	13	9 a(3), e(1)
			[233]	15	14 a(1)
			[234]	4	4
			[235]	64	64
			[236]	15	7 a(8)
			[237]	10	10
	Diastaliashan Dlutdrusk		[000]	172	154 g(4)
	Diastolischer Blutdruck	DBP	[238]	29	22 D(3) a(4)
			[107]	55	1 D(29) a(0)
			[109]	11	37 a(14)
			[239]	2	2
			[2+0]	158	$\frac{2}{94}$ q(6) h(2)
	Systolischer Blutdruck	SBP	[238]	29	22 b(4) a(3)
	-,		[239]	24	23 f(1)
			[107]	4	1 a(3)
			[240]	3	3
			[109]	54	37 a(17)
				114	81 g(5)
	Blutdruck- Allgemein	GBP	[238]	29	29
	Blutdruckamplitude	PP	[239]	42	39 f(3)
			[240]	4	4
				44	42 g(1)
	Hypertonie	HIN	[238]	29	10 e(19)
			[107]	10	1 b(9)
Dlutlinida	Low Donaity Linenratain		[044]	59	59
ыцприсе	Low Density Lipoprotein		[241]	30	00 10
			[242]	77	19 76 g(1)
	High Density Lipoprotein	ноі	[2/1]	71	70 g(1)
		TIDE	[241]	35	35
				106	99 q(6) h(1)
	Gesamtcholesterin	тс	[241]	74	74
			[242]	19	18 a(1)
				93	89 g(3)
	Triglyceride	TG	[241]	40	39 e(1)
			[242]	33	33 `´
				73	66 g(6)

a: bereits in anderen Studien enthalten (es wurden SNP-Werte der Studie mit niedrigerem P-Wert verwendet) b: SNP nicht genomweit signifikant c: korrelierte SNPs, r2 > 0,5 d: SNP auf X-Chromosom

e: Imputationsqualität < 0,3

f: Ausschluss Exomstudien

g: SNP im GTEx Datensatz fehlend

h: fehlende Genotypinformationen im GTEx Datensatz (mehr als 5% der Genotypdaten fehlend)

Anhangtabelle 2: Merkmalspaare Lebergewebe

Überschneidungen zwischen den Genlisten bezüglich des Gewebes Leber. Die Spalten **Merkmal 1** und **Gene Merkmal 1** respektive **Merkmal 2** und **Gene Merkmal 2** geben das betrachtete Merkmal sowie die Anzahl der Gene der jeweiligen Genliste an. Die Spalte **Gemeinsame Gene** zeigt die Anzahl der sich überschneidenden Gene der beiden Genlisten. Die Spalten **P-Wert** und **FDR** geben die jeweiligen Signifikanzwerte der Überschneidungen an, welche mit dem exakten Test nach Fisher berechnet wurden.

		Gene	Gene	Gemeinsame		
Merkmal 1	Merkmal 2	Merkmal 1	Merkmal 2	Gene	P-Wert	FDR
LDL	TC	22	29	13	4.89E-32	5.13E-30
ALP	LEP	5	9	5	1.47E-17	7.72E-16
DBP	SBP	14	10	6	2.77E-17	9.68E-16
GGT	LEP	4	9	4	4.69E-14	1.23E-12
GBP	SBP	4	10	4	7.81E-14	1.64E-12
DBP	GBP	14	4	4	3.72E-13	6.52E-12
GBP	HTN	4	2	2	4.72E-08	7.09E-07
CAD	LDL	40	22	4	2.41E-07	3.16E-06
HTN	SBP	2	10	2	3.54E-07	4.13E-06
CAD	ТС	40	29	4	7.72E-07	7.37E-06
DBP	HTN	14	2	2	7.16E-07	7.37E-06
HDL	TC	14	29	3	1.94E-06	1.70E-05
ALP	LDL	5	22	2	1.81E-05	0.00015
ALP	ТС	5	29	2	3.19E-05	0.00024
LDL	LEP	22	9	2	6.51E-05	0.00046
TC	LEP	29	9	2	0.00011	0.00075
ALP	CAD	5	40	1	0.01249	0.07712
HDL	LDL	14	22	1	0.01916	0.11176
CAD	LEP	40	9	1	0.02237	0.12360
AMD	HDL	83	14	1	0.07051	0.37019

Anhangtabelle 3: Merkmalspaare Blutgewebe

Überschneidungen zwischen den Genlisten bezüglich des Gewebes Blut. Die Spalten **Merkmal 1** und **Gene Merkmal 1** respektive **Merkmal 2** und **Gene Merkmal 2** geben das betrachtete Merkmal sowie die Anzahl der Gene der jeweiligen Genliste an. Die Spalte **Gemeinsame Gene** zeigt die Anzahl der sich überschneidenden Gene der beiden Genlisten. Die Spalten **P-Wert** und **FDR** geben die jeweiligen Signifikanzwerte der Überschneidungen an, welche mit dem exakten Test nach Fisher berechnet wurden.

