Molecular genetic analyses to elucidate the biological basis of angiotensin-converting enzyme inhibitor-induced angioedema

Doctoral thesis

to obtain a doctorate (PhD)

from the Faculty of Medicine

of the University of Bonn

Carina Margit Mathey
from Erlangen, Germany
2024

Written with authorization of

the Faculty of Medicine of the University of Bonn

First reviewer: Prof. Dr. Markus M. Nöthen

Second reviewer: Prof. Dr. Sven Cichon

Day of oral examination: 19.02.2024

From the Institute of Human Genetics

Director: Prof. Dr. Markus M. Nöthen

The central results of this dissertation have been published in two original articles:

Mathey CM, Maj C, Scheer AB, Fazaal J, Wedi B, Wieczorek D, Amann PM, Löffler H, Koch L, Schöffl C, Dickel H, Ganjuur N, Hornung T, Forkel S, Greve J, Wurpts G, Hallberg P, Bygum A, von Buchwald C, Karawajczyk M, Steffens M, Stingl J, Hoffmann P, Heilmann-Heimbach S, Mangold E, Ludwig KU, Rasmussen ER, Wadelius M, Sachs B*, Nöthen MM*, Forstner AJ* (2022) **Molecular genetic screening in patients with ACE inhibitor/ angiotensin receptor blocker-induced angioedema to explore the role of hereditary angioedema genes.** *Frontiers in Genetics*, 13, 914376.

DOI: 10.3389/fgene.2022.914376
PMID: 35923707

Mathey CM, Maj C, Eriksson N, Krebs K, Westmeier J, David FS, Koromina M, Scheer AB, Szabo N, Wedi B, Wieczorek D, Amann PM, Löffler H, Koch L, Schöffl C, Dickel H, Ganjuur N, Hornung T, Buhl T, Greve J, Wurpts G, Aygören Pürsün E, Steffens M, Herms S, Heilmann-Heimbach S, Hoffmann P, Schmidt B, Mavarani L, Andresen T, Sørensen SB, Andersen V, Vogel U, Landén M, Bulik CM, Estonian Biobank research team, DBDS Genomic consortium, Bygum A, Magnusson PKE, von Buchwald C, Hallberg P, Ostrowski SR, Sørensen E, Pedersen OB, Ullum H, Erikstrup C, Bundgaard H, Milani L, Rasmussen ER, Wadelius M, Ghouse J*, Sachs B*, Nöthen MM*, Forstner AJ* (2024) **Meta-analysis of ACE inhibitor-induced angioedema identifies novel risk locus.** *The Journal of Allergy and Clinical Immunology*, online ahead of print.

DOI: 10.1016/j.jaci.2023.11.921
PMID: 38300190

Table of content

List of abbreviations

ACE ..... Angiotensin-converting enzyme
ACEi ..... Angiotensin-converting enzyme inhibitor
ACEi-AE ..... Angiotensin-converting enzyme inhibitor-induced angioedema
ACMG ..... American College of Medical Genetics
ACSS2 ..... Acyl-CoA synthetase short chain family member 2
ADR ..... Adverse drug reaction
ANGPT1 ..... Angiopoietin-1
APP ..... Aminopeptidase P
ARB ..... Angiotensin-II-receptor blocker
ARB-AE ..... Angiotensin-II-receptor blocker-induced angioedema
ATP1B1 ..... ATPase Na+/K+ transporting subunit beta 1
B1-R ..... Bradykinin receptor B1
B2-R ..... Bradykinin receptor B2
BDKRB1 ..... Bradykinin receptor B1
BDKRB2 ..... Bradykinin receptor B2
BK ..... Bradykinin
BK-AE ..... Bradykinin-induced angioedema
bp ..... Base pair
C1-INH ..... C1-esterase inhibitor
C1-INH-HAE ..... HAE with a C1-INH deficiency/dysfunction
C1orf112 ..... Chromosome 1 open reading frame 112
CADD ..... Combined annotation-dependent depletion
CCDC181 ..... Coiled-coil domain containing 181
CHB-CVDC/DBDS .... Copenhagen Hospital Biobank-Cardiovascular Disease Cohort/Danish Blood Donor Study
ChIP-seq ..... Chromatin immunoprecipitation sequencing
dbSNP ..... Single nucleotide polymorphism database
ddNTP ..... Dideoxynucleotide triphosphate
DNA ..... Deoxyribonucleic acid
dNTP ..... Deoxynucleotide triphosphate
dsDNA ..... Double-stranded DNA
DVT ..... Deep vein thrombosis

EDEM2 ..... Endoplasmic reticulum degradation enhancing alpha-mannosidase like protein 2
EIF6 ..... Eukaroytic translation initiation factor 6
ENCODE ..... Encyclopedia of DNA elements
EPCR ..... Endothelial protein C receptor
eQTL ..... Expression quantitative trait locus
EstBB ..... Estonian Biobank
ETV6 ..... ETS variant transcription factor 6
F12 ..... Coagulation factor XII
F5 ..... Coagulation factor V
FAM83C ..... Family with sequence similarity 83 member C
FDR ..... False discovery rate
FXII ..... Factor XII
GATK ..... Genome Analysis Toolkit
gDNA ..... Genomic DNA
GGT7 ..... Gamma-glutamyltransferase 7
gnomAD ..... Genome aggregation database
GSA ..... Global screening array
GTEx ..... Genotype-Tissue Expression project
GWAS ..... Genome-wide association study
HAE ..... Hereditary angioedema
HAE-1 ..... HAE due to C1-INH deficiency
HAE-2 ..... HAE due to C1-INH dysfunction
HAE-ANGPT1 ..... HAE with mutation in the angiopoietin-1 gene
HAE-FXII ..... HAE with mutation in the factor XII gene
HAE-HS3ST6 ..... HAE with mutation in the heparan 3-O-sulfotransferase 6 gene
HAE-KNG1 ..... HAE with mutation in the kininogen 1 gene
HAE-MYOF ..... HAE with mutation in the myoferlin gene
HAE-PLG ..... HAE with mutation in the PLG gene
HAE-UNK ..... HAE due to unknown mutations
Het*P* ..... Heterogeneity *p*-value
hg19 ..... Human reference genome build 19
HGMD ..... Human Gene Mutation Database
HK ..... High molecular weight kininogen
HNR ..... Heinz Nixdorf Recall
HS3ST6 ..... Heparan 3-O-sulfotransferase 6

HWE ..... Hardy-Weinberg equilibrium
Indel ..... Insertion/deletion
kb ..... Kilobase
KCNMA1 ..... Potassium calcium-activated channel subfamily M alpha 1
KLKB1 ..... Kallikrein B1
KNG1 ..... Kininogen 1
LD ..... Linkage disequilibrium
LK ..... Low molecular weight kininogen
MAF ..... Minor allele frequency
MAP1LC3A ..... Microtubule associated protein 1 light chain 3 alpha
METTL18 ..... Methyltransferase like 18
MME ..... Membrane metalloendopeptidase (NEP)
MMP24 ..... Matrix metallopeptidase 24
MYH7B ..... Myosin heavy chain 7B
MYOF ..... Myoferlin
nC1-INH-HAE ..... HAE with normal C1-INH
NGS ..... Next generation sequencing
NME7 ..... NME/NM23 family member 7
OR ..... Odds ratio
P_{adjust} ..... Bonferroni adjusted *p*-value
PC ..... Principal component
PCA ..... Principal component analysis
PCR ..... Polymerase chain reaction
PIP ..... Posterior inclusion probability
PLG ..... Plasminogen
PRKCQ ..... Protein kinase C theta
PROCR ..... Protein C receptor
PRS ..... Polygenic risk score
QC ..... Quality control
RefSeq ..... NCBI Reference sequences
RNA ..... Ribonucleic acid
SELE ..... Selectin E
SERPING1 ..... Serpin family G member 1
SLC19A2 ..... Solute carrier family 19 member 2
smMIP ..... Single-molecule Molecular Inversion Probe

SNP ..... Single nucleotide polymorphism
TIE-2 ..... Tie-2 tyrosin kinase receptor
TMEM119 ..... Transmembrane protein 119
TRPC4AP Transient receptor potential cation channel subfamily C member 4 associated protein
UKB ..... UK Biobank
VTE ..... Venous thromboembolism
WES ..... Whole exome sequencing
WGS ..... Whole genome sequencing
XPNPEP2 ..... X-prolyl aminopeptidase 2

1 Abstract

Angioedema is a rare, potentially life-threatening adverse reaction to angiotensin-converting enzyme inhibitors (ACEi) and angiotensin-II-receptor blockers (ARB), two commonly prescribed drug classes with an indication in antihypertensive therapy. Like hereditary angioedema (HAE), angioedema induced by an ACEi (ACEi-AE) or ARB (ARB-AE) is mediated by bradykinin and thus non-allergic in nature. Findings from previous research suggest that individual susceptibility to ACEi/ARB-AE is shaped by a genetic predisposition and non-genetic risk factors. Several non-genetic risk factors have been reported and include for example, female sex, smoking and a history of seasonal allergies. In addition, recent genetic studies have identified the first genetic loci involved in ACEi/ARB-AE susceptibility. However, understanding of the exact pathophysiology remains limited.

To further elucidate the genetic factors underlying ACEi/ARB-AE risk, this dissertation employed two strategies: (i) exploration of the potential involvement of pathogenic HAE-associated variants and genes, and (ii) genome-wide investigation of the contribution of common variants. As groundwork for the abovementioned genetic studies, the vARIANCE study, the largest case collection of German/Austrian ACEi/ARB-AE patients to date, encompassing genetic material and comprehensive phenotypic data, was established.

The first study was based on the observation that the most recently discovered HAE subtypes with normal C1-inhibtor levels share certain clinical features with ACEi/ARB-AE, raising the possibility of an inaccurate diagnosis. To investigate this and to determine whether the genetic factors involved might converge in the same genes, targeted re-sequencing of five then-known HAE-associated genes (*SERPING1*, *F12*, *PLG*, *ANGPT1*, and *KNG1*) was performed in a large cohort of 197 ACEi/ARB-AE patients. The study revealed that none of the patients carried a known causal HAE variant. In addition, no other common or rare variant in these five genes showed a significant association with ACEi/ARB-AE. These findings suggest that an underlying pathogenic HAE-associated variant in patients initially diagnosed with ACEi/ARB-AE is infrequent at best.

The second study addressed the contribution of common variants by the means of a large-scale genome-wide association study meta-analysis that included, for the first time,

more than 1,000 ACEi-AE patients. Three genome-wide significant loci were identified, including a novel risk locus on chromosome 20q11.22. The other two loci mapped to 1q24.2 and 14q32.2 and are consistent with two loci previously reported with exome- and genome-wide significance. Integrative secondary analyses highlighted previously reported genes (*BDKRB2*, *F5*) and biologically plausible novel candidate genes (*PROCR*, *EDEM2*). Cross-ancestry analyses involving European and African-American individuals revealed shared common variants with concordant effect directions and sizes, in particular at the three genome-wide significant risk loci. As such, the present findings indicate an involvement of bradykinin signaling, coagulation and fibrinolysis pathways in ACEi-AE pathophysiology. Interestingly, two of the three identified risk loci corresponded to loci associated with venous thromboembolism, an observation that should be further investigated in future studies. In addition, preliminary evidence suggested the relevance of the three genome-wide significant loci across different ancestries, thus underscoring their role in the pathophysiology of this adverse drug reaction.

In conclusion, the present dissertation contributed considerably to the further understanding of the genetic background of ACEi/ARB-AE and provided important groundwork for future research in this field, not least *via* the establishment of a deeply phenotyped cohort of ACEi/ARB-AE patients (vARIANCE study).

2 Introduction

2.1 Angioedema and its classification

The first mention of an acute localized edema dates back to 1848. However, it was German physician Heinrich Quincke's report of a case series of swelling symptoms in 1882 after which "angioedema" ("angioneurotic edema" or "Quincke's edema") was for the first time recognized as a distinct entity (Bruun, 1953). Since then, angioedema research has been characterized by periodic "leaps" of interest. Yet, overall, tremendous progress has been made, e.g., by the unraveling of mediators and signaling pathways involved (Reshef et al., 2016).

Table 1 | Angioedema classification.

Bradykinin-induced AE	C1-INH deficiency/dysfunction	Inherited	**C1-INH-HAE**
		Acquired	AAE-C1-INH
	C1-INH normal	Inherited	**nC1-INH-HAE**
		Acquired	**ACEi-AE/ARB-AE** other drug-induced AE*
Mast cell mediator-induced AE	IgE-mediated	AE with anaphylaxis AE with/without wheals (urticaria)	
	Non-IgE mediated	AE with/without wheals (urticaria)	
Unknown mediator	Idiopathic AE		

*e.g., AE induced by gliptins or neprilysin inhibitors. Abbreviations: AAE-C1-INH, acquired angioedema due to C1-INH deficiency | ACEi-AE, angiotensin-converting enzyme inhibitor-induced angioedema | ARB-AE, angiotensin-II-receptor blocker-induced angioedema | AE, angioedema | C1-INH-HAE, hereditary angioedema with a deficiency/dysfunction of C1-INH | nC1-INH-HAE, hereditary angioedema with normal C1-INH. Table adapted from Maurer et al. (2022).

Clinically, angioedema refers to a transient, limited swelling of subcutaneous or submucosal tissue that is caused by a rapid increase in blood vessel permeability, which in turn leads to local plasma extravasation (Kaplan and Greaves, 2005). Nowadays, an angioedema is most commonly classified according to its presumed mediator (Table 1). Thus, "allergic" angioedema triggered by mast cell mediators such as histamine is distinguished from "non-allergic" angioedema mediated by bradykinin (BK) and angioedema with an unknown mediator (Sharma et al., 2021; Maurer et al., 2022).

The focus of the present dissertation is on the rare form of BK-induced angioedema (BK-AE). Specifically, the acquired types of angioedema induced by the intake of an

angiotensin-converting enzyme inhibitor (ACEi) or an angiotensin-II-receptor blocker (ARB) and, to some extent, the familial, hereditary forms of angioedema (HAE) are addressed. Therefore, the clinical features, underlying disease mechanisms known to date, and genetic findings of these types of angioedema will be discussed in more detail in the following parts of the introduction.

2.2 Bradykinin-induced angioedema

In general, BK-AE comprises inherited and acquired forms; moreover, they can be classified according to their antigenic and functional C1-esterase inhibitor (C1-INH) levels (Table 1; (Maurer et al., 2022)). Most commonly affected by such angioedema are the skin, the gastrointestinal tract, and the upper respiratory tract, with slight subtype-specific patterns regarding the frequency of affected sites (Cicardi et al., 2014). If the head and neck region is affected, as shown in Figure 1, the angioedema can even progress into airway obstruction and thus become a life-threatening event (Caballero et al., 2011). BK-AE are in general not accompanied by urticaria and/or pruritus, whereas angioedema triggered by mast cell mediators usually is (Cicardi et al., 2014). Given the non-allergic etiology underlying BK-AE, they do not respond to treatment with antihistamines, corticosteroids, and/or epinephrine, but rather require targeted therapy that addresses BK synthesis or BK receptor activity (Cicardi et al., 2016; Maurer et al., 2022).

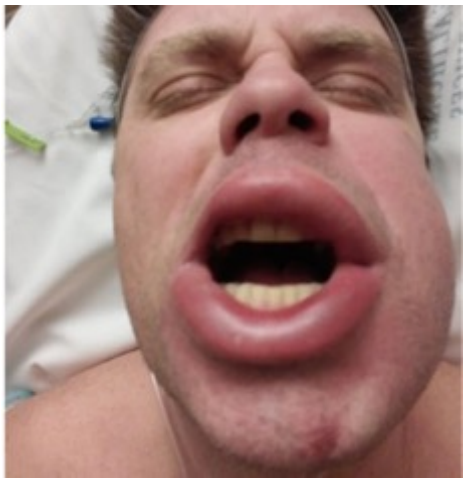

Figure 1 | Typical clinical picture of a bradykinin-induced angioedema.*

2.2.1 Bradykinin and its related pathways

BK is a vasoactive peptide of the kinin family that has various physiological effects, including regulation of blood pressure and vascular permeability and the promotion of inflammatory symptoms like vasodilation and pain (Maurer et al., 2011). Most of these biological effects are exerted *via* activation of the bradykinin receptor B2 (B2-R), which has a high affinity for native kinins such as BK and is constitutively expressed on, for instance, endothelial and smooth muscle cells (Blais et al., 2000; Leeb-Lundberg, 2005).

* derived from de Maat et al. (2018) under the terms of the CC-BY-NC-ND License

In contrast, expression of the second type of bradykinin receptor, the bradykinin receptor B1 (B1-R), is induced by states of tissue injury or inflammation. Moreover, B1-R is thought to interact mainly with the active metabolite des-Arg^9-BK, rather than with BK itself (Moreau et al., 2005).

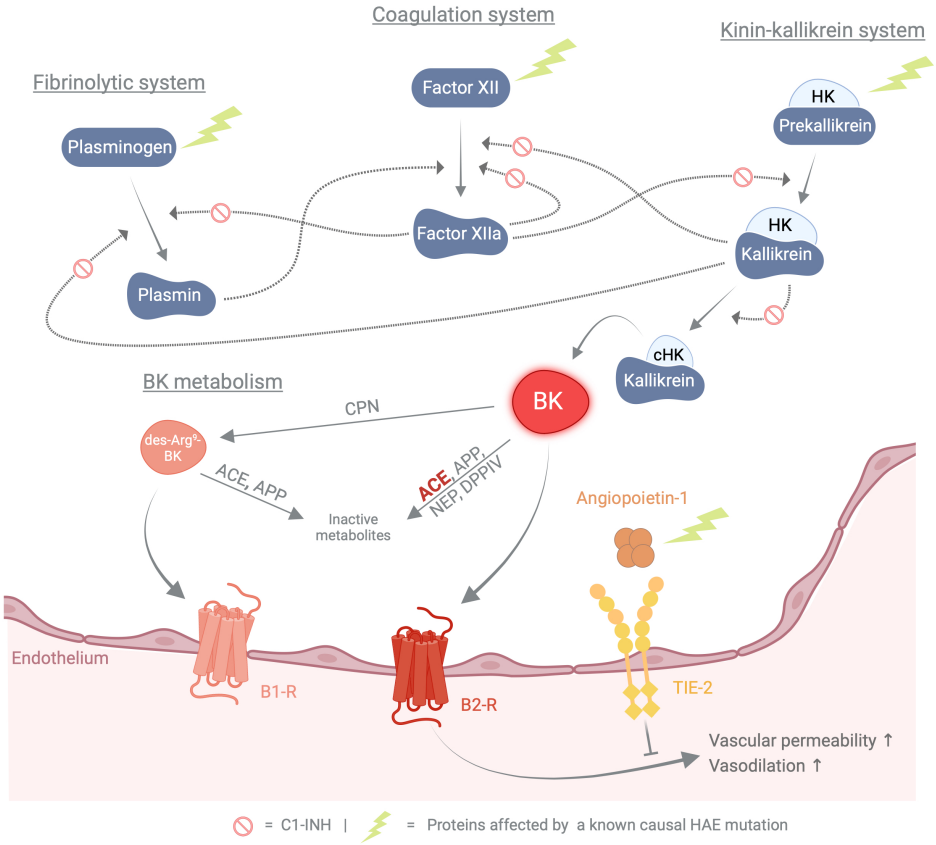

Figure 2 | Main components and pathways involved in bradykinin formation, metabolism and signaling.

Abbreviations: ACE, angiotensin-converting enzyme | APP, aminopeptidase P | B1-R, bradykinin receptor B1 | B2-R, bradykinin receptor B2 | BK, bradykinin | cHK, cleaved high molecular weight kininogen | CPN, carboxypeptidase N | DPPIV, dipeptidyl peptidase-4 | C1-INH, C1-esterase inhibitor | HAE, hereditary angioedema | HK, high molecular weight kininogen| NEP, neutral endopeptidase (neprilysin) | TIE-2, tie-2 tyrosine kinase receptor. Figure created with BioRender.com.

A schematic overview of the interrelationships and major components involved in the formation, signaling, and metabolism of BK is provided in Figure 2. Briefly, components of several plasma enzyme cascades, such as contact activation/coagulation, kinin-kallikrein, and the fibrinolytic system, are involved in the formation of BK. The three most important components are coagulation factor XII (FXII), prekallikrein and high molecular weight kininogen (HK), which are collectively referred to as the "bradykinin-forming

cascade" (Kaplan and Ghebrehiwet, 2010). Within this cascade, activated FXII will activate prekallikrein to form kallikrein, which eventually cleaves BK from HK (Schmaier, 2016). FXII is initially activated by slow auto-activation, after which the generated kallikrein leads to rapid feedback activation of FXII (Kaplan and Ghebrehiwet, 2010). In addition, plasmin can act as an FXII activator (Hofman et al., 2016). Another key component is the C1-INH, which controls BK formation as an irreversible inhibitor of FXII and kallikrein (Cicardi et al., 2005). BK itself has a very short plasma half-life (~15-17s) and is rapidly metabolized by plasma enzymes and enzymes along the surface of endothelial cells (Kaplan and Ghebrehiwet, 2010). Enzymes involved in the degradation of BK are the angiotensin-converting enzyme (ACE), aminopeptidase P (APP), carboxypeptidase N and to some extent, neutral endopeptidase and dipeptidyl peptidase-4. However, ACE is the main metabolic enzyme in the degradation of BK (Blais et al., 2000; Moreau et al., 2005). At the endothelial cell, BK modulates among others vascular permeability and vasodilation *via* the B2-R (Maurer et al., 2011; Blaes and Girolami, 2013); processes that are counterbalanced by the angiopoietin-1/TIE-2 axis and its stabilizing effect on the vascular endothelium (Baffert et al., 2006; Ngok et al., 2012).

In the context of BK-AE, perturbations in different pathways involved in the regulation of endothelial permeability might ultimately result in the tissue swelling observed in angioedema patients (Debreczeni et al., 2021). Such dysregulations may result from inhibition of BK degradation, as is thought to be the case in ACEi-induced angioedema (ACEi-AE), or may be the consequence of uncontrolled BK production or imbalanced BK signaling at the endothelial cell, as has been shown to be the case in HAE (Montinaro and Cicardi, 2020; Veronez et al., 2021).

2.2.2 ACEi/ARB-induced angioedema

The first report of an ACEi-AE was published two years after the world-wide launch of captopril, the first ACEi (Jett, 1984). Nowadays, angioedema is a recognized rare adverse drug reaction (ADR) of this class of drugs. The estimated 12-month population prevalence of ACEi-AE ranges from 0.004% to 0.026%, depending on the general population under study (Aygören-Pürsün et al., 2018), while the annual incidence rate of ACEi-AE is reported to be up to 0.7% in patients taking these drugs (Miller et al., 2008; Banerji et al., 2017). However, despite this rather low incidence, ACEi-AE is thought to be

the most common form of non-hereditary BK-AE. For instance, a study in the US reported that approximately one-third of all angioedema cases in the emergency department were caused by an ACEi (Banerji et al., 2008). This quite high number is attributable to the widespread use of ACEi, following their broad range of indications in the treatment of hypertension and other hypertensive-related disorders such as heart failure and chronic kidney disease (Williams et al., 2018; Unger et al., 2020). In addition, angioedema has also been observed as an adverse reaction to ARBs, another class of antihypertensive drugs. However, in contrast to ACEi-AE, ARB-induced angioedema (ARB-AE) presents with lower incidence rates (Makani et al., 2012; Rasmussen et al., 2019).

Clinically, ACEi-AE is most commonly observed in the head and neck region, with frequent reports of lip or tongue swellings (Kostis et al., 2005; Rasmussen et al., 2017), whereas the gastrointestinal tract is rarely affected (Benson et al., 2013). Symptoms of ACEi-AE are usually mild, but life-threatening and even fatal cases caused by upper airway obstruction have been reported (Banerji et al., 2008). A particular challenge in ACEi-AE is the variable time-to-symptom onset, which can range from months to years after treatment initiation (Miller et al., 2008; Norman et al., 2013), thereby often obscuring and potentially delaying diagnosis. Of note, a delayed diagnosis is problematic as symptoms have been reported to worsen over time, especially if the ACEi is not discontinued (Brown et al., 1997; Mahmoudpour et al., 2015). The clinical data on ARB-AE are not as comprehensive as those on ACEi-AE and are only based on single case reports (Chiu et al., 2001; Lo, 2002; Nykamp and Winter, 2007). However, on the basis of these data, it can be assumed that the clinical picture is quite similar to that seen for ACEi-AE.

The main cause underlying ACEi/ARB-AE is thought to be an accumulation of BK, as increased BK levels have been found in patients taking either of the two drug classes (Campbell et al., 2005; Nussberger et al., 2002; Nussberger et al., 1998). ACEi and ARBs both target the renin-angiotensin system, which is closely connected to the kinin-kallikrein system and, to some degree, counterbalances its effects (Bas et al., 2007). ACE is the enzyme that links both systems. While in the renin-angiotensin system it is the key enzyme responsible for the conversion of angiotensin I to the biologically active angiotensin II, it is also the major enzyme for the degradation of BK generated within the kinin-kallikrein system (Bas et al., 2015). Consequently, ACE inhibition in the context of

antihypertensive therapy with an ACEi leads to an increase in plasma levels of BK. In contrast, ARBs target the signaling of angiotensin II by blocking angiotensin-II-receptor type 1 activation. As such, they do not directly interfere with BK metabolism, and the exact mechanism underlying the observed BK accumulation during ARB intake has not yet been fully understood. However, it was suggested that, for example, activation of angiotensin-II-receptor type 2 may play a role (Bas, 2017).

Considering the variable time-to-onset and the fact that not all patients receiving an ACEi/ARB develop an angioedema, an elevation in BK plasma levels cannot be the sole cause for the development of an ACEi/ARB-AE. In fact, research suggests that individual susceptibility to ACEi/ARB-AE is dependent on additional risk factors. One of the major risk factors for the development of an ACEi-AE seems to be the patient's ancestry. Several epidemiological studies have reported individuals of African ancestry to be at a 3-6 times higher risk for ACEi-AE compared to individuals of European ancestry (Brown et al., 1996; Kostis et al., 2005; Miller et al., 2008; Banerji et al., 2017; Reichman et al., 2017), thereby suggesting the involvement of a genetic predisposition. Other factors epidemiological studies have associated with a higher risk for ACEi-AE include female sex (Miller et al., 2008; Rasmussen et al., 2017; Kostis et al., 2018), advanced age (Kostis et al., 2005; Stauber et al., 2014; Mahmoudpour et al., 2016), a history of drug rash and seasonal allergies (Kostis et al., 2005; Mahmoudpour et al., 2016) or a concomitant coronary artery disease (Miller et al., 2008; Banerji et al., 2017). In contrast, it was found that patients with diabetes were less likely to be affected (Kostis et al., 2005; Byrd et al., 2008; Miller et al., 2008).

2.2.3 Hereditary angioedema

HAE is a group of rare, heterogeneous genetic disorders with a complex pathophysiology (Germenis and Speletas, 2016). Patients suffering from HAE present with chronically recurrent episodes of angioedema with often great variability in clinical expression and severity (Caballero et al., 2011). To date, two types of HAE can be biochemically distinguished: HAE patients presenting with a deficiency/dysfunction in the C1-INH (C1-INH-HAE) and HAE patients presenting with normal antigenic and functional C1-INH levels (nC1-INH-HAE) (Maurer et al., 2022). An overview of all HAE (sub)types known to date and their estimated frequency in relation to one another is provided in Figure 3.

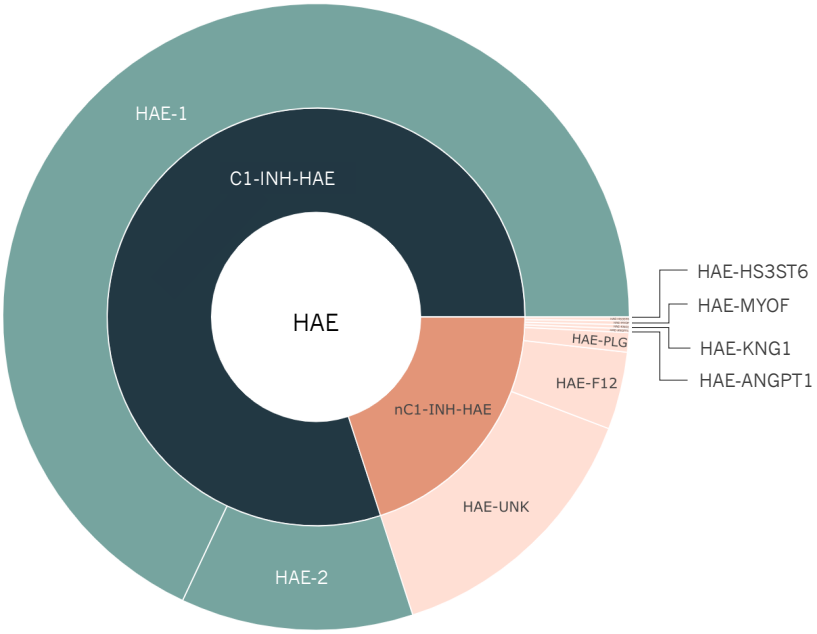

Figure 3 | HAE (sub)types and their relative frequencies known to date.

About 80% of HAE patients are C1-INH-HAE cases (green). Of these, 85% are due to C1-INH deficiency (HAE-1), while the remaining 15% have reduced C1-INH function (HAE-2) (Ponard et al., 2020). Patients with normal C1-INH levels and function (nC1-INH-HAE; orange) account for approximately 20% of all HAE cases (Marcelino-Rodriguez et al., 2019). Pathogenic variants in the genes *F12* and *PLG* are the underlying cause of HAE-FXII and HAE-PLG. Both have been identified in several affected families (> 400 and > 100 patients, respectively) and are collectively found in about one third of all nC1-INH-HAE cases (Lepelley et al., 2020; Maurer and Magerl, 2021). Pathogenic variants in the genes *ANGPT1*, *KNG1*, *MYOF* and *HS3ST6* are the underlying cause of HAE-ANGPT1, HAE-KNG1, HAE-MYOF and HAE-HS3ST6. To date, they are each associated with only single familial findings (Germenis et al., 2021) and likely explain less than 5% of the remaining nC1-INH-HAE cases (Marcelino-Rodriguez et al., 2019). Therefore, for the majority of nC1-INH-HAE cases the underlying genetic cause remains yet to be identified (HAE-UNK).

C1-INH-HAE is the most common form of HAE, affects around 1 in 50,000 individuals worldwide (Maurer 2022) and is caused by several different pathogenic variants in *SERPING1* (see Section 2.4.2; (Ponard et al., 2020)). Depending on the functional consequences of the underlying pathogenic variant, two different C1-INH-HAE subtypes are distinguishable. HAE-1 is characterized by low C1-INH concentration and function, whereas HAE-2 shows reduced C1-INH function but normal or even increased antigenic C1-INH levels (Caballero et al., 2011).

An even rarer disorder in the spectrum of inherited BK-AE is nC1-INH-HAE. The clinical picture of nC1-INH-HAE very much resembles that of C1-INH-HAE (Maurer and Magerl, 2021), although factors such as gender, age or hormones were found to influence the phenotype (Bork et al., 2020). In the absence of biomarkers for HAE beyond C1-INH,

genetic testing is so far the only reliable method to diagnose nC1-INH-HAE and distinguish between its subtypes (Germenis et al., 2020). At the time of writing, variants in six different genes had been associated with nC1-INH-HAE (see Section 2.4.2), and depending on the affected gene, the following subtypes were distinguished: HAE-FXII, HAE-PLG, HAE-KNG1, HAE-ANGPT1, HAE-MYOF and HAE-HS3ST6. Besides these subtypes, there is a substantial proportion of nC1-INH-HAE cases for which the underlying genetic cause has not yet been identified (HAE-UNK; (Veronez et al., 2021)). Pathophysiologically, HAE disease-causing variants interfere with BK formation or signaling at the endothelial cell (see Figure 2). Briefly, in HAE-1/2 the absence or dysfunction of C1-INH, caused by variants in *SERPING1*, leads to inadequate control of plasma kallikrein and FXII, and consequently to uncontrolled BK formation (Cicardi et al., 2005). Variants causal for HAE-F12, HAE-PLG, and HAE-KNG1 lead to mutant FXII, plasminogen and HK proteins, ultimately resulting in excessive BK formation (Cichon et al., 2006; Bork et al., 2018, 2019). The pathogenic *HS3ST6* variant is thought to lead to a disruption of the heparan sulfate biosynthesis, resulting in reduced endocytosis of HK and thus excessive amounts of HK becoming available for BK formation (Bork et al., 2021). In contrast, HAE-ANGPT1 and HAE-MYOF, are caused by altered BK signaling at the vascular endothelium. The *ANGPT1* variant disrupts the endothelium stabilizing effects of the angiopoietin-1/TIE-2 axis, leading to uncontrolled BK signaling (Cordisco et al., 2019). The variant identified in *MYOF* is suspected to cause an excessive increase in intracellular vascular endothelial growth factor signaling due to impaired ubiquitination and degradation of VEGF2-receptors (Ariano et al., 2020).

2.3 Genetic basis and identification of disease-associated variants

2.3.1 Mutational spectrum of bradykinin-induced angioedema

ACEi-AE is characterized by a multifactorial pathogenesis (Liau et al., 2019) and the same is likely to be true for ARB-AE. As such, genetic factors in various combinations with non-genetic personal and/or environmental factors, are likely to be involved in the development of ACEi/ARB-AE (Ludwig et al., 2019). The genetic architecture of ACEi/ARB-AE is assumed to encompass the entire allelic spectrum from common variants with low to moderate penetrance (Figure 4, lower right corner) to rare, highly penetrant variants (Figure 4, upper left corner) (Ludwig et al., 2019).

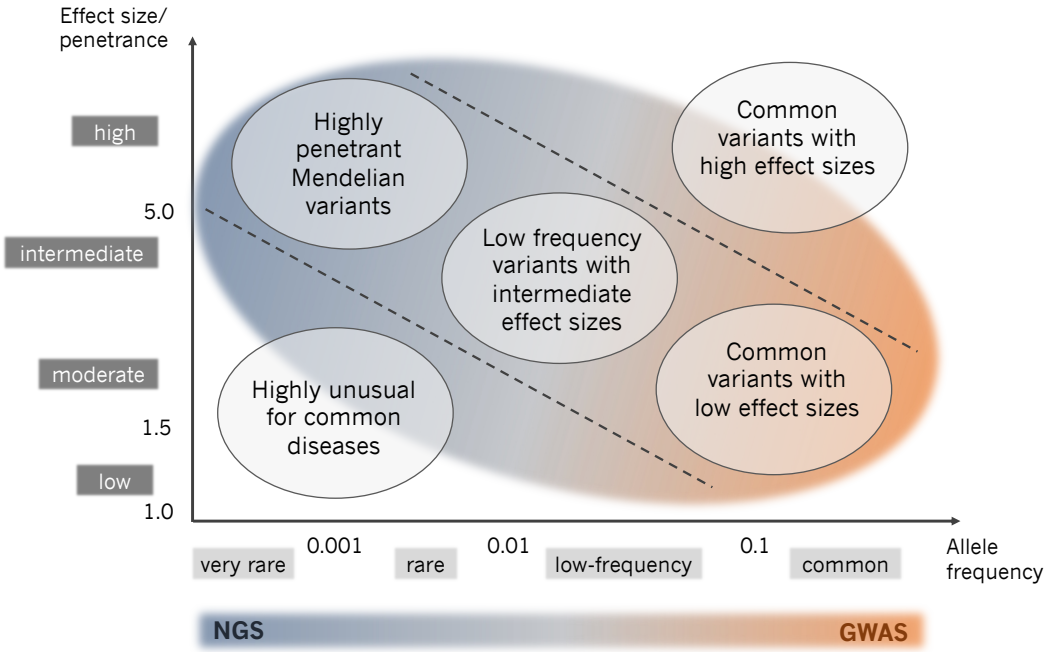

Figure 4 | The spectrum of disease-associated variants.