		Gene	Gene	Gemeinsame		
Merkmal 1	Merkmal 2	Merkmal 1	Merkmal 2	Gene	P-Wert	FDR
DBP	SBP	19	15	14	1.70E-43	1.78E-41
LDL	тс	14	19	9	8.45E-25	4.44E-23
ALP	LEP	4	6	4	5.15E-15	1.80E-13
HDL	TG	19	19	5	1.41E-11	3.71E-10
GBP	SBP	4	15	3	2.54E-09	5.33E-08
DBP	GBP	19	4	3	5.41E-09	7.10E-08
HDL	тс	19	19	4	5.11E-09	7.10E-08
тс	TG	19	19	4	5.11E-09	7.10E-08
GGT	LEP	2	6	2	1.14E-07	1.32E-06
HDL	LDL	19	14	2	0.00012	0.00112
LDL	TG	14	19	2	0.00012	0.00111
AMD	HDL	20	19	2	0.00024	0.00172
AMD	тс	20	19	2	0.00024	0.00172
AMD	TG	20	19	2	0.00024	0.00172
GBP	HTN	4	1	1	0.00025	0.00172
HTN	SBP	1	15	1	0.00092	0.00605
DBP	HTN	19	1	1	0.00117	0.00722
ALP	LDL	4	14	1	0.00344	0.02007
ALP	HDL	4	19	1	0.00467	0.02227
ALP	ТС	4	19	1	0.00467	0.02227
ALP	TG	4	19	1	0.00467	0.02227
GBP	HDL	4	19	1	0.00467	0.02227
LDL	LEP	14	6	1	0.00516	0.02354
HDL	LEP	19	6	1	0.00699	0.02824
тс	LEP	19	6	1	0.00699	0.02824
TG	LEP	19	6	1	0.00699	0.02824
CAD	GBP	48	4	1	0.01176	0.04409
CAD	PP	48	4	1	0.01176	0.04409
HDL	SBP	19	15	1	0.01739	0.06298
DBP	HDL	19	19	1	0.02198	0.07695
CAD	SBP	48	15	1	0.04340	0.14700
CAD	DBP	48	19	1	0.05466	0.17935

Anhangtabelle 4: Gemeinsame Gene aufgrund gemeinsamer SNPs in Lebergew.

Gemeinsame Gene in Bezug auf das Gewebe Leber, deren Gemeinsamkeit zwischen den Genlisten durch die Effekte auf die Genexpression desselben SNPs zu erklären sind. Die Spalten **Gensymbol** und **SNP** geben das Gen an, welches durch den entsprechenden SNP in seiner Expression beeinflusst wird. Die Spalten **Merkmal 1** bzw. **Merkmal 2** geben die beiden Merkmale an, in deren Genliste das Gen gefunden wurde. Die Spalten **eQTL Effektgröße Merkmal 1** respektive **eQTL Effektgröße Merkmal 2** zeigen das Ausmaß der Beeinflussung der Genexpression bezüglich des jeweiligen Merkmals und SNPs.

Gensymbol	SNP	Merkmal	Merkmal	eQTL	eQTL
		1	2	Effektgröße	Effektgröße
				Merkmal 1	Merkmal 2
WBSCR28	rs579459	ALP	CAD	0.39843	-0.39842
FYB	rs314253	ALP	LDL	-0.12094	0.12094
NBPF3	rs1976403	ALP	LEP	0.20542	0.20542
ABHD12	rs7267979	ALP	LEP	-0.05107	-0.05107
FYB	rs314253	ALP	LEP	-0.12094	-0.12094
IL6ST	rs10819937	ALP	LEP	0.07131	0.07131
WBSCR28	rs579459	ALP	LEP	0.39843	0.39843
FYB	rs314253	ALP	тс	-0.12094	0.12094
WBSCR28	rs579459	CAD	LEP	-0.39843	0.39843
HLA.DRB5	rs805303	DBP	GBP	0.24672	0.24672
MEI1	rs1327235	DBP	GBP	-0.11202	-0.11202
LIMS3L	rs4373814	DBP	GBP	-0.40654	-0.40654
HLA.DRB5	rs805303	DBP	HTN	0.24672	0.24672
HLA.DRB5	rs805303	DBP	SBP	0.24672	0.24672
MEI1	rs1327235	DBP	SBP	-0.11202	-0.11202
SCGB1D2	rs4247374	DBP	SBP	-0.69335	-0.69335
MYOT	rs1620668	DBP	SBP	-0.30980	-0.30980
PPAPDC3	rs11105354	DBP	SBP	0.16488	0.16488
LIMS3L	rs4373814	DBP	SBP	-0.40654	-0.40654
HLA.DRB5	rs805303	GBP	HTN	0.24672	0.24672
PPAPDC3	rs17249754	GBP	HTN	0.16488	0.16488
HLA.DRB5	rs805303	GBP	SBP	0.24672	0.24672
MEI1	rs1327235	GBP	SBP	-0.11202	-0.11202
LIMS3L	rs4373814	GBP	SBP	-0.40654	-0.40654
MLIP	rs9296736	GGT	LEP	-0.36747	-0.36747
MTRNR2L6	rs1335645	GGT	LEP	0.52493	0.52493
MUC3A	rs2073398	GGT	LEP	-0.26650	-0.26650
EFHD1	rs2140773	GGT	LEP	0.16728	0.16728
SIT1	rs9987289	HDL	LDL	0.44604	0.44604
RP11.1105G2.3	rs1800961	HDL	тс	1.09757	1.09757
SIT1	rs9987289	HDL	тс	0.44604	0.44604
RANBP17	rs1800961	HDL	ТС	0.78169	0.78169
HLA.DRB5	rs805303	HTN	SBP	0.24672	0.24672
FYB	rs314253	LDL	LEP	0.12094	-0.12094
PSRC1	rs629301	LDL	ТС	-0.43645	-0.43645

SORT1	rs629301	LDL	тс	-0.40659	-0.40659
CELSR2	rs629301	LDL	тс	-0.27941	-0.27941
EFCAB13	rs7206971	LDL	тс	0.11582	0.11582
RHD	rs12027135	LDL	TC	-0.35140	-0.35140
ST3GAL4	rs11220462	LDL	ТС	0.11948	0.11948
FYB	rs314253	LDL	TC	0.12094	0.12094
SIT1	rs9987289	LDL	TC	0.44604	0.44604
WBSCR28	rs9411489	LDL	TC	-0.44148	-0.44148
SPTB	rs4253772	LDL	TC	0.12523	0.12523
PRKCSH	rs2030746	LDL	TC	-0.03541	-0.03541
CDK5RAP2	rs2030746	LDL	TC	-0.05911	-0.05911
DNAH11	rs12670798	LDL	TC	0.22763	0.22763
FYB	rs314253	тс	LEP	0.12094	-0.12094

Anhangtabelle 5: Gemeinsame Gene aufgrund anderer Effekte in Lebergew.

Gemeinsame Gene in Bezug auf das Gewebe Leber, deren Gemeinsamkeit zwischen den Genlisten nicht durch die Effekte auf die Genexpression desselben SNPs zu erklären sind. Die Spalten **Gensymbol** sowie **Merkmal 1** und **Merkmal 2** geben hierbei die Merkmale an, in deren Genlisten das entsprechende Gen gefunden wurde. Die Spalten **SNP Merkmal 1** bzw. **SNP Merkmal 2** zeigen die SNPs, welche die Expression des entsprechenden Gens in Bezug auf Merkmal 1 bzw. Merkmal 2 beeinflussen. Die Spalten **eQTL Effektgröße Merkmal 1** bzw. **eQTL Effektgröße Merkmal 2** zeigen das Ausmaß der Beeinflussung der Genexpression bezüglich der entsprechenden Merkmale und SNPs. Die Spalte **r² SNPs** gibt die die Höhe des r² des entsprechenden SNP Paares an.

Gensymbol	Merkmal	Merkmal	SNP	eQTL	SNP	eQTL	r ²
	1	2	Nerkmal 1	Merkmal 1	Nerkmai 2	Merkmal 2	SNPS
WBSCR28	ALP	LDL	rs579459	0.39843	rs9411489	-0.44147	0.83
WBSCR28	ALP	тс	rs579459	0.39843	rs9411489	-0.44148	0.83
INTS12	AMD	HDL	rs943080	-0.05241	rs4650994	-0.05470	0.0004
PSRC1	CAD	LDL	rs602633	-0.40459	rs629301	-0.43645	0.86
SORT1	CAD	LDL	rs602633	-0.37420	rs629301	-0.40659	0.86
CELSR2	CAD	LDL	rs602633	-0.25753	rs629301	-0.27941	0.86
WBSCR28	CAD	LDL	rs579459	-0.39843	rs9411489	-0.44148	0.83
PSRC1	CAD	тс	rs602633	-0.40459	rs629301	-0.43645	0.86
SORT1	CAD	тс	rs602633	-0.37420	rs629301	-0.40659	0.86
CELSR2	CAD	тс	rs602633	-0.25753	rs629301	-0.27941	0.86
WBSCR28	CAD	тс	rs579459	-0.39843	rs9411489	-0.44148	0.83
PPAPDC3	DBP	GBP	rs11105354	0.16488	rs17249754	0.16488	1
PPAPDC3	DBP	HTN	rs11105354	0.16488	rs17249754	0.16488	1
PPAPDC3	GBP	SBP	rs17249754	0.16488	rs11105354	0.16488	1
PPAPDC3	HTN	SBP	rs17249754	0.16488	rs11105354	0.16488	1
WBSCR28	LDL	LEP	rs9411489	-0.44148	rs579459	0.39843	0.83
WBSCR28	ТС	LEP	rs9411489	-0.44148	rs579459	0.39843	0.83

Anhangtabelle 6: Merkmalspaare Lebergewebe

Merkmalspaare bezüglich des Gewebes Leber, zu welchen gemeinsame Gene zwischen den Genlisten identifiziert wurden. Die Spalten **Merkmal 1** und **Merkmal 2** geben die beiden Merkmale des jeweiligen Merkmalspaares an. In den Spalten **Gemeinsame SNPs** bzw. **Gemeinsame Gene** ist die Anzahl an Überschneidungen zwischen den zur Analyse verwendeten SNPs bzw. den Genlisten der Merkmale des jeweiligen Merkmalspaares aufgeführt. Die Spalte **Gemeinsame Gene aufgrund gemeinsamer SNPs** zeigt die Anzahl an Genen, deren Gemeinsamkeit zwischen den Genlisten der beiden Merkmale auf den Effekten desselben SNPs beruht, während die Spalte **Gemeinsame Gene aufgrund anderer Effekte** die Anzahl an Genen angibt, deren Gemeinsamkeit zwischen den Genlisten der kmale nicht auf den Effekten desselben SNPs beruht. Die Spalte **Gemeinsame Gene aufgrund anderer Effekte** die Anzahl an Genen angibt, deren Gemeinsamkeit zwischen den Genlisten der beiden Merkmale nicht auf den Effekten desselben SNPs beruht. Die Spalte **Gemeinsame Gene aufgrund anderer Effekte** nicht in **LD**($r^2 < 0.5$) zeigt die Anzahl an gemeinsamen Genen, deren Gemeinsamkeit zwischen den SNPs, die der Gemeinsamkeit zwischen den Genlisten der beiden Merkmale nicht auf den Effekten desselben SNPs beruht und bei denen die beiden SNPs, die der Gemeinsamkeit zwischen den Genlisten zugrunde liegen, sich nicht in LD befinden und somit unabhängig sind.

Merkmal 1	Merkmal 2	Gemeinsame SNPs	Gemeinsame Gene	Gemeinsame Gene aufgrund gemeinsamer SNPs	Gemeinsame Gene aufgrund anderer Effekte	Gemeinsame Gene aufgrund anderer Effekte:nicht in LD(r ² < 0.5)
LDL	тс	47	13	13	0	-
ALP	LEP	14	5	5	0	-
DBP	SBP	45	6	6	0	-
GGT	LEP	25	4	4	0	-
GBP	SBP	22	4	3	1	-
DBP	GBP	23	4	3	1	-
GBP	HTN	11	2	2	0	-
CAD	LDL	2	4	0	4	-
HTN	SBP	9	2	1	1	-
CAD	ТС	1	4	0	4	-
DBP	HTN	8	2	1	1	-
HDL	ТС	13	3	3	0	-
ALP	LDL	1	2	1	1	-
ALP	ТС	1	2	1	1	-
LDL	LEP	1	2	1	1	-
ТС	LEP	2	2	1	1	-
ALP	CAD	1	1	1	0	-
HDL	LDL	8	1	1	0	-
CAD	LEP	1	1	1	0	-
AMD	HDL	0	1	0	1	1

Anhangtabelle 7: Gemeinsame Gene aufgrund gemeinsamer SNPs in Blutgew.