A common concept is to distinguish disease-associated variants based on their allele frequency and penetrance/effect size. The majority of these variants is likely to be found within the diagonal dashed lines and can be identified using NGS-based methods (rare, high penetrant variants, underlying rare Mendelian disorders) or GWAS (common variants with low to moderate effect size). Figure adapted from Manolio et al. (2009). Abbreviations: GWAS, genome-wide association study | NGS, next generation sequencing.

This assumption is substantiated by the findings of the most recent genome-wide association studies (GWAS), which identified the first genetic risk loci for ACEi/ARB-AE harboring variants with moderate to low penetrance (Rasmussen et al., 2020; Ghouse et al., 2021). In addition, the results of a recent exome sequencing study suggested the contribution of rare deleterious variants in this ADR (Maroteau et al., 2020).

In contrast, HAE is a rare genetic disorder, with an autosomal dominant Mendelian inheritance. The subtypes of HAE characterized to date (Figure 3) are caused by single rare variants (Figure 4, upper left corner) identified in several different genes. However, common variants have been shown to alter the phenotypic expression of C1-INH-HAE (Speletas et al., 2015; Gianni et al., 2017), while factors such as age or hormones have been found to influence certain nC1-INH-HAE subtypes (Bork et al., 2020).

The identification of risk variants along the entire allelic spectrum now requires different approaches and methods. While GWAS have proven to be a highly effective tool in the

identification of common variants with low to moderate effect sizes, the discovery of rare, high penetrant variants usually requires the application of next generation sequencing (NGS) methods (see Figure 4).

2.3.1.1 Common variants and GWAS

Over the past decade, GWAS have revolutionized the identification of common genetic risk loci and variants, enabling a deeper understanding of the underlying biology of complex diseases and traits as well as translation towards new diagnostics and therapeutics (Visscher et al., 2017). In the field of pharmacogenomics, GWAS contributed to the identification of loci that influence drug response or susceptibility to an ADR (Daly, 2010).

In a GWAS, the allele frequencies of a set of genome-wide distributed markers, so-called single nucleotide polymorphisms (SNPs), are systematically examined to detect markers associated with a specific phenotype. Given the genome-wide scan, GWAS have the advantage of not needing any prior assumptions regarding the potential pathomechanisms underlying the phenotype under investigation. GWAS are usually carried out on the basis of whole-genome genotyping data. Following the results and data derived from the 1000 Genomes project (Auton et al., 2015) and the International HapMap Project (The International HapMap Consortium, 2005), whole-genome genotyping exploits the haplotype block structure of the human genome. A haplotype block is a genomic region characterized by a reduced occurrence of recombination events within its boundaries (Daly et al., 2001). This means that SNPs within the same haplotype block are usually inherited together, resulting in the so-called linkage disequilibrium (LD). LD is defined as the non-random association of alleles at neighboring loci within a population, meaning that SNPs with high LD harbor redundant information. As a result, the genotyping of relatively few so-called *tag*SNPs is sufficient to capture most of the genetic variation present within a certain population (The International HapMap Consortium, 2005). For example, around 550,000 *tag*SNPs are sufficient to capture ~ 90% of the genetic variation in the European population (Steemers and Gunderson, 2007). The non-random selection of informative *tag*SNPs together with available information on haplotype structure and LD patterns from reference data sets eventually enables the downstream imputation of SNPs not captured

by the array (Yun et al., 2009). Due to the ever decreasing sequencing costs, GWAS are nowadays increasingly performed on the basis of exome- or genome-wide sequence data, which offer, for example, the advantage of investigating rare variants in the context of a GWAS (Uffelmann et al., 2021).

GWAS can be employed to examine both quantitative (e.g., height or body mass index) and binary (categorical) phenotypes (e.g., type 2 diabetes or ADRs). In the case of a binary phenotype, such as the ACEi-AE phenotype studied in this dissertation, GWAS are conducted using large case-control samples in which the allele frequency of each genetic variant is compared between unrelated cases and controls. Alleles with significantly different frequencies in cases and controls are considered to tag regions ("risk loci") that are associated or involved in the etiology of the phenotype. Since allele frequencies are known to be highly dependent on the ethnicity of an individual, the use of an ethnically matched cohort is essential to avoid population stratification (systematic differences in allele frequencies between sub-populations due to non-random mating) (Freedman et al., 2004).

In the scope of a GWAS, single marker association tests are performed for each individual genetic variant. For each of these tests, a *p*-value is returned that indicates the probability of a true association between the tested variant and the phenotype. However, given the large number of statistical tests performed there is a high probability that a significant result will be found only by chance. Therefore, a correction for multiple testing is required. In a GWAS of European individuals, approximately one million independent tests are performed, and thus the threshold for "genome-wide significance" is commonly defined as $5x10^{-8}$ (The International HapMap Consortium, 2005; Pe'er et al., 2008).

To date, the NHGRI-EBI GWAS Catalog, a curated compilation of published GWAS in humans, contains roughly 400,000 SNP-trait associations derived from more than 45,000 individual GWAS (Sollis et al., 2023). The vast majority of these associations are found in non-coding regions (Maurano et al., 2012; Visscher et al., 2017). In addition, the identified loci usually include multiple correlated variants, often spanning several genes, rendering identification of the actual causal SNP or gene within a GWAS risk locus quite challenging. Here, a commonly used approach is the integration of correlated GWAS SNPs with other available information such as regional LD patterns and functional genomic data sets generated by large-scale "*-omics*" sequencing projects such as

ENCODE (Encyclopedia of DNA elements; (Dunham et al., 2012)), Epigenome RoadMap (Roadmap Epigenomics Consortium et al., 2015), or GTEx (Genotype-Tissue Expression project; (Lonsdale et al., 2013)). The integration of additional data layers can be helpful to identify disease-relevant tissues or effects on gene expression, thereby ultimately facilitating the biological interpretation of GWAS results (Visscher et al., 2017; Cano-Gamez and Trynka, 2020). Moreover, GWAS data can be used for secondary analyses such as the estimation of the SNP-based heritability (the proportion of phenotypic variance explained by the additive effects of common variants; (Yang et al., 2017)) of a given trait/disease or the investigation of shared genetic factors between different traits and diseases (Tam et al., 2019).

Owing to the polygenic nature of complex traits and diseases, variants detected by GWAS usually have a small to moderate effect size. Thus, to achieve sufficient statistical power to detect even the most subtle effects, a very large number of individuals is required (Visscher et al., 2017). Combining multiple GWAS data sets in a meta-analysis has been proven to be an effective way to overcome this issue.

Since the individual variants identified in a GWAS usually only contribute to a very small extent to the overall disease risk of the phenotype under study, the clinical implication of GWAS findings is limited. The calculation of a so-called polygenic risk score (PRS), on the other hand, reflecting the sum of risk alleles carried by an individual, allows the definition of individual genetic risk profiles. In the future, such genetic profiles may, for example, play a role in predicting an individual's disease risk (Khera et al., 2018).

2.3.1.2 Rare variants and NGS

Prior to the advent of NGS methods, linkage analyses were used to map genomic loci that co-segregated in pedigrees of families with Mendelian disorders. Subsequent Sanger sequencing of genes within the mapped loci was a quite effective, albeit tedious, method for identifying rare disease-causing variants that underlie Mendelian disorders (Teare and Koref, 2014). The completion of the Human Genome Project (Lander et al., 2001; Venter et al., 2001) and the subsequent seminal publication of the first fully annotated human genome in 2003 (The International HapMap Consortium, 2003) paved the way for the development of the more advanced NGS technologies. NGS is thereby an umbrella term for a variety of sequencing methods that can not only be used to sequence

DNA (whole exome sequencing (WES), whole genome sequencing (WGS)), but also to investigate regulatory mechanisms of the genome (e.g., chromatin immunoprecipitation sequencing (ChIP-seq)) or the transcriptome (ribonucleic acid (RNA)-seq, single-cell RNA-seq) (Goodwin et al., 2016). Today, these high-throughput sequencing technologies are the methods of choice for the systematic identification of the full spectrum of rare disease-associated variants in research and clinical practice (Claussnitzer et al., 2020).

In NGS, millions of short read sequences are generated in a massively parallel and multiplexed manner, allowing the sequencing of entire genomes. Furthermore, the application of NGS-based targeted enrichment methods, such as single-molecule Molecular Inversion Probes (smMIPs) (see Section 3.5.2; (Hiatt et al., 2013)), enables cost-effective (re-)sequencing of single candidate genes or gene panels in large cohorts (Kanzi et al., 2020).

Given their high-throughput nature, NGS methods generate large amounts of sequencing data that require bioinformatics tools and pipelines to convert raw sequence data into annotated variants. Several genetic resources and databases are available for the annotation of obtained sequence variants. These include databases such as gnomAD (genome aggregation database; (Karczewski et al., 2020)) and dbSNP (single nucleotide polymorphism database; (Sherry et al., 2001)) which provide information on variant frequencies in the general population, whereas ClinVar (Landrum et al., 2020) and HGMD (Human Gene Mutation Database; (Stenson et al., 2020)) are disease-specific databases collecting information on sequence variants in relation to human disease. To assess the functional consequences of the identified variants, different *in silico* prediction tools are available. Here, a commonly used one is the score CADD (combined annotation-dependent depletion; (Rentzsch et al., 2021)), which integrates several different annotations to assess variant deleteriousness. To enable a comprehensive and consistent interpretation of variants, especially in the context of clinical genetic testing but also in research, the American College of Medical Genetics (ACMG) has published a guideline with standards for the interpretation of sequence variants. This guideline suggests several lines of evidence that should be considered when classifying sequence variants (Richards et al., 2015).

2.4 Genetic variation in bradykinin-induced angioedema

2.4.1 Findings in ACEi/ARB-induced angioedema

Over the last two decades, several genetic studies have sought to shed light on the underlying pharmacogenetics of ACEi/ARB-AE. At the time of writing, these included several candidate gene/SNP studies, three GWAS, and one exome sequencing study.

Early candidate SNP studies focused primarily on the role of polymorphisms in genes involved in BK degradation as well as known polymorphisms of the B2-R. The first gene linked to ACEi-AE risk was *XPNPEP2*, which encodes APP, the main BK-degrading enzyme during ACE inhibition (Blais et al., 2000). A regulatory variant (c.-2399C>A, rs3788853) in the upstream region of *XPNPEP2* was found to be significantly associated with lower APP activity and ACEi-AE (Duan et al., 2005) and later confirmed by two independent studies (Woodard-Grice et al., 2010; Cilia La Corte et al., 2011). Furthermore, one of the studies identified two additional variants upstream of *XPNPEP2* (c.-1612G>T, rs205011; c.-393G>A, rs2235444) that form a functional haplotype with rs3788853 which explains more variance in plasma APP levels and was associated with a greater increased risk for ACEi-AE (4.9-fold vs 3.3-fold) than rs3788853 alone (Cilia La Corte et al., 2011). An insertion/deletion (Indel) variant in the gene ACE (ACE I/D, rs4646994), known to alter the activity of the ACE (Rigat et al., 1990) showed no association with ACEi-AE in three independent studies (Gulec et al., 2008; Bas et al., 2010; Moholisa et al., 2013). A 9 base pair (bp) tandem repeat in exon 1 of the *BDKRB2* gene, previously reported to affect BK-induced vasodilation during ACE inhibition, showed inconsistent results in two studies (Bas et al., 2010; Moholisa et al., 2013). These two studies further examined a regulatory SNP in the promoter region of *BDKRB2* (c.-58C>T, rs1799722), previously associated with ACEi-induced cough (Mukae et al., 2002), but did not identify a significant association between rs1799722 and ACEi-AE. The most comprehensive candidate study so far examined 33 candidate SNPs in 17 different genes with *a priori* evidence and identified a SNP in *MME* (rs989692) that showed a significant and replicated association (both $P < 0.05$) for an increased risk of ACEi-AE (Pare et al., 2013).

Pare and colleagues (2013) were also the first to investigate the contribution of common genetic variation to ACEi-AE within the scope of a GWAS. In their study they examined

175 European and African-American ACEi-AE patients and 489 ethnically matched, ACEi-exposed controls yet they could not identify any SNPs at the level of genome-wide significance. An independent replication of SNPs identified at $P < 10^{-4}$ revealed, however, two nominally significant SNPs ($P < 0.05$) in genes involved in immune regulation. Namely, these were rs500766 in *PRKCQ* and rs2724635 in *ETV6* which were associated with decreased and increased risk of ACEi-AE, respectively. Since then, two additional GWAS have been conducted, each identifying one genome-wide significant risk locus. The second GWAS included 173 ACEi/ARB-AE patients and 4,890 population-based controls, both of Swedish ancestry, and revealed a genome-wide significant association between intronic variants in the gene *KCNMA1* and ACEi/ARB-AE risk (Rasmussen et al., 2020). The third GWAS examined 462 Danish ACEi-AE patients and 53,391 Danish ACEi-treated controls and identified a genome-wide significant association between ACEi-AE risk and variants upstream of the gene *BDKRB2* (Ghouse et al., 2021). Furthermore, based on the data of this latest study, the SNP-based heritability for ACEi-AE was estimated to be 21.7%.

The contribution of rare variants to ACEi/ARB-AE was recently assessed in an exome-wide sequencing study. Based on WES data from 408 ACEi/ARB-AE patients and 658 ACEi-tolerant controls, the study identified common and rare variants in the gene *F5* to be significantly associated with ACEi/ARB-AE. Notably, the strongest association was found for rs6025, also known as “Factor V Leiden” mutation (Maroteau et al., 2020).

2.4.2 HAE-associated genes and pathogenic variants

Pathogenic variants in *SERPING1* are the underlying genetic cause of C1-INH-HAE (Stoppa-Lyonnet et al., 1987). To date, around 750 causative variants have been identified, underlining the great allelic heterogeneity of this HAE subtype (Ponard et al., 2020). The HAE-associated *SERPING1* variants are mainly comprised of small Indels (36.2%), missense variants (32.1%), and variants leading to splicing defects (13.9%), followed by nonsense variants (9%), large deletions and duplications (8.2%), and variants affecting the 5' or 3' untranslated regions (0.9%) (Ponard et al., 2020). Beyond that, two deep intronic variants were recently discovered (Hujová et al., 2020; Vatsiou et al., 2020). Moreover, the disease expression of C1-INH-HAE has been found to be influenced by common variants in BK-related genes. In particular, a functional polymorphism in the

promoter region of the *F12* gene (-46C/T, rs1801020) and a functional variant in the *KLKB1* gene (-428G/A, rs3733402) were significantly associated with age of disease onset either alone or in combination (Speletas et al., 2015; Gianni et al., 2017).

Table 2 | Genes and pathogenic variants associated with nC1-INH-HAE.

HAE type	Gene	Pathogenic variant	Amino acid change	Initial description
HAE-FXII	*F12*	c.1032C>A	p.T328K	Cichon et al., 2006; Dewald and Bork, 2006
		c.1032C>G	p.T328R	
		c.971_1018+24del72	Indel	Bork et al., 2014
		c.892_909dup	p.P298_P303dup	Kiss et al., 2013
HAE-PLG	*PLG*	c.988A>G	p.K330E	Bork et al., 2018
HAE-ANGPT1	*ANGPT1*	c.807G>T	p.A119S	Bafunno et al., 2018
HAE-KNG1	*KNG1*	c.1136T>A	p.M379K	Bork et al., 2019
HAE-MYOF	*MYOF*	c.651G>T	p.R217S	Ariano et al., 2020
HAE-HS3ST6	*HS3ST6*	c.430A>T	p.T114S	Bork et al., 2021

The group of nC1-INH-HAE is much more heterogeneous in terms of its underlying genetic causes. So far, pathogenic variants in six different genes have been associated with the phenotype (Table 2). Four different variants in the *F12* gene are linked to HAE-FXII, including two missense variants, one Indel and one duplication, all located within exon 9 or at its boundaries (Stieber et al., 2017). Moreover, recent WES studies led to the identification of five missense variants in the genes *PLG*, *ANGPT1*, *KNG1*, *MYOF*, and *HS3ST6* each found to be associated with a different nC1-INH-HAE subtype (Veronez et al., 2021).

Notably, several case reports indicated that some patients initially diagnosed with ACEi-AE might actually be undiagnosed HAE cases as genetic screening subsequently revealed them to be carriers of pathogenic HAE-associated variants in the genes *F12* (Veronez et al., 2017) and *PLG* (Germenis et al., 2018; Yakushiji et al., 2018).

2.5 Aim of the dissertation

The overall aim of the present dissertation was to further elucidate the genetic basis of ACEi/ARB-AE. As a resource for the genetic studies, a large, comprehensively phenotyped cohort of ACEi/ARB-AE patients (vARIANCE cohort) was recruited.

The first study intended to systematically investigate the presence of monogenic HAE subtypes in a cohort of ACEi/ARB-AE patients. In addition, the potential contribution of other variants in the HAE-associated genes to ACEi/ARB-AE susceptibility was examined. To this end, targeted re-sequencing of the five – at that time known – HAE-associated genes was performed in a large ACEi/ARB-AE cohort. Subsequently, the ACEi/ARB-AE patients were systematically screened for known HAE disease-causing variants, while in addition, single variant association analyses, and enrichment analyses of rare, potentially functional variants were carried out in a case-control setting.

The second study aimed to identify common variants involved in ACEi-AE susceptibility. Here, GWAS data from the vARIANCE cohort were combined with seven other independent European case-control data sets in the most comprehensive GWAS meta-analysis to date. Subsequently, bioinformatics follow-up analyses were performed based on the meta-analysis data. First, the consistency of the ACEi-AE phenotype definition across cohorts was examined in PRS analyses. Next, to obtain additional insights into the molecular mechanisms underlying ACEi-AE, statistical fine-mapping, integration of expression quantitative trait loci (eQTL) and chromatin interaction data, and gene- and pathway-based analyses were carried out. LD score regression (Bulik-Sullivan et al., 2015) was performed to (i) estimate the SNP-based heritability, and (ii) examine the genetic overlap between ACEi-AE and a selection of associated traits and diseases. Finally, to investigate the extent of shared genetics and, in particular, the contribution of the associated loci to ACEi-AE across different ancestries, exploratory cross-ancestry analyses, including an additional GWAS cohort of African-American individuals were conducted.

3 Material and methods

All equipment, chemicals, buffers, solutions, reagents, enzymes, commercial kits, consumables, primer, software and databases used throughout this dissertation are listed in APPENDIX A1.

The studies included in the present work, investigated different patient and control cohorts, which are described in the following. With the exception of the CHB-CVDC/DBDS individuals, who provided scientific ethical approval, meaning they were informed about their samples being used for research purposes and have been given the possibility to opt out, all investigated individuals provided written informed consent prior to their inclusion. Studies described in this work were approved by the respective institutional ethic committees.

3.1 Patients and controls with available DNA samples

3.1.1 vARIANCE patients

The vARIANCE study (Angioedema Risk under Angiotensin Converting Enzyme Inhibitors, www.variance-studie.info) is a case collection of ACEi/ARB-AE patients that was initiated in 2018 as a collaborative project between the Institute of Human Genetics of the University Hospital Bonn, Germany and the German Federal Institute for Drugs and Medical Devices (BfArM). The ongoing patient recruitment for the study was developed and implemented as part of this dissertation, and currently involves two strategies:

(i) Recruitment *via* collaborating physicians

Since 2018, collaborations with dermatology, ear-nose-throat, or angioedema-specialized departments at different hospitals in Germany and Austria have been established (Figure 5). In addition, a small fraction of patients was enrolled by office-based physicians. The general recruitment process is as follows: patients with a suspected diagnosis of ACEi/ARB-AE are clinically assessed by experienced physicians on site. If found to be eligible and if they consent to participate in the study, patients provide a DNA sample (EDTA blood or saliva) and complete a questionnaire that asks questions about the occurred angioedema, the suspected causal ACEi/ARB, possible other triggers, medical and family history as well as basic anamnestic data of the patient.

(ii) Recruitment *via* the study website/office in Bonn

To inform additional patients about the study and the possibility to participate, the vARIANCE study website (www.variance-studie.info) was launched in May 2020. Patients, with a corresponding angioedema diagnosis and an interest in participating in the study can now get in direct contact with the study office in Bonn. Trained staff in the study office evaluates the inquiries and enrolls eligible patients into the study. For this purpose, study documents and DNA collection material (saliva kits) are provided per mail, while the study questionnaire as well as any other potential questions from the patients are addressed in a telephone interview.

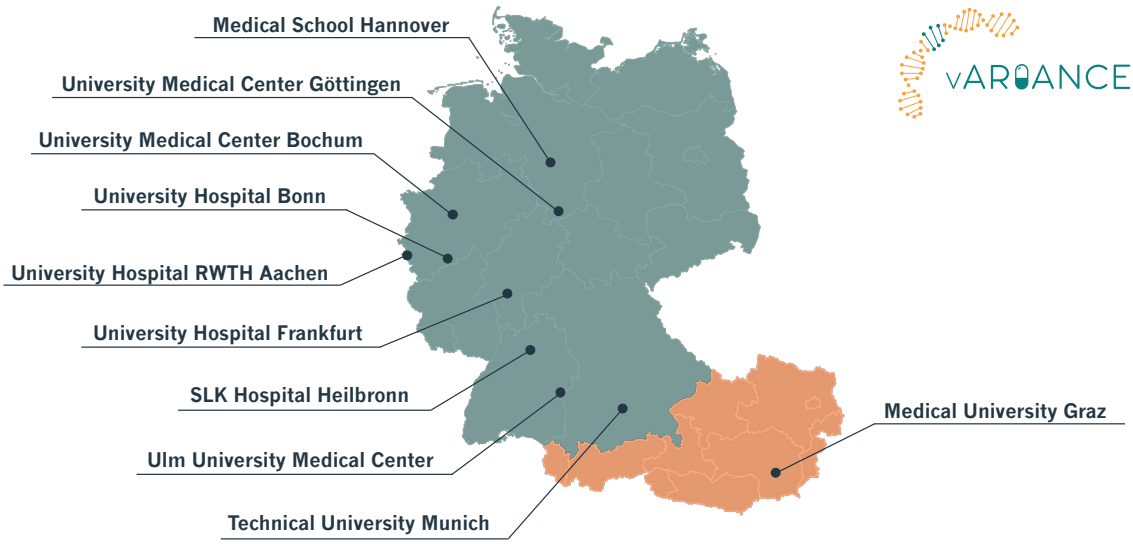

Figure 5 | German and Austrian hospitals involved in patient recruitment for the vARIANCE study. *

As of May 2023, the vARIANCE study had enrolled 153 ACEi/ARB-AE patients, of which 133 were recruited *via* collaborating physicians and 20 were recruited *via* the study website/office (Table 3). For the two studies discussed in this dissertation, sub-cohorts were drawn from the vARIANCE cohort at two different time points. The candidate gene analysis included in total 67 ACEi/ARB-AE patients who were subjected to targeted sequencing and genotyping, while the vARIANCE GWAS cohort comprised genome-wide

*The German/Austrian map was created using the R packages *raster* and *ggplot2*.

genotype data of 106 ACEi-AE patients. Inclusion criterion for both sub-cohorts was a "certain", "probable/likely", or "possible" causal relationship between the experienced angioedema and the intake of an ACEi or ARB as assigned by the enrolling physician. In case of a study inclusion via the study office in Bonn, causality was first self-assessed by the patient and later re-assessed by a clinical expert. A concomitant urticaria (self-reported by the patient) at the time of the angioedema was considered an exclusion criterion.

Table 3 | vARIANCE patients stratified by recruitment approach, gender and angioedema triggering drug class.

	Recruitment *via* physicians			Recruitment *via* the study website/office			Overall recruitment		
Drug class	**Female**	**Male**	**Total**	**Female**	**Male**	**Total**	**Female**	**Male**	**Total**
ACEi	49	59	**108**	5	6	**11**	54	65	**119**
ARB	8	10	**18**	3	3	**6**	11	13	**24**
ACEi/ARB	3	3	**6**	3	0	**3**	6	3	**9**
NA	0	1	**1**	0	0	**0**	0	1	**1**
Total	**60**	**73**	**133**	**11**	**9**	**20**	**71**	**82**	**153**

NA, triggering drug class (ACEi or ARB) not known.

3.1.2 Danish and Swedish patients

DNA material and phenotypic data of Danish and Swedish ACEi/ARB-AE patients were kindly provided by Eva Rye Rasmussen and Mia Wadelius (Table 4). Detailed information on the recruitment and phenotype definitions can be found elsewhere (Rasmussen et al. 2020).

Table 4 | Danish and Swedish patients stratified by gender and angioedema triggering drug class.

	Danish patients			Swedish patients		
Drug class	**Female**	**Male**	**Total**	**Female**	**Male**	**Total**
ACEi	33	27	**60**	27	42	**69**
ARB	4	3	**7**	3	5	**8**
NA	0	1	**1**	**0**	**0**	**0**
Total	**37**	**31**	**68**	**30**	**47**	**77**

NA, triggering drug class (ACEi or ARB) not known.

Briefly, Danish patients were recruited from multiple Danish hospitals, as well as *via* a collaborating general practitioner. Swedish patients were drawn from Swedegene

(www.swedegene.se), a nationwide biobank that contains clinical data and blood samples of various ADR patients.

For both cohorts, the clinical diagnosis of an ACEi/ARB-AE was initially established by the enrolling physician and later on re-assessed and adjudicated by a clinical expert in allergology using pre-defined phenotype standardization criteria (Wadelius et al., 2014). According to these criteria, patients with a swelling in the head or neck region during ACEi/ARB treatment judged by a physician to be an angioedema were considered cases. Exclusion criteria comprised: (i) a concomitant urticaria, (ii) other suspected causes for the angioedema, and (iii) a history of HAE or acquired C1-INH deficiency.

Patients in both cohorts underwent targeted sequencing and genome-wide genotyping. All 68 Danish and 77 Swedish patients were part of the candidate gene study, while the Danish and Swedish GWAS cohorts included 51 and 44 ACEi-AE patients, respectively. The other 9 Danish and 25 Swedish ACEi-AE patients were already part of other GWAS cohorts (CHB-CVDC/DBDS or Swedegene) and thus (in light of the downstream meta-analysis) not included in the Danish or Swedish GWAS cohort, respectively.

3.1.3 Control individuals

For the candidate gene analysis control individuals (N = 352; “Buffy controls”) were drawn from an in-house cohort derived from volunteer blood donors recruited in Bonn, Germany (Birnbaum et al., 2009). As part of the study, all individuals were subjected to targeted sequencing, while genome-wide genotype data had already been generated previously. The selected Buffy controls were not screened for a potential diagnosis of an ACEi/ARB-AE or the intake of an ACEi/ARB.

For the vARIANCE GWAS, genome-wide genotype data of 4,249 participants of the Heinz Nixdorf Recall (HNR) study (Erbel et al., 2012), which had been generated previously, were used as control data. The HNR individuals were not screened for an ACEi-AE diagnosis or the intake of an ACEi.

3.2 Patients and controls with available individual-level genotypes

3.2.1 VanMar cohort

Genome-wide genotype data of 172 ACEi-AE patients and 485 controls from the “PGRN-RIKEN: Identification of Genetic Predictors of ACE Inhibitor-Associated Angioedema

study" were obtained from dbGaP (study accession: phs000438.v1.p1). Details on the recruitment as well as the phenotype definition are published elsewhere (Pare et al., 2013). Briefly, cases were defined as individuals who developed a swelling of the lips, pharynx, or face while taking an ACEi. Furthermore, cases ought never to have had an angioedema while not taking an ACEi. Treatment-matched individuals (intake of an ACEi for at least six months) who never developed an angioedema were considered controls. As the data set comprised cases and controls from different ancestries, individuals were stratified into two independent cohorts prior to the GWAS, resulting in (i) 107 cases and 330 controls with a reported European ancestry ($VanMar_{EUR}$) and (ii) 65 cases and 155 controls with a reported African-American ancestry ($VanMar_{AFR}$).

3.2.2 UK Biobank cohort

The UK Biobank (UKB) is a large-scale longitudinal biomedical database that collects comprehensive genetic and phenotypic information from approximately 500,000 UK citizens aged 40 to 69 years at recruitment (Bycroft et al., 2018).

For the UKB GWAS, individuals with a suspected ACEi-AE diagnosis as well as treatment-matched controls were extracted from the whole UKB data set. A comprehensive overview on the applied filter strategy and the data fields used is provided in Figure 6. Briefly, filtering of potential ACEi-AE patients was performed using hospital in-patient data on ICD-10 diagnoses (UKB field 41270) as well as self-reported data on (i) ancestry (UKB field 21000); and (ii) the use of antihypertensive medication/ACEi (UKB fields 6153/6177 and 20003). To appraise causality, all individuals with a suspected ACEi-AE diagnosis were further evaluated to determine whether the reported use of the ACEi preceded the angioedema diagnosis in time by matching the respective angioedema diagnosis date (UKB field 41280) and all assessment center visit dates (UKB field 53) for which an ACEi use was reported ("causality filter").

Filtering of the UKB data set resulted in a total of 90 individuals with the diagnosis of an angioedema (T78.3 "angioneurotic edema") and a chronologically related use of an ACEi. Subsequently, treatment-matched controls (N = 360) were selected on the basis of age and sex from the remaining individuals with a reported intake of an ACEi. All analyses were performed under the approved UKB project number 60928.

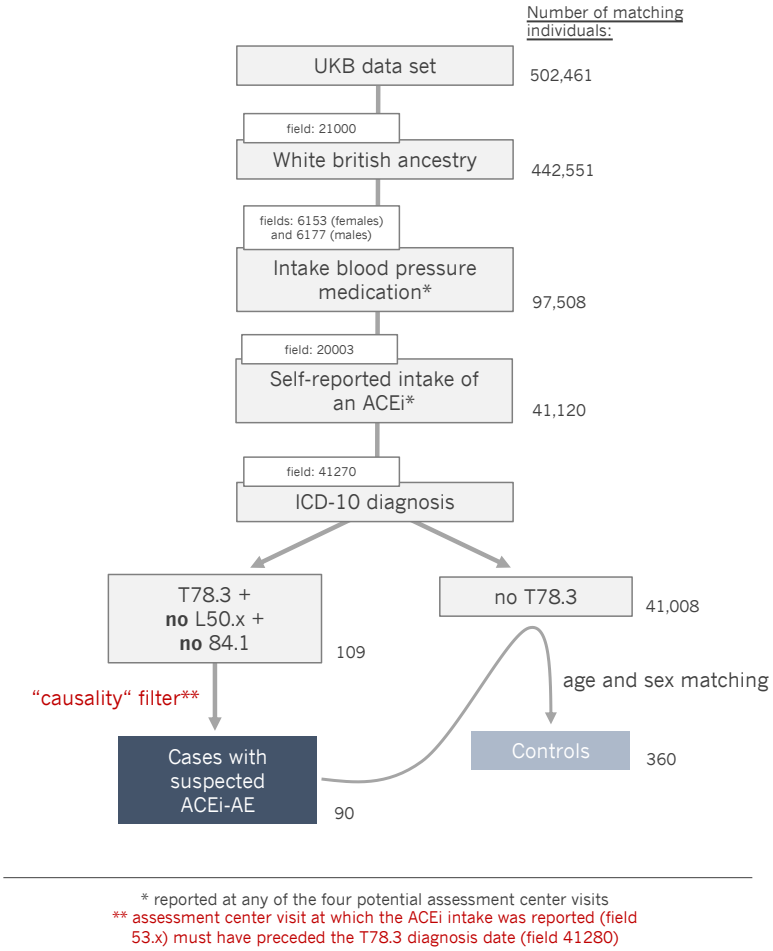

Figure 6 | Workflow to extract individuals with a suspected ACEi-AE diagnosis and treatment-matched controls from the UKB data set.

Individuals with a reported T78.3 ("angioneurotic angioedema") diagnosis and no diagnosis of any form of urticaria (L50.1-L50.9) or a defect in the complement system (D84.1) were considered as potential cases, while individuals without these case defining diagnoses were considered as potential controls. To estimate causality between the ACEi intake and the reported angioedema diagnosis, the potential cases were filtered for those whose reported ACEi intake preceded the angioedema diagnosis in time. Finally, the four-fold number of treatment-matched controls was matched on the basis of age and sex for all eligible patients.

3.2.3 Control individuals

The genome-wide genotype data of 1,628 healthy Danish blood donors was kindly provided by Vibeke Andersen and Signe Bek Sørensen and used as control data for the Danish GWAS. For the Swedish GWAS, genome-wide genotype data of 1,033 Swedish population-based controls (ANGI-SE(Community) sample), recruited in the context of the Anorexia Nervosa Genetics Initiative (Thornton et al., 2018)), were kindly provided by

Cynthia Bulik and Mikael Landén and used as a control data. None of these GWAS control cohorts was screened for an ACEi-AE diagnosis or the intake of an ACEi.

3.3 GWAS summary statistics

3.3.1 Swedegene

The Swedegene cohort represents an ACEi-AE stratified version of the second GWAS published on ACEi/ARB-AE (Rasmussen et al., 2020), whose summary statistics were kindly provided by Mia Wadelius and Eva Rye Rasmussen. The cohort comprised 142 ACEi-AE patients and 1,345 ACEi-treated controls, both of Swedish ancestry. More detailed information on the study cohort and phenotype definitions can be found in the original publication (Rasmussen et al., 2020). Briefly, ACEi-AE patients were drawn from the Swedegene database (www.swedegene.se) and subsequently screened and adjudicated using published phenotype standardization criteria (Wadelius et al., 2014). Control subjects had received at least two ACEi prescriptions and had not been diagnosed with angioedema or laryngeal angioedema.

3.3.2 CHB-CVDC/DBDS

The GWAS summary statistics of the third published GWAS on ACEi-AE, including 462 ACEi-AE cases and 53,391 treatment-matched controls, both of Danish ancestry, were kindly provided by Jonas Ghouse. Details on the study cohort and the phenotype definitions can be found in the original publication (Ghouse et al., 2021). Briefly, individuals of the cohort were drawn from the Copenhagen Hospital Biobank-Cardiovascular Disease Cohort/Danish Blood Donor Study (CHB-CVDC/DBDS; (Laursen et al., 2021)). Cases were defined as individuals who had been diagnosed with angioedema, while being treated with an ACEi within at least 180 days before the angioedema event. Individuals who had been treated with an ACEi for at least 2 years and had no history of angioedema were defined as controls.