Gemeinsame Gene in Bezug auf das Gewebe Blut, deren Gemeinsamkeit zwischen den Genlisten durch die Effekte auf die Genexpression desselben SNPs zu erklären sind. Die Spalten **Gensymbol** und **SNP** geben das jeweilige Gen an, welches durch den entsprechenden SNP in seiner Expression beeinflusst wird. Die Spalten **Merkmal 1** bzw. **Merkmal 2** geben die beiden Merkmale an, in deren Genliste das Gen gefunden wurde. Die Spalten **eQTL Effektgröße Merkmal 1** respektive **eQTL Effektgröße Merkmal 2** zeigen das Ausmaß der Beeinflussung der Genexpression bezüglich des jeweiligen Merkmals und SNPs.

Gensymbol	SNP	Merkmal 1	Merkmal 2	eQTL Effektgröße Merkmal 1	eQTL Effektgröße Merkmal 2
NBPF3	rs1976403	ALP	LEP	0.18332	0.18331
FADS2	rs174601	ALP	LEP	0.12753	0.12753
PYGB	rs7267979	ALP	LEP	-0.05069	-0.05069
GNRH2	rs281377	ALP	LEP	-0.21495	-0.21495
ZNF461	rs17608766	CAD	GBP	-0.08052	-0.08052
ZNF461	rs17608766	CAD	SBP	-0.08052	-0.08052
ULK4	rs3774372	DBP	GBP	0.10600	0.10600
LY6G5C	rs805303	DBP	GBP	0.11288	0.11288
SCUBE2	rs13107325	DBP	GBP	0.23141	0.23141
SCUBE2	rs13107325	DBP	HDL	0.23141	0.23141
LY6G5C	rs805303	DBP	HTN	0.11288	0.11288
LY6G5C	rs805303	DBP	SBP	0.11288	0.11288
WDYHV1	rs6271	DBP	SBP	0.14741	0.14741
RNF216	rs751984	DBP	SBP	0.04799	0.04799
HHLA3	rs6271	DBP	SBP	0.16697	0.16697
SCUBE2	rs13107325	DBP	SBP	0.23141	0.23141
SNAP25	rs943037	DBP	SBP	0.39924	0.39924
STAP2	rs6271	DBP	SBP	0.23228	0.23228
LTBP1	rs6271	DBP	SBP	-0.15655	-0.15655
CDC42EP5	rs751984	DBP	SBP	-0.27268	-0.27268
РРВР	rs6271	DBP	SBP	-0.23427	-0.23427
GP1BA	rs6271	DBP	SBP	-0.13584	-0.13584
RBM48	rs751984	DBP	SBP	0.05944	0.05944
RFPL3S	rs6271	DBP	SBP	0.15300	0.15300
KIF1A	rs943037	DBP	SBP	0.30254	0.30254
SCUBE2	rs13107325	GBP	HDL	0.23141	0.23141
LY6G5C	rs805303	GBP	HTN	0.11288	0.11288
LY6G5C	rs805303	GBP	SBP	0.11288	0.11288
SCUBE2	rs13107325	GBP	SBP	0.23141	0.23141
ZNF461	rs17608766	GBP	SBP	-0.08052	-0.08052
TSC22D1	rs4074793	GGT	LEP	0.09487	0.09487
DEGS2	rs1335645	GGT	LEP	0.21889	0.21889
FADS2	rs174546	HDL	LDL	-0.13862	-0.13862
SERPINB10	rs12748152	HDL	LDL	0.29271	-0.29271
SCUBE2	rs13107325	HDL	SBP	0.23141	0.23141

FADS2	rs174546	HDL	тс	-0.13862	-0.13862
AGAP8	rs970548	HDL	TC	0.13843	0.13843
SLFN13	rs1532085	HDL	TC	-0.07171	-0.07171
HIST1H2BF	rs1532085	HDL	TC	-0.15253	-0.15253
FADS2	rs174546	HDL	TG	-0.13862	0.13862
SERPINB10	rs12748152	HDL	TG	0.29271	-0.29271
SLFN13	rs1532085	HDL	TG	-0.07171	-0.07171
SIGLEC14	rs731839	HDL	TG	0.15578	-0.15578
HIST1H2BF	rs1532085	HDL	TG	-0.15253	-0.15253
LY6G5C	rs805303	HTN	SBP	0.11288	0.11288
RHD	rs12027135	LDL	тс	-0.35395	-0.35395
FADS2	rs174546	LDL	тс	-0.13862	-0.13862
PLEC	rs11136341	LDL	тс	-0.06094	-0.06094
STXBP4	rs1800562	LDL	тс	0.21535	0.21535
SPEG	rs12670798	LDL	тс	0.08796	0.08796
HLA.DQB2	rs3177928	LDL	тс	-0.27277	-0.27277
FAM221A	rs1800562	LDL	тс	0.13735	0.13735
HSPB3	rs7640978	LDL	тс	-0.43333	-0.43333
ARMCX1	rs11563251	LDL	TC	0.09314	0.09314
FADS2	rs174546	LDL	TG	-0.13862	0.13862
SERPINB10	rs12748152	LDL	TG	-0.29271	-0.29271
FADS2	rs174546	ТС	TG	-0.13862	0.13862
IL33	rs1495741	ТС	TG	-0.19644	-0.19644
SLFN13	rs1532085	ТС	TG	-0.07171	-0.07171
HIST1H2BF	rs1532085	тс	TG	-0.15253	-0.15253

Anhangtabelle 8: Gemeinsame Gene aufgrund anderer Effekte in Blutgewebe

Gemeinsame Gene in Bezug auf das Gewebe Blut, deren Gemeinsamkeit zwischen den Genlisten nicht durch die Effekte auf die Genexpression desselben SNPs zu erklären sind. Die Spalten **Gensymbol** sowie **Merkmal 1** und **Merkmal 2** geben hierbei die Merkmale an, in deren Genlisten das entsprechende Gen gefunden wurde. Die Spalten **SNP Merkmal 1** bzw. **SNP Merkmal 2** zeigen die SNPs, welche die Expression des entsprechenden Gens in Bezug auf Merkmal 1 bzw. Merkmal 2 beeinflussen. Die Spalten **eQTL Effektgröße Merkmal 1** bzw. **eQTL Effektgröße Merkmal 2** zeigen das Ausmaß der Beeinflussung der Genexpression bezüglich der entsprechenden Merkmale und SNPs. Die Spalte **r² SNPs** zeigt die die Höhe des r² des entsprechenden SNP Paares.

Gensymbol	Merkmal 1	Merkmal 2	SNP Merkmal 1	eQTL Effektgröße	SNP Merkmal 2	eQTL Effektgröße	r² SNPs
				Merkmal 1		Merkmal 2	
FADS2	ALP	HDL	rs174601	0.12753	rs174546	-0.13862	0.82
FADS2	ALP	LDL	rs174601	0.12753	rs174546	-0.13862	0.82
FADS2	ALP	тс	rs174601	0.12753	rs174546	-0.13862	0.82
FADS2	ALP	TG	rs174601	0.12753	rs174546	0.13862	0.82
HIST1H2BF	AMD	HDL	rs2043085	0.15090	rs1532085	-0.15253	0.95
SLFN13	AMD	HDL	rs2043085	0.06900	rs1532085	-0.07171	0.95
HIST1H2BF	AMD	ТС	rs2043085	0.15090	rs1532085	-0.15253	0.95
SLFN13	AMD	тс	rs2043085	0.06900	rs1532085	-0.07171	0.95
HIST1H2BF	AMD	TG	rs2043085	0.15090	rs1532085	-0.15253	0.95
SLFN13	AMD	TG	rs2043085	0.06900	rs1532085	-0.07171	0.95
IBSP	CAD	DBP	rs17080091	-0.28789	rs17080093	-0.27718	0.89
SF3A3	CAD	PP	rs61776719	-0.03205	rs4360494	-0.03166	0.96
FADS2	HDL	LEP	rs174546	-0.13862	rs174601	0.12753	0.82
FADS2	LDL	LEP	rs174546	-0.13862	rs174601	0.12753	0.82
FADS2	ТС	LEP	rs174546	-0.13862	rs174601	0.12753	0.82
FADS2	TG	LEP	rs174546	0.13862	rs174601	0.12753	0.82

Anhangtabelle 9: Merkmalspaare Blutgewebe

Merkmalspaare bezüglich des Gewebes Blut, zu welchen gemeinsame Gene zwischen den Genlisten identifiziert wurden. In den Spalten Gemeinsame SNPs bzw. Gemeinsame Gene ist die Anzahl an Überschneidungen zwischen den zur Analyse verwendeten SNPs bzw. den Genlisten der Merkmale des jeweiligen Merkmalspaares aufgeführt. Die Spalte Gemeinsame Gene aufgrund gemeinsamer SNPs zeigt die Anzahl an Genen, deren Gemeinsamkeit zwischen den Genlisten der beiden Merkmale auf den Effekten desselben SNPs beruht, während die Spalte Gemeinsame Gene aufgrund anderer Effekte die Anzahl an Genen angibt, deren Gemeinsamkeit zwischen der beiden Merkmale nicht auf den Effekten desselben SNPs beruht. Die Spalte Gemeinsame Gene aufgrund anderer Effekte in LD (r² <0.5) zeigt die Anzahl an gemeinsamen Genen, deren Gemeinsamkeit zwischen den Genlisten der beiden Merkmale nicht auf den Effekten desselben SNPs beruht. Die Spalte Gemeinsame Gene aufgrund anderer Effekte: nicht in LD (r² <0.5) zeigt die Anzahl an gemeinsamen Genen, deren Gemeinsamkeit zwischen den Genlisten der beiden Merkmale nicht auf den Effekten desselben SNPs beruht. Die Spalte Gemeinsamkeit zwischen den Genlisten der beiden Merkmale nicht auf den Effekten desselben SNPs beruht. Die Spalte Gemeinsamkeit zwischen den Genlisten der beiden Merkmale nicht auf den Effekten desselben SNPs beruht. Die Spalte Gemeinsamkeit zwischen den Genlisten der beiden Merkmale nicht auf den Effekten desselben SNPs beruht und bei denen die beiden SNPs, die der Gemeinsamkeit zwischen den Genlisten zugrunde liegen, sich nicht in LD befinden und somit unabhängig sind.