3.3.3 Estonian Biobank

The Estonian Biobank (EstBB) is a population-based biobank that today covers about 20% of the adult population of Estonia (https://genomics.ut.ee/en/content/estonian-biobank). Besides different levels of genetic data, the biobank withholds information on health status, lifestyle and diet. In addition, it is linked to different national registries like

hospital databases, and the database of the national health insurance fund. Leveraging the data of the EstBB, Lili Milani, Kristi Krebs and colleagues conducted a GWAS in 82 ACEi-AE cases and 15,787 treatment-matched controls, whose summary statistics they kindly provided us with. For the GWAS, individuals diagnosed with an ICD-10 code T78.3 ("angioneurotic edema") after treatment with an ACEi (C09A*) were defined as cases, whereas individuals with at least three ACEi purchases and no case-defining ICD-10 diagnosis were considered controls.

3.4 Basic molecular biological methods

The following molecular biological methods were performed in the context of the present dissertation. A detailed overview of which patient/control cohort with available DNA samples (see Section 3.1) underwent which method can be found in Table S3 in APPENDIX A2.

3.4.1 Isolation and organisation of DNA samples

Lymphocyte DNA was extracted from EDTA anti-coagulated venous blood using the Chemagic DNA Blood 10k Kit H12 and the Chemagic Magnetic Separation Module I according to the manufacturer's instructions. In case of saliva samples, DNA was manually isolated from saliva lymphocytes and epithelial cells using the Oragene prepIT L2P kit. Subsequently, two working solutions (20ng/µl and 100ng/µl) were prepared from each DNA stock and managed using the FluidX system, which is based on a scannable 2D barcode linking the dilution to its specific DNA stock.

3.4.2 Determination of DNA quantity and quality

The FluidX working solutions were prepared based on DNA concentration measurements performed with a NanoDrop spectrophotometer. Here, 1.5 µl was applied for each DNA sample. The calculation of the DNA concentration was performed automatically by the NanoDrop software based on a modified version of the Beer-Lambert Equation and purity of the DNA samples was evaluated based on purity ratios. Samples with a 260/280 ratio of ~ 1.8 and a 260/230 ratio between 1.8 – 2.2 are generally considered as "pure" DNA. Deviations from these values are indicative of DNA contamination by, for example, proteins or salts. While DNA quantification based on UV absorption, as performed with the NanoDrop, is sufficient for applications such as polymerase chain reaction (PCR; see

Section 3.4.3) and genotyping (see Section 3.5.3), it is too insensitive for NGS methods which require uniform amounts of high-quality double-stranded DNA (dsDNA). DNA samples subjected to sequencing using smMIPs (see Section 3.5.2) were thus re-measured using the Quant-iT PicoGreen dsDNA reagent, which relies on a DNA-binding fluorescent dye that specifically binds to dsDNA. Depending on the number of measured samples the assay read out was performed with either a Qubit 2.0 fluorometer (single samples) or the Tecan GENios Pro Microplate Reader (96-well plates) using standard fluorescein wavelengths with an excitation of 485 nm and an emission of 535 nm. The actual dsDNA concentration of the sample was then calculated using a standard curve of lambda DNA dilutions with known concentrations.

3.4.3 Polymerase chain reaction

PCR is a widely used *in vitro* method for the specific amplification of DNA segments. The method is based on repetitive thermal cycles, that enable denaturation of the DNA double strand, primer annealing to the DNA single strands and enzyme-driven DNA replication (Mullis et al., 1986). Given that this reaction is repetitive, the amount of PCR product within the reaction mixture increases exponentially with each subsequent cycle. The specificity of the amplification is ensured by the use of two primers (commonly 18 – 25 bp), that are flanking the region of interest and have a complementary sequence to the template DNA strand. The primer used in the scope of this dissertation were designed using the publicly available online tool Primer3 (Untergasser et al., 2012). A standard PCR mix was generated using Taq DNA Polymerase and 10x Extra Buffer (see Table S4 in APPENDIX A2). Moreover, thermal cycling was performed using a touchdown PCR, to increase sensitivity, specificity and efficiency of the PCR (Korbie and Mattick, 2008).

3.4.4 Electrophoretic separation of DNA fragments

Gel electrophoresis is a well-established method for separation and fragment length determination of DNA molecules and thus commonly used to verify the efficiency and specificity of a PCR reaction. It exploits the negative charge of the DNA molecules, which migrate through a polysaccharide matrix towards the anode (positive pole) when an electric field is applied. The migration velocity depends on the polysaccharide concentration, the applied voltage and ultimately the size of the DNA fragments. Thus, the DNA fragments migrate anti-proportional to their size. In the present dissertation,

2% agarose matrices with 1% ethidium bromide were used. Ethidium bromide, a substance that intercalates into the DNA double helix and is excited by UV light, was added to later visualize the DNA bands in the intas GelStick imager. The fragment size of the amplified product was determined using a DNA ladder (100 bp or 1 kilobase (kb)) as a length standard.

3.5 DNA sequencing and genotyping methods

In the framework of the present dissertation, genetic data sets were acquired by genome-wide genotyping as well as different sequencing methods, which are described in more detail in the following. A detailed overview of which patient/control cohorts with available DNA samples (see Section 3.1) underwent which method can be found in Table S3 in APPEDNDIX A2.

3.5.1 Sanger sequencing

Dating back to the late 1970's, Sangers' chain-termination sequencing method is a still widely used technology for detecting the exact base sequence of DNA segments (Sanger et al. 1977). For Sanger sequencing performed in the context of the present dissertation, the DNA was first amplified using standard PCR (Table S4 in APPENDIX A2), and PCR products were purified using Agencourt AMPure XP magnetic beads. Following that, cycle sequencing of the purified PCR products was performed according to standard procedures using the 5X Big Dye Terminator Cycle Sequencing Kit 3.1, which contained deoxynucleotide triphosphates (dNTPs) and fluorescent-labeled dideoxynucleotide triphosphates (ddNTPs) (Table S5 in APPENDIX A2). During a standard PCR reaction, the individual nucleotides get linked to each other *via* their 3'-hydroxyl group. As labeled ddNTPs lack these 3'-hydroxyl groups, their incorporation leads to the termination of the PCR. Statistically, this occurs at least once at each position of the DNA fragment, resulting in a mixture of fragments of different lengths, each carrying a fluorescently labeled ddNTP at the end. The cycle sequencing products were then purified using the Agencourt CleanSEQ Kit and, after dilution (1.5 µl product + 10 µl HPLC water), loaded onto a SeqStudio Genetic Analyzer. In the analyzer, fragments were separated according to their length by capillary electrophoresis and the last incorporated nucleotide of each fragment was identified by a laser. Using the analyzer's built-in software, the detected signals were then converted into sequence information.

3.5.2 Next generation sequencing – single-molecule Molecular Inversion Probes

In the last 20 years, NGS methods have been adopted into research and clinical practice. These methods have the ability to sequence many millions of DNA fragments in a massively parallel fashion, thereby enabling capture of the entire exome or even genome (Goodwin et al., 2016). Beyond that, NGS-based enrichment methods, such as smMIPs, which were used in the scope of the candidate gene analysis in the present dissertation, provide an easily scalable and cost-effective tool for targeted (re-)sequencing studies (Hiatt et al., 2013).

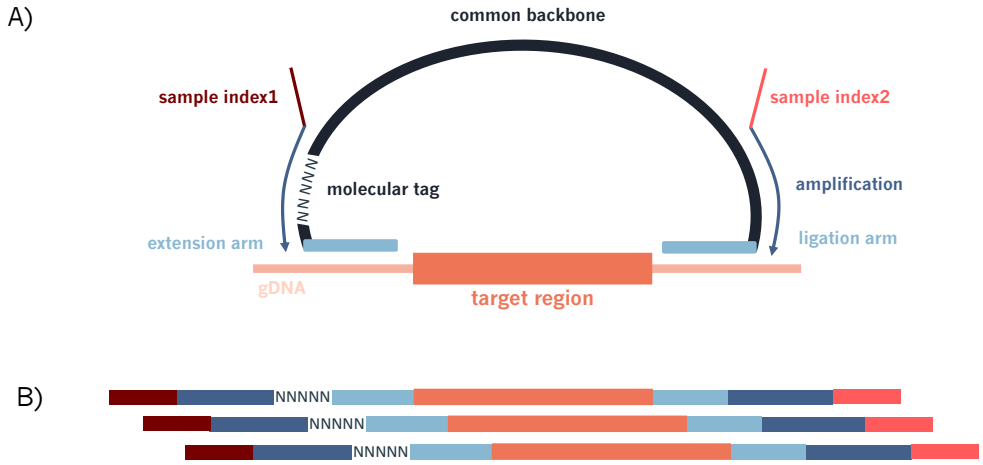

Figure 7 | Schematic overview of the single-molecule Molecular Inversion Probe method.

(A) An smMIP comprises two DNA complementary targeting arms (extension and ligation arm) that are connected by a common backbone. The backbone incorporates two PCR primer sites alongside a molecular tag that is unique to the individual probe and enables the downstream detection of PCR artefacts. (B) PCR amplified DNA sequences with two added sample-specific barcodes. Abbreviation: gDNA, genomic DNA.

The targeting arms of an smMIP are complementary to the DNA, flanking the region of interest, so that upon hybridization of the probes to the DNA specific parts (e.g., the exons of a gene) are targeted (Figure 7A). The obtained gaps are filled using a polymerase/ligase reaction, resulting in fragments that are amplified in a PCR and eventually sequenced on an NGS platform (Hiatt et al., 2013). Pooling hundreds to thousands of smMIPs and the use of up to two sample-specific indices during the amplification step eventually enables a multiplex reaction and massively parallel sequencing of large cohorts (Figure 7B).

3.5.3 Genome-wide genotyping using Illumina BeadArray technology

Another breakthrough technology in the field of genomics was whole-genome genotyping, which emerged around 2004 and paved the way for GWAS. A highly efficient method for the genotyping of SNPs is the Infinium BeadArray technology. The basis of this technology is the so-called BeadChip, which harbors randomly immobilized beads. Each bead is linked to a 50-mer locus-specific primer that captures one SNP and contains 30 bases which are used for positional mapping on the array (decoding) (Gunderson et al., 2006). Moreover, to increase the overall coverage per SNP, the beads on an array are redundant (Steemers and Gunderson, 2007).

The genotyping reaction itself involves three main steps (Figure 8). First, the amplified genomic DNA (gDNA) is hybridized to the bead-connected probes. Then, enzymatic primer extension is performed, either *via* allele-specific primer extension or single base extension. The use of two different primer extension assays is necessary to (i) ensure adequate coverage of all potential polymorphisms, including AT and GC polymorphisms, and (ii) increase the overall efficiency of the array. Finally, after extension of each target sequence, the labeled nucleotides are detected by immunohistochemical sandwich staining (Steemers and Gunderson, 2007).

For the present dissertation, genome-wide genotyping was performed using the Illumina Global Screening Array (GSA) v2.0 and v3.0, the Infinium HTS Assay and the BeadArray platform. The Infinium workflow was performed according to the manufacturer's 3-day protocol using 200ng DNA as input and comprised the following steps: (i) gDNA amplification (day 1); (ii) enzymatic fragmentation and precipitation of the gDNA, resuspension and hybridization to the BeadChip (day 2); and (iii) primer extension and immunohistochemical staining (day 3). Read out of the fluorescents signals was performed using the Illumina iScan system. Finally, Illumina's GenomeStudio v2 software and a manifest file of the respective array were used to convert the raw intensity files into SNP genotype calls. For further processing, the obtained SNP genotypes were exported in PLINK format (Chang et al., 2015).

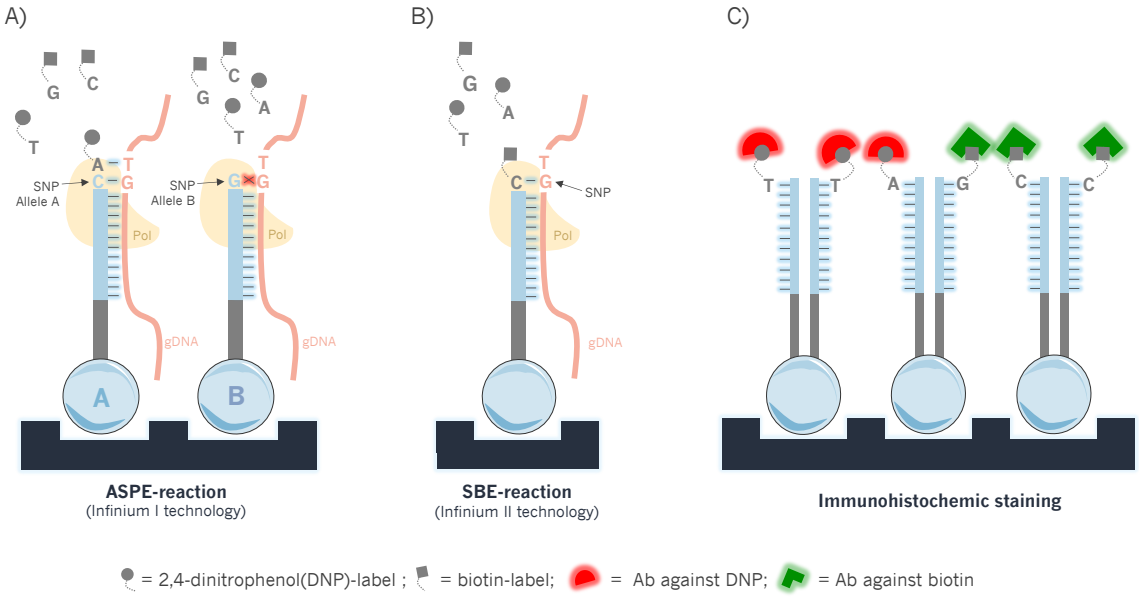

Figure 8 | Schematic overview of the Infinium BeadArray-technology.

(A) The ASPE reaction involves two allele-specific SNP probes, thereby enabling the detection of all SNP classes, including G/C and A/T SNPs, as labeled nucleotides are only incorporated in the case of a complementary base pairing ("perfect match"). (B) In the SBE reaction, the two alleles are detected with the same probe. Due to indifferent labeling of A and T as well as G and C nucleotides, SBE reactions cannot distinguish G/C and A/T SNPs. (C) Immunohistochemical sandwich assay for the detection of the incorporated nucleotides. The measured intensities eventually enable conclusions regarding the allelic distribution in each SNP. Abbreviations: Ab, antibody | ASPE, allele-specific primer extension | gDNA, genomic DNA | Pol, polymerase | SNP, single nucleotide polymorphism | SBE, single base extension.

3.6 ACEi/ARB-AE candidate gene analysis

3.6.1 Candidate gene selection and study cohort

For the candidate gene study, genes carrying previously described pathogenic HAE variants were identified using the professional version of the HGMD (query mid-2019; Stenson et al., 2020). Moreover, a literature search was performed to identify any additional genes with reported HAE-causing variants that at the time of the query might not (yet) have been listed in the HGMD. Eventually, five genes were selected: *SERPING1*, *F12*, *PLG*, *ANGPT1* and *KNG1*.

The study cohort comprised 212 ACEi/ARB-AE patients from Germany/Austria (Section 3.1.1), Denmark and Sweden (Section 3.1.2), and 352 German controls (Section 3.1.3).

3.6.2 Pre-sequencing QC: removal of relatives, ethnic, and population outliers

Kinship analysis was used to identify related individuals, while principal component analyses (PCA) were performed to remove ethnic and population outliers in order to avoid any potential biases due to population stratification in the downstream analyses. Both, kinship analysis and PCA, were performed on the basis of genome-wide genotype data as the in-study gene panel did not cover enough markers. As patients and controls of the candidate gene study were genotyped on different versions of the Illumina GSA (v2.0 and v3.0, respectively), all pre-sequencing analyses were carried out using the overlapping SNP content only (in total, 701,178 SNPs). Of all kinship pairs up to the 3rd degree (kinship coefficient > 0.0442) as identified by KING (Kinship-based INference for GWAS; Manichaikul et al., 2010), one individual was excluded from further analyses. If applicable for the pairs, controls were removed for the benefit of patients. For the subsequent PCA, SNPs were filtered for: (i) a call rate > 0.98; (ii) Hardy-Weinberg equilibrium (HWE) $P < 1x10^{-3}$; and (iii) a minor allele frequency (MAF) > 0.01, and then pruned ($r^2 < 0.2$) using PLINK1.9 (Chang et al., 2015). The ancestry of the study cohort was inferred using KING and the 1000 Genomes (Auton et al., 2015) as reference data. After the removal of individuals with an assigned Non-European ancestry, an in-sample PCA was performed using PLINK1.9 (Chang et al., 2015) to account for potential population outliers. Here, all individuals deviating more than six standard deviations from the mean in any of the first ten principal components (PC) were removed.

3.6.3 SmMIP design, library preparation and sequencing

The smMIPs for the five genes were designed in the context of a larger HAE candidate gene panel (in total 29 genes, Table S6 in APPENDIX A3) using an in-house pipeline (Thieme et al., 2021) based on the MIPgen software (Boyle et al., 2014). Target region were all coding exons (±6 bp flanking sequence) based on the human reference genome build 19 (hg19) and chromosomal positions according to NCBI Reference Sequence (RefSeq) definitions (O'Leary et al., 2016). Obtained smMIPs were evaluated for (i) the presence of SNPs with a MAF > 0.01 (according to gnomAD v2.1.1; Karczewski et al., 2020; Non-Finnish Europeans) in either of the two target arms, (ii) their individual logistic score, and (iii) the product of their target arm copy number. For smMIPs with SNPs in either target arm, ambiguous coding was used to obtain a primer mixture that captures

all possible alleles at the respective base. SmMIPs with a low logistic score (< 0.6) and/or a high copy number (> 20) were subjected to a re-design. To ensure proper exon coverage the final smMIP design was visually validated using the UCSC Genome Browser (Kent et al., 2002). The sequences of the 116 smMIPs capturing the five genes investigated in the present candidate gene study are listed in Table S1 of APPENDIX A1. For the smMIPs library preparation DNA concentration of the individual samples was measured using a Quant-iT PicoGreen dsDNA Assay (see Section 3.4.2). Final input was 50ng dsDNA per sample, resulting in an smMIP to DNA ratio of 1600:1. The library preparation was performed as described elsewhere (Eijkelenboom et al., 2016; Thieme et al., 2021). Briefly, this included pooling and phosphorylation of smMIPs, hybridization of smMIPs to the DNA, exonuclease treatment, and PCR amplification with the introduction of sample-specific barcodes (see Tables S7.1-4 in APPENDIX A3 for reaction mixes and thermocycler programs). To adjust for over- and underperforming smMIPs within the panel, two (re-)balancing runs were performed in six and two test samples, respectively, using an Illumina MiSeq platform. The results of these (re-)balancing runs are depicted in Figures S1 and S2 in APPENDIX A3. Upon amplification, the PCR products were pooled in a plate-wise manner and purified using 0.8x volume of Agencourt AMPure XP beads. The purity and fragment length were validated on an Agilent 2200 TapeStation system using High Sensitivity D1000 ScreenTapes. The individual purified pools were combined into two mega-pools containing 111 samples (patients only) and 440 samples (92 patients, 348 controls), respectively. Eventually, paired-end sequencing was performed in two NextSeq550 runs (2 x 150 bp; mid-output mode) using costume sequencing primers (O'Roak et al., 2012).

3.6.4 SmMIP analysis pipeline and data quality control

The sequencing base call files were first demultiplexed and converted to FASTQ files using the bcl2fastq conversion software. The obtained FASTQ files were then subjected to an in-house smMIP analysis pipeline (Thieme et al., 2021), which involved several bioinformatics steps as described elsewhere (Hiatt et al., 2013). In brief, paired-end reads were merged and aligned to the hg19 reference genome using PEAR (Paired-End reAd mergeR; Zhang et al., 2014) and Burrows-Wheeler Aligner (BWA-MEM; Li and Durbin, 2009). Following that, the reads were trimmed and collapsed using available MIPgen

scripts (Boyle et al., 2014). In this step, molecular tags, introduced as five degenerate bases during the synthesis of the smMIPs, were used to collapse individual reads carrying the same molecular tag. As a result, reads with high consensus sequences were generated, while PCR artifacts were reduced. Variants were called using the Genome Analysis Toolkit (GATK) UnifiedGenotyper (Van der Auwera et al., 2013) and then annotated using ANNOVAR (Wang et al., 2010). The threshold for the homozygous and heterozygous occurrence of a variant was set to alternative alleles occurring in > 75% or > 25% of reads, respectively.

Quality control (QC) of the sequencing data was carried out using data on the bp-wise coverage of the whole target region, derived from BAM files generated by the analysis pipeline. The QC included two stages and was performed separately for patients and controls. In stage one, low-quality samples (mean coverage < 50x and/or total coverage ≤ 15x in > 10% of the target region) and low-quality variants (GATK hard filter criteria: QUAL < 30 and QD < 10) were excluded. In addition, the coverage of all known pathogenic HAE variants, as listed in the professional version of the HGMD (version 04.2020; Stenson et al., 2020), was examined. For known HAE-associated variants with a mean coverage < 20x in either first run or second run patients or controls, re-sequencing was performed using Sanger sequencing (see Section 3.6.5). In stage two, an additional variant filter was applied (total coverage of > 15x in ≥ 90% of sequenced individuals) to only obtain high-confidence variants. Finally, all post-QC variants were merged between patients and controls for the subsequent analyses. To ensure comparability between the two sequencing runs, the mean coverages were evaluated using a Kruskal-Wallis rank sum test (kruskal.test() function of the R stats package; R Core Team, 2020).

3.6.5 Re-sequencing of low coverage regions using Sanger sequencing

Sanger sequencing (see Section 3.5.1) was performed for two exons (exon 9 of *F12* and exon 3 of *SERPING1*). In total, 195 of 197 patients and all 346 controls which remained after sample QC were subjected to re-sequencing. Due to a lack of DNA material, two Danish patients could not be re-sequenced. The sequences of the used primer are listed in Table S2 of APPENDIX A1. All steps subsequent to the standard PCR (Table S4 in APPENDIX A2) were performed by GENEWIZ' Sanger sequencing service (GENEWIZ

Europe by Azenta Life Science, Leipzig Germany). The generated sequences were analyzed using the SeqMan II software. All identified variants were aligned to the hg19 reference genome using the BLAT tool of the UCSC Genome Browser and then jointly analyzed with the variants obtained *via* smMIP sequencing.

3.6.6 Variant evaluation, prioritization and statistical analysis

To screen for known pathogenic HAE variants, all variants identified in patients after QC stage one were individually queried in the HGMD (version 04.2020; Stenson et al., 2020). Variants remaining after QC stage two were further prioritized for rare (MAF < 0.05), potentially functional (CADD ≥ 10) variants. Fisher exact tests were calculated using the fisher.test() function of the R Stats package (R Core Team, 2020) to (i) test all QC stage two variants individually for a statistically significant association with ACEi/ARB-AE, and (ii) test whether the number of samples carrying at least one prioritized variant was significantly different between patients and controls. For the latter, the enrichment of the prioritized variants was tested at the single gene level and for all five genes combined. *P*-values < 0.05 were considered nominally significant. In order to correct for multiple testing, adjusted *p*-values (P_{adjust}) were calculated using the Bonferroni method taking into account the number of individual variants that were identified (n = 85) or the number of investigated genes plus the entire gene panel (n = 5 + 1 = 6), respectively.

3.6.7 Power and sample size calculation

The R package genpwr (Moore et al., 2020) was used to (i) assess the overall power in the final study cohort (N_{total} = 543, case rate = 0.36) to detect rare variants (MAF < 0.05 or < 0.01) within a range of different effect sizes, and (ii) determine the number of individuals required to detect variants with a MAF < 0.05 and an odds ratio (OR) below the one achievable with the present study cohort. All analyses were based on an additive genetic model with a significance level of 0.05.

3.7 ACEi-AE GWAS meta-analysis

The main meta-analysis investigated eight independent GWAS cohorts of European ancestry, totaling 1,060 ACEi-AE patients and 77,799 controls ($meta_{EUR}$). Moreover, a second cross-ancestry meta-analysis ($meta_{ALL}$) was conducted, which included an

additional African-American cohort. Hence, the $meta_{ALL}$ analysis comprised 1,123 ACEi-AE patients and 77,948 controls. A summary of all GWAS cohorts is provided in Table 5.

Table 5 | GWAS meta-analysis cohorts.

Total sample size of the $meta_{EUR}$ analysis was N = 78,859 individuals; the $meta_{ALL}$ analysis comprised in total N = 79,071 individuals.

Cohort	Recruitment site	Ancestry	Patients	Controls	Number of SNPs	Available data level	Meta-analysis
vARIANCE	Germany/ Austria	European	95	4,135	9,418,075	Genotype data	$meta_{EUR}$ (Σ patients/controls = 1,060/77,799)
Denmark	Denmark	European	45	1,489	9,415,505	Genotype data	
Sweden	Sweden	European	42	975	9,472,108	Genotype data	
VanMar$_{EUR}$*	USA	European	106	321	9,456,972	Genotype data	
UKB*	UK	European	86	356	7,620,921	Imputed genotype data	
Swedegene*	Sweden	European	142	1,345	7,523,168	Summary statistics	
EstBB*	Estonia	European	82	15,787	8,734,929	Summary statistics	
CHB-CVDC/DBDS*	Denmark	European	462	53,391	7,490,822	Summary statistics	
VanMar$_{AFR}$*	USA	African-American	63	149	16,156,332	Genotype data	$meta_{ALL}$ (Σ patients/controls = 1,123/77,948)

The number of individuals and SNPs refers to those available post-QC. *Cohorts with treatment-matched controls. Abbreviations: GWAS, genome-wide association study | SNP, single nucleotide polymorphism.

3.7.1 Genome-wide genotyping, quality control, imputation, and association analysis

Given the availability of individual-level genotype data, the GWAS for six of the nine cohorts (vARIANCE, Denmark, Sweden, VanMar$_{EUR}$, VanMar$_{AFR}$ and UKB) were conducted in the scope of this dissertation. The analysis of these data is described in more detail in the following. For the other three cohorts, GWAS were performed externally and results were provided as summary statistics. For the Swedegene and CHB-CVDC/DBDS cohorts more detailed information is provided in the respective original publication (Rasmussen et al., 2020; Ghouse et al., 2021). The analysis of the EstBB cohort is described in APPENDIX A4.

Genome-wide genotyping

Genotyping of the DNA samples was performed using the Illumina GSA v2 and v3 (vARIANCE, Denmark and Sweden), the Illumina 610Quadv1.B Array (VanMar$_{EUR}$, VanMar$_{AFR}$) or the Affymetrix UK Biobank Axiom Array (UKB). The vARIANCE cohort and the Danish and Swedish patients were genotyped as described in Section 3.5.3 at the

Institute of Human Genetics of the University Hospital Bonn, Germany. Genotype-level data of the Danish and Swedish controls were obtained through collaborations. The $VanMar_{EUR}$ and $VanMar_{AFR}$ genotype data as well as the already imputed UKB genotype data were obtained upon applications to dbGaP and UKB, respectively.

QC and imputation

QC was performed separately for each cohort using PLINK1.9 (Chang et al., 2015). During genotype QC, variants with a (i) low call rate (< 0.95), (ii) MAF < 0.01, (iii) high difference in missingness between cases and controls (> 0.02), or (iv) deviation from HWE ($P < 10^{-10}$ for patients, $P < 10^{-6}$ for controls) were excluded. QC at the sample level comprised the exclusion of individuals with a low call rate (< 0.98) or a deviation in autosomal heterozygosity (FHET ± 0.20). In addition, the relatedness and population structure within each cohort were analyzed using PLINK1.9 (Chang et al., 2015), KING (Manichaikul et al., 2010) and an independent subset of SNPs subjected to LD pruning ($r^2 < 0.2$). Hereafter, one member of each pair of cryptically related individuals (up to 3rd degree, kinship coefficient > 0.0442) was removed. This was done randomly or, if applicable in the pairs, controls were removed before patients. Ancestral outliers were identified in a PCA by projecting each cohort against the five superpopulations of the 1000 Genomes Project (Auton et al., 2015), while population outliers were identified by two iterations of in-sample PCA. Only individuals with an assigned European ancestry and individuals who deviated no more than six standard deviations from the mean in any of the first ten PCs were retained for further analysis. To generate the PCs used as covariates in the association analysis, the PCA was repeated with the final set of individuals.

To harmonize the SNP set across the individual studies and facilitate downstream meta-analysis, all post-QC genotypes – with exception for the already imputed UKB genotypes - were imputed. To this end, the genotype data from each cohort were phased using Eagle (Loh et al., 2016) and imputed against the 1000 Genomes reference panel using minimac3 (Das et al., 2016). Details on the imputation of the UKB cohort can be found in the original publication (Bycroft et al., 2018).

In summary, QC led to the exclusion of 356 individuals (26 patients, 330 controls); the final number of patients and controls per cohort is given in Table 5. More detailed

information on the number of excluded individuals per cohort as well as the reason behind their exclusion is provided in Table S8 in APPENDIX A4.

Both, pre-imputation QC and imputation, were performed by Dr. Carlo Maj (Institute for Genomic Statistics and Bioinformatics, University of Bonn, Germany).

Single marker association analysis

For each individual cohort, association analysis was performed using SAIGE (Zhou et al., 2018), which accounts for inflated type I error rates in the presence of unbalanced case-control ratios using a generalized mixed model. Prior to the analysis, all post-imputation variants were filtered for a MAF ≥ 0.01 and an imputation score ≥ 0.3. Sex and the first four PCs were included as covariates into the model.

3.7.2 Meta-analysis ($meta_{EUR}$)

Preceding the meta-analysis, all GWAS summary statistics were harmonized. This included filtering of SNPs for a MAF ≥ 0.01 and an imputation score ≥ 0.3, the removal of duplicated and multiallelic SNPs, and the alignment of SNP IDs. Subsequently, a fixed-effects meta-analysis was conducted using METAL (V.2011-02-25; Willer et al., 2010), by weighting the effect size estimates with the inverse of the corresponding standard error under genomic control correction. For all downstream analyses, only SNPs analyzed in the patients of the CHB-CVDC/DBDS cohort and the patients of at least four other cohorts were considered. Consequently, all variants contained in the $meta_{EUR}$ analysis were investigated in at least 67.1% of all patients and 92.5% of all controls. Genome-wide significance was defined at $P < 5x10^{-8}$. Loci that reached a *p*-value of less than $1x10^{-5}$ were considered suggestive (Lander and Kruglyak, 1995). Heterogeneity between the meta-analysis cohorts was assessed using the heterogeneity scores Cochrans Q and I^2 index as implemented in METAL. Significant heterogeneity between studies is reflected in a heterogeneity *p*-value (Het*P*) < 0.05.

3.7.3 Leave-one-out polygenic risk score analysis

PRS were calculated for all cohorts with available individual-level genotype data (Table 5). Given the lack of a large independent GWAS sample, the analyses were performed in a leave-one-out setting. This means that the PRS for each individual of a given cohort (target data) was based on effect size estimates derived from a meta-analysis of the remaining cohorts (base data).

QC of target and base data was performed according to published parameters (Choi et al., 2020). In short, for the effect size estimation, a set of highly informative SNPs was derived by filtering variants for (i) a MAF ≥ 0.01, (ii) an imputation score ≥ 0.8, and (iii) by removing duplicated and ambiguous SNPs. Finally, the remaining SNPs were clumped by discarding SNPs within a 500 kb window that showed LD (r^2 > 0.1) with another more significant marker. QC of the target data comprised (i) standard GWAS QC, (ii) removal of duplicated SNPs, and (iii) removal of related samples (up to 3rd degree kinship, kinship coefficient > 0.0442). PRS were calculated using the "C+T" approach (SNP clumping/*p*-value thresholding) as implemented in PRSice-2 (v2.3.3 (2020-08-05); Choi and O'Reilly, 2019). According to this approach, PRS were calculated at ten different *p*-value thresholds ($5x10^{-8}$, $1x10^{-6}$, $1x10^{-4}$, 0.001, 0.01, 0.05, 0.1, 0.2, 0.5 and 1.0) based on the weighted sum of clumped variants. To test the association between the calculated PRS and ACEi-AE case-control status, standard logistic regression including the same covariates as in the single marker association testing (sex, PC1-4) was performed. Finally, to calculate the proportion of variance explained (Nagelkerke's R^2), scores from a full model (covariates and PRS) and a reduced model (covariates only) were compared.

3.7.4 Identification of risk loci, functional annotation, and gene prioritization

The FUMA platform (FUnctional Mapping and Annotation of GWAS; version 1.4.1; Watanabe et al., 2017) was used to define genomic risk loci, functionally annotate the SNPs within these loci, and prioritize the most likely causal genes.

First, independent significant SNPs were defined as those that were genome-wide significant ($P < 5x10^{-8}$) and independent at $r^2 < 0.6$. For this, LD structures derived from pre-calculated LD scores of the European 1000 Genome reference population (Auton et al., 2015) were used. In a second step, all SNPs with an $r^2 > 0.6$ in relation to any of the identified independent SNPs were included into the risk locus ("candidate variants"). Subsequently, independent significant SNPs that were independent at $r^2 < 0.1$ were defined as independent lead SNPs within the locus. Moreover, LD blocks of independent significant SNPs less than 250 kb apart from each other were merged into one risk locus. Suggestive risk loci were defined accordingly except for the *p*-value thresholds which were set to $P < 1x10^{-5}$ for independent significant SNPs and $P < 0.05$ for candidate SNPs. Functional annotation of all SNPs within the pre-defined genome-wide significant

and suggestive risk loci was performed using ANNOVAR (Wang et al., 2010), CADD scores (scores > 12.37 indicate suggestive deleterious SNPs; Kircher et al., 2014; Amendola et al., 2015), RegulomeDB scores (lower scores indicate a higher potential of a regulatory function; Boyle et al., 2012), and chromatin state annotations based on five chromatin marks for 127 cell/tissue types in the Roadmap Epigenomics Project (scores ≤ 7 indicate higher accessibility of the genomic regions, Roadmap Epigenomics Consortium, 2015). Gene prioritization at the genome-wide significant risk loci was performed employing three different approaches embedded in FUMA: (i) positional mapping, (ii) eQTL mapping, and (iii) 3D chromatin interaction mapping. Positional mapping was performed on the basis of potential deleterious SNPs using ANNOVAR annotations and CADD scores. In the case of intergenic SNPs, a maximum distance of 10 kb was set to assign SNPs to genes. Using eQTL information from 44 tissue types (GTEx v8; Lonsdale et al., 2013) SNPs within the risk loci were assigned to genes within a distance of up to one megabase (*cis*-eQTLs). Here, SNP-gene pairs with a false discovery rate (FDR) < 0.05 were considered significant. Lastly, 3D DNA-DNA interactions between SNPs at the risk loci and other regions (genes) in the genome were explored using Hi-C data derived from 21 different tissue and cells type (Schmitt et al., 2016). The significance threshold here was set at an FDR < $1x10^{-6}$.