				Gemeinsame Gene	Gemeinsame Gene	Gemeinsame Gene aufgrund
Merkmal	Merkmal	Gemeinsame	Gemeinsame	aufgrund	aufgrund	anderer
1	2	SNPs	Gene	gemeinsamer	anderer	Effekte: nicht
				SNPs	Effekte	in LD (r ² < 0.5)
DBP	SBP	45	14	14	0	-
LDL	тс	47	9	9	0	-
ALP	LEP	14	4	4	0	-
HDL	TG	17	5	5	0	-
GBP	SBP	22	3	3	0	-
DBP	GBP	23	3	3	0	-
HDL	тс	13	4	4	0	-
ТС	TG	12	4	4	0	-
GGT	LEP	25	2	2	0	-
HDL	LDL	8	2	2	0	-
LDL	TG	10	2	2	0	-
AMD	HDL	0	2	0	2	-
AMD	тс	0	2	0	2	-
AMD	TG	0	2	0	2	-
GBP	HTN	11	1	1	0	-
HTN	SBP	9	1	1	0	-
DBP	HTN	8	1	1	0	-
ALP	LDL	1	1	0	1	-
ALP	HDL	0	1	0	1	-
ALP	тс	1	1	0	1	-
ALP	TG	0	1	0	1	-
GBP	HDL	1	1	1	0	-
LDL	LEP	1	1	0	1	-
HDL	LEP	0	1	0	1	-
тс	LEP	2	1	0	1	-
TG	LEP	1	1	0	1	-
CAD	GBP	2	1	1	0	-
CAD	PP	1	1	0	1	-
HDL	SBP	1	1	1	0	-
DBP	HDL	1	1	1	0	-
CAD	SBP	3	1	1	0	-
CAD	DBP	2	1	0	1	-

Anhangtabelle 10: Pleiotrope Gene mit Übersicht zu aktuellem Wissensstand und bekannten Assoziationen

Die Spalten **Gen**, **Gensymbol**, **Gewebe** und **Merkmal** geben sowohl das jeweilige pleiotrope Gen mit Gensymbol an als auch in den Genlisten welcher Merkmale das Gen gefunden wurde und bezüglich welchen Gewebes. In der Spalte **Anzahl PubMed Studien zum Gen** ist die Anzahl der Studien in der PubMed Gen Datenbank zum entsprechenden Gen aufgetragen. Die Spalte **Bekannte Merkmalsassoziationen mit Gen** gibt die Merkmale und Erkrankungen an, zu welchen in den PubMed Studien Zusammenhänge mit dem Gen beschrieben werden.

				Anzahl PubMed	
Gen	Gensymbol	Gewebe	Merkmale	Gen	Bekannte Merkmalsassoziationen mit Gen
ENSG00000175877	WBSCR28	Leber	ALP; CAD; LDL; LEP; TC	3	Williams-Beuren Syndrom e[148]
ENSG0000082074	FYB	Leber	LDL; TC	59	Thrombozytopenie b[243]; Lupus Erythematodes e[244]
ENSG00000142794	NBPF3	Blut; Leber	ALP; LEP	23	ALP b[245]; VitaminB6 Spiegel b[246]; Neuroblastom b[247]; NSCLC b[247]
ENSG00000100997	ABHD12	Leber	ALP; LEP	22	PHARC b[248,249]
ENSG00000134352	IL6ST	Leber	ALP; LEP	251	Hepatozelluläre Adenome e[250,251]; CAD e[252]; Diabetische Retinopathie b[253]; Chronisch spontane Urtikaria b[254]; Asthma b[255]; Herzinsuffizienz b[256]; NASH e[257]; Multiples Myelom e[258]; Nichtalkoholische Fettleber NAFLD e[257]; Metabolisches Syndrom b[259]; Rheumatoide Arthritis b[260]; Nasopharyngeales Carcinom b[261]; Blutdruck e[262] Alzheimer b[263,264]; System. Lupus Erythematodes b[265]; Multiple Sklerose e[266] b[267]; Keloidwachstum b[268]; Rheumatische Herzerkrankung b[269]; Frontotemporale Demenz b[270]; Hämorrhagisches
ENSG00000198502	HLA.DRB5	Leber	DBP; GBP; HTN; SBP	202	Fieber mit renalem Syndrom b[271]; Diabetes b[272]
ENSG00000167077	MEI1	Leber	DBP; GBP; SBP	8	Azoospermie b[273]
ENSG00000257207	LIMS3L	Leber	DBP; GBP; SBP	Fehlt in PubMed Gene Database	-
ENSG0000124935	SCGB1D2	Leber	DBP; SBP	16	-
ENSG00000120729	МҮОТ	Leber	DBP; SBP	45	Myopathie b[274]
ENSG00000160539	PPAPDC3	Leber	DBP; GBP; HTN; SBP	7	Myoblastendifferenzierung e[275]

ENSG00000146147	MLIP	Leber	GGT; LEP	16	Kardiomyopathie e[276]
ENSG00000270672	MTRNR2L6	Leber	GGT; LEP	10	Alkoholbedingte Pankreatitis b[277]
ENSG00000169894	MUC3A	Leber	GGT; LEP	21	Morbus Crohn b[278]; Colitis ulzerosa b[278]; Nierenkarzinom b[279];
					GGT b[230]; AST b[280]; ALT b[230,280]; Kolorektales
ENSG00000115468	EFHD1	Leber	GGT; LEP	22	Karzinom b[281]
ENSG00000137078	SIT1	Leber	HDL; LDL; TC	20	-
	RP11.1105G2.			Fehlt in PubMed	
ENSG00000258365	3	Leber	HDL; TC	Gene Database	-
ENSG00000204764	RANBP17	Leber	HDL; TC	14	Viszeralfett b[282]
ENSG00000134222	PSRC1	Leber	CAD; LDL; TC	81	CAD b[86–88] e[89]; LDL b[87,90,91] e[89]; TC e[89]; HDL e[87,89]
ENSG00000134243	SORT1	Leber	CAD: LDL; TC	163	CAD b[89,92]; LDL b[89,93–95,97] e[89,96,185]; TC b[89,97] e[89]; TG e[89]; Atherosklerose e[185,186,283] b[284];Chron. Nierenversagen b[285]; Brustkrebsagressivität e[286]; Alzheimer b[287]; Depression b[288]
ENSG00000143126	CELSR2	Leber		55	CAD b[98,99]; LDL b[93,100,101]; TC b[100–102]; HDL b[100]; Aortenaneurysma b[289] ; CRP b[100]; Skoliose b[290]
ENSG00000178852	FECAB13	Leher		6	-
ENSG00000187010	RHD	Blut: Leber		158	Blutgruppe(Bhesus) b[291–294]
					LDL b[97,103]; TC b[97,103]; ALP b[295]; Plasma-VWF b[296]; FVIII Aktivität b[296]; ApoB b[103]; H5N1 b[297]; Thrombozytopenie e[298]; Diabetische Retinopathie
ENSG00000110080	ST3GAL4	Leber	LDL; TC	51	e[299]; Magenkarzinom b[300] e[301]
ENSG00000070182	SPTB	Leber	LDL; TC	99	Sphärozytose b[302,303]; Ischämischer Insult b[304]; Knochendichte b[305]
ENSG00000130175	PRKCSH	Leber	LDL; TC	75	PCLD b[306,307]
ENSG00000136861	CDK5RAP2	Leber	LDL; TC	76	Gehirnvolumen b[308]; Rheumatoide Arthritis b[309]
ENSG00000105877	DNAH11	Leber	LDL; TC	30	TC b[104]; LDL b[104]; HDL b[104]; TG b[104];