3.7.5 Gene-based tests, gene-set and tissue expression analyses

To analyze the joint effects of the genetic markers detected in the GWAS meta-analysis gene- and gene-set analyses were performed using MAGMA (Multi-maker Analysis of GenoMic Annotation v1.08; de Leeuw et al., 2015) as implemented FUMA. In data pre-processing, SNPs were assigned to genes within a window size of 35 kb upstream and 10 kb downstream of a protein-coding gene to also include regulatory regions (O'Dushlaine et al., 2015). SNPs mapping within the pre-defined window were then used to compute gene-based *p*-values (SNP-wise mean model). Taking into account the number of tested protein-coding genes (n = 18,983) the genome-wide significance threshold for the gene-based test was defined at a *p*-value of $2.63x10^{-6}$ (= 0.05/18,983). Following that and based on the obtained gene-level *p*-values, genes were further aggregated to the level of gene sets and using a competitive gene-set analysis jointly tested for association. Gene-sets included in the analysis were obtained from MSigDB

(Molecular signatures database v7.0; Subramanian et al., 2005; Liberzon et al., 2011) and comprised "curated gene sets" (n = 5,500) as well as "GO (gene ontology) terms" (n = 9,996). For the gene-set analysis a Bonferroni-corrected *p*-value of 3.23×10^{-6} (= 0.05/15,496) was considered genome-wide significant. In addition, to determine if the assigned genes showed tissue-specific expression levels, a gene property analysis in 53 GTEx tissues (v8 release; Lonsdale et al., 2013) was performed.

3.7.6 Statistical fine-mapping analysis

Summary statistics based fine-mapping was performed using SuSiE (Sum of single effects; Wang et al., 2020; Zou et al., 2022), as implemented in PolyFun (Polygenic functionally informed fine-mapping; Weissbrod et al., 2020). As individual-level genotype data were not available for all meta-analysis cohorts, fine-mapping was performed using two different LD correlation matrices for comparison. These were (i) pre-computed LD scores derived from a large-scale UKB data set (N = 337K un-related British ancestry individuals; as provided by PolyFun), and (ii) LD scores derived from the Danish GWAS cohort (N = 1,542; obtained from PLINK files using LDstore 2.0), as this ethnicity is representative for the majority of all meta-analysis individuals. The model was based on a fine-mapping window of one megabase surrounding the respective lead SNP and N = 5 as the maximum number of causal SNPs per locus.

3.7.7 LD score regression analyses

LD score regression (Bulik-Sullivan et al., 2015) was used to (i) estimate the SNP-based heritability of ACEi-AE, and (ii) evaluate the genetic overlap between ACEi-AE and a selection of associated traits and diseases. Both analyses were performed using LDSC (v.1.0.1; Bulik-Sullivan et al., 2015) and were based on an ACEi-AE meta-analysis without genomic control correction, additionally filtered for high confidence variants (imputation score ≥ 0.8).

Heritability estimates were determined on the liability scale taking into account the sample prevalence as well as the lower and upper limit of the 12-month population prevalence estimate for ACEi-AE (0.004% to 0.026%; Aygören-Pürsün et al. 2018). Cross-trait LD score regression was performed using publicly available GWAS summary statistics (Table S9 in APPENDIX A4). Overall, nine traits were tested including five ACEi-AE associated traits and diseases (hypertension, asthma, blood clot in the leg (deep vein

thrombosis (DVT)), blood clot in the lung, intake of renin angiotensin-agents), three reported risk factors (smoking (Morimoto et al., 2004), hay fever/allergic rhinitis (Kostis et al., 2005; Mahmoudpour et al., 2016) and coronary artery disease (Miller et al., 2008)) and one reported protective factor (diabetes (Kostis et al., 2005; Byrd et al., 2008; Miller et al., 2008)). To address the potential influence of hypertensive or ACEi-related association signals, arising from the population-based controls present in three of the GWAS cohorts (~ 8.5% of all controls), (nominally) significant results were re-analyzed using a stratified $meta_{EUR}$ analysis that included only cohorts with treatment-matched controls (Ncase/Ncontrol = 878/71,200; Table 5). To account for multiple testing, the obtained *p*-values were Bonferroni corrected for the number of investigated trait pairs (n = 9) and the re-analysis of three nominally significantly associated traits (P_{adjust} = 0.05/12 = 0.0041).

3.7.8 Cross-ancestry comparison and meta-analysis ($meta_{ALL}$)

The available African-American GWAS cohort ($VanMar_{AFR}$) was used to explore whether ACEi-AE shares common genetic variants across different ancestries.

First, the collective contribution of common variants in African-American individuals was assessed by PRS. For this, PRS were calculated as described in Section 3.7.3 using effect size estimates derived from the European individuals ($meta_{EUR}$ data). Next, a fixed-effects meta-analysis under genomic control correction was performed between the $VanMar_{AFR}$ and $meta_{EUR}$ data using METAL ($meta_{ALL}$). Subsequently, the obtained variants were filtered and grouped into genomic risk loci as has been done for the $meta_{EUR}$ data (see Sections 3.7.2 and 3.7.4). Finally, the effect sizes and effect allele frequencies of the $meta_{EUR}$ lead SNPs were compared between the two ancestries. To this end, effect alleles were aligned and Pearson correlation coefficients were calculated.

4 Results

4.1 Candidate gene analysis in an ACEi/ARB-AE cohort

The results of this study have been published as an original article in *Frontiers in Genetics* (Mathey et al., 2022).

4.1.1 Pre-sequencing, sample and variant QC

To ensure a homogeneous study cohort with comparable genetic variant profiles (e.g., in terms of allele frequencies), kinship, ancestry and population stratification were investigated.

The kinship analysis revealed five pairs of first-degree relatives (one pair of patients, four pairs of controls), for each of which one individual was excluded from the study cohort. The subsequent PCA revealed four patients with an assigned Non-European ancestry (Figure 9), and an additional four patients were identified as population outliers within the study cohort (Figure S3 in APPENDIX B1).

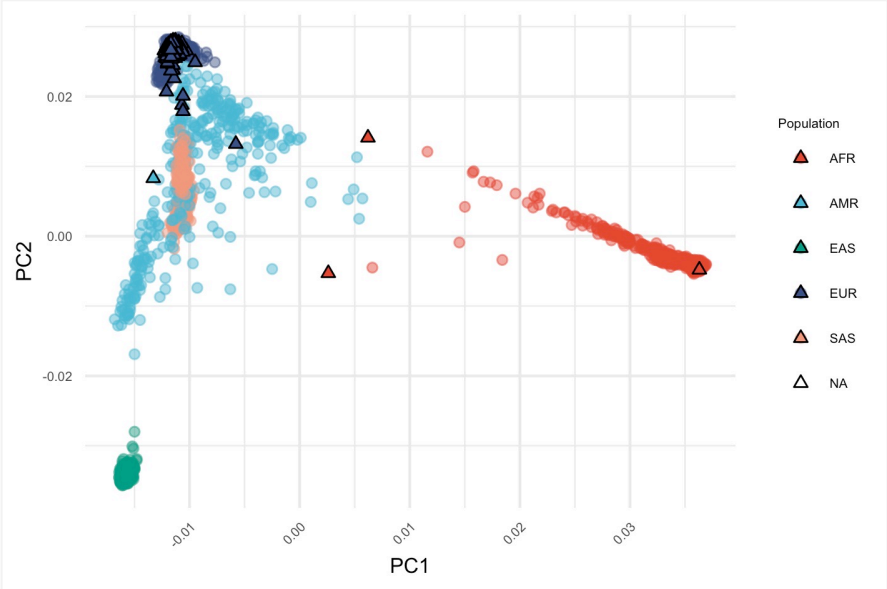

Figure 9 | Ancestry inference using 1000 Genomes as a reference data set.

Individuals of the 1000 Genomes data set are represented by circles and colored according to their assigned superpopulation. Individuals of the study cohort are represented by black triangles filled with the color corresponding to their inferred ancestry. In total, four individuals were identified as outlying: three individuals with African ancestry (red triangles) and one individual with Ad Mixed American ancestry (light blue triangle). Abbreviations: AFR, African | AMR, Ad Mixed American | EAS, East Asian | EUR, European | PC, principal component | SAS, South Asian | NA, not available.

Removal of all outlying individuals resulted in a total number of 203 patients and 348 controls which ultimately underwent smMIP sequencing.

Sequencing was performed in two separate runs which achieved mean coverages of 245x and 239x, respectively. A stratified comparison of the mean coverages revealed no significant inter-run differences between patients and controls sequenced in different runs. However, patients sequenced in the first and controls sequenced in the second run showed a significantly higher coverage compared to the patients sequenced in the second run (Figure S4 in APPENDIX B1). Given that the mean coverage in each of the three groups was well above 220x, this should not have impacted downstream variant detection.

The sample-level QC criteria were met by 543 individuals (> 98%, 197 patients and 346 controls). Variant annotation was performed separately for patients and controls and resulted in a total of 74 identified variants in patients and 104 identified variants in controls. During variant-level QC 24 variants in patients (18 in QC stage one, six in QC stage two) and 31 variants in controls were excluded. Finally, merging of the post-QC variants resulted in a total of 85 independent variants.

4.1.2 Clinical characteristics of the final analysis cohort

The final study cohort included 197 ACEi/ARB-AE patients (46.5% females, 53.5% males) and 346 controls (55.8% females, 44.2% males). The mean age for the development of the angioedema (age at event) was 65.1 years (54-71 years, interquartile range). In over two thirds of all patients (73.1%), more than one year elapsed between the first intake of an ACEi/ARB and the onset of the angioedema (time-to-onset).

The angioedema was caused by an ACEi in 87% of all investigated patients, whereas ARBs were reported as causal in 13% of patients. However, the most frequently reported ACEi/ARB differed between the cohorts. While Ramipril was the most frequently suspected drug in the vARIANCE patients (66% of all ACEi-AE), it was Enalapril in the Danish and Swedish patients (67.3% and 92% of all ACEi-AE, respectively). Among all ARB-AE cases, Candesartan was the most frequently reported drug in the vARIANCE and Swedish patients (45.5% and 62.5% of all ARB-AE, respectively), while in the Danish patients it was Losartan (83.3% of all ARB-AE). A detailed overview of the baseline and clinical characteristics of the three patient cohorts is given in Table 6.

Table 6 | Baseline and clinical characteristics of all post-QC patients stratified per cohort.

		vARIANCE (N = 64)	**Denmark** (N = 62)	**Sweden** (N = 71)
Female : Male (%)		50.0 : 50.0	51.6 : 48.4	38.0 : 62:0
Age at event (years, IQR)		63 (54-70)	65 (55-75)	67 (63-71)
Time-to-onset (N, %)	1-3d	2 (3.2%)	3 (9.1%)	4 (6.5%)
	4-14d	4 (6.5%)	-	3 (4.8%)
	2w-2m	2 (3.2%)	-	3 (4.8%)
	2m-1y	5 (8.1%)	3 (9.1%)	16 (25.8%)
	> 1y	49 (79.0%)	17 (81.8%)	36 (58.1%)
Suspected ACEi, N (%)		**Total, 50 (82.0):** Ramipril, 33 (66.0) Lisinopril, 10 (20.0) Enalapril, 7 (14.0)	**Total, 55 (90.2):** Enalapril, 37 (67.3) Ramipril, 8 (14.5) Trandolapril, 4 (7.3) Lisinopril, 3 (5.5) Perindopril, 1 (1.8) NA, 1 (1.8)	**Total, 63 (88.7):** Enalapril, 58 (92.1) Ramipril, 5 (7.9)
Suspected ARB, N (%)		**Total, 11 (18.0):** Candesartan, 5 (45.5) Valsartan, 3 (27.3) Telmisartan, 2 (18.2) Irbesartan, 1 (9.1)	**Total, 6 (9.8):** Losartan, 5 (83.3) Irbesartan, 1 (16.7)	**Total, 8 (11.3):** Candesartan, 5 (62.5) Losartan, 3 (37.5)

Age at event was not available for one patient of the vARIANCE cohort and two patients of the Denmark cohort. For two vARIANCE, 29 Danish and nine Swedish patients, no data on the time-to-onset was available. Percentages were therefore calculated on the basis of N = 62, N = 33 and N = 62 for the vARIANCE, Danish and Swedish patients, respectively. Data on the suspected angioedema triggering ACEi/ARB are only based on N = 61 patients for the vARIANCE and Danish cohort, as for one Danish patient no information on the suspected drug was available and for three vARIANCE patients two different suspected drugs were reported. Abbreviations: ACEi, angiotensin-converting enzyme inhibitor | ARB, angiotensin-II-receptor blocker | d, days | IQR = interquartile range | m = months | N, number of individuals | NA = not available | w = weeks | y = years.

4.1.3 ACEi/ARB-AE patients: screening for pathogenic HAE-associated variants

Overall, 56 variants were identified in the ACEi/ARB-AE patients after QC stage one. Evaluation of these variants in the HGMD revealed that none of them represents a known pathogenic HAE variant. In fact, only one missense variant identified in exon 6 of the *F12* gene (rs35515200; NM_000505.4:c.418C>G; p.L140V), was listed in the HGMD and reported as associated with deep vein thrombosis (Lotta et al., 2012). In the present study, this particular variant was identified in a heterozygous state in two patients and three controls and showed no significant association ($P > 0.999$) with ACEi/ARB-AE. Moreover, a re-evaluation of the variant according to the ACMG guidelines (Richards et

al., 2015) indicated the variant to be of uncertain significance. As such, the relevance of this finding is not readily apparent.

4.1.4 Patients vs. controls: single variant association and enrichment analyses

Of the 85 independent variants identified in patients and controls, 15 were common (MAF ≥ 0.05) and 70 were rare (MAF < 0.05). With the exception of two variants, all rare variants were found in a heterozygous state. One variant in the *F12* gene and one variant in the *PLG* gene, however, occurred homozygously in one patient and in one control, respectively.

A Fisher's exact test at the single variant level revealed that none of the identified variants, common or rare, showed a significant association ($P < 0.05$) with ACEi/ARB-AE. The top five rare variants, which showed the lowest *p*-value are listed in Table 8.

The further prioritization of rare (MAF < 0.05), potentially functional (CADD ≥ 10) variants, resulted in 42 variants. Across all genes, the number of individuals who carried at least one prioritized variant was statistically higher in controls than in patients ($P = 0.015$). However, this finding was no longer significant after correction for multiple testing ($P_{adjust} = 0.091$). The same trend was observed at the level of single genes. The difference, though, was only nominally significant for the *F12* gene ($P = 0.033$) and again this finding did not withstand correction for multiple testing ($P_{adjust} = 0.199$). The complete results of the enrichment analyses are summarized in Table 7.

Table 7 | Enrichment analyses of rare, potentially functional variants.

	Patients carrying ≥ 1 prioritized variant		Controls carrying ≥ 1 prioritized variant			
Gene	**N**	**%**	**N**	**%**	***P***	***P*$_{adjust}$**
ANGPT1	8	4.1	22	6.4	0.330	>0.999
F12	4	2.0	21	6.1	**0.033**	0.199
KNG1	21	10.7	57	16.5	0.075	0.448
PLG	19	9.6	37	10.7	0.770	>0.999
SERPING1	1	0.5	5	1.4	0.425	>0.999
Whole panel	47	23.9	118	34.1	**0.015**	0.091

Nominal significant values are written in bold. Abbreviations: N, number of individuals |*P*, *p*-value | P_{adjust}, Bonferroni adjusted *p*-value.

Table 8 | Top five rare variants identified in re-sequencing.

Gene	**Chr**	**Pos** (hg19)	**Ref/Alt**	**AA change**	**rsID** (dbSNP155)	**Consequence**	**HGVS nomenclature**	**CADD** (v1.6)	**AC cases/ ctrls**	***P***	P_{adjust}
PLG	6	161152905	C/T	R523W	rs4252129	Non-synonymous SNV	NM_000301.5: c.1567C>T	19.54	0/8	0.057	>0.999
F12	5	176831826	C/G	A207P	rs17876030	Non-synonymous SNV	NM_000505.4: c.619G>C	13.18	2/14	0.064	>0.999
KNG1	3	186461524	C/T	R376X, R412X	rs76438938	Stopgain	NM_001166451.2: c.1126C>T, NM_000893.4: c.1234C>T	16.28	6/24	0.081	>0.999
PLG	6	161134069	G/A	R153R	rs144153702	Synonymous SNV	NM_000301.5: c.459G>A	11.19	2/0	0.131	>0.999
PLG	6	161137790	G/A	R261H	rs4252187	Non-synonymous SNV	NM_000301.5: c.782G>A	27.70	4/3	0.263	>0.999

Abbreviations: AA, amino acid | AC cases/ctrls, allele count in the cases/controls (total sample = 197 cases, 346 controls) | Alt, alternative allele | Chr, chromosome | CADD, Phred-scaled combined annotation dependent depletion score; version 1.6 | HGVS, human genome variant society | *P*, *p*-value | P_{adjust}, Bonferroni corrected *p*-value | Pos, genomic position according to the Genome Reference Consortium human build 37 (GRCh37/hg19) | Ref, reference allele | rsID, reference SNP cluster ID | SNV, single nucleotide variant.

4.1.5 Power and sample size analyses

The candidate gene analysis included all ACEi/ARB-AE patients available to us at that time. Given the non-significant results, a power analysis was performed to determine its overall power to detect variants with a MAF < 0.05 or < 0.01. The analysis revealed 80% power to detect variants with an OR of 2.2 and a MAF < 0.05. For even rarer variants (MAF < 0.01) an OR of only 6.2 would have been detectable (Figure 10A). In turn, the detection of variants with a MAF < 0.05 and an OR below 2.2 would have required a substantial increase in sample size. For instance, with an approximately 2-fold increase in total sample size (Ntotal = 1,082), variants with an OR of ~ 1.75 would have been in the detectable range (Figure 10B).

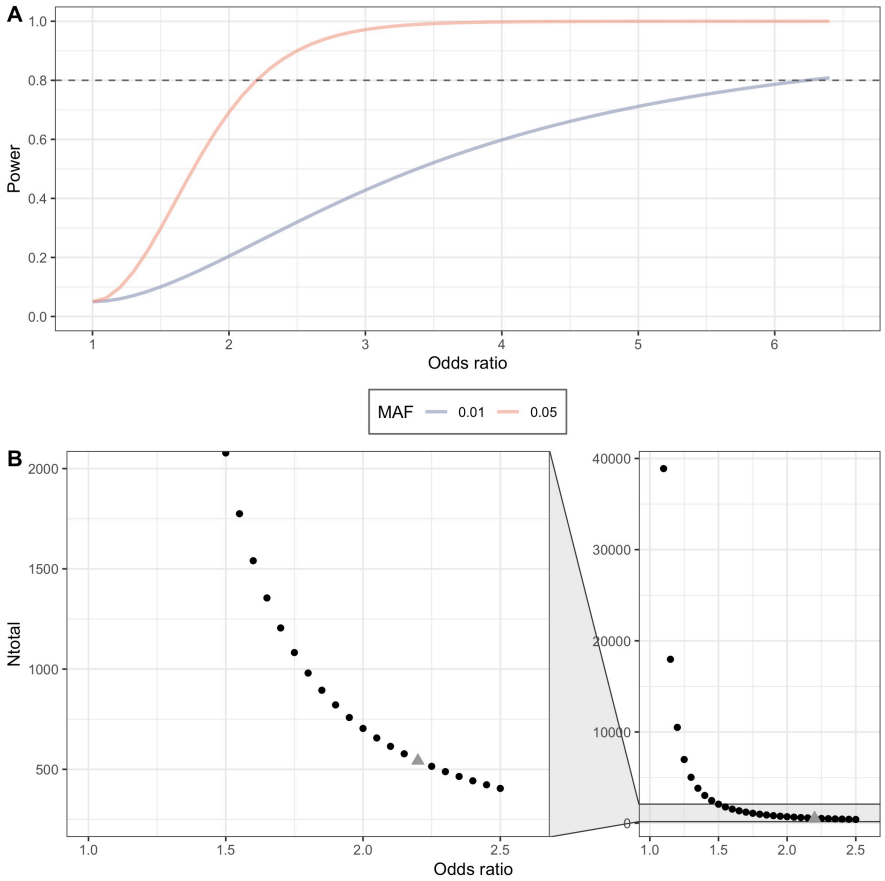

Figure 10 | Power analysis and sample size calculation results.

(A) Power calculation on the basis of the final sample size (Ncase/Ncontrol = 197/346), a significance level of 0.05 and an additive genetic model. (B) Calculation of the total number of samples (Ntotal) that would have been required to detect variants with odds ratios lower than the odds ratios that could be detected with the present study cohort. The model was based on a case rate of 0.36, variants with a MAF < 0.05 and assumes an additive genetic model, a significance level of 0.05 as well as a power of 80%. The sample size of the present study cohort is represented by the grey triangle. Abbreviations: MAF, minor allele frequency | N, number of individuals.

4.2 Genome-wide meta-analysis of ACEi-AE

The results of this study have been published as an original article in *The Journal of Allergy and Clinical Immunology* (Mathey et al., 2024).

4.2.1 ACEi-AE risk loci: single marker association results ($meta_{EUR}$)

The $meta_{EUR}$ analysis included a total of 1,060 ACEi-AE patients and 77,799 controls from eight independent GWAS cohorts, all of European ancestry (Table 5). By analyzing ~ 6.99 million markers, the meta-analysis identified three genome-wide significant risk loci, mapping to 1q24.2, 14q32.2, and 20q11.22 (Figure 11, Table 9). The locus on chromosome 20q11.22 represent a novel risk locus for ACEi-AE, whereas the other two loci have been described previously. The 14q32.2 locus was first identified as genome-wide significant in the GWAS by Ghouse et al. (2021), whose discovery cohort was part of the present meta-analysis. The 1q24.2 locus was identified in the study by Maroteau et al. (2020), although the association in this previous study was exome-wide significant only ($P < 1x10^{-6}$).

The most significant association in the present analysis, that is the locus with the lowest *p*-value, was identified at chromosome 14q32.2 (top SNP rs35136400, $P = 1.28x10^{-12}$, OR = 1.50). SNP rs35136400 is located approximately 50kb upstream of the gene *BDKRB2* (Figure 12B). Moreover, the SNP was found to be in near-perfect LD with rs34485356 ($r^2 = 0.971$), the most significant SNP in the GWAS in which this locus was first reported (Ghouse et al., 2021).

The second most significant association was found on chromosome 1q24.2. Top SNP at this locus was rs6687813 ($P = 2.67x10^{-10}$, OR = 1.70), an intergenic variant approximately 5.8 kb downstream of the gene *F5* (Figure 12A). In the previous study the coding variant rs6025, also known as Factor V Leiden mutation, was reported as the lead SNP at this locus (Maroteau et al., 2020). The SNP rs6025 was also among the genome-wide significant SNPs in the present meta-analysis ($P = 5.81x10^{-9}$) and was found to be only in low LD with rs6687813 ($r^2 = 0.172$).

The third most significant association mapped to chromosome 20q11.22, a locus that had not yet been reported as associated with ACEi-AE risk. Top SNP at this locus was rs6060237 ($P = 3.47x10^{-8}$, OR = 0.70), an intergenic variant approximately 9 kb downstream of the gene *EDEM2* (Figure 12C).

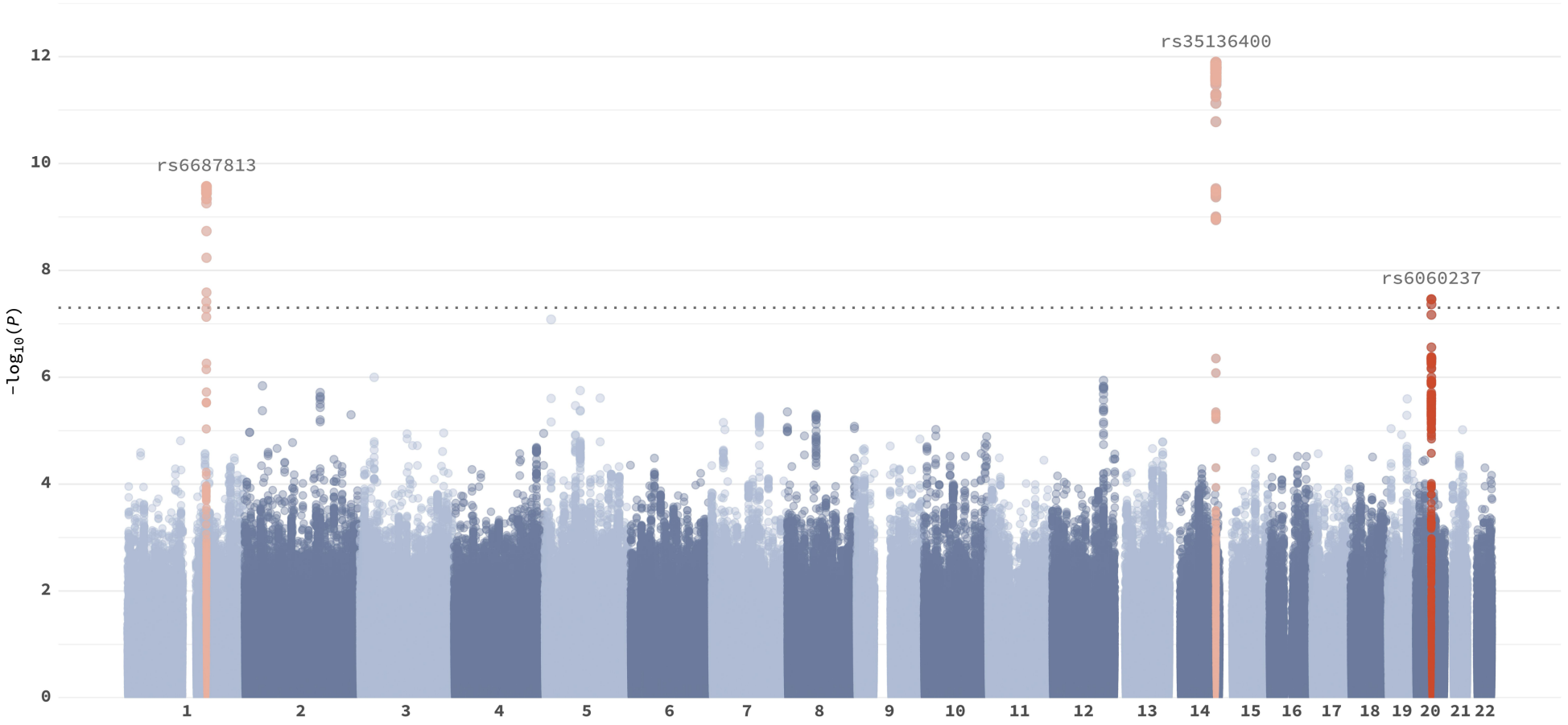

Figure 11 | Manhattan plot of the meta$_{EUR}$ analysis.

Manhattan plot displaying the -log10 association *p*-values (vertical axis) against the genomic position (horizontal axis) of all analyzed SNPs. The threshold for genome-wide significance ($P = 5x10^{-8}$) is indicated by the gray dotted line. Loci reaching genome-wide significance are highlighted (lead SNP ±500 kb) in red (novel locus) and orange (previously reported loci). Abbreviations: kb, kilobase | SNP, single nucleotide polymorphism.

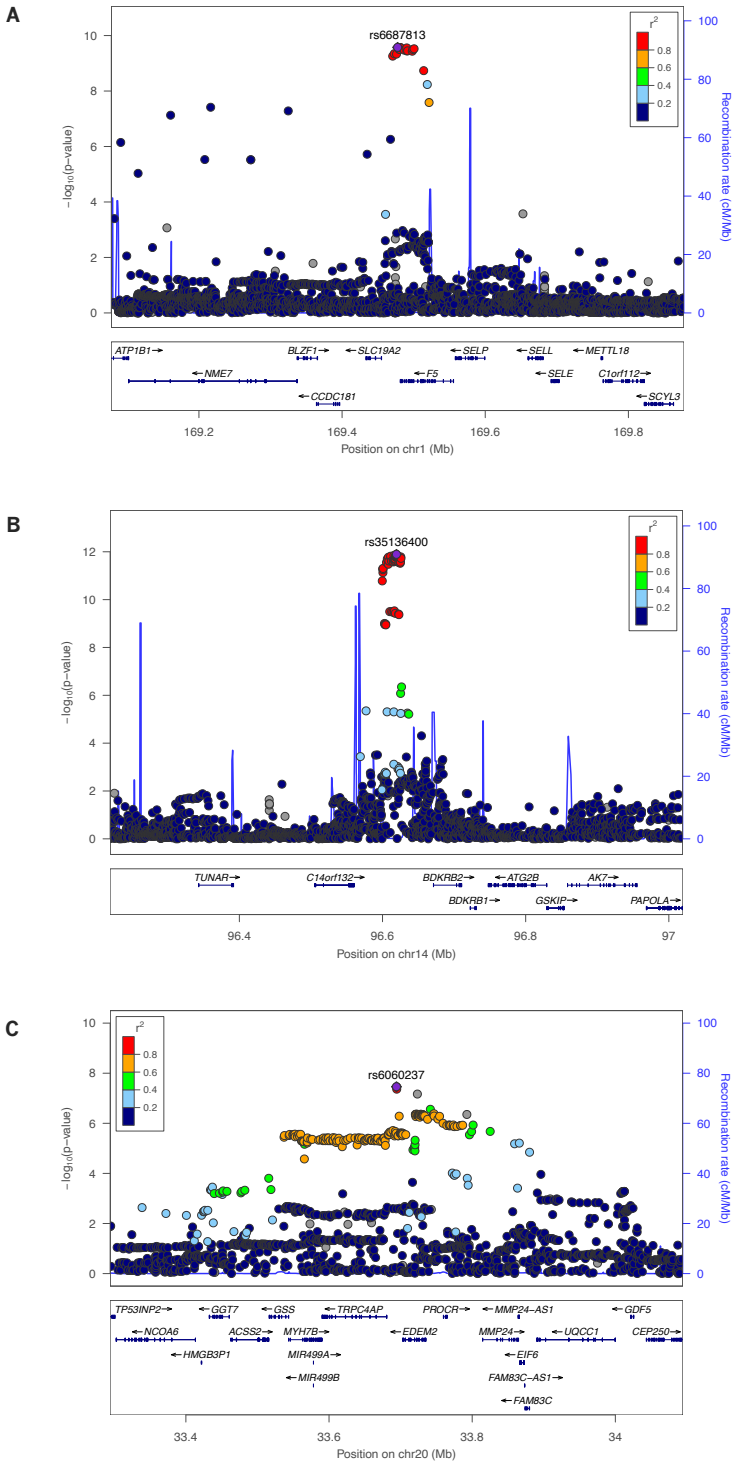

Figure 12 | LocusZoom plots for the 1q24.2 (A), 14q32.2 (B), and 20q11.22 (C) risk loci.

Regional association plots displaying the genomic region (±400 kb) surrounding the top SNP at each of the genome-wide significant loci. In each plot, the top SNP is denoted by the purple diamond, all other SNPs are colored according to their linkage disequilibrium (r^2) with the respective top SNP: red ($0.8 < r^2 < 1$), orange ($0.6 < r^2 \leq 0.8$), green ($0.4 < r^2 \leq 0.6$), light blue ($0.2 < r^2 \leq 0.4$), and dark blue ($r^2 \leq 0.2$). Abbreviations: kb, kilobase | SNP, single nucleotide polymorphism.

The risk loci on chromosomes 1q24.2 and 20q11.22 showed effect sizes and effect directions that were consistent across studies (Het*P* = 0.723 and Het*P* = 0.201, respectively), whereas a nominally significant in-between study heterogeneity (Het*P* = 0.024) was observed at the previously reported 14q32.2 locus. This is probably attributable to the opposite effect direction observed in one of the eight GWAS cohorts (Figure S5 in APPENDIX B2).

Table 9 | Genome-wide significant risk loci identified in the meta$_{EUR}$ analysis.

Lead SNP	Chr	Pos	A1/A2	FreqA1	OR	95% CI	*P*	Het I^2	Het*P*
rs6687813	1	169477574	A/C	0.083	1.70	1.54-1.87	$2.67x10^{-10}$	0	0.723
rs35136400	14	96619480	A/G	0.774	1.50	1.39-1.61	$1.28x10^{-12}$	56.5	0.024
rs6060237	20	33694210	A/G	0.855	0.70	0.57-0.83	$3.47x10^{-8}$	28.5	0.201

Abbreviations: Chr, chromosome | Pos, genomic position (hg19) | A1/A2, effect allele/other allele | FreqA1, effect allele frequency in the combined case-control cohort | OR, odds ratio of the effect allele | CI, confidence interval | *P*, *p*-value |Het I^2, heterogeneity I^2| Het*P*, heterogeneity *p*-value.

Besides the three genome-wide significant loci, 20 additional loci showed a borderline association ($P < 1x10^{-5}$) with ACEi-AE. These loci mapped to chromosome 2 (2p22.2, 2q24.1, 2q36.1), chromosome 3 (3p24.1), chromosome 5 (5p15.2, 5q12.3, 5q13.3, 5q23.1), chromosome 7 (7p15.2, 7p15.1, 7q22.1), chromosome 8 (8p23.3, 8q12.2, 8q24.3), chromosome 10 (10p12.1), chromosome 12 (12q23.3), chromosome 19 (19p13.2, 19q13.2 (2x)) and chromosome 21 (21q22.12). More detailed information on the lead SNPs of these suggestive loci are provided in Table S9 in APPENDIX B2.

4.2.2 Leave-one-out polygenic risk score analysis

The consistency of the ACEi-AE phenotype across the meta$_{EUR}$ cohorts was assessed in leave-one-out PRS analyses. Naturally, this was only possible for cohorts with available individual-level genotype data, i.e., for five out of the eight GWAS meta-analysis cohorts (see Table 5).

Overall, the PRS analyses revealed a positive direction of the polygenic signal alongside a significant prediction of the ACEi-AE case-control status in all five investigated GWAS cohorts (Figure 13, left panel). The maximum variance explained by the PRS ranged from 1.10% (vARIANCE) to 5.37% (VanMar$_{EUR}$).

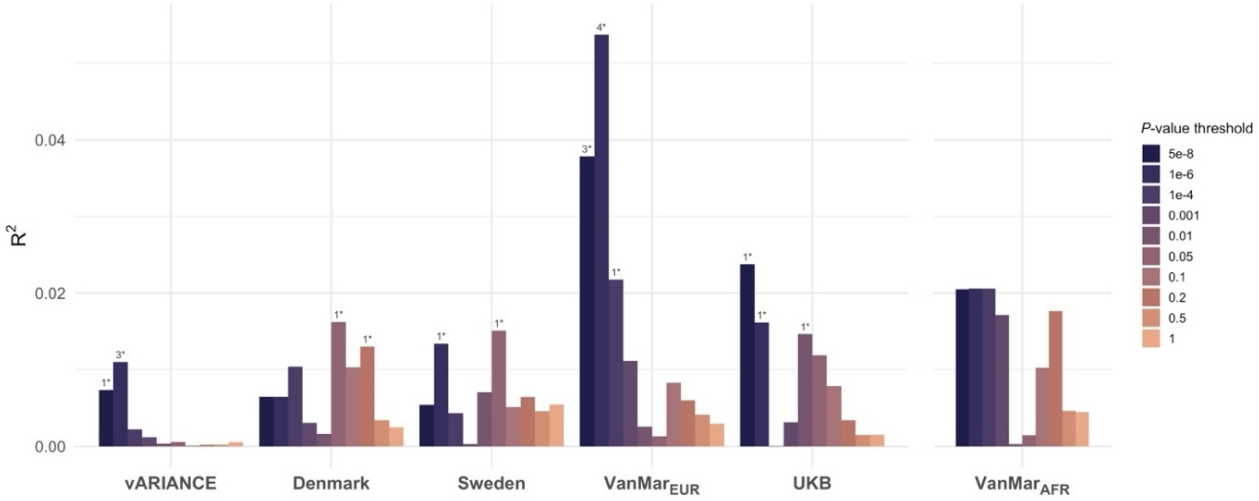

Figure 13 | Polygenic risk score analysis results.