					Kartagener Syndrom b[310,311] e[312]
					TC b[97,104]; LDL b[97,104,105]; HDL b[95,97,104]
					TG b[95,97,104]; CAD b[313]; Plasma-Fettsäuren
ENSG00000134824	FADS2	Blut	HDL; LDL; TC; TG	155	b[314,315]; Asthma b[316]; Bipolare Störung b[317]
					ALP b[230]; Prostatakarzinom e[318]; Osteosarkom
ENSG00000100994	PYGB	Blut	ALP; LEP	50	e[319]; Myokardinfarkt b[320]
					Ovarialkarzinom e[321] b[322]; Endometriumkarzinom
ENSG00000125787	GNRH2	Blut	ALP; LEP	57	e[321] b[322]
ENSG00000197808	ZNF461	Blut	CAD; GBP; SBP	9	-
					DBP b[106–109] ; Hypertension b[110,111];
					Aortendissektion b[323]; Multiples Myelom b[324];
ENSG00000168038	ULK4	Blut	DBP; GBP	30	Schizophrenie e[325]
ENSG00000204428	LY6G5C	Blut	DBP; GBP; HTN; SBP	5	-
					Tumorsupression e[166]; Angiogenese e[167]; Gliome
ENSG00000175356	SCUBE2	Blut	DBP; GBP; HDL; SBP	26	e[326]; Mammakarzinom e[327]
ENSG00000156795	WDYHV1	Blut	DBP; SBP	16	-
					Gordon-Holmes-Syndrom e[328]; Kolorektales Karzinom
ENSG0000011275	RNF216	Blut	DBP; SBP	45	e[329]; Gewichtszunahme b[330]
ENSG00000197568	HHLA3	Blut	DBP; SBP	7	-
					Herzversagen b[331]; Insulinsekretion e[332];
					Bipolare Störung b[333]; ADHS b[334];
					Schizophrenie b[335,336] e[337]; Autismus b[338];
					Glykämieparameter b[339]; Depression b[336,340];
ENSG00000132639	SNAP25	Blut	DBP; SBP	211	Persönlichkeitsmerkmale b[341]; Intelligenz b[342]
ENSG00000178078	STAP2	Blut	DBP; SBP	30	Mammakarzinom e[343]; Prostatakarzinom e[344]
					HCC b[345]; CAD b[346] ; SVD e[347]; Alkoholkonsum
ENSG0000049323	LTBP1	Blut	DBP; SBP	101	b[348]; Gliom e[349]; Ovarialkarzinom b[350]
ENSG00000167617	CDC42EP5	Blut	DBP; SBP	9	-
					Atherosklerose b[351] ; CAD b[352]; Nierenzellkarzinom
ENSG00000163736	PPBP	Blut	DBP; SBP	102	b[353]; Parodontitis b[354]
ENSG00000185245	GP1BA	Blut	DBP; SBP	325	Bernard-Soulier-Syndrom b[355]; CAD b[356,357];

					Myokardinfarkt b[358–360]; Apoplex b[361];
					Thrombozytengröße e[362]; Mundhöhlenkarzinom
					b[363]; Aspirinsensibilität b[364]; Lungenembolie b[365];
					Lungenmetastasen e[366]
ENSG00000127993	RBM48	Blut	DBP; SBP	12	-
ENSG00000205853	RFPL3S	Blut	DBP; SBP	3	-
					Spastische Paraplegie b[367–370]; PEHO-Syndrom b[371];
ENSG00000130294	KIF1A	Blut	DBP; SBP	58	Mammakarzinom b[372]; HNSCC b[373]; HSN2C b[374]
					Zervixkarzinom e[375]; Speicheldrüsenkarzinom e[376];
					diabetische Mikroangiopathie b[377]; Prostatakarzinom
ENSG00000102804	TSC22D1	Blut	GGT; LEP	47	b[378]
					Kognitive Fähigkeiten bei Schizophrenie b[379];
ENSG00000168350	DEGS2	Blut	GGT; LEP	8	Herzstillstand[380]
ENSG00000242550	SERPINB10	Blut	HDL; LDL; TG	14	Prostatakarzinom b[381]
				Fehlt in PubMed	
ENSG00000174194	AGAP8	Blut	HDL; TC	Gene Database	-
ENSG00000154760	SLFN13	Blut	AMD; HDL; TC; TG	11	-
ENSG00000277224	HIST1H2BF	Blut	AMD; HDL; TC; TG	26	-
ENSG00000254415	SIGLEC14	Blut	HDL; TG	9	Plasminogenspiegel b[382]; COPD Exazerbation b[383]
					Epidermolysis-Bullosa-Simplex b[384–386];
					Vorhofflimmern b[387]; Pankreaskarzinom e[388,389];
					Hodenkarzinom b[390]; Muskeldystrophie b[391,392];
					Kongenitales-Myasthenes-Syndrom b[393];
					Plasmafibrinogen b[394]; Ruheherzfrequenz b[395];
ENSG00000178209	PLEC	Blut	LDL; TC	170	Arthrose b[396]; Ösophaguskarzinom b[397]
					Insulinsekretion e[398]; Zelluläre Glukoseaufnahme
					e[399,400]; Diabetes Typ2 b[401]; DNA-Methylierung
					b[402]; Plattenepithelkarzinom e[403]; Mammakarzinom
ENSG0000166263	STXBP4	Blut	LDL; TC	33	b[404]
ENSG0000072195	SPEG	Blut	LDL; TC	26	Kardiomyopathie b[405] e[406]; Atherosklerose b[407]
					Rheumatoide Arthritis b[408]; Chronische Hepatitis B
ENSG00000232629	HLA.DQB2	Blut	LDL; TC	68	b[409,410]; Lymphom b[411]; Kawasaki-Syndrom b[412]