Leave-one-out PRS results of the five European GWAS cohorts (left panel) in comparison to the PRS results of the African-American GWAS cohort (right panel), which were based on the effect sizes derived from the $meta_{EUR}$ data. For each cohort, PRS across the ten tested *p*-value thresholds are displayed. The statistical significance of the variance explained (R^2) by the PRS is depicted above each bar: 1* = $P < 0.05$ | 2* = $P < 0.01$ | 3* = $P < 0.005$ | 4* = $P <$ 1e-4. Abbreviations: GWAS, genome-wide association study | PRS, polygenic risk score.

4.2.3 Bioinformatics follow-up analyses

4.2.3.1 Functional annotation of lead and candidate SNPs

Overall, 231 SNPs with a *p*-value < $1x10^{-5}$ were found to be in moderate to high LD ($r^2 > 0.6$) with a lead SNP and thus considered "candidate SNPs" at the three genome-wide significant loci. As is typical for GWAS findings, SNPs within the risk loci were almost exclusively found to be located in the non-coding regions of the genome. In addition, the functional annotation results of several SNPs indicated their potential involvement in gene regulation and/or regulatory effects on transcription factor binding.

Only five of the SNPs mapped to exonic regions of the genome. Of these five SNPs, three were found to have high CADD scores thereby indicating a potential deleterious effect on protein function: rs6025 (“Factor V Leiden”; located in exon 10 of *F5*, CADD = 18.92); rs867186 (located in exon 4 of *PROCR*, CADD = 16.65); and rs80109502 (located in exon 38 of *MYH7B*, CADD = 17.02). The annotation results of all candidate SNPs at the genome-wide significant loci are provided in Table S10 in APPENDIX B2.

4.2.3.2 Gene prioritization

To pinpoint the most likely causal genes at the three genome-wide significant loci, all lead and candidate SNPs within these loci were assigned to genes based on their genomic location and deleteriousness, effect on gene expression and 3D chromatin interactions.

Altogether, 84 genes were prioritized at the three risk loci. Broken down by locus, 32 genes were mapped at 1q24.2, 21 genes were mapped at 14q32.2, and 31 genes were mapped at 20q11.22. Broken down by method, five genes were prioritized by deleterious SNPs in their physical distance, for 21 genes significant eQTL effects were identified, while 80 genes were found to have significant chromatin interactions with the tested SNPs. A total of 21 genes were supported by at least two lines of evidence (Figure 14, Table S11 in APPENDIX B2).

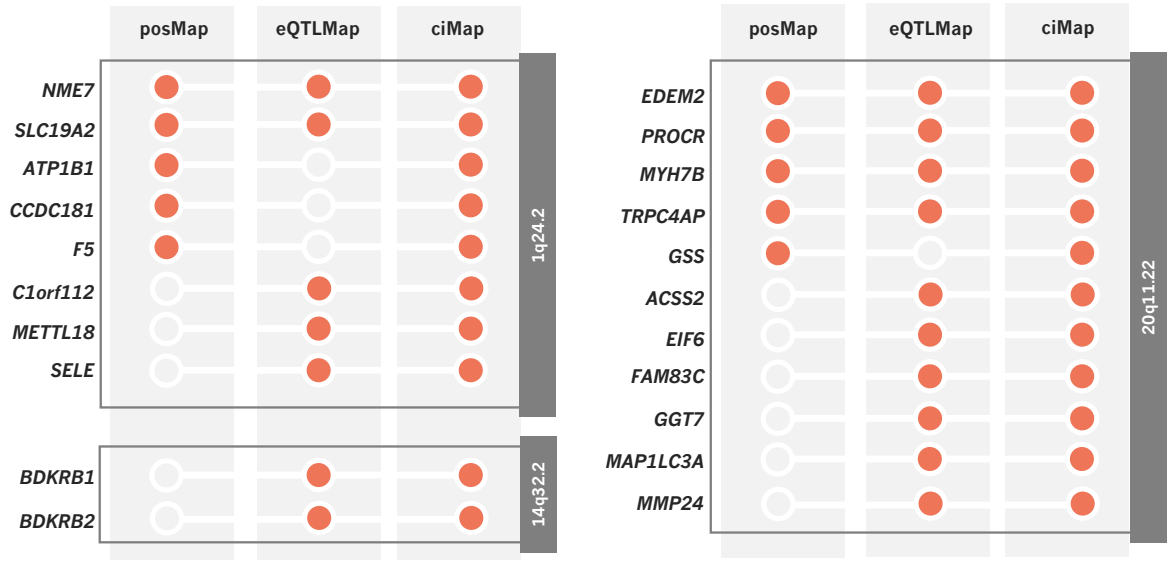

Figure 14 | Prioritized genes at the genome-wide significant loci.

Locus-wise overview of all genes that were prioritized by at least two mapping strategies: positional mapping of deleterious SNPs (posMap), effects on gene expression (eQTLMap), or chromatin interactions (ciMap).

At the 1q24.2 locus, two genes (*NME7*, *SLC19A2*) were prioritized by all three mapping strategies. Moreover, six additional genes were supported by two strategies: eQTL effects and chromatin interactions were found for three genes (*C1orf112*, *METTL18*, *SELE*), while another three genes (*ATP1B1*, *CCDC181*, *F5*) were tagged by positional and chromatin

interaction mapping. Gene mapping at 14q32.2 revealed two genes (*BDKRB1, BDKRB2*), which were prioritized by two lines of evidence, eQTL and chromatin interactions. At the 20q11.22 locus, four genes (*EDEM2*, *MYH7B*, *PROCR*, *TRPC4AP*) were prioritized by all three strategies, while another six genes (*ACSS2*, *EIF6*, *FAM83C*, *GGT7*, *MAP1LC3A*, *MMP24*) were supported by the mapping of eQTL effects and chromatin interactions.

4.2.3.3 Fine-mapping analysis

Statistical fine-mapping revealed one 95% credible set at each genome-wide significant locus. The 95% credible sets comprised 15, 40, and 120 SNPs at the 1q24.2, 14q32.2 and 20q11.22 locus, respectively.

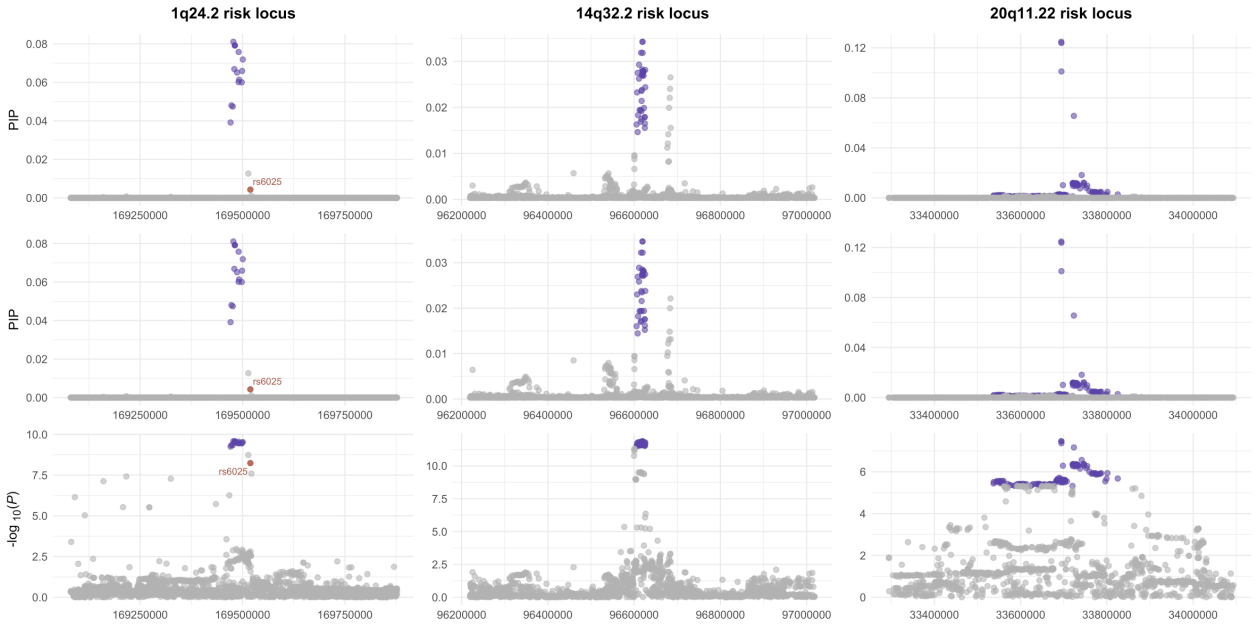

Figure 15 | Fine-mapping results of the genome-wide significant loci.

Results of the fine-mapping analysis using (i) an external UK Biobank sample (upper panels), or (ii) the Danish GWAS cohort (middle panels) as LD reference. In the lower plots the GWAS association results are displayed. SNPs contained in the 95% credible set are highlighted in purple. Abbreviations: GWAS, genome-wide association study | LD, linkage disequilibrium | PIP, posterior inclusion probability | *P*, *p*-value.

Irrespective of the LD reference used – pre-calculated UKB LD scores or LD scores derived from the internal Danish GWAS cohort – the exact same SNPs (with the exception of one SNP at the 20q11.22 locus) were retrieved. Moreover, the posterior inclusion probabilities (PIPs) obtained with either LD reference panel were virtually identical with only subtle differences in decimal places (Figure 15). The retrieved PIPs were in general

rather low. For instance, the maximum PIP was 12.5%, detected at the 20q11.22 locus, while the other two loci showed a maximum PIP of only 8.1% (1q24.2 locus) and 3.5% (14q32.2 locus).

4.2.4 Gene-based tests, gene-set, and tissue enrichment analyses

In aggregating the association results to the level of genes, two genes were found to be significantly associated with ACEi-AE after correction for multiple testing (Figure 16). These were *TMEM119* (P = 7.66x10^{-6}), and *EDEM2* (P = 2.39x10^{-6}).

Moreover, the 50 genes with the lowest p-value (Table S12 in APPENDIX B2) included two genes previously reported in the context of ACEi-AE: *BDKRB2* (P = 8.74x10^{-5}); and *F5* (P = 3.08x10^{-4}). In addition, another potentially interesting gene, namely *KNG1* (P = 1.65x10^{-3}), was identified among the top 50 genes. *KNG1* is known to harbor a pathogenic variant associated with an HAE subtype (HAE-KNG1, see Table 2) and may therefore be a biologically plausible candidate gene for ACEi-AE.

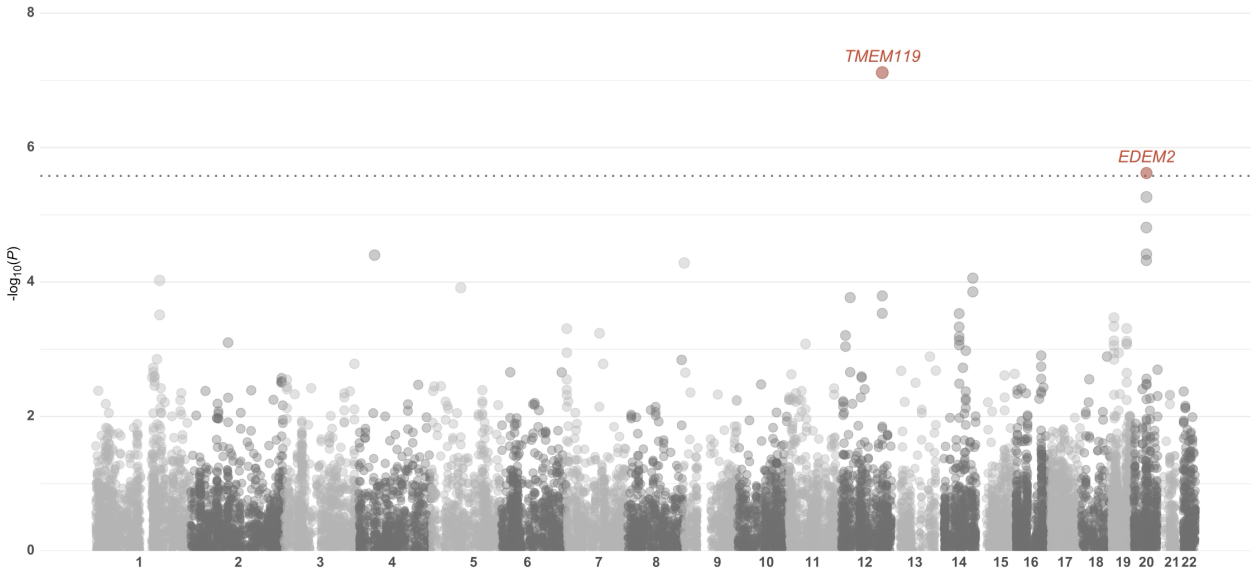

Figure 16 | Gene-based analyses results.

Manhattan plot displaying the -log10 association p-values of the gene-based test (vertical axis) against the genomic position of the respective gene (horizontal axis). The dotted line indicates the Bonferroni corrected threshold for genome-wide significance (P = 2.63x10^{-6}).

The gene-set enrichment analysis did not reveal any significantly associated gene-sets. However, 607 gene-sets with a nominally significant enrichment were identified,

including biologically plausible pathways such as "go_endothelial_cell_activation". Table S13 in APPENDIX B2 lists all gene-set that reached a *p*-value < 10^{-3}.
Similarly, a tissue enrichment analysis of the genes identified by MAGMA only revealed nominally significant associations for eight of the 53 tested GTEx tissue types (Figure S6 in APPENDIX B2).

4.2.5 SNP-based heritability

Using LD score regression (Bulik-Sullivan et al., 2015), the liability-scale SNP-based heritability for ACEi-AE was estimated to range between 4.2% (±2.6%) and 5.2% (±3.2%), taking into account the lower and upper bounds of the population prevalence estimate for ACEi-AE (Aygören-Pürsün et al., 2018).

4.2.6 Genetic correlation analyses

Cross-trait LD score regression (Bulik-Sullivan et al., 2015) revealed no significant genetic correlation between ACEi-AE and any of the four tested previously reported risk or protective factors.

Table 10 | Genetic correlation analyses results.

	$meta_{EUR}$			Stratified $meta_{EUR}$		
Trait	**r_g**	**SE**	**P_{LDSC}**	**r_g**	**SE**	**P_{LDSC}**
Hypertension	0.268	0.122	**0.028**	0.160	0.099	0.107
Asthma	0.419	0.200	**0.036**	0.409	0.214	0.056
Blood clot leg (DVT)	0.028	0.257	0.912	-	-	-
Blood clot lung	0.107	0.267	0.689	-	-	-
Intake of renin-angiotensin agents	0.281	0.127	**0.027**	0.145	0.105	0.168
Coronary artery disease	0.147	0.195	0.453	-	-	-
Hayfever/Allergic rhinitis	0.080	0.162	0.621	-	-	-
Smoking	0.121	0.088	0.166	-	-	-
Type 2 diabetes	0.197	0.110	0.073	-	-	-

Genetic correlation was tested for all nine traits using the $meta_{EUR}$ data ($N_{case}/N_{control}$ = 1,060/77,799). Further, all nominally significant associated traits were re-assessed using a stratified $meta_{EUR}$ analysis comprising only cohorts with treatment-matched controls ($N_{case}/N_{control}$ = 878/71,200). The displayed *p*-values are uncorrected. Nominal significant results are shown in bold font. Abbreviations: DVT, deep vein thrombosis | r_g, genetic correlation | P_{LDSC}, *p*-value obtained from LDSC | SE, standard error.

Among the other five investigated traits nominally significant, positive genetic correlations were found for ACEi-AE and hypertension (r_g = 0.268, SE = 0.122), asthma

(r_g = 0.419, SE = 0.200), and the intake of renin-angiotensin agents (r_g = 0.281, SE = 0.127). However, these nominally significant results did not withstand a re-evaluation using a stratified $meta_{EUR}$ analysis that included only treatment-matched controls (Table 10, Figure S7 in APPENDIX B2). Especially for hypertension and the intake of renin-angiotensin agents, the genetic correlations were substantially lower in the re-analysis. In contrast, asthma showed a comparable correlation using either the $meta_{EUR}$ or stratified $meta_{EUR}$ data. As such, the correlations observed for hypertension and the intake of renin-angiotensin agents might have been confounded by underlying hypertension-related genetic factors given the unscreened controls present in the $meta_{EUR}$ data (~ 8.5% of all controls; see Table 5). In turn, the stable genetic correlation seen for asthma might be reflective of actual shared genetic factors between those two traits.

4.2.7 Explorative cross-ancestry comparison and meta-analysis ($meta_{ALL}$)

The PRS results of the $VanMar_{AFR}$ cohort showed an overall positive signal with a maximum variance explained of 2.1% (Figure 13, right panel). As such, the variance explained by the PRS is comparable to that found in the European cohorts, despite the non-significant prediction of the ACEi-AE case-control status.

The comparable and highly correlated effect size estimates (R = 0.7, Figure 17A) of the $meta_{EUR}$ lead SNPs in the $meta_{EUR}$ and $VanMar_{AFR}$ data, respectively, suggested that the risk loci contribute to ACEi AE risk in both ancestries. Moreover, a near-perfect positive correlation was observed for the lead SNPs effect allele frequencies (R = 0.99, Figure 17B).

The meta-analysis of the $meta_{EUR}$ data and the $VanMar_{AFR}$ cohort ($meta_{ALL}$) totaled 1,123 ACEi-AE patients and 77,948 controls and identified three genome-wide significant loci (Table 11, Figure 18). These three loci were concordant with those identified in the $meta_{EUR}$ analysis, however, two of them were marked by a different lead SNP. The top SNP at the 14q32.2 locus was rs12888576 (P = $3.53x10^{-13}$, OR = 1.50), while the 20q22.11 locus was marked by the SNP rs141521143 (P = $2.32x10^{-8}$, OR = 0.67).

Table 11 | Lead SNPs identified in the meta$_{EUR}$ and meta$_{ALL}$ analysis.

				meta$_{EUR}$							meta$_{ALL}$						
Lead SNP	**Chr**	**Pos**	**A1/A2**	**FreqA1**	**OR**	**95% CI**	**Direction**	***P***	**Het I²**	**Het*P***	**FreqA1**	**OR**	**95% CI**	**Direction**	***P***	**Het I²**	**Het*P***
rs6687813	1	169477575	A/C	0.083	1.70	1.54 – 1.87	++++++++	$2.67x10^{-10}$	0	0.723	0.086	1.70	1.54 – 1.86	+++++++++	$1.03x10^{-10}$	0	0.808
rs35136400°	14	96611391	A/G	0.774	1.50	1.39 – 1.61	+++++++-	$1.28x10^{-12}$	56.5	0.024	0.769	1.47	1.36 – 1.58	+++++++-+	$3.63x10^{-12}$	55.9	0.020
rs12888576°°	14	96619480	A/G	0.775	1.50	1.38 – 1.61	+++++++-	$1.53x10^{-12}$	56.4	0.025	0.776	1.50	1.39 – 1.61	+++++++-+	$3.53x10^{-13}$	50.3	0.041
rs6060237°	20	33694210	A/G	0.855	0.70	0.57 – 0.83	-+-----+	$3.47x10^{-8}$	28.5	0.201	0.854	0.71	0.58 – 0.83	-+-----+-	$5.65x10^{-8}$	25.8	0.215
rs141521143°°	20	33723455	A/AATAAT	0.890	0.66	0.51 – 0.81	-------?	$6.79x10^{-8}$	0	0.838	0.879	0.67	0.52 – 0.81	-------?-	$2.32x10^{-8}$	0	0.904

Note: At the risk locus on chromosome 1, the lead SNP was the same in both analyses. Otherwise, the lead SNPs are marked according to the respective meta-analysis: °meta$_{EUR}$ lead SNP; °°meta$_{ALL}$ lead SNP. The effect directions are denoted as positive (+), negative (-), or SNP not present (?). The order of the GWAS cohorts is as follows; meta$_{EUR}$: vARIANCE, Denmark, Sweden, VanMar$_{EUR}$, Swedegene, CHB-CVDC/DBDS, UKB, EstBB; meta$_{ALL}$: vARIANCE, Denmark, Sweden, VanMar$_{EUR}$, Swedegene, CHB-CVDC/DBDS, UKB, EstBB and VanMar$_{AFR}$. Abbreviations: Chr, chromosome | Pos, genomic position (hg19) | A1/A2, effect allele/other allele | FreqA1, effect allele frequency in the combined case-control cohort | OR, odds ratio | CI, confidence interval | *P*, *p*-value |Het I², heterogeneity | Het*P*, heterogeneity *p*-value.

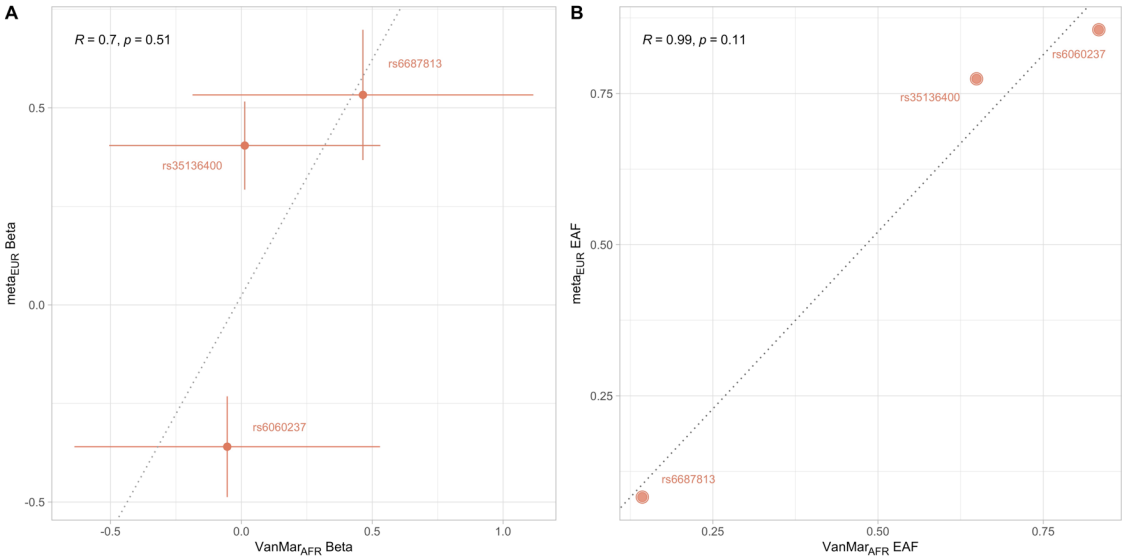

Figure 17 | Comparison of effect estimates (A) and effect allele frequencies (B) between the European ($meta_{EUR}$) and African-American ($VanMar_{AFR}$) data.

Each dot represents a lead SNP that reached genome-wide significance in the $meta_{EUR}$ analysis. Horizontal and vertical lines in the left panel represent 95% confidence intervals of the effect estimates. Abbreviations: EAF, effect allele frequency | *p*, *p*-value | R, Pearson correlation coefficient | SNP, single nucleotide polymorphism.

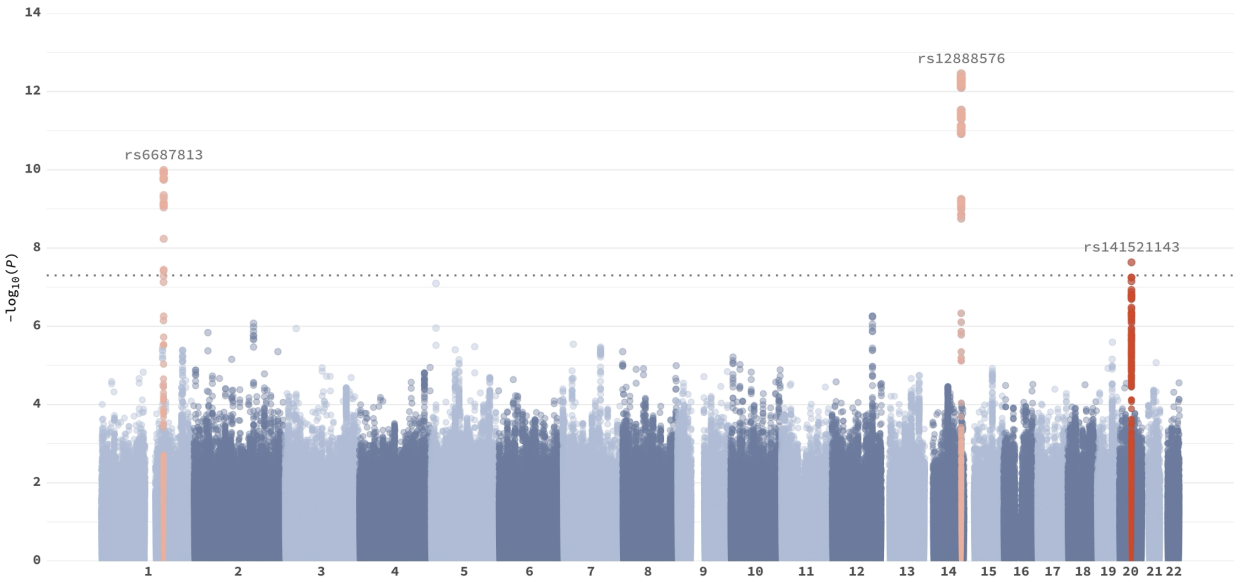

Figure 18 | Manhattan plot of the $meta_{ALL}$ analysis.

Manhattan plot displaying the -log10 association *p*-values (vertical axis) against the genomic position (horizontal axis) of all analyzed SNPs. The threshold for genome-wide significance ($P = 5x10^{-8}$) is indicated by the grey dotted line. Loci that reached genome-wide significance are highlighted (lead SNP ±500 kb) in red (novel locus) and orange (previously reported loci). Abbreviations: kb, kilobases | SNP, single nucleotide polymorphism.

5 Discussion

5.1 The role of HAE-associated genes and variants in ACEi/ARB-AE

ACEi/ARB-AE and HAE are two distinct types of BK-AE that may, however, overlap in certain clinical symptoms. In particular, recently characterized subtypes of nC1-INH-HAE, such as HAE-PLG and HAE-ANGPT1, are often associated with a negative family history and an advanced age of onset and may thus hamper differential diagnosis (Bork et al., 2020). In the absence of any biomarkers to clearly distinguish between nC1-INH-HAE and ACEi/ARB-AE (Kaplan and Maas, 2017; Bindke et al., 2021), genetic testing is nowadays the only reliable way to detect such missed HAE cases, as illustrated by the example of a recent case report of an HAE-FXII case originally diagnosed as ACEi-AE (Veronez et al., 2017). Besides that, case reports indicated ACEi as a trigger factor in the case of a HAE-PLG carrier state (Germenis et al., 2018; Yakushiji et al., 2018), which can result in the inaccurate diagnosis of an ACEi-AE.

To further investigate this, systematic molecular genetic screening of HAE-associated variants in a large cohort of ACEi/ARB-AE patients was performed. Overall, five genes (*SERPING1*, *F12*, *PLG*, *ANGPT1*, and *KNG1*), for which at least one pathogenic variant was known at the time of the present study, were investigated.

Evaluation of known pathogenic HAE variants in ACEi/ARB-AE patients

A systematic evaluation of the identified variants in the HGMD revealed that none of the 197 investigated ACEi/ARB-AE patients carried an HAE-associated variant.

This confirms the findings of a previous, smaller study conducted by Carucci et al. (2020) who screened 33 ACEi-AE patients for known HAE disease-causing variants in *SERPING1*, *ANGPT1*, *F12* (p.T328K only) and *PLG* (p.K330E only) and did not identify any such variants as well. Overall, the present results together with the findings of the previous study suggest that pathogenic HAE variants appear to be, at best, a very rare cause of ACEi/ARB-AE.

Involvement of HAE-associated genes in ACEi/ARB-AE susceptibility

Although the investigated patients were not found to carry a known pathogenic HAE variant, other (rare) variants in these genes might still contribute to the pathophysiology of ACEi/ARB-AE. This hypothesis was further investigated in a case-control setting.

An association analysis at the individual variant level did not reveal any variants that showed a significant difference between patients and controls. This was true for both, rare (MAF < 0.05) and common (MAF ≥ 0.05) variants. As such the present results seem to be in line with those of a previous exome sequencing study, which did not report variants in the genes studied here among the top 20 associated ones (Maroteau et al., 2020). Yet, given the limited sample size of the present study the absence of any statistically significant findings might very well be attributable to a lack in statistical power (see limitations of the candidate gene analysis).

Overall, a nominally significant enrichment for rare, potentially functional variants (MAF < 0.05, CADD ≥ 10) was found in controls compared to patients. This was true in a combined enrichment analysis across all five genes and at the single-gene level for *F12*. Although neither of these associations withstood a multiple testing correction, they might still indicate the existence of protective variants in the context of ACEi/ARB-AE.

Additional evidence for the role of protective effects might be provided by a missense variant in *PLG* that showed the lowest *p*-value in the single variant association tests. The variant (rs4252129; NM_00301.5:c.1567C>T; p.R523W) was identified exclusively in controls. Similar to the HAE-associated *PLG* variant (p.K330E), p.R523W is located in a kringle domain of *PLG*, which facilitates the binding of plasminogen to large surfaces, such as fibrinogen, while additionally contributing to the arrangement and thus protein function of plasminogen (Castellino and Ploplis, 2005). In a previous GWAS, rs4252129 has been associated with altered plasma plasminogen levels in as much as the variant's minor T allele was associated with a decrease in plasminogen levels (average decrease of 14.6% per T allele; Ma et al., 2014). Plasminogen is the precursor of the protein plasmin, which is involved in the activation of the kinin-kallikrein system and subsequently also plays a role in the formation of BK (Figure 2; De Maat et al., 2018). Therefore, in view of the results of the above-mentioned GWAS, a plausible hypothesis in the context of ACEi/ARB-AE might be that individuals carrying this particular variant are protected from BK accumulation during ACEi/ARB treatment because their BK levels are naturally lower due to lower plasminogen levels. It must be noted, however, that aside from the non-significant results for this variant, the present study was additionally limited by the use of unscreened controls (see limitations of the candidate gene analysis).

Therefore, further studies are warranted to investigate the variant's presence in treatment-matched controls.

Interestingly, the top five identified variants also included a stopgain variant in *KNG1* (rs76438938; NM_00893:c.1234C>T; p.R412X). The gene *KNG1* encodes kininogen, the precursor for HK and low molecular weight kininogen (LK) which are produced *via* alternative splicing. HK and LK share the first four N-terminal domains, encoding the heavy chains (exon 1-9) and the 9 bp BK sequence (parts exon 10). The remaining bp of exon 10 encode the unique light chain of HK which is critical for the "contact activation" properties of HK, whereas the unique light chain of LK is encode by exon 11 (Kaplan et al., 2022). The HAE-associated *KNG1* variant (p.M379K) is located in exon 10 of the gene and directly affects the N-terminal cleavage sites of BK located between the amino acids p.Lys380/Arg381 (Bork et al., 2019). In contrast, the here identified variant p.R412X is located within exon 11 and thus affects the light chain of LK, whose exact function is not yet fully understood (Colman and Schmaier, 1997). Any possible effects of p.R412X on protein function are therefore not readily apparent and would require further functional studies.

Limitations of the candidate gene analysis

Overall, the candidate gene analysis had four major limitations.

First and foremost, the sample size of the study was limited. Although it represents the largest study to systematically investigate HAE-associated genes in an ACEi/ARB-AE cohort to date, it was underpowered in the detection of statistically significant associations for variants with small to moderate effects. As illustrated in Figure 10, a substantially larger sample size would have been required to detect such variants.

Second, the five genes studied here do not explain the entire HAE phenotype. In fact, pathogenic variants in two additional genes (*MYOF*, (Ariano et al., 2020) and *HS3ST6*, (Bork et al., 2021)) have been associated with two new nC1-INH-HAE subtypes during the course of this study. Besides that, there is a significant proportion of HAE patients for whom the underlying genetic cause has not yet been identified (HAE-UNK, see Figure 3). In addition, the present study solely focused on exonic and flanking (±6 bp) regions. Thus, deep intronic or regulatory variants, which have recently been reported for the *SERPING1* gene (Hujová et al., 2020; Ponard et al., 2020; Vatsiou et al., 2020), would

have been missed. Besides the genes harboring disease-causing variants, variants in, for example, *KLKB1* have been shown to influence certain aspects of the phenotypic variation observed in HAE (Gianni et al., 2017). While such disease-modifying genes might be interesting candidate genes with regard to ACEi/ARB-AE susceptibility, their analysis was beyond the scope of the present investigation.

Third, an unscreened and nontreatment-matched control cohort was used. Considering that the 12-month population prevalence of ACEi/ARB-AE was estimated to be less than 1% (Aygören-Pürsün et al., 2018), the use of unscreened controls should not have had a major impact on the overall power of the present analyses (Moskvina et al., 2005). Nonetheless, the possibility that individuals in the control cohort may be susceptible to angioedema when taking an ACEi or ARB cannot be completely ruled out. In addition, as noted above, the use of nontreatment-matched controls hampers the interpretation of certain results, e.g., with regard to protective effects.

Fourth, the sequencing data of the present analysis were not comprehensive enough to investigate possible population stratification within the generated data itself. For this reason, the corresponding analyses were based on available genome-wide genotyping data and thus common variants. However, this approach may not have fully accounted for all ancestry-related differences in the study cohort, as stratification patterns may differ between common and rare variants (Mathieson and McVean, 2012).

Pathogenic HAE variants in ACEi/ARB-AE are rare; findings beyond need further follow-up

In summary, targeted genetic screening of a large ACEi/ARB-AE cohort (N = 197 patients and 346 controls) revealed no patients who carried a known pathogenic HAE-associated variant in *SERPING1*, *F12*, *PLG*, *ANGPT1* or *KNG1*. Although described in the literature (Veronez et al., 2017), the diagnosis of an HAE as an ACEi/ARB-AE seems to be an extremely rare event, at best. Moreover, none of the identified individual variants – common or rare – showed a significant association with ACEi/ARB-AE.

However, a missense variant in *PLG* (p.R523W) identified only in controls might be an interesting candidate variant as it has been associated with decreased plasma plasminogen levels in an earlier study (Ma et al., 2014) and thus may potentially indicate a protective effect against ACEi/ARB-AE. Future studies should therefore encompass

larger sample sizes to increase statistical power and ideally include screened and treatment-matched controls to allow definite conclusions about the role of rare and in particular protective variants in HAE genes.

5.2 The role of common variants in ACEi-AE

By 2022, two genome-wide significant risk loci for ACEi-AE had been identified (Rasmussen et al., 2020; Ghouse et al., 2021). Moreover, one exome-wide significant risk locus marked by the common Factor V Leiden mutation was reported (Maroteau et al., 2020). However, only one of these loci could be replicated so far (Ghouse et al., 2021) and the overall understanding of the pathophysiology of ACEi-AE is still limited.

To gain further insights into the underlying genetics of ACEi-AE a comprehensive GWAS meta-analysis was conducted within the scope of the present dissertation. By combining GWAS data from eight independent European cohorts, totaling for the first time over 1,000 ACEi-AE patients, the present meta-analysis identified three genome-wide significant risk loci (Figure 11). These loci included a novel risk locus located on chromosome 20q11.22 (Figure 12C), and two previously reported loci on chromosomes 1q24.2 (Figure 12A) and 14q32.2 (Figure 12B), respectively.

Genetic factors implicated in ACEi-AE pathophysiology

In recent years, research has shown that dysregulation in endothelial cell permeability is a key factor in the formation of BK-AE (Debreczeni et al., 2021). While BK itself is involved in the regulation of vascular permeability and has been proven to be elevated during the intake of an ACEi (Nussberger et al., 2002; Nussberger et al., 1998), its increased plasma levels cannot be the sole cause for the development of an ACEi-AE. This is evident from the fact that not all patients taking these drugs are affected. Rather it seems plausible that other, additional factors – presumably also involved in the regulation of endothelial cell permeability – are contributing to ACEi-AE susceptibility. The risk loci identified in the present meta-analysis now provide (further) insights into these possible factors.

The strongest association, i.e., the lowest *p*-value, was identified on chromosome 14q32.2, a risk locus that has already been described for ACEi-AE (Ghouse et al., 2021). All SNPs at this locus were located in non-coding regions mapping upstream of the genes encoding the B2-R and B1-R. Based on two lines of evidence (eQTL and chromatin interactions, Figure 14), it seems plausible that regulatory effects involving the genes

BDKRB2 or *BDKRB1* are what's underlying the association signal at this locus. This is in line with the findings of the previous GWAS, in which the authors concluded that variation at this locus may lead to increased BK sensitivity and/or increased expression of B2-R (Ghouse et al., 2021).

The second most significant association was found at the 1q24.2 locus, which corresponds to the *F5* locus that was previously reported with exome-wide significance by Maroteau et al. (2020). In the present study this locus has now been identified as genome-wide significant ($P < 2.67x10^{-10}$), thereby replicating its association with ACEi-AE. While the results of the exome sequencing study suggested an association of common and rare variants within *F5* and in particular the Factor V Leiden mutation, the results of the present meta-analysis are not as conclusive. Although the Factor V Leiden mutation was retrieved among the genome-wide significant SNPs at the locus (rs6025, $P = 5.81x10^{-9}$, OR = 1.97), statistical fine-mapping of the data did not identify rs6025 within the 95% credible SNP set of likely causal SNPs (Figure 15). It must be noted, however, that the fine-mapping analyses in general did not reveal any clearly prioritized variants (PIP > 0.5), probably due to insufficient statistical power or the unavailability of an ethnically more appropriate LD reference panel. As such, the results must be interpreted with caution. Similarly, the *F5* gene was prioritized as a likely candidate gene at the locus, but other genes were supported with similar or even higher levels of evidence (Figure 14). Among those highly prioritized genes, *NME7*, *SLC19A2*, and *ATP1B1* were retrieved. Variation in all three genes has been associated with venous thromboembolism (VTE) as well. In addition, the genes have been shown to form a haploblock with Factor V Leiden/the *F5* gene (Heit et al., 2012), complicating the biological interpretation of the GWAS signal identified at this locus.

Lastly, the locus identified at chromosome 20q11.22 represents a novel locus which has not yet been reported in relation to ACEi-AE. Notably, 20q11.22 is – like the *F5* locus – a known risk locus for VTE (Lindström et al., 2019). In the present analysis, the locus encompassed several SNPs in high LD with the lead SNP, and spanned several genes. Based on the integration of additional data layers four highly prioritized genes emerged: *EDEM2*, *PROCR*, *MYH7B* and *TRPC4AP*. From a biological point of view, *PROCR* was the most plausible candidate gene among them. The gene *PROCR* encodes the endothelial protein C receptor (EPCR), a key component in the protein C pathway, inasmuch as its

membrane-bound form significantly enhances the activation of protein C (Stearns-Kurosawa et al., 1996). Activated protein C, in turn, is involved not only in anticoagulation/fibrinolysis (Dahlbäck and Villoutreix, 2005) but also in the stabilization of the endothelial barrier. The latter is presumably mediated by the angiopoietin-1/TIE-2 axis (Minhas et al., 2010, 2017), which, when disrupted by specific pathogenic variants, has already been shown to cause another type of BK-AE, namely HAE-ANGPT1 (Bafunno et al., 2018; Cordisco et al., 2019). The effect of variants in or near *PROCR* on protein C levels has been demonstrated by several GWAS (Tang et al., 2010; Athanasiadis et al., 2011; Oudot-Mellakh et al., 2012). Notably, among the candidate variants at the locus a well-known coding variant in *PROCR* was identified (p.Ser219Gly, rs867186, $P < 1.22x10^{-6}$). In the literature, this variant has been reported as a risk variant for VTE (Dennis et al., 2012; Medina et al., 2014), while it has further been shown to be associated with elevated levels of soluble EPCR (Reiner et al., 2008; Pintao et al., 2011), the latter being associated with an attenuated protein C activation (Kurosawa et al., 1997). To add to this, another gene at the locus, *EDEM2*, has been reported as associated with altered protein C levels (Tang et al., 2010). *EDEM2* was both among the most highly prioritized genes at the 20q11.22 locus and further identified as significantly associated in the gene-based tests. Assuming that a disrupted endothelial cell barrier may be a trigger of ACEi-AE, a reasonable hypothesis might be that variation at the 20q11.22 locus impairs *PROCR* and/or *EDEM2*, leading to decreased protein C activation and downstream loss of endothelial integrity, thereby ultimately promoting the development of angioedema.

Heritability of ACEi-AE and consistency of the phenotype

Based on the data from the previous GWAS by Ghouse et al. (2021) the SNP-based heritability of ACEi-AE was estimated to be approximately 20%. The common variants investigated in the present study, however, explained only 4.2 – 5.2% of the phenotypic variance of ACEi-AE. This considerable difference can possibly be explained by two reasons. First, different methods were used to obtain the heritability estimates. The previous estimate was derived from genotype-level data (Ghouse et al., 2021). Given that individual-level genotypes were not available for all individuals of the present meta-analysis, summary-level data and the tool LDSC (Bulik-Sullivan et al., 2015) were used

to derive the SNP-based heritability estimate. However, based on heritability estimates derived for thousands of phenotypes using the UKB dataset, it was found that an effective sample size of at least 10,000 is required for LDSC to obtain reasonably reliable heritability estimates (rkwalters and Palmer, 2022). The rather low estimate obtained with the data of the present analysis may therefore possibly be the result of the small effective sample size of the GWAS meta-analysis, which was only 4,381. Second, given the setting of a meta-analysis, the present heritability estimate might have been influenced by slight differences in the ACEi-AE phenotype definitions, resulting in a more heterogeneous phenotype overall. This phenomenon – decreasing estimates of heritability in spite of larger samples in a meta-analysis – has been observed previously in, for example, psychiatric phenotypes (Anttila et al., 2018).

As indicated by the positive signal and significant prediction of ACEi-AE case-control status in the PRS analyses (Figure 13), however, the phenotype definitions appear to be generally quite comparable. At least this is true for the five GWAS cohorts that could be investigated in the PRS analyses. For the individuals in the Swedegene, CHB-CVDC/DBDS, and EstBB cohorts, no conclusions can be drawn on a genetic basis regarding their phenotype definitions.

Of note, the variance explained by the PRS in the UKB cohort, which was derived solely on the basis of data on ICD-10 codes and self-reported medication use, was comparable to that of other cohorts, which included clinically assessed patient cohorts. This is an encouraging finding given the increasing availability of biobank-based patient data, which are generally a powerful tool for the enlargement of genetic data sets. However, one should keep in mind that the use of biobank-derived data may come at the expense of "minimal phenotyping", potentially leading to the identification of non-specific genetic factors (Cai et al., 2020).

ACEi-AE and its genetic correlation with other traits

Epidemiological studies have linked ACEi-AE risk to factors such as smoking, a history of hay fever/allergic rhinitis, a concomitant coronary artery disease or diabetes (see Section 2.2.2). The results of the present genetic correlation analyses indicate that these observed relationships are most likely not due to shared genetic factors.

Additionally, the evaluation of the genetic correlation between ACEi-AE and an additional five selected traits from the spectrum of cardiovascular, allergic, and blood coagulation traits as well as medication use revealed only nominally significant genetic correlations. Furthermore, these nominally significant results were not stable in a re-analysis using only treatment-matched controls. This suggests that, in general, the results based on the $meta_{EUR}$ data may have been confounded by genetic effects attributable to factors associated with hypertension. As such, future analyses should be based on GWAS including only treatment-matched controls to obtain more robust results in this regard. Interestingly, asthma showed an almost equal genetic correlation in both analyses (41.9% and 40.9%). This could indicate a true genetic overlap, maybe due to BK-related factors implicated in the pathophysiology of both ACEi-AE and asthma (Ricciardolo et al., 2018). Future, larger studies will allow more definitive conclusions in this regard.
Overall, the present analysis was the first to systematically investigate genetic factors that ACEi-AE might share with other phenotypically associated traits and diseases. Although the present results did not reveal significant correlations between ACEi-AE and the nine traits/diseases studied, the existence of single overlapping pathways cannot be excluded. Given the large standard errors, it may be that the obtained results reflect, at least to some extent, the limited power of the LDSC instrument given the small sample sizes (Bulik-Sullivan et al., 2015), and not a general absence of shared genetic factors. Future studies should therefore ideally encompass larger samples and also examine other traits/diseases to eventually obtain a comprehensive picture of possible overlapping genetic factors of ACEi-AE and other traits/diseases.

Cross-population effects of ACEi-AE associated risk loci

Although meaningful GWAS in diverse populations are becoming more frequent (e.g., Shrine et al., 2023; Tcheandjieu et al., 2022), the field of GWAS is still largely dominated by an abundance of European-based studies. That's despite the fact that studies involving individuals from multiple ancestries have proven to be very valuable, e.g., in terms of (i) an increased power to detect causal alleles that harbor great differences in frequencies across populations, (ii) the advantage of leveraging distinct LD patterns in fine-mapping of risk loci, and (iii) the possibility to assess PRS based on effect sizes from different ancestries thereby improving its predictive power (Martin et al., 2019).

The present GWAS meta-analysis combined all ACEi-AE GWAS cohorts known at the time. Consistent with the general picture seen in GWAS, the vast majority of these cohorts were of European ancestry, while only one smaller cohort of African-American ancestry was available. This cohort was used to perform an exploratory cross-ancestry comparison and meta-analysis ($meta_{ALL}$) to gain initial insights into generalizability of the identified ACEi-AE risk loci as well as the general contribution of shared common variants across different ancestries.

The identified lead SNPs at the three risk loci showed matching effect directions (Table 11) and comparable effect sizes (Figure 17A), thereby suggesting that the risk loci identified in the European cohorts contribute to ACEi-AE susceptibility in African-American patients as well. Moreover, the positive and comparable polygenic signal that was observed for the $VanMar_{AFR}$ cohort in the PRS analysis (Figure 13) indicated the presence of further shared common variants beyond the already genome-wide significant loci. The subsequent cross-ancestry meta-analysis did not reveal any additional identified risk loci; however, it further supported the three loci identified in the $meta_{EUR}$ analysis. In this regard, the inability to identify any additional identified risk loci might very well be due to the relatively small sample size of the $VanMar_{AFR}$ cohort (Ncase/Ncontrol = 63/149).

Limitations of the GWAS meta-analysis

The present GWAS meta-analysis bears three major limitations.

First, although this was the largest analysis to date to examine common variants associated with ACEi-AE, it had limited power to discover additional risk loci. In particular, risk loci with small effect sizes. Similarly, the non-significant results observed in some of the follow-up analyses, e.g., the pathway-based and genetic correlation analyses, are probably reflective of the still relatively small sample size of the $meta_{EUR}$ analysis.

Second, the functional relevance of the presented findings was only supported by bioinformatic evidence. In order to determine the exact biological function and underlying mechanisms of the identified risk loci *in vitro* studies in relevant cell types, e.g., endothelial cells, would have been required.

Third, the present meta-analysis focused exclusively on ACEi-AE. As such, the results do not equally extend to ARB-AE. In fact, it is controversial whether ACEi-AE and ARB-AE

can or should be considered and studied as one or rather two separate phenotypes. In recent genetic studies, one finds studies examining ACEi-AE alone (Pare et al., 2013; Ghouse et al., 2021) as well as those examining ACEi-AE and ARB-AE together (Maroteau et al., 2020; Rasmussen et al., 2020). Based on personal communication with Jonas Ghouse, who observed an attenuation of his GWAS results in the presence of ARB-AE cases, and based on the fact that several GWAS studies included in the present meta-analysis were already limited to ACEi-AE, the present analysis was determined to focus solely on ACEi-AE. Provided that larger patient cohorts with an ideally reasonably even distribution of ACEi-AE and ARB-AE cases will become available in the future, stratified analyses are feasible and desirable to clarify whether and to what extent those two types of angioedema share a common genetic basis.

GWAS meta-analysis provides further insights into ACEi-AE pathophysiology; functional follow-up studies are warranted

To the extent of our current knowledge, the present GWAS meta-analysis represents the largest analysis of common variants in ACEi-AE to date. By investigating more than 1,000 ACEi-AE patients, three genome-wide significant risk loci could be identified. While two of the loci represent previously associated loci, the present analysis was the first to identify 20q11.22 as a risk locus for ACEi-AE. Overall, the present results contributed to a deeper understanding of the pathophysiology underlying ACEi-AE by revealing novel, biologically plausible candidate genes (*PROCR*, *EDEM2*), implicating for the first time the involvement of the fibrinolysis pathway in the development of this ADR. The retrieval of previously reported candidate genes (*BDKRB2, F5*) further underscored the implicated role of bradykinin signaling and the coagulation pathway in ACEi-AE. Interestingly, two of the three risk loci coincide with known VTE risk loci, a finding that warrants further investigation using, for example, PRS analyses. Furthermore, the cross-ancestry analyses provided some initial evidence that Europeans and African-Americans share common variants involved in ACEi-AE susceptibility. In particular, the analyses suggested that the three risk loci identified in the European individuals also contribute to ACEi-AE risk in individuals of African-American ancestry. At this point, functional studies need to follow to further elucidate the molecular mechanisms underlying the identified risk loci.

5.3 Conclusion and future perspectives

The work and studies conducted within the scope of the present dissertation contributed significantly to the field of ACEi/ARB-AE research.

First of all, in a major collaborative effort (Figure 5) the vARIANCE cohort, the largest German/Austrian cohort of ACEi/ARB-AE patients to date (Table 3), has been established. This cohort encompasses not only genetic material and data but also comprehensive phenotypic data, making it a valuable resource for any future studies to be conducted in this field.

Prior to the start of this work, susceptibility to ACEi/ARB-AE was already considered to depend on a genetic predisposition alongside non-genetic risk factors. However, little was known about the actual genetic factors involved in ACEi-AE pathophysiology, and the few known associations were derived from candidate gene or candidate SNP studies, the majority of which included only rather small sample sizes. Moreover, the degree to which ACEi/ARB-AE and HAE, another type of BK-AE with a strong genetic etiology, may share a common genetic basis had not been systematically investigated at that time.

In performing systematic targeted re-sequencing of five genes harboring pathogenic HAE-associated variants, the present work did not identify any significantly associated variants – known or novel. Interestingly, however, *KNG1*, a gene sequenced in the candidate gene analyses, was identified as one of the top 50 genes in the gene-based analyses. Thus, for now it cannot be completely ruled out that the genetic factors underlying ACEi/ARB-AE and HAE converge, at least to some extent, in the same genes.

While the studies included in this dissertation were being conducted, other large-scale studies led to the identification of two genome- and one exome-wide significant risk locus for ACEi/ARB-AE (Maroteau et al., 2020; Rasmussen et al., 2020; Ghouse et al., 2021), providing initial insights into the genetics underlying this ADR. Through the identification of another genome-wide significant risk locus the present GWAS meta-analysis now not only provides further profound insights into the biology/pathways possibly involved in ACEi-AE susceptibility but also emphasized its multifactorial etiology. In addition, initial findings obtained in the present study revealed for the first time the presence of common variants associated with ACEi-AE risk that are shared across different ancestral groups.

Nevertheless, the current (genetics) knowledge regarding ACEi/ARB-AE is merely a first step on a long road to ultimately be able to predict whether or not a patient taking an

ACEi or ARB will develop an angioedema. To continue down this path, additional genetic factors associated with ACEi/ARB-AE alongside gene x environment and eventually gene x gene interactions need to be explored to obtain a comprehensive picture of the pathophysiology underlying ACEi/ARB-AE. Future studies will be able to address this with different strategies.

Of utmost importance for all these future studies will be the expansion of the cohorts under study. Face-to-face recruitment of patients within the clinical setting and the use of comprehensive questionnaires will allow for an in-depth phenotyping of patients. In particular, such cohorts will enable the identification of additional environmental risk factors alongside the investigation of specific sub-phenotypic aspects, e.g., severity and localization of the angioedema. Furthermore, they are very well suited for the investigation of gene x environment interactions. However, such recruitment approaches tend to be quite tedious and time-consuming, making it very difficult to achieve a statistically meaningful number of patients within a reasonable timeframe. This is well exemplified by the recruitment statistics of our in-house vARIANCE cohort, which despite the joint efforts of several physicians only comprises about 150 patients to date. Additional innovative recruitment approaches, such as the online recruitment established for the vARIANCE study, can be very effective in increasing recruitment numbers but require consistent follow-up, e.g., through explicit advertising (Davies et al., 2019). Besides, and as demonstrated by the $meta_{EUR}$ analysis, the continued use of available biobank data and the aggregation of the ever-increasing individual genomic data sets will greatly contribute to the enlargement of future study collectives. However, it is important to keep in mind that cohorts obtained through online recruitment and biobank data may come at the expense of "minimal phenotyping" and may thus lead to the identification of non-specific genetic factors (Cai et al., 2020).

In GWAS, the continuous expansion of the study cohorts will allow the identification of further associated loci, in particular those with small effect sizes, alongside the uncovering of the most relevant biological pathways involved. Here, another key aspect will be the investigation of diverse populations as it will not only advance the discovery and elucidation of risk loci but will further allow the investigation of population-specific genetic factors (Liu et al., 2015). The latter is particularly interesting in light of the increased risk for ACEi/ARB-AE that has been reported for African individuals in

epidemiological studies (Brown et al., 1996; Kostis et al., 2005; Miller et al., 2008; Banerji et al., 2017; Reichman et al., 2017). In addition to the identification of novel genetic risk factors, the already known associated loci will have to be further elucidated. Here, advanced statistical methods and combination of GWAS data with, among others, eQTL, transcriptomics and/or epigenomics data are promising approaches to gain a deeper understanding of the biological mechanisms at play (Akiyama, 2021). In addition, the integration of such data sets can help to refine the number of likely causal variants/genes at a given locus which can then be more easily followed-up in functional studies. The use of even more sophisticated datasets, such as eQTLs derived from single-cell rather than bulk RNA sequencing data, promises to eventually provide an even more detailed picture, e.g., in terms of single cell types or cellular processes relevant in ACEi/ARB-AE (Jagadeesh et al., 2022; Zhang et al., 2022).

To date, the investigation of rare sequence variants in the context of ACEi/ARB-AE is still in its infancy, with only one published WES study (Maroteau et al., 2020). Thus, to complement the GWAS findings, further studies encompassing the whole exome or even better whole genome are required to identify rare variants associated with ACEi/ARB-AE risk in a systematic manner.

Finally, to unambiguously determine the actual causal variants, genes, regulatory elements, or molecular mechanisms underlying the associated risk loci functional studies are required. To this end, a wide variety of methods, such as CRISPR screens and analyses using animal models or induced pluripotent stem cells, are nowadays available (Rao et al., 2021; Bock et al., 2022).

Additional insights into the underlying biology of ACEi/ARB-AE can be gained by a more comprehensive screening of genetic factors that ACEi/ARB-AE may share with other traits or diseases. Upon robust genetic correlation results, subsequent application of innovative methods such as summary-based Mendelian randomization will be possible and eventually provide deeper insights into the causal relationships underlying these shared genetic factors (Benn and Nordestgaard, 2018).

Provided the availability of large enough sample sizes, genetic correlation analyses could also be used to quantify the genetic overlap between ACEi-AE and ARB-AE. This can be a potential first step in the clarification of whether or to what extent these two types of drug-induced angioedema actually share a common pathophysiology and thus should

be studied as one or rather two separate entities in future (genetic) studies (see limitations of the GWAS meta-analysis).

One of the key considerations in ACEi/ARB-AE is the individual susceptibility and the extent to which genetic and non-genetic risk factors contribute to the development of this adverse effect. Here, the application of PRS may be a means to identify patients at increased risk. At present, only a very small portion of the observed phenotypic variance in ACEi/ARB-AE is explained by the PRS. However, with the integration of additional (common) variants, rare variants, and non-genetic risk factors, the predictive power of PRS is likely to increase, possibly making it a useful screening tool in the near future (Lewis & Vassos, 2020). In the long term, the further identification of the molecular genetic factors involved, together with functional studies, will unravel the precise mechanisms underlying ACEi/ARB-AE susceptibility. Ideally, this will ultimately result in the identification of molecular targets and pave the way for the development of novel prevention or intervention strategies for ACEi/ARB-AE.

6 List of figures

7 List of tables

8 References

Akiyama, M. (2021). Multi-omics study for interpretation of genome-wide association study. *J Hum Genet* 66, 3–10. doi: 10.1038/s10038-020-00842-5.

Amendola, L. M., Dorschner, M. O., Robertson, P. D., Salama, J. S., Hart, R., Shirts, B. H., et al. (2015). Actionable exomic incidental findings in 6503 participants: Challenges of variant classification. *Genome Res* 25, 305–315. doi: 10.1101/gr.183483.114.

Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., Duncan, L., et al. (2018). Analysis of shared heritability in common disorders of the brain. *Science (1979)* 360. doi: 10.1126/science.aap8757.

Ariano, A., D'Apolito, M., Bova, M., Bellanti, F., Loffredo, S., D'Andrea, G., et al. (2020). A myoferlin gain-of-function variant associates with a new type of hereditary angioedema. *Allergy: European Journal of Allergy and Clinical Immunology* 75, 2989–2992. doi: 10.1111/all.14454.

Athanasiadis, G., Buil, A., Souto, J. C., Borrell, M., López, S., Martinez-Perez, A., et al. (2011). A genome-wide association study of the protein C anticoagulant pathway. *PLoS One* 6. doi: 10.1371/journal.pone.0029168.

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., et al. (2015). A global reference for human genetic variation. *Nature* 526, 68–74. doi: 10.1038/nature15393.

Aygören-Pürsün, E., Magerl, M., Maetzel, A., and Maurer, M. (2018). Epidemiology of Bradykinin-mediated angioedema: A systematic investigation of epidemiological studies. *Orphanet J Rare Dis*. doi: 10.1186/s13023-018-0815-5.

Baffert, F., Le, T., Thurston, G., McDonald, D. M., and McDonald Angiopoietin-, D. M. (2006). Angiopoietin-1 decreases plasma leakage by reducing number and size of endothelial gaps in venules. *Am J Physiol Heart Circ Physiol* 290, 107–118. doi: 10.1152/ajpheart.00542.2005.-Angiopoietin-1.

Bafunno, V., Firinu, D., D'Apolito, M., Cordisco, G., Loffredo, S., Leccese, A., et al. (2018). Mutation of the angiopoietin-1 gene (ANGPT1) associates with a new type of hereditary angioedema. *Journal of Allergy and Clinical Immunology*. doi: 10.1016/j.jaci.2017.05.020.

Banerji, A., Blumenthal, K. G., Lai, K. H., and Zhou, L. (2017). Epidemiology of ACE Inhibitor Angioedema Utilizing a Large Electronic Health Record. *Journal of Allergy and Clinical Immunology: In Practice* 5, 744–749. doi: 10.1016/j.jaip.2017.02.018.

Banerji, A., Clark, S., Blanda, M., LoVecchio, F., Snyder, B., and Camargo, C. A. (2008). Multicenter study of patients with angiotensin-converting enzyme inhibitor-induced angioedema who present to the emergency department. *Annals of Allergy, Asthma and Immunology* 100, 327–332. doi: 10.1016/S1081-1206(10)60594-7.

Bas, M. (2017). The Angiotensin-Converting-Enzyme-Induced Angioedema. *Immunol Allergy Clin North Am* 37, 183–200. doi: 10.1016/j.iac.2016.08.011.

Bas, M., Adams, V., Suvorava, T., Niehues, T., Hoffmann, T. K., and Kojda, G. (2007). Nonallergic angioedema: Role of bradykinin. *Allergy: European Journal of Allergy and Clinical Immunology* 62, 842–856. doi: 10.1111/j.1398-9995.2007.01427.x.

Bas, M., Greve, J., Strassen, U., Khosravani, F., Hoffmann, T. K., and Kojda, G. (2015). Angioedema induced by cardiovascular drugs: New players join old friends. *Allergy: European Journal of Allergy and Clinical Immunology* 70, 1196–1200. doi: 10.1111/all.12680.

Bas, M., Hoffmann, T. K., Tiemann, B., Dao, V. T. V., Bantis, C., Balz, V., et al. (2010). Potential genetic risk factors in angiotensin-converting enzyme-inhibitor-induced angio-oedema. *Br J Clin Pharmacol* 69, 179–186. doi: 10.1111/j.1365-2125.2009.03567.x.

Benn, M., and Nordestgaard, B. G. (2018). From genome-wide association studies to Mendelian randomization: Novel opportunities for understanding cardiovascular disease causality, pathogenesis, prevention, and treatment. *Cardiovasc Res* 114, 1192–1208. doi: 10.1093/cvr/cvy045.

Benson, B. C., Smith, C., and Laczek, J. T. (2013). Angiotensin Converting Enzyme Inhibitor-induced Gastrointestinal Angioedema A Case Series and Literature Review. Available at: www.jcge.com.

Bindke, G., Gehring, M., Wieczorek, D., Kapp, A., Buhl, T., and Wedi, B. (2021). Identification of novel biomarkers to distinguish bradykinin-mediated angioedema from mast cell-/histamine-mediated angioedema. *Allergy: European Journal of Allergy and Clinical Immunology*. doi: 10.1111/all.15013.

Binkley, K. E. (2010). Factor XII mutations, estrogen-dependent inherited angioedema, and related conditions. *Allergy, Asthma & Clinical Immunology* 6. doi: 10.1186/1710-1492-6-16.

Birnbaum, S., Ludwig, K. U., Reutter, H., Herms, S., Steffens, M., Rubini, M., et al. (2009). Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nat Genet* 41, 473–477. doi: 10.1038/ng.333.

Blaes, N., and Girolami, J. P. (2013). Targeting the "Janus face" of the B2-bradykinin receptor. *Expert Opin Ther Targets* 17, 1145–1166. doi: 10.1517/14728222.2013.827664.

Blais, C., Marceau, F., Rouleau, J.-L., and Adam, A. (2000). The kallikrein-kininogen-kinin system: lessons from the quantification of endogenous kinins.

Bock, C., Datlinger, P., Chardon, F., Coelho, M. A., Dong, M. B., Lawson, K. A., et al. (2022). High-content CRISPR screening. *Nature Reviews Methods Primers* 2. doi: 10.1038/s43586-021-00093-4.

Bork, K., Machnig, T., Wulff, K., Witzke, G., Prusty, S., and Hardt, J. (2020). Clinical features of genetically characterized types of hereditary angioedema with normal C1 inhibitor: a systematic review of qualitative evidence. *Orphanet J Rare Dis* 15. doi: 10.1186/s13023-020-01570-x.

Bork, K., Wulff, K., Hardt, J., Witzke, G., and Lohse, P. (2014). Characterization of a partial exon 9/intron 9 deletion in the coagulation factor XII gene (F12) detected in two Turkish families with hereditary angioedema and normal C1 inhibitor. *Haemophilia* 20, e372–e375. doi: 10.1111/hae.12519.

Bork, K., Wulff, K., Möhl, B. S., Steinmüller-Magin, L., Witzke, G., Hardt, J., et al. (2021). Novel hereditary angioedema linked with a heparan sulfate 3-O-sulfotransferase 6 gene mutation. *Journal of Allergy and Clinical Immunology* 148, 1041–1048. doi: 10.1016/j.jaci.2021.01.011.

Bork, K., Wulff, K., Rossmann, H., Steinmüller-Magin, L., Brænne, I., Witzke, G., et al. (2019). Hereditary angioedema cosegregating with a novel kininogen1 gene mutation changing the N-terminal cleavage site of bradykinin. *Allergy*, all.13869. doi: 10.1111/all.13869.

Bork, K., Wulff, K., Steinmüller-Magin, L., Braenne, I., Staubach-Renz, P., Witzke, G., et al. (2018). Hereditary angioedema with a mutation in the plasminogen gene. *Allergy* 73, 442–450. doi: 10.1111/all.13270.

Boyle, A. P., Hong, E. L., Hariharan, M., Cheng, Y., Schaub, M. A., Kasowski, M., et al. (2012). Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 22, 1790–1797. doi: 10.1101/gr.137323.112.

Boyle Evan A., O'Roak, B. J., Martin, B. K., Kumar, A., and Shendure, J. (2014). MIPgen: Optimized modeling and design of molecular inversion probes for targeted resequencing. *Bioinformatics* 30, 2670–2672. doi: 10.1093/bioinformatics/btu353.

Brown, N. J., Ray, W. A., Snowden, M., Griflin, M. R., and Nashville, M. (1996). Black Americans have an increased rate of angiotensin converting enzyme inhibitor-associated angioedema.

Brown, N. J., Snowden, M., Griffin, M. R., and Medicine Griffin, P. (1997). Recurrent Angiotensin-Converting Enzyme Inhibitor\p=m-\Associated Angioedema. Available at: http://jama.jamanetwork.com/.

Browning, S. R., and Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* 81, 1084–1097. doi: 10.1086/521987.

Bruun, E. (1953). The so-called angioneurotic edema. *J Allergy* 24, 97–105.

Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P. R., et al. (2015a). An atlas of genetic correlations across human diseases and traits. *Nat Genet* 47, 1236–1241. doi: 10.1038/ng.3406.

Bulik-Sullivan, B., Loh, P. R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., et al. (2015b). LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 47, 291–295. doi: 10.1038/ng.3211.

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562, 203–209. doi: 10.1038/s41586-018-0579-z.

Byrd, J. B., Touzin, K., Sile, S., Gainer, J. V., Yu, C., Nadeau, J., et al. (2008). Dipeptidyl peptidase IV in angiotensin-converting enzyme inhibitor-associated angioedema. *Hypertension*. doi: 10.1161/HYPERTENSIONAHA.107.096552.

Caballero, T., Baeza, M., Cabañas, R., Campos, A., Cimbollek, S., Gómez-Traseira, C., et al. (2011). Consensus Statement on the Diagnosis, Management, and Treatment of Angioedema Mediated by Bradykinin. Part I. Classification, Epidemiology, Pathophysiology, Genetics, Clinical Symptoms, and Diagnosis Spanish Study Group on Bradykinin-Induced Angioedema (SGBA) (Grupo Español de Estudio del Angioedema mediado por Bradicinina: GEAB). *Spanish Consensus on Bradykinin-Induced Angioedema (I) J Investig Allergol Clin Immunol* 21, 333–347.

Cai, N., Revez, J. A., Adams, M. J., Andlauer, T. F. M., Breen, G., Byrne, E. M., et al. (2020). Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat Genet* 52, 437–447. doi: 10.1038/s41588-020-0594-5.

Campbell, D. J., Krum, H., and Esler, M. D. (2005). Losartan increases bradykinin levels in hypertensive humans. *Circulation* 111, 315–320. doi: 10.1161/01.CIR.0000153269.07762.3B.

Cano-Gamez, E., and Trynka, G. (2020). From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. *Front Genet* 11. doi: 10.3389/fgene.2020.00424.

Carucci, L., Bova, M., Petraroli, A., Ferrara, A. L., Sutic, A., de Crescenzo, G., et al. (2020). Angiotensin-converting enzyme inhibitor– associated angioedema: From bed to bench. *J Investig Allergol Clin Immunol* 30, 272–280. doi: 10.18176/jiaci.0458.

Castellino, F. J., and Ploplis, V. A. (2005). Structure and function of the plasminogen/plasmin system. *Thromb Haemost* 93, 647–654. doi: 10.1160/TH04-12-0842.

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* 4. doi: 10.1186/s13742-015-0047-8.

Chiu, A. G., Krowiak, E. J., and Deeb, Z. E. (2001). Angioedema Associated With Angiotensin II Receptor Antagonists: Challenging Our Knowledge of Angioedema and Its Etiology.

Choi, S. W., Mak, T. S. H., and O'Reilly, P. F. (2020). Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc* 15, 2759–2772. doi: 10.1038/s41596-020-0353-1.

Choi, S. W., and O'Reilly, P. F. (2019). PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience* 8. doi: 10.1093/gigascience/giz082.

Cicardi, M., Aberer, W., Banerji, A., Bas, M., Bernstein, J. A., Bork, K., et al. (2014). Classification, diagnosis, and approach to treatment for angioedema: consensus report from the Hereditary Angioedema International Working Group. *Allergy* 69, 602–616. doi: 10.1111/all.12380.

Cicardi, M., Suffritti, C., Perego, F., and Caccia, S. (2016). Novelties in the diagnosis and treatment of angioedema. *J Investig Allergol Clin Immunol*. doi: 10.18176/jiaci.0087.

Cicardi, M., Zingale, L., Zanichelli, A., Pappalardo, E., and Cicardi, B. (2005). C1 inhibitor: Molecular and clinical aspects. *Springer Semin Immunopathol*. doi: 10.1007/s00281-005-0001-4.

Cichon, S., Martin, L., Hennies, H. C., Müller, F., Driessche, K. Van, Karpushova, A., et al. (2006). Increased Activity of Coagulation Factor XII (Hageman Factor) Causes Hereditary Angioedema Type III. Available at: www.ajhg.org.

Cilia La Corte, A. L., Carter, A. M., Rice, G. I., Duan, Q. L., Rouleau, G. A., Adam, A., et al. (2011). A functional XPNPEP2 promoter haplotype leads to reduced plasma aminopeptidase P and increased risk of ACE inhibitor-induced angioedema. *Hum Mutat* 32, 1326–1331. doi: 10.1002/humu.21579.

Claussnitzer, M., Cho, J. H., Collins, R., Cox, N. J., Dermitzakis, E. T., Hurles, M. E., et al. (2020). A brief history of human disease genetics. *Nature* 577, 179–189. doi: 10.1038/s41586-019-1879-7.