ENSG00000188732	FAM221A	Blut	LDL; TC	6	Parkinson b[413]; Alzheimer b[414]
ENSG00000169271	HSPB3	Blut	LDL; TC	11	-
ENSG00000126947	ARMCX1	Blut	LDL; TC	16	Karzinome b[415]; Metastasenbildung e[416]; Zervixkarzinom b[417]; Mammakarzinom e[418];
ENSG00000137033	IL33	Blut	TC; TG	416	Adipositas b[419]; Asthma b[420–422]; Allergie b[423,424]; Endometriose b[425]; beiges Fettgewebe e[426]; Depression b[427,428]; Kolorektales Karzinom e[429]; Nierenzellkarzinom e[430]; Tumoraggressivität b[431]; Rhinosinusitis b[432,433]; Still-Syndrom b[434]; Endometriumkarzinom b[435]; Cholesterolstoffwechsel e[436]; Polyzystisches Ovarialsyndrom b[437]; Abort b[438]; Colitis ulcerosa b[439]; Systemische Sklerose b[440]; Atherosklerose b[441]; Ösophagitis e[442]; Myokardinfarkt b[443]; COPD b[444]; Mammakarzinom e[445]; CAD b[446,447]; Riesenzellarteriitis b[448]; primär biliäre Cholangitis b[449]; pulmonale Hypertonie e[450]; Rheumatoide Arthritis b[451]; Lungenkarzinom b[452]; Leberfibrose b[453]
ENSG00000138785	INTS12	Leber	AMD; HDL	24	Lungenfunktion b[209]
ENSG0000029559	IBSP	Blut	CAD; DBP	52	Knochenmetastasen b[454,455] e[456]; Knochenmineraldichte b[457,458]; Osteoarthritis e[459]; Gliom b[460]; Karies b[461]
ENSG00000183431	SF3A3	Blut	CAD; PP	79	CAD b[235]; PP b[239]

e: experimentelles Studiendesign

b: Korrelatives oder bioinformatisches Studiendesign

hellgrau hinterlegt: Gene mit vermutlich pleiotropen Wirkungen auf Merkmale aus unterschiedlichen Merkmalsgruppen

rot: Beziehung zwischen Gen und Merkmal wurde nur in derselben Studie beschreiben, aus welcher die für GRS Berechnung verwendeten GWAS SNPs stammen

orange: Beziehung zwischen Gen und Merkmal wurde in von GRS Berechnung unabhängiger GWAS Studie beschrieben oder Beziehung wurde in Korrelationsstudie beschrieben, die sich auf solche GWAS bezieht

grün: Beziehung zwischen Gen und Merkmal wurde in von GWAS unabhängiger Studie beschrieben

Danksagung

An dieser Stelle möchte ich meinen herzlichsten Dank nachstehenden Personen entgegenbringen, ohne deren Hilfe diese Promotionsschrift niemals zustande gekommen wäre:

Mein Dank gilt zuallererst Herrn Prof. Dr. Bernhard Weber dafür, diese Arbeit am Institut für Humangenetik durchführen zu können sowie an einem spannenden und unterhaltsamen Forschungsgruppen Retreat teilnehmen zu dürfen. Die Tätigkeit in Ihrem Institut habe ich wirklich sehr genossen.

Herzlichen Dank zudem an Frau Prof. Dr. Birte Kehr für die freundliche Übernahme der Funktion des Zweitgutachters.

Ganz besonders möchte ich Dr. Tobias Strunz und Dr. Christina Kiel für die tolle Unterstützung danken. Ohne euren Rat hätte ich diese Arbeit niemals durchführen können. Danke, dass ihr euch immer die Zeit genommen habt mir Dinge zu erklären und auch Verständnis hattet, wenn ihr manche Fragen nicht nur einmal beantworten musstet.

Des Weiteren möchte ich mich bei Juliane, Verena, Fabiola und natürlich auch allen weiteren Kollegen für die netten Unterhaltungen in den Pausen und die gemeinsame Zeit bedanken.

Danke auch an meine Schwester Anna, welche mir während der Anfertigung dieser Arbeit sowie in meiner gesamten bisherigen Lebenszeit eine gute Freundin und wertvolle Unterstützung war, besonders auch in schwierigen Zeiten.

Abschließend gilt ein ganz besonderer Dank meinen Eltern. Ohne ihre Unterstützung in jeglicher Hinsicht wäre mein Studium und auch mein bisheriger Lebensweg nicht möglich gewesen. Ihnen widme ich diese Arbeit.

Selbstständigkeitserklärung

Ich, Simon Stelzl, geboren am 16.04.1993 in Roding, erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.

Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Insbesondere habe ich nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeit erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Regensburg, den 24. Januar 2022

Simon Stelzl