Colman, R. W., and Schmaier, A. H. (1997). Contact System: A Vascular Biology Modulator With Anticoagulant, Profibrinolytic, Antiadhesive, and Proinflammatory Attributes.

Cordisco, G., Santacroce, R., Margaglione, M., d'Apolito, M., Colia, A. L., and Maffione, A. B. (2019). Angiopoietin-1 haploinsufficiency affects the endothelial barrier and causes hereditary angioedema. *Clinical & Experimental Allergy*. doi: 10.1111/cea.13349.

Dahlbäck, B., and Villoutreix, B. O. (2005). The anticoagulant protein C pathway. *FEBS Lett* 579, 3310–3316. doi: 10.1016/j.febslet.2005.03.001.

Daly, A. K. (2010). Genome-wide association studies in pharmacogenomics. *Nat Rev Genet* 11, 241–246. doi: 10.1038/nrg2751.

Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J., and Lander, E. S. (2001). High-resolution haplotype structure in the human genome. *Nat Genet* 29, 229–232. Available at: http://genetics.nature.com.

Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A. E., Kwong, A., et al. (2016). Next-generation genotype imputation service and methods. *Nat Genet* 48, 1284–1287. doi: 10.1038/ng.3656.

Davies, M. R., Kalsi, G., Armour, C., Jones, I. R., McIntosh, A. M., Smith, D. J., et al. (2019). The Genetic Links to Anxiety and Depression (GLAD) Study: Online recruitment into the largest recontactable study of depression and anxiety. *Behaviour Research and Therapy* 123. doi: 10.1016/j.brat.2019.103503.

de Leeuw, C. A., Mooij, J. M., Heskes, T., and Posthuma, D. (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput Biol* 11. doi: 10.1371/journal.pcbi.1004219.

De Maat, S., Hofman, Z. L. M., and Maas, C. (2018). Hereditary angioedema: The plasma contact system out of control. *Journal of Thrombosis and Haemostasis*. doi: 10.1111/jth.14209.

Debreczeni, M. L., Németh, Z., Kajdácsi, E., Farkas, H., and Cervenak, L. (2021). Molecular Dambusters: What Is Behind Hyperpermeability in Bradykinin-Mediated Angioedema? *Clin Rev Allergy Immunol* 60, 318–347. doi: 10.1007/s12016-021-08851-8.

Dennis, J., Johnson, C. Y., Adediran, A. S., de Andrade, M., Heit, J. A., Morange, P.-E., et al. (2012). The endothelial protein C receptor (PROCR) Ser219Gly variant and risk of common thrombotic disorders: a HuGE review and meta-analysis of evidence from observational studies. doi: 10.1182/blood-2011-10.

Dewald, G., and Bork, K. (2006). Missense mutations in the coagulation factor XII (Hageman factor) gene in hereditary angioedema with normal C1 inhibitor. *Biochem Biophys Res Commun* 343, 1286–1289. doi: 10.1016/j.bbrc.2006.03.092.

Duan, Q. L., Nikpoor, B., Dubé, M.-P., Molinaro, G., Meijer, I. A., Dion, P., et al. (2005). A Variant in XPNPEP2 Is Associated with Angioedema Induced by Angiotensin I–Converting Enzyme Inhibitors. *The American Journal of Human Genetics* 77, 617–626. doi: 10.1086/496899.

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. doi: 10.1038/nature11247.

Eijkelenboom, A., Kamping, E. J., Kastner-van Raaij, A. W., Hendriks-Cornelissen, S. J., Neveling, K., Kuiper, R. P., et al. (2016). Reliable Next-Generation Sequencing of Formalin-Fixed, Paraffin-Embedded Tissue Using Single Molecule Tags. *Journal of Molecular Diagnostics* 18, 851–863. doi: 10.1016/j.jmoldx.2016.06.010.

Erbel, R., Eisele, L., Moebus, S., Dragano, N., Möhlenkamp, S., Bauer, M., et al. (2012). Die Heinz Nixdorf Recall Studie. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 55, 809–815. doi: 10.1007/s00103-012-1490-7.

Farsetti, A., Misiti, S., and Citarella, F. (1995). Molecular Basis of Estrogen Regulation of Hageman Factor XII Gene Expression*. *Endocrinology* 136. doi: 10.1210/endo.136.11.7588244.

Freedman, M. L., Reich, D., Penney, K. L., McDonald, G. J., Mignault, A. A., Patterson, N., et al. (2004). Assessing the impact of population stratification on genetic association studies. *Nat Genet* 36, 388–393. doi: 10.1038/ng1333.

Germenis, A. E., Loules, G., Zamanakou, M., González-Quevedo, T., and Speletas M. (2018). On the pathogenicity of the plasminogen K330E mutation for hereditary angioedema. doi: 10.1111/all.13324.

Germenis, A. E., Margaglione, M., Pesquero, J. B., Farkas, H., Cichon, S., Csuka, D., et al. (2020). International Consensus on the Use of Genetics in the Management of Hereditary

Angioedema. *Journal of Allergy and Clinical Immunology: In Practice* 8, 901–911. doi: 10.1016/j.jaip.2019.10.004.

Germenis, A. E., Rijavec, M., and Veronez, C. L. (2021). Leveraging Genetics for Hereditary Angioedema: A Road Map to Precision Medicine. *Clin Rev Allergy Immunol* 60, 416–428. doi: 10.1007/s12016-021-08836-7.

Germenis, A. E., and Speletas, M. (2016). Genetics of Hereditary Angioedema Revisited. *Clin Rev Allergy Immunol*. doi: 10.1007/s12016-016-8543-x.

Ghouse, J., Ahlberg, G., Andreasen, L., Banasik, K., Brunak, S., Schwinn, M., et al. (2021). Association of Variants Near the Bradykinin Receptor B2 Gene With Angioedema in Patients Taking ACE Inhibitors. *J Am Coll Cardiol* 78, 696–709. doi: 10.1016/j.jacc.2021.05.054.

Gianni, P., Loules, G., Zamanakou, M., Kompoti, M., Csuka, D., Psarros, F., et al. (2017). Genetic determinants of C1 inhibitor deficiency angioedema age of onset. *Int Arch Allergy Immunol* 174, 200–204. doi: 10.1159/000481987.

Goodwin, S., McPherson, J. D., and McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nat Rev Genet* 17, 333–351. doi: 10.1038/nrg.2016.49.

Gulec, M., Caliskaner, Z., Tunca, Y., Ozturk, S., Bozoglu, E., Gul, D., et al. (2008). The role of ace gene polymorphism in the development of angioedema secondary to angiotensin converting enzyme inhibitors and angiotensin II receptor blockers. *Allergol Immunopathol (Madr)* 36, 134–140. doi: 10.1016/s0301-0546(08)72537-0.

Gunderson, K. L., Steemers, F. J., Ren, H., Ng, P., Zhou, L., Tsan, C., et al. (2006). Whole-Genome Genotyping. *Methods Enzymol* 410, 359–376. doi: 10.1016/S0076-6879(06)10017-8.

Heit, J. A., Armasu, S. M., Asmann, Y. W., Cunningham, J. M., Matsumoto, M. E., Petterson, T. M., et al. (2012). A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24.2 and 9q. *Journal of Thrombosis and Haemostasis* 10, 1521–1531. doi: 10.1111/j.1538-7836.2012.04810.x.

Hiatt, J. B., Pritchard, C. C., Salipante, S. J., O'Roak, B. J., and Shendure, J. (2013). Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. *Genome Res* 23, 843–854. doi: 10.1101/gr.147686.112.

Hofman, Z., de Maat, S., Hack, C. E., and Maas, C. (2016). Bradykinin: Inflammatory Product of the Coagulation System. *Clin Rev Allergy Immunol*. doi: 10.1007/s12016-016-8540-0.

Hujová, P., Souček, P., Grodecká, L., Grombiříková, H., Ravčuková, B., Kuklínek, P., et al. (2020). Deep Intronic Mutation in SERPING1 Caused Hereditary Angioedema Through Pseudoexon Activation. *J Clin Immunol* 40, 435–446. doi: 10.1007/s10875-020-00753-2.

Jagadeesh, K. A., Dey, K. K., Montoro, D. T., Mohan, R., Gazal, S., Engreitz, J. M., et al. (2022). Identifying disease-critical cell types and cellular processes by integrating single-cell RNA-sequencing and human genetics. *Nat Genet* 54, 1479–1492. doi: 10.1038/s41588-022-01187-9.

Jett, G. K. (1984). Captopril-Induced Angioedema. *Ann Emerg Med* 13, 489–490.

Kanzi, A. M., San, J. E., Chimukangara, B., Wilkinson, E., Fish, M., Ramsuran, V., et al. (2020). Next Generation Sequencing and Bioinformatics Analysis of Family Genetic Inheritance. *Front Genet* 11. doi: 10.3389/fgene.2020.544162.

Kaplan, A. P., and Ghebrehiwet, B. (2010). The plasma bradykinin-forming pathways and its interrelationships with complement. *Mol Immunol* 47, 2161–2169. doi: 10.1016/j.molimm.2010.05.010.

Kaplan, A. P., and Greaves, M. W. (2005). Angioedema. *J Am Acad Dermatol*. doi: 10.1016/j.jaad.2004.09.032.

Kaplan, A. P., Joseph, K., and Ghebrehiwet, B. (2022). The complex role of kininogens in hereditary angioedema. *Frontiers in Allergy*.

Kaplan, A. P., and Maas, C. (2017). The Search for Biomarkers in Hereditary Angioedema. *Front Med (Lausanne)*. doi: 10.3389/fmed.2017.00206.

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581, 434–443. doi: 10.1038/s41586-020-2308-7.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., et al. (2002). The Human Genome Browser at UCSC. *Genome Res* 12, 996–1006. doi: 10.1101/gr.229102.

Khera, A. V., Chaffin, M., Aragam, K. G., Haas, M. E., Roselli, C., Choi, S. H., et al. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 50, 1219–1224. doi: 10.1038/s41588-018-0183-z.

Kircher, M., Witten, D. M., Jain, P., O'roak, B. J., Cooper, G. M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 46, 310–315. doi: 10.1038/ng.2892.

Kiss, N., Barabás, E., Várnai, K., Halász, A., Varga, L. Á., Prohászka, Z., et al. (2013). Novel duplication in the F12 gene in a patient with recurrent angioedema. *Clinical Immunology* 149, 142–145. doi: 10.1016/j.clim.2013.08.001.

Korbie, D. J., and Mattick, J. S. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat Protoc* 3, 1452–1456. doi: 10.1038/nprot.2008.133.

Kostis, J. B., Kim, H. J., Rusnak, J., Casale, T., Kaplan, A., Corren, J., et al. (2005). Incidence and Characteristics of Angioedema Associated With Enalapril. *Arch Intern Med* 165, 16371642. doi: doi: 10.1001/archinte.165.14.1637.

Kostis, W. J., Shetty, M., Chowdhury, Y. S., and Kostis, J. B. (2018). ACE Inhibitor-Induced Angioedema: a Review. doi: 10.1007/s11906-018-0859-x.

Kurosawa, S., Stearns-Kurosawa, D. J., Hidari, N., and Esmon, C. T. (1997). Identification of functional endothelial protein C receptor in human plasma. *Journal of Clinical Investigation* 100, 411–418. doi: 10.1172/JCI119548.

Lander, E., and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11, 241–247. doi: 10.1038/ng1195-241.

Lander, S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome International Human Genome Sequencing Consortium* The Sanger Centre: Beijing Genomics Institute/Human Genome Center. Available at: www.nature.com.

Landrum, M. J., Chitipiralla, S., Brown, G. R., Chen, C., Gu, B., Hart, J., et al. (2020). ClinVar: Improvements to accessing data. *Nucleic Acids Res* 48, D835–D844. doi: 10.1093/nar/gkz972.

Laursen, I. H., Banasik, K., Haue, A. D., Petersen, O., Holm, P. C., Westergaard, D., et al. (2021). Cohort profile: Copenhagen Hospital Biobank - Cardiovascular Disease Cohort (CHB-CVDC): Construction of a large-scale genetic cohort to facilitate a better understanding of heart diseases. *BMJ Open* 11. doi: 10.1136/bmjopen-2021-049709.

Leeb-Lundberg, L. M. F. (2005). International Union of Pharmacology. XLV. Classification of the Kinin Receptor Family: from Molecular Mechanisms to Pathophysiological Consequences. *Pharmacol Rev*. doi: 10.1124/pr.57.1.2.

Lepelley, M., Bernardeau, C., Defendi, F., Crochet, J., Mallaret, M., and Bouillet, L. (2020). Update on bradykinin-mediated angioedema in 2020.

Lewis, C. M., and Vassos, E. (2020). Polygenic risk scores: From research tools to clinical instruments. *Genome Med* 12. doi: 10.1186/s13073-020-00742-5.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324.

Liau, Y., Chua, I., Kennedy, M. A., and Maggo, S. (2019). Pharmacogenetics of angiotensin-converting enzyme inhibitor-induced angioedema. *Clinical and Experimental Allergy* 49, 142–154. doi: 10.1111/cea.13326.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J. P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740. doi: 10.1093/bioinformatics/btr260.

Lindström, S., Wang, L., Smith, E. N., Gordon, W., van Hylckama Vlieg, A., de Andrade, M., et al. (2019). Genomic and transcriptomic association studies identify 16 novel susceptibility loci for venous thromboembolism. *Blood* 123, 1645–1657. doi: 10.1182/blood.2019000435.

Liu, J. Z., Van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., Takahashi, A., et al. (2015). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 47, 979–986. doi: 10.1038/ng.3359.

Lo, K. S. (2002). Angioedema associated with candesartan. *Pharmacotherapy* 22, 1176–1179. doi: 10.1592/phco.22.13.1176.33516.

Loh, P. R., Danecek, P., Palamara, P. F., Fuchsberger, C., Reshef, Y. A., Finucane, H. K., et al. (2016). Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet* 48, 1443–1448. doi: 10.1038/ng.3679.

Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., et al. (2013). The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 45, 580–585. doi: 10.1038/ng.2653.

Lotta, L. A., Wang, M., Yu, J., Martinelli, I., Yu, F., Passamonti, S. M., et al. (2012). Identification of genetic risk variants for deep vein thrombosis by multiplexed next-generation sequencing of 186 hemostatic/pro-inflammatory genes. *BMC Med Genomics* 5. doi: 10.1186/1755-8794-5-7.

Ludwig, K. U., Degenhardt, F., and Nöthen, M. M. (2019). The role of rare variants in common diseases. *Medizinische Genetik* 31, 212–221. doi: 10.1007/s11825-019-0246-2.

Ma, Q., Ozel, A. B., Ramdas, S., Mcgee, B., Khoriaty, R., Siemieniak, D., et al. (2014). Genetic variants in PLG, LPA, and SIGLEC 14 as well as smoking contribute to plasma plasminogen levels Key Points. *Blood* 124. doi: 10.1182/blood-2014-03.

Mahmoudpour, S. H., Asselbergs, F. W., Terreehorst, I., Souverein, P. C., De Boer, A., and Maitland-Van Der Zee, A. H. (2015). Continuation of angiotensin converting enzyme inhibitor therapy, in spite of occurrence of angioedema. *Int J Cardiol* 201, 644–645. doi: 10.1016/j.ijcard.2015.08.185.

Mahmoudpour, S. H., Baranova, E. V., Souverein, P. C., Asselbergs, F. W., de Boer, A., and Maitland-van der Zee, A. H. (2016). Determinants of angiotensin-converting enzyme inhibitor (ACEI)

intolerance and angioedema in the UK Clinical Practice Research Datalink. *Br J Clin Pharmacol* 82, 1647–1659. doi: 10.1111/bcp.13090.

Makani, H., Messerli, F. H., Romero, J., Wever-Pinzon, O., Korniyenko, A., Berrios, R. S., et al. (2012). Meta-analysis of randomized trials of angioedema as an adverse event of renin-angiotensin system inhibitors. *American Journal of Cardiology* 110, 383–391. doi: 10.1016/j.amjcard.2012.03.034.

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., and Chen, W. M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics* 26, 2867–2873. doi: 10.1093/bioinformatics/btq559.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., et al. (2009). Finding the missing heritability of complex diseases. *Nature* 461, 747–753. doi: 10.1038/nature08494.

Marcelino-Rodriguez, I., Callero, A., Mendoza-Alvarez, A., Perez-Rodriguez, E., Barrios-Recio, J., Garcia-Robaina, J. C., et al. (2019). Bradykinin-Mediated Angioedema: An Update of the Genetic Causes and the Impact of Genomics. *Front Genet* 10. doi: 10.3389/fgene.2019.00900.

Maroteau, C., Siddiqui, M. K., Veluchamy, A., Carr, F., White, M., Cassidy, A. J., et al. (2020). Exome Sequencing Reveals Common and Rare Variants in F5 Associated With ACE Inhibitor and Angiotensin Receptor Blocker–Induced Angioedema. *Clin Pharmacol Ther* 108, 1195–1202. doi: 10.1002/cpt.1927.

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., and Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* 51, 584–591. doi: 10.1038/s41588-019-0379-x.

Mathieson, I., and McVean, G. (2012). Differential confounding of rare and common variants in spatially structured populations. *Nat Genet* 44, 243–246. doi: 10.1038/ng.1074.

Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., et al. (2012). Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. *Science (1979)* 337. Available at: https://www.science.org.

Maurer, M., Bader, M., Bas, M., Bossi, F., Cicardi, M., Cugno, M., et al. (2011). New topics in bradykinin research. *Allergy: European Journal of Allergy and Clinical Immunology*. doi: 10.1111/j.1398-9995.2011.02686.x.

Maurer, M., and Magerl, M. (2021). Differences and Similarities in the Mechanisms and Clinical Expression of Bradykinin-Mediated vs. Mast Cell–Mediated Angioedema. *Clin Rev Allergy Immunol* 61, 40–49. doi: 10.1007/s12016-021-08841-w.

Maurer, M., Magerl, M., Betschel, S., Aberer, W., Ansotegui, I. J., Aygören-Pürsün, E., et al. (2022). The international WAO/EAACI guideline for the management of hereditary angioedema—The 2021 revision and update. *Allergy: European Journal of Allergy and Clinical Immunology* 77, 1961–1990. doi: 10.1111/all.15214.

Medina, P., Navarro, S., Bonet, E., Martos, L., Estellés, A., Bertina, R. M., et al. (2014). Functional analysis of two haplotypes of the human endothelial protein C receptor gene. *Arterioscler Thromb Vasc Biol* 34, 684–690. doi: 10.1161/ATVBAHA.113.302518.

Miller, D. R., Oliveria, S. A., Berlowitz, D. R., Fincke, B. G., Stang, P., and Lillienfeld, D. E. (2008). Angioedema incidence in US veterans initiating angiotensin-converting enzyme inhibitors. *Hypertension* 51, 1624–1630. doi: 10.1161/HYPERTENSIONAHA.108.110270.

Minhas, N., Xue, M., Fukudome, K., and Jackson, C. J. (2010). Activated protein C utilizes the angiopoietin/Tie2 axis to promote endothelial barrier function. *The FASEB Journal* 24, 873–881. doi: 10.1096/fj.09-134445.

Minhas, N., Xue, M., and Jackson, C. J. (2017). Activated protein C binds directly to Tie2: possible beneficial effects on endothelial barrier function. *Cellular and Molecular Life Sciences* 74, 1895–1906. doi: 10.1007/s00018-016-2440-6.

Mitt, M., Kals, M., Pärn, K., Gabriel, S. B., Lander, E. S., Palotie, A., et al. (2017). Improved imputation accuracy of rare and low-frequency variants using population-specific high-coverage WGS-based imputation reference panel. *European Journal of Human Genetics* 25, 869–876. doi: 10.1038/ejhg.2017.51.

Moholisa, R. R., Rayner, B. R., Patricia Owen, E., Schwager, S. L. U., Stark, J. S., Badri, M., et al. (2013). Association of B2Receptor Polymorphisms and ACE Activity With ACE Inhibitor-Induced Angioedema in Black and Mixed-Race South Africans. *J Clin Hypertens* 15, 413–419. doi: 10.1111/jch.12104.

Montinaro, V., and Cicardi, M. (2020). ACE inhibitor-mediated angioedema. *Int Immunopharmacol* 78. doi: 10.1016/j.intimp.2019.106081.

Moore, C. M., Jacobson, S. A., and Fingerlin, T. E. (2020). Power and Sample Size Calculations for Genetic Association Studies in the Presence of Genetic Model Misspecification. *Hum Hered* 84, 256–271. doi: 10.1159/000508558.

Moreau, M. E., Garbacki, N., Molinaro, G., Brown, N. J., Marceau, F., and Adam, A. (2005). The Kallikrein-Kinin System: Current and Future Pharmacological Targets. *Journal of Pharmacological Sciences J Pharmacol Sci* 99, 6–38.

Morimoto, T., Gandhi, T. K., Fiskio, J. M., Seger, A. C., So, J. W., Cook, E. F., et al. (2004). An evaluation of risk factors for adverse drug events associated with angiotensin-converting enzyme inhibitors. *J Eval Clin Pract* 10, 499–509. doi: 10.1111/j.1365-2753.2003.00484.x.

Moskvina, V., Holmans, P., Schmidt, K. M., and Craddock, N. (2005). Design of case-controls studies with unscreened controls. *Ann Hum Genet* 69, 566–576. doi: 10.1111/j.1529-8817.2005.00175.x.

Mukae, S., Itoh, S., Aoki, S., Iwata, T., Nishio, K., Sato, R., et al. (2002). Association of polymorphisms of the renin-angiotensin system and bradykinin B2 receptor with ACE-inhibitor-related cough. *J Hum Hypertens* 16, 857–863. doi: 10.1038/sj.jhh.1001486.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., and Horn, G. (1986). Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction.

Ngok, S. P., Geyer, R., Liu, M., Kourtidis, A., Agrawal, S., Wu, C., et al. (2012). VEGF and angiopoietin-1 exert opposing effects on cell junctions by regulating the Rho GEF Syx. *Journal of Cell Biology* 199, 1103–1115. doi: 10.1083/jcb.201207009.

Norman, J. L., Holmes, W. L., Bell, W. A., and Finks, S. W. (2013). Life-threatening ACE inhibitor-induced angioedema after eleven years on lisinopril. *J Pharm Pract* 26, 382–388. doi: 10.1177/0897190012465990.

Nussberger, J., Cugno, M., Amstutz, C., Cicardi, M., Pellacani, A., and Agostoni, A. (1998). Plasma bradykinin in angio-oedema. *The Lancet* 351, 1693–1697. doi: 10.1016/S0140-6736(97)09137-X.

Nussberger, J., Cugno, M., and Cicardi, M. (2002). Bradykinin-mediated angioedema. *N Engl J Med* 347, 621–622. doi: 10.1056/NEJM200208223470820.

Nykamp, D., and Winter, E. E. (2007). Olmesartan medoxomil-induced angioedema. *Annals of Pharmacotherapy* 41, 518–520. doi: 10.1345/aph.1H566.

O'Dushlaine, C., Rossin, L., Lee, P. H., Duncan, L., Parikshak, N. N., Newhouse, S., et al. (2015). Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci* 18, 199–209. doi: 10.1038/nn.3922.

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 44, D733–D745. doi: 10.1093/nar/gkv1189.

O'Roak, B. J., Vives, L., and Fu, W. (2012). Multiplex Targeted Sequencing Identifies Recurrently Mutated Genes in Autism Spectrum Disorder. *Science (1979)* 338, 1616–1619. doi: 10.1126/science.1227764.

Oudot-Mellakh, T., Cohen, W., Germain, M., Saut, N., Kallel, C., Zelenika, D., et al. (2012). Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: The MARTHA project. *Br J Haematol* 157, 230–239. doi: 10.1111/j.1365-2141.2011.09025.x.

Pare, G., Kubo, M., Byrd, J. B., McCarty, C. A., Woodard-Grice, A., Teo, K. K., et al. (2013). Genetic variants associated with angiotensin-converting enzyme inhibitor-associated angioedema. *Pharmacogenet Genomics*. doi: 10.1097/FPC.0b013e328363c137.

Pe'er, I., Yelensky, R., Altshuler, D., and Daly, M. J. (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol* 32, 381–385. doi: 10.1002/gepi.20303.

Pintao, M. C., Roshani, S., de Visser, M. C. H., Tieken, C., Tanck, M. W. T., Wichers, I. M., et al. (2011). High levels of protein C are determined by PROCR haplotype 3. *Journal of Thrombosis and Haemostasis* 9, 969–976. doi: 10.1111/j.1538-7836.2011.04256.x.

Ponard, D., Gaboriaud, C., Charignon, D., Ghannam, A., Wagenaar-Bos, I. G. A., Roem, D., et al. (2020). SERPING1 mutation update: Mutation spectrum and C1 Inhibitor phenotypes. *Hum Mutat* 41, 38–57. doi: 10.1002/humu.23917.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rao, S., Yao, Y., and Bauer, D. E. (2021). Editing GWAS: experimental approaches to dissect and exploit disease-associated genetic variation. *Genome Med* 13. doi: 10.1186/s13073-021-00857-3.

Rasmussen, E. R., Hallberg, P., Baranova, E. V., Eriksson, N., Karawajczyk, M., Johansson, C., et al. (2020). Genome-wide association study of angioedema induced by angiotensin-converting enzyme inhibitor and angiotensin receptor blocker treatment. *Pharmacogenomics J*. doi: 10.1038/s41397-020-0165-2.

Rasmussen, E. R., Pottegård, A., Bygum, A., von Buchwald, C., Homøe, P., and Hallas, J. (2019). Angiotensin II receptor blockers are safe in patients with prior angioedema related to angiotensin-converting enzyme inhibitors – a nationwide registry-based cohort study. *J Intern Med* 285, 553–561. doi: 10.1111/joim.12867.

Rasmussen, E. R., von Buchwald, C., Wadelius, M., Prasad, S. C., Kamaleswaran, S., Ajgeiy, K. K., et al. (2017). Assessment of 105 Patients with Angiotensin Converting Enzyme-Inhibitor Induced Angioedema. *Int J Otolaryngol* 2017, 1–7. doi: 10.1155/2017/1476402.

Reichman, M. E., Wernecke, M., Graham, D. J., Liao, J., Yap, J., Chillarige, Y., et al. (2017). Antihypertensive drug associated angioedema: effect modification by race/ethnicity. *Pharmacoepidemiol Drug Saf* 26, 1190–1196. doi: 10.1002/pds.4260.

Reiner, A. P., Carty, C. L., Jenny, N. S., Nievergelt, C., Cushman, M., Stearns-Kurosawa, D. J., et al. (2008). PROC, PROCR and PROS1 polymorphisms, plasma anticoagulant phenotypes, and risk of cardiovascular disease and mortality in older adults: The Cardiovascular Health Study. *Journal of Thrombosis and Haemostasis* 6, 1625–1632. doi: 10.1111/j.1538-7836.2008.03118.x.

Rentzsch, P., Schubach, M., Shendure, J., and Kircher, M. (2021). CADD-Splice—improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Med* 13. doi: 10.1186/s13073-021-00835-9.

Reshef, A., Kidon, M., and Leibovich, I. (2016). The Story of Angioedema: from Quincke to Bradykinin. *Clin Rev Allergy Immunol*. doi: 10.1007/s12016-016-8553-8.

Ricciardolo, F. L. M., Folkerts, G., Folino, A., and Mognetti, B. (2018). Bradykinin in asthma: Modulation of airway inflammation and remodelling. *Eur J Pharmacol* 827, 181–188. doi: 10.1016/j.ejphar.2018.03.017.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine* 17, 405–424. doi: 10.1038/gim.2015.30.

Rigat, B., Hubert, C., Alhenc-Gelas, F., Cambien, F., Corvol, P., and Soubrier, F. (1990). An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. *Journal of Clinical Investigation* 86, 1343–1346. doi: 10.1172/JCI114844.

rkwalters, and Palmer, D. (2022). Nealelab/UKBB_ldsc: v2.0.0 (Round 2 GWAS update). doi: 10.5281/ZENODO.7186871.

Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* 518, 317–329. doi: 10.1038/nature14248.

Schmaier, A. H. (2016). The contact activation and kallikrein/kinin systems: Pathophysiologic and physiologic activities. *Journal of Thrombosis and Haemostasis*. doi: 10.1111/jth.13194.

Schmitt, A. D., Hu, M., Jung, I., Xu, Z., Qiu, Y., Tan, C. L., et al. (2016). A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep* 17, 2042–2059. doi: 10.1016/j.celrep.2016.10.061.

Sharma, J., Jindal, A. K., Banday, A. Z., Kaur, A., Rawat, A., Singh, S., et al. (2021). Pathophysiology of Hereditary Angioedema (HAE) Beyond the SERPING1 Gene. *Clin Rev Allergy Immunol* 60, 305–315. doi: 10.1007/s12016-021-08835-8.

Sherry, S. T., Ward, M.-H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., et al. (2001). dbSNP: the NCBI database of genetic variation. Available at: http://www.ncbi.nlm.nih.gov/SNP.

Shrine, N., Izquierdo, A. G., Chen, J., Packer, R., Hall, R. J., Guyatt, A. L., et al. (2023). Multi-ancestry genome-wide association analyses improve resolution of genes and pathways influencing lung function and chronic obstructive pulmonary disease risk. *Nat Genet* 55, 410–422. doi: 10.1038/s41588-023-01314-0.

Sollis, E., Mosaku, A., Abid, A., Buniello, A., Cerezo, M., Gil, L., et al. (2023). The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res* 51, D977–D985. doi: 10.1093/nar/gkac1010.

Speletas, M., Szilágyi, Csuka, D., Koutsostathis, N., Psarros, F., Moldovan, D., et al. (2015). F12-46C/T polymorphism as modifier of the clinical phenotype of hereditary angioedema. *Allergy: European Journal of Allergy and Clinical Immunology* 70, 1661–1664. doi: 10.1111/all.12714.

Stauber, T., Confino-Cohen, R., and Goldberg, A. (2014). Life-threatening angioedema induced by angiotensin-converting enzyme inhibitors: Characteristics and risk factors. *Am J Rhinol Allergy*. doi: 10.2500/ajra.2014.28.3989.

Stearns-Kurosawa, D. J., Kurosawa, S., Mollica, J. S., Ferrellt, G. L., and Esmon, C. T. (1996). The endothelial cell protein C receptor augments protein C activation by the thrombin-thrombomodulin complex. Available at: https://www.pnas.org.

Steemers, F. J., and Gunderson, K. L. (2007). Whole genome genotyping technologies on the BeadArray™ platform. *Biotechnol J* 2, 41–49. doi: 10.1002/biot.200600213.

Stenson, P. D., Mort, M., Ball, E. V., Chapman, M., Evans, K., Azevedo, L., et al. (2020). The Human Gene Mutation Database (HGMD®): optimizing its use in a clinical diagnostic or research setting. *Hum Genet* 139, 1197–1207. doi: 10.1007/s00439-020-02199-3.

Stieber, C., Cichon, S., Magerl, M., and Nöthen, M. M. (2017). Clinical Utility Gene Card for hereditary angioedema with normal C1 inhibitor (HAEnC1). *European Journal of Human Genetics* 25, e1–e4. doi: 10.1038/ejhg.2017.104.

Stoppa-Lyonnet, D., Tosi, M., Laurent, J., Sobel, A., Lagrue, G., and Meo, T. (1987). Altered C1 Inhibitor Genes in Type I Hereditary Angioedema. *New England Journal of Medicine* 317, 1–6. doi: 10.1056/NEJM198707023170101.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *PNAS* 102, 15545–15550. doi: 10.1073/pnas.0506580102.

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat Rev Genet* 20, 467–484. doi: 10.1038/s41576-019-0127-1.

Tang, W., Basu, S., Kong, X., Pankow, J. S., Aleksic, N., Tan, A., et al. (2010). Genome-wide association study identifies novel loci for plasma levels of protein C: The ARIC study. *Blood* 116, 5032–5036. doi: 10.1182/blood-2010-05-283739.

Tcheandjieu, C., Zhu, X., Hilliard, A. T., Clarke, S. L., Napolioni, V., Ma, S., et al. (2022). Large-scale genome-wide association study of coronary artery disease in genetically diverse populations. *Nat Med* 28, 1679–1692. doi: 10.1038/s41591-022-01891-3.

Teare, M. D., and Koref, M. F. S. (2014). Linkage analysis and the study of mendelian disease in the era of whole exome and genome sequencing. *Brief Funct Genomics* 13, 378–383. doi: 10.1093/bfgp/elu024.

The International HapMap Consortium (2003). The International HapMap Project. *Nature* 426, 789–796. Available at: http://www.hapmap.

The International HapMap Consortium (2005). A haplotype map of the human genome. *Nature* 437, 1299–1320. doi: 10.1038/nature04226.

Thieme, F., Henschel, L., Hammond, N. L., Ishorst, N., Hausen, J., Adamson, A. D., et al. (2021). Extending the allelic spectrum at noncoding risk loci of orofacial clefting. *Hum Mutat* 42, 1066–1078. doi: 10.1002/humu.24219.

Thornton, L. M., Munn-Chernoff, M. A., Baker, J. H., Juréus, A., Parker, R., Henders, A. K., et al. (2018). The Anorexia Nervosa Genetics Initiative (ANGI): Overview and methods. *Contemp Clin Trials* 74, 61–69. doi: 10.1016/j.cct.2018.09.015.

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., et al. (2021). Genome-wide association studies. *Nature Reviews Methods Primers* 1, 59. doi: 10.1038/s43586-021-00056-9.

Unger, T., Borghi, C., Charchar, F., Khan, N. A., Poulter, N. R., Prabhakaran, D., et al. (2020). 2020 International Society of Hypertension Global Hypertension Practice Guidelines. *Hypertension* 75, 1334–1357. doi: 10.1161/HYPERTENSIONAHA.120.15026.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Res* 40. doi: 10.1093/nar/gks596.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., et al. (2013). From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics*. doi: 10.1002/0471250953.bi1110s43.

Vatsiou, S., Zamanakou, M., Loules, G., Psarros, F., Parsopoulou, F., Csuka, D., et al. (2020). A novel deep intronic SERPING1 variant as a cause of hereditary angioedema due to C1-inhibitor deficiency. *Allergology International* 69, 443–449. doi: 10.1016/j.alit.2019.12.009.

Venter, C. J., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., et al. (2001). The Sequence of the Human Genome. *Science (1979)* 291, 1304–1351.

Veronez, C. L., Csuka, D., Sheikh, F. R., Zuraw, B. L., Farkas, H., and Bork, K. (2021). The Expanding Spectrum of Mutations in Hereditary Angioedema. *Journal of Allergy and Clinical Immunology: In Practice* 9, 2229–2234. doi: 10.1016/j.jaip.2021.03.008.

Veronez, C. L., Serpa, F. S., and Pesquero, J. B. (2017). A rare mutation in the F12 gene in a patient with ACE inhibitor-induced angioedema. *Annals of Allergy, Asthma and Immunology* 118, 743–745. doi: 10.1016/j.anai.2017.04.014.

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., et al. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* 101, 5–22. doi: 10.1016/j.ajhg.2017.06.005.

Wadelius, M., Marshall, S. E., Islander, G., Nordang, L., Karawajczyk, M., Yue, Q. Y., et al. (2014). Phenotype standardization of angioedema in the head and neck region caused by agents acting on the angiotensin system. *Clin Pharmacol Ther* 96, 477–481. doi: 10.1038/clpt.2014.138.

Wang, G., Sarkar, A., Carbonetto, P., and Stephens, M. (2020). A simple new approach to variable selection in regression, with application to genetic fine mapping. *J R Stat Soc Series B Stat Methodol* 82, 1273–1300. doi: 10.1111/rssb.12388.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38. doi: 10.1093/nar/gkq603.

Watanabe, K., Taskesen, E., van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 8. doi: 10.1038/s41467-017-01261-5.

Weissbrod, O., Hormozdiari, F., Benner, C., Cui, R., Ulirsch, J., Gazal, S., et al. (2020). Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat Genet* 52, 1355–1363. doi: 10.1038/s41588-020-00735-5.

Willer, C. J., Li, Y., and Abecasis, G. R. (2010). METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 26, 2190–2191. doi: 10.1093/bioinformatics/btq340.

Williams, B., Mancia, G., Spiering, W., Rosei, E. A., Azizi, M., Burnier, M., et al. (2018). 2018 ESC/ESH Guidelines for themanagement of arterial hypertension. *Eur Heart J* 39, 3021–3104. doi: 10.1093/eurheartj/ehy339.

Woodard-Grice, A. V., Lucisano, A. C., Byrd, J. B., Stone, E. R., Simmons, W. H., and Brown, N. J. (2010). Sex-dependent and race-dependent association of XPNPEP2 C-2399A polymorphism with angiotensin-converting enzyme inhibitor-associated angioedema. *Pharmacogenet Genomics* 20, 532–536. doi: 10.1097/FPC.0b013e32833d3acb.

Yakushiji, H., Hashimura, C., Fukuoka, K., Kaji, A., Miyahara, H., Kaname, S., et al. (2018). A missense mutation of the plasminogen gene in hereditary angioedema with normal C1 inhibitor in Japan. *Allergy: European Journal of Allergy and Clinical Immunology* 73, 2244–2247. doi: 10.1111/all.13550.

Yang, J., Zeng, J., Goddard, M. E., Wray, N. R., and Visscher, P. M. (2017). Concepts, estimation and interpretation of SNP-based heritability. *Nat Genet* 49, 1304–1310. doi: 10.1038/ng.3941.

Yun, L., Willer, C., Sanna, S., and Abecasis, G. (2009). Genotype imputation. *Annu Rev Genomics Hum Genet* 10, 387–406. doi: 10.1146/annurev.genom.9.081307.164242.

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30, 614–620. doi: 10.1093/bioinformatics/btt593.

Zhang, M. J., Hou, K., Dey, K. K., Sakaue, S., Jagadeesh, K. A., Weinand, K., et al. (2022). Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data. *Nat Genet* 54, 1572–1580. doi: 10.1038/s41588-022-01167-z.

Zhou, W., Nielsen, J. B., Fritsche, L. G., Dey, R., Gabrielsen, M. E., Wolford, B. N., et al. (2018). Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet* 50, 1335–1341. doi: 10.1038/s41588-018-0184-y.

Zou, Y., Carbonetto, P., Wang, G., and Stephens, M. (2022). Fine-mapping from summary data with the “Sum of Single Effects” model. *PLoS Genet* 18. doi: 10.1371/journal.pgen.1010299.

APPENDIX A | Material and methods

APPENDIX A1: List of equipment, chemicals, buffers, solutions, reagents, enzymes, commercial kits, consumables, primer, software and databases

Equipment

Balances:

- TE3102S (Sartorius, Göttingen, Germany)

Centrifuges:

- Centrifuge 5920 R (Eppendorf SE, Hamburg, Germany)
- Centrifuge 5720 (Eppendorf SE, Hamburg, Germany)
- Mikrozentrifuge MiniStar silverline (VWR International, Radnor, PA, US)

DNA isolation:

- Chemagic Magnetic Separation Module I (Perkin Elmer Chemagen Technology, Baesweiler, Germany)

DNA quantification and adjustment:

- Agilent 2200 TapeStation (Agilent Technologies, Santa Clara, CA, US)
- Agilent 4200 TapeStation (Agilent Technologies, Santa Clara, CA, US)
- Concentrator plus (Eppendorf SE, Hamburg, Germany)
- NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, US)
- NanoDrop® 8000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, US)
- Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, US)
- Tecan GEnios Pro Microplate Reader (Tecan, Männedorf, Switzerland)

Gel electrophoresis:

- intas GelStick imager (intas Science Imaging Instruments GmbH, Göttingen, Germany)
- SubCell Model 96 (Bio-Rad Laboratories GmbH, Feldkirchen, Germany)
- PowerBac Basic Power Supply (Bio-Rad Laboratories GmbH, Feldkirchen, Germany)

Genotyping platform:

- iScan System (Illumina, San Diego, CA, US)

Other:

- AF 80 Flake ice machine (Scotsman Ice Systems, Milan, Italy)
- Milli-Q Q-Pod® (Merck KGaA, Darmstadt, Germany)

Pipetting devices:

- Eppendorf research® plus, 1-channel, variable (Eppendorf SE, Hamburg, Germany)
- Transferpette® S-8, variable (Brand GmbH + CO KG, Wertheim, Germany)
- CellMate II Pipet-boy S1 (Thermo Fisher Scientific, Waltham, MA, US)
- Sartorius eLINE® Single Channel Electronic Pipettor (Sartorius, Göttingen, Germany)

Sample management:

- Impression Whole Rack 2D & 1D Code Scanner (Azenta Life Sciences, Griesheim, Germany)
- FluidX 1.9ml Tri-Coded Tube, 48-format (Azenta Life Sciences, Griesheim, Germany)
- FluidX LidLock rack, 48-format (Azenta Life Sciences, Griesheim, Germany)
- FluidX LidLock low base rack, 96-format (Azenta Life Sciences, Griesheim, Germany)
- FludiX 0.3ml, Dual-coded Tube (Azenta Life Sciences, Griesheim, Germany)
- Fluidx TPE Septum Cap (Azenta Life Sciences, Griesheim, Germany)
- SmartScan SoloTM 2D Barcode Reader (Thermo Fisher Scientific, Waltham, Massachusetts)

Sequencing platforms:

- MiSeq System (Illumina, San Diego, CA, US)
- NextSeq 550 System System (Illumina, San Diego, CA, US)
- SeqStudio Genetic Analyzer (Thermo Fisher Scientific, Waltham, MA, USA)

Shakers and Heaters:

- High-speed microplate shaker (Illumina, San Diego, CA, US)
- IKA MS3 Vortexer (IKA-Werke GmbH & Co. KG, Staufen, Germany)
- Are Heating Magnetic Stir (VELP Scientifica, Usumate, Italy)
- Microwave MW 9625 (Severin, Sundern, Germany)
- Vortex Genie 2 (Scientific Industries Inc., Bohemia, NY, US)

Thermocycler:

- Vapo.protect Mastercycler pro (Eppendorf SE, Hamburg, Germany)

Chemicals, buffers, solutions, reagents and enzymes

- 2x iProof HF Master Mix (Bio-Rad Laboratories GmbH, Feldkirchen, Germany)
- Ampligase 10x Reaction Buffer (LGC Biosearch Techologies, Hoddesdon, UK)
- Ampligase® DNA Ligase 100U/µl (LGC Biosearch Techologies, Hoddesdon, UK)
- Biozym LE Agarose (Biozym Scientific, Hessisch Oldendorf, Germany)
- Dimethyl Sulfoxide (DMSO) 100% (Bio-Rad Laboratories GmbH, Feldkirchen, Germany)
- dNTP mix (dATP, dGTP, dCTP, dTTP) 10mM (Labomedic, Bonn, Germany)
- Buffer EB (Qiagen, Hilden, Germany)
- Ethanol (EtOH) absolute (Th. Geyer, Renningen, Germany)
- Ethidium Bromide (EtBr) 1% (Merck KGaA, Darmstadt, Germany)
- Exonuclease I; 20000 U/ml (New England Biolabs, Frankfurt am Main, Germany)

- Exonuclease III; 100000 U/ml (New England Biolabs, Frankfurt am Main, Germany)
- Extra Buffer 10x; 15mM $MgCl_2$ (VWR International, Radnor, PA, US)
- Hemo KlenTaq® DNA Polymerase (New England Biolabs, Frankfurt am Main, Germany)
- Loading Dye: 10 ml 10x TBE, 10 ml 0.1% bromophenol blue, 40 ml 20% Ficoll, 40 ml H_2O bidest.
- T4 DNA Ligase Reaction Buffer 10x with 10mM ATP (New England Biolabs, Frankfurt am Main, Germany)
- T4 Polynucleotid Kinase (New England Biolabs, Frankfurt am Main, Germany)
- Taq DNA Polymerase (VWR International, Radnor, PA, US)
- TBE Buffer 10x (AppliChem, Darmstadt Germany)
- Ultra-pure water (H_2O) (Merck KGaA, Darmstadt, Germany)

Commercial kits

- 5X Big Dye Terminator Cycle Sequencing Kit 3.1 (Thermo Fisher Scientific, Waltham, MA, US)
- Agencourt AMPure XP (Beckman Coulter, Krefeld, Germany)
- Agencourt CleanSEQ (Beckman Coulter, Krefeld, Germany)
- Chemagic DNA Blood 10k Kit H12 (Perkin Elmer Chemagen Technology, Baesweiler, Germany)
- DNA ladder; 100bp/1kb (AppliChem, Darmstadt, Germany)
- Infinium HTS Assay Kit (Illumina, San Diego, CA, US)
- MiSeq Reagent Nano Kit v2; 300 cycles (Illumina, San Diego, CA, US)
- NextSeq Mid Output v2.5; 300 cycles (Illumina, San Diego, CA, US)
- TapeStation D1000/High Sensitivity D1000 reagents (Agilent Technologies, Santa Clara, CA, US)
- TapeStation D1000/High Sensitivity D1000 ScreenTapes (Agilent Technologies, Santa Clara, CA, US)
- Oragene DNA OG-500 (DNA Genotek, Ottawa, Canada)
- Oragene prepIT L2P kit (DNA Genotek, Ottawa, Canada)
- Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific, Waltham, MA, US)
- Qubit dsDNA BR/HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, US)

Primer

All primer used in the scope of this dissertation were ordered from IDT (Leuven, Belgium).

Table S1 | SmMIP sequences of the five candidate genes.

Primer Sequence	Primer Name
GTGGAGACGGGAGGCGGGATGGTGCGGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAGGGTGGGAGCTGGCT	SERPING1_001
ACTACGAGGCACAGTCCTCTAACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCCTCCAGGATGGGTTCAACGAAT	SERPING1_002
GACCTTCCCTTCGCCTCTGTCTTGCAAACCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTGGGTTGGGTGGTGG	SERPING1_003
GCTCTGTGGTGGGTTGTGTGGTGGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAAATCTACCAAAGCATCC	SERPING1_004
CCCAACACGGCCTCTGTTGAATGACTCTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCCAACAAATGACCTGG	SERPING1_005
CACGACCAAAGGTGTCACCTCAGTCTCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCCCCAACCCTCATTC	SERPING1_006
GTCCTGAGAGGACTCTGAAGGGGGACCCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAACACCAAAACAAACC	SERPING1_007
GTCCTGAACACGTCGTTCTTTCCATGAGCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTTGTTGGTGTTCTTG	SERPING1_008
GTTGGCGTCACTGTTGTTGCTTAGGACTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAAGAAGGACCCAGAATG	SERPING1_009

GGAAGAGGTGGGAGGGTTGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGAGAAGGAAAGGTTAAGAAC	SERPING1_010
GTCAATGAAATGGGCCACAGGGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAGGAGAAAAGATAGGGTGG	SERPING1_011
GCCATCATGGAGAAACTGGAGATGTCCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAGGAGAGAGATGCGGT	SERPING1_012
CAGAGGAGAAAGGGGGGATCCCTAAGATGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCCCTTCTGTTTTCAAG	SERPING1_013
GCAGGGTGCGGGCCACAGAGATGGCGGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGAGGCTGGAGAGGTAG	SERPING1_014
GCATCGCAGAAACCTGAAGATCTGGGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCCATGAAGACAGGGAA	SERPING1_015
CAGAGTCAGAAGCCAGCATGATACCCTCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGAGATGGCGGAGGCTG	SERPING1_016
GTTTCCGACCCAGGGTGATTCCGGAGGCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNATCCTCCATCCTCCCC	F12_001
GCGTGCCAGGTGAGCTCTTAGCCCGGTTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGGGTGTGAAGAAGG	F12_002
GCAACAAGCCAGGCGTCTACACCGATGTGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCAAATCTCAGGTCCAC	F12_003
GTGATTCCGCAGTGAGAGAGTGGCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGCTGGTGTGTGAGGAC	F12_004
GAATGGGTGGCGCTGACCTGATGGGTTGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCACGCTCTGCCAGGTG	F12_005
CTGCCGTCCGCATCCTCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCTTCTTCCGCCTAACCCAGTGATCA	F12_006
CCGGGGCCCCAAGCTCTCTTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGCAGAGCGTGGTCTCGGAGGGT	F12_007
GCGGTAGGAGCGCACGGCCAACGTCTGGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGAAGGCGCAACAGAG	F12_008
GCCACAAGCGTTCTGGGGAAGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGCTGGTAGCTGACGGGCGAG	F12_009
GCCAACGACGCGGGTCATCGAAGACAGACCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCTGGGGCGGCTCTGGG	F12_010
GCACCCCTACATCGCCGCGCTGTACTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCTGCTCCTCCACAGCC	F12_011
GCGGCTCCGCAAGAGTCTGTCTTCGATGACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAGGAAGTGGGGGGGGG	F12_012
GCTGCGCGGGCATGAGTGGGACATGAAGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGCAAGGCTGTGGAGG	F12_013
GCCTGGGTTGGGGTCTGGCACTGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCTCTCGGCTCCTCCTT	F12_014
GCCACACGACGGGGCGCCGTTAGAGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGCATGAGTGGGACATG	F12_015
GCTGGGAGTACTGCGACCTGGCACACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGAACTGGGGACTGGG	F12_016
GCCTTGGTGTCTGAGGAGAAAGGGGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGCAGAAGGCGTGGCC	F12_017
TGATGGCCGCGGGCTCAGCTACCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTGGGTGAGTGAGGGTCTGGGG	F12_018
SCCTTCTGCGACGTGGGTGAGTGAGGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGAGAGCTCTCTGGGG	F12_019_SNP
GATGAGAGGGAGGCAGGAGAGCCCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGCCTAGAGGTGGAGGG	F12_020
GTTGGGAACGGGCCAGGGAGGAGCGTCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTATCCCTCTTTGTCCC	F12_021
GGTATAGAACTGAGCAAGCAGCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGGCGGGGTGCTGGGGG	F12_022
GCTGCAGTGGTCTGAGAGATGGACATGGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGTAGGCCCAGGGTTG	F12_023
GGGACCACTCCTTCCCAGAACTCTCCCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGCAGGGGCTGTGTTT	F12_024
GAGACAAGGCTTCCCTGCTCTACCCAGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGTGTGTGGGGTCTGG	F12_025
GCCAGGCCCTCAGCCCTGGTAAGACTACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTTTCCTGACCAGACC	F12_026
GGGTAGAGCAGGGAAGCCTTGTCTCTTTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNACAAATGTACCCACAA	F12_027
GTCTAGTCTAGTGCCTACCTGGTGCTAGGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTCCCTGCCTTCTTCT	F12_028
GCTGTGGGAACCAGGATTGTCCCAGGATTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCAGGAGGGCAGCTTG	F12_029
GTTTGTGGGTGGGGGTGAAATGAAGAGCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCAGATCAATAGGACTG	F12_ERE_030
CAATCTCCCTCTAGGAGCTGAGGGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCTGGTTGTTACTTTGGTTTTG	KNG1_001
GCATCCACAGCTTTAAATAAATCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGTGGTATGTGTGTGTGT	KNG1_002
GTATTGGCCATTCTTGGGCCTTCTGTYTTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGGTCAGTTGGATGA	KNG1_003_SNP
GCCAGGAACACAATCTTGACCAGGCTCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGATCTCTTTCTTTTCT	KNG1_004
GATCGCAATSAGCATTCGCATACACTGCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGTTGAGTGTTGTTGT	KNG1_005_SNP
GCCGTGTCTCAGAATGGGCTCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAAAACACCACCAGCCATGCAA	KNG1_006
GTCCCTTTGGAATGGTGTAAGTAGGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCAGCGAATAATGTTTAAAC	KNG1_007
GCTGGGTGGGAAGACTGTCACGAAAAGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCTCAATTGTGCAAACGA	KNG1_008
GGTCCCAGACAACTGGCTGAGTCTTTTCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTTAACTGAGCACTTA	KNG1_009
GAAGTAGCAGCCTGGCCTAGCATCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTTGCGTTATTCTCTGCATT	KNG1_010
GCAGCCCACGCAAATCTTGGTAGGTGGTTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTGTTACTGCTTTTGT	KNG1_011
CATGCACCTGTCTACTTTTTCACTGGAAGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGCAAATATTTTTAAGC	KNG1_012
GCTCATTCTGAAAATCCATATTTGGGGGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAACACTGTCTCTCTTTC	KNG1_013
GTATTACTGCAAAAATCATGCTATTGATGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCCTTGTTCTTTTCCTG	KNG1_014
GCAGGTGCCATGGAAGTGTGGGGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGCCATGTCCTCTTTGGTG	KNG1_015
GCCCTTGGTCACGTTCATGTTTATGGCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCATTCTTTTTGCCTTT	KNG1_016
GTTTTCCGTGGCCATGACCATGCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGCTTGGCTAGGGAAGGG	KNG1_017
GTCTTCTCTTGTGTCTGTGCAGAAGGTGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTGGGTCTATCTGGAT	KNG1_018
GTRTGGGAGCTGGTGATATAGGAGGCATCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTGTGTGGTTGGATTA	KNG1_019_SNP
GGAGGTCGTGTCTGGAAAATCTGATATTGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGAGAGAGGGATATTG	KNG1_020
GTATAGTAAAACGCAGTTAATATGATGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCTTGTCCTTCCTCTC	KNG1_021
ACATTCCATTTAGATTGGAGGGGCCACCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCTCCGGATTTCTTTGTTG	ANGPT1_001

CTTGCTTGTTCTTGTTATGCTTATTTGCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGTTTTGTCCCGCAG	ANGPT1_002
ATTGGTTTGGGGCTTAAGGTTTCTTATCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTTAAAAGGTCACACT	ANGPT1_003
GCAAATGTGCCCTCATGTTAACAGGAGGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGGTTCTGTTATTCT	ANGPT1_004
GTGTAGTGTTCGACTACCTTTTACCTAGCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGAATATTGGCTGGGGR	ANGPT1_005_SNP
GGAAGGGAACCGAGCCTATTCACAGTATGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGCTGAACATGAAAAGA	ANGPT1_006
GSAAGAATTTATGGTGCTTTTTGGTGTTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNATGTTTTTTAAAGTAG	ANGPT1_007_SNP
GCCAGATCCCAGTTGAATTGCTGGACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGCAGTTTTACTAAAGGGAGG	ANGPT1_008
GCAGATGTATATCAAGCTGGTTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTGCGATATTGACATTTGTGTGG	ANGPT1_009
ATGTCTTCCTCACTTTGGTATTGTTAATACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCCTTCAAGGCTTGGTT	ANGPT1_010
CAGAGCTACCACCAACAACAGTGTCCTTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTTTTATTTTCACTTC	ANGPT1_011
GTTTCTGTGTGTACTTATTTATATCCTTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAATCAAACTTCTCGAC	ANGPT1_012
GAGATACAGCTGCTGGAGAATTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGCGCCCGGCGGTAAACCATTT	ANGPT1_013
CGACTTCATGTTTTCCACAATGTAATTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTCCTTGTTGAGTCTG	ANGPT1_014
AGACACCGCTGGCAAATCAGCCATCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGTAGCCGTGTGGTTCTG	ANGPT1_015
GCACGGACCTTTTTCCTTTCTTGCACTGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCGTGAGAGTACGACAGA	ANGPT1_016
CAACACAAACGCTCTGCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTTCCTTTGCTTTCCTCGCTGCCATTC	ANGPT1_017
GACTCACATAGGGTGCAGCAATCAGCGCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTTGAGGGGGAAAGAG	ANGPT1_018
CATTTAATTTTTGATTCATGAAACTTCCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCTGGGCACTGCTGGC	PLG_001
GCAAAATGTGAGGAGGACGAAGAATTCACCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTTACTGACCATTTAT	PLG_002
GAGGAAAGAGAAATTTATGGAGCCAGAGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGGGAGCAGGAAGTATA	PLG_003
TCCTCCCCATCCTCCCRCTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTACTTATTGGATTTCCTGCTTCGTT	PLG_004_SNP
CACAAGACACCACATGAAGGTCTGCACAGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTGGGGAGAAGTGGAA	PLG_005
GTARTTCTTTCCATTCCCAGTCTTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAACCAATCCCTCACAGACACA	PLG_006_SNP
GCGACATTCTTGAGTGTGAAGGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGCTGATTTTTAGAATATAGTCT	PLG_007
GCCCCTTCCCACAGGGATGTTATTACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAACTACTGCAGGAATCCAGA	PLG_008
ACATTCCATGTTTAATTAAGGCTCTGCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCCTTCCTTCCCACTCT	PLG_009
GTCAGTGCCTGAGTGCAGCCTCTGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAGAATTACTGTCGTAACCCCG	PLG_010
GATTCCAGGATTTGGACCTGCCCTGTTCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAAAATCTTKCTTGTCC	PLG_011_SNP
GACCCCTCACACACATAACACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCTCCGTCTCAAAAAATATATATATT	PLG_012
CRGTCTCATTCTGCTGCTATGGAATGTGACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCTTCTGGTCCCACCT	PLG_013_SNP
CAAAAAGAAAAAAGTCTAGGGAACCACGCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTCCGTGGATACTGGGG	PLG_014
GCACTTGGCTGTTGGTTGTATGGCACCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTACAAAGCTACTGTA	PLG_015
CAATTACTGAAAAAAAAGAAGCATGAAGCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGCCGGTGTGGTGTCA	PLG_016
GTGGTGGTGGTGGAGGATGTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTCCTAGGAAGTTGGCTTGAAG	PLG_017
CGAGTGTTGTAGCACCTCCGCCTGTTGTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTTCTCCCACCTCTTGT	PLG_018
GCAGAAACCTTCCATGCTACACGAGAAACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAACCTGAAAAAATGCTC	PLG_019
ACAGAGACCCAGGATGATATGGAATCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTGGATTTGTCTCTGG	PLG_020
GTCTATGGGGCTCCTGGGCAGCCCAGTCCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCCCCCATTACAAAAAA	PLG_021
GAGGGAGAAGGTGTTCCAAGGCTCACRCACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTACCAAAAAGAAGGCA	PLG_022_SNP
GAAACGATTTATACTGTCCCTCCACGTAACTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNATTTCTTTCCCACCTT	PLG_023
GTTAAATTGATACTTTGTTCTGCTCCATTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAGAAGGACAGAAAAAG	PLG_024
YGTCAAGAGGAAAATATGGTCCAGCCCCTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNGTACCTGCTTAGCTTT	PLG_025_SNP
GTGTGCACCCAGGATGACCTTGTAGGATGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAAACCCAGACATAAAG	PLG_026
GCATCAGCAGTTATGTTTGACTGCTCTGCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAGGTCAAAACCAATTC	PLG_027
GGCAACTGCACCCAAAACACCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTGGCAACTGTCAGTGCCTCCGGC	PLG_028
GCACAGAGTTCGGTGGATTGGACTCTTCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNTAAACAACACTTAGAC	PLG_029
GTTAGGCTGCCTGCCTTTTATTATGGGATCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNAACAAACCTTGAAACA	PLG_030
ACCAGAGGMCCTCCACTGTCACCCTATACCTTCAGCTTCCCGATATCCGACGGTAGTGTNNNNNCGTCACTCTGTCTCCC	PLG_031_SNP

Bases representative of known variants (MAF > 1%) within the sequence of the ligation or extension arm are highlighted in blue and encoded by ambiguous coding to obtain primers that comprise a mixture of reference and alternative allele(s) at the respective base pair position. The NNNNN sequence represents five degenerate bases (molecular tag) that are unique to each smMIP and incorporated during its synthesis.

Table S2 | Sequences of the primers used in Sanger sequencing.

Primer Sequence	Primer Name
CCACACCTTCTCTTCCTGCT	SERPING1_Ex3_F
CCAGAGGCATGGCTTTGTAA	SERPING1_Ex3_R
ACGTGACTGCCGAGCAAG	F12_Ex9_F
CCTCTCGGCTCCTCCTTC	F12_Ex9_R

Consumables

- 8-well PCR tube strip (4titude, Berlin, Germany)
- BD Falcon™ conical centrifuge tubes; 15ml/50ml (Corning Inc., Corning, NY, US)
- Brand® pipette tips, racked, TipStack™; 0,5-20µl (Merck KGaA, Darmstadt, Germany)
- BrandTech™ BRAND™ BIO-CERT™ Filterspitzen; 10µl/100µl/1000µl (Thermo Fisher Scientific, Waltham, MA, US)
- CoStar® Stripette®; 5ml/10ml/25ml/50ml (Corning Inc., Corning, NY, US)
- Eppendorf Tubes ®, 5.0ml (Eppendorf SE, Hamburg, Germany)
- FrameStar® 96-well skirted PCR plate (4titude, Berlin, Germany)
- Gloves Peha-Soft nitrile (Hartmann, Heidenheim, Germany)
- Gloves MICRO-TOUCH Nitra-Tex EP (Ansell, Brussels, Belgium)
- Greiner Bio-one pipette tips; 200-1000µl (Thermo Fisher Scientific, Waltham, MA, US)
- Microplate Seal (4titude, Berlin, Germany)
- Optical tube strips + caps (Agilent Technologies, Santa Clara, CA, US)
- PCR Seal (4titude, Berlin, Germany)
- Pipette tips; 200 µl (Sarstedt AG & Co. KG, Nümbrecht, Germany)
- Reaction tubes; 1.5ml/ 2ml (Greiner Bio-One, Frickenhausen, Germany)
- SafeSeal SurPhob tips; 300µl steril (Biozym Scientific, Hessisch Oldendorf, Germany)

Software and databases

- 1000 Genomes Browser (https://www.internationalgenome.org/1000-genomes-browsers/index.html)
- 2D CYPHER Pilot Database (Thermo Fisher Scientific, Waltham, MA, US)
- Agilent TapeStation Controller Software (Agilent Technologies, Santa Clara, CA, US)
- Agilent TapeStation Analysis Software (Agilent Technologies, Santa Clara, CA, US)
- ANNOVAR (https://annovar.openbioinformatics.org/en/latest/)
- Burrows-Wheeler Aligner (BWA-MEM; https://sourceforge.net/projects/bio-bwa/)
- bcl2fastq conversion software (Illumina, San Diego, CA, US)
- CADD (https://cadd.gs.washington.edu)
- dbSNP (https://www.ncbi.nlm.nih.gov/snp/)
- Eagle 2 (https://www.hsph.harvard.edu/alkes-price/software/)
- FUMA (https://fuma.ctglab.nl)
- GATK UnifiedGenotyper (https://gatk.broadinstitute.org/hc/en-us)
- GenomeStudio v2.0 (Illumina, San Diego, CA, US)
- GnomAD (https://gnomad.broadinstitute.org)
- GTEx Portal v8 Release (https://gtexportal.org/home/)
- GO (http://geneontology.org)
- HGMD (https://www.hgmd.cf.ac.uk/ac/index.php)

- KING (https://www.kingrelatedness.com)
- LDSC v1.0.1 (https://github.com/bulik/ldsc/wiki)
- LocusZoom (http://locuszoom.org/genform.php?type=yourdata)
- MAGMA v1.08 (Multi-marker analysis of Genomic Annotation; https://ctg.cncr.nl/software/magma)
- METAL (https://genome.sph.umich.edu/wiki/METAL_Documentation)
- Minimac3 (https://genome.sph.umich.edu/wiki/Minimac3)
- MIPGen (http://shendurelab.github.io/MIPGEN/)
- MSigDB v7.0 (https://www.gsea-msigdb.org/gsea/msigdb/)
- NanoDrop ND-1000 software (Thermo Fisher Scientific, Waltham, MA, US)
- NanoDrop 8000 software (Thermo Fisher Scientific, Waltham, MA, US)
- PEAR (https://cme.h-its.org/exelixis/web/software/pear/doc.html)
- PLINK 1.9 (https://www.cog-genomics.org/plink/)
- PolyFun (https://github.com/omerwe/polyfun)
- Primer3 (https://primer3.ut.ee)
- PRSice-2 (https://choishingwan.github.io/PRSice/)
- R Version 3.5.1 and 4.1.3 (https://www.r-project.org)
- RegulomeDB score (https://regulomedb.org/regulome-search/)
- Roadmap Epigenomics Project (https://www.ncbi.nlm.nih.gov/geo/roadmap/epigenomics/)
- RefSeq (https://www.ncbi.nlm.nih.gov/refseq/)
- SAIGE (https://github.com/weizhouUMICH/SAIGE/blob/master/README.md)
- SeqMan II (DNASTAR, Madison, WI, USA)
- SuSiE (https://stephenslab.github.io/susieR/articles/finemapping.html)
- UCSC Genome Browser (http://genome.ucsc.edu)

APPENDIX A2: Supplementary Tables for Sections 3.4 and 3.5

Table S3 | Overview of methods used in patient/control cohorts with available DNA samples.

Cohort	DNA isolation	DNA quantification		Sanger sequencing (PCR only*)	SmMIP library preparation	Genome-wide genotyping
		NanoDrop	Quant-iT			
vARIANCE patients	X	X	X	X	X	X
Danish patients		X	X	X	X	X
Swedish patients		X	X	X	X	X
Buffy controls	X	X	X	X	X	X
HNR individuals	X	X				X

After all PCR steps (in either Sanger sequencing or the smMIP library preparation), successful amplification was checked for each 96-well plate in a gel-electrophoresis using a selection of 3-4 samples per plate. *Apart from test runs with a small number of samples, the subsequent cycle sequencing of the PCR products was performed using the GENEWIZ Sanger Sequencing Service. Abbreviations: PCR, polymerase chain reaction | smMIP, single-molecule Molecular Inversion Probe.

Table S4 | Standard PCR: mix and thermocycler program.

		Thermocycler Program (touchdown PCR):			
	MM (µl)	**PCR step**	**Temperature**	**Time (min)**	
10x Extra Buffer	2.5	initial denaturation	95°C	05:00	
DMSO (5%)	1.25	denaturation	95°C	00:30	2 cylces per annealing temperature (1°C steps)
dNTPs (10µM)	0.5	annealing	63°C-56°C	00:30	
Primer_FWD (10µM)	1.0	elongation	72°C	01:00	
Primer_REV (10µM)	1.0	denaturation	95°C	00:30	15 cycles
Taq DNA Polymerase	2.0	annealing	55°C	00:30	
ddH_2O	16.55	elongation	72°C	01:00	
DNA template (20ng/µl)	2.0	final elongation	72°C	05:00	
Total	**25.0**		12°C	hold	

Abbreviations: ddH_20, double distilled water | dNTPs, dideoxynucleotide triphosphates | DMSO, dimethyl sulfoxide | FWD, forward | min, minutes | MM, master mix | PCR, polymerase chain reaction | REV, reverse.

Table S5 | Cycle-sequencing of PCR products: mix and thermocycler program.

		Thermocycler Program:			
	MM (µl)	**PCR Step**	**Temperature**	**Time (min)**	
5x Big Dye V3.1	2.5	initial denaturation	96°C	01:00	
5x Buffer for Big Dye V3.1	3.75	denaturation	96°C	00:10	24 cycles
Primer (FWD or REV; 3.2 pmol/µl)	1.0	annealing	50°C	00:05	
ddH_2O	13.75	elongation/termination	60°C	02:00	
Template PCR reaction	1.0		12°C	hold	
Total	**20.0**				

Abbreviations: ddH_20, double distilled water | FWD, forward | min, minutes | MM, master mix | PCR, polymerase chain reaction | REV, reverse.

APPENDIX A3: Supplementary Tables and Figures for Section 3.6

Table S6 | Gene content of the HAE panel.

ACE	***ANGPT1***	*BDKRB1*	*BDKRB2*	*CPM*	*CPN1*
CPN2	*ENPEP*	***F12****	*HSP90AB1*	*KLK1*	*KLKB1*
KNG1	*MME*	*NOS1*	*NOS2*	*NOS3*	*PLAT*
PLAU	*PLAUR*	***PLG***	*PRCP*	*SERPINB2*	*SERPINE1*
SERPINF2	***SERPING1***	*TAC1*	*TEK*	*XPNPEP2*	

The HAE gene panel comprised 29 genes from pathways involved in the formation, metabolism and signaling of bradykinin. Genes written in bold were part of the candidate genes analysis discussed in this dissertation. *The smMIP design for the *F12* gene, included an estrogen-response-element localized in its 5' UTR (Farsetti et al., 1995), since HAE-F12 was found to be an estrogen-dependent HAE subtype (Binkley, 2010).

Table S7.1 | SmMIP phosphorylation: mix and thermocycler program.

		Themocycler program:	
	MM (µl)	**Temperature**	**Time (min)**
smMIP pool	50.0*	37°C	45:00
T4 PNK (1µl per 25µl smMIPs)	2.0	65°C	20:00
H_20 (nuclease-free)	2.0	4°C	hold
10x Buffer T4 DNA Ligase 10mM ATP (10% of total volume)	6.0		
Total	**60.0**		

*Example smMIP pool volume, the actual volume depends on the number of smMIPs within the gene panel. Abbreviations: ATP, adenosine triphosphate | min, minutes | MM, master mix |PNK, polynucleotide kinase | smMIP, single-molecule Molecular Inversion Probe.

Table S7.2 | SmMIP hybridization: mix and thermocycler program.

		Thermocycler program:	
	MM (µl)	**Temperature**	**Time**
10x Ampligase DNA Ligase Buffer	2.5	95°C	10:00 min
smMIP working dilution	x*	60°C	22-24 h
dNTP mix (0.25mM)	0.032		
Hemo Klentaq (10U/µl)	0.32		
Amligase DNA Ligase (100U/µl)	0.01		
H_20 (nuclease-free)	x**		
DNA template (20ng/µl)	2.5		
Total	**25.0**		

*The actual volume of the smMIP working dilution to be used depends on the smMIP concentration of the phosphorylated smMIP pool and must be determined individually for each assay. **Dependent on the volume of the smMIP working dilution. Abbreviations: dNTP, deoxynucleotide triphosphate | h, hours | min, minutes | MM, master mix | smMIP, single-molecule Molecular Inversion Probe.

Table S7.3 | Exonuclease treatment: mix and thermocycler program.

		Thermocycler program:	
	MM (µl)	**Temperature**	**Time (min)**
Exonuclease I	0.5	37°C	45:00
Exonuclease III	0.5	95°C	2:00
10x Ampligase DNA Ligase Buffer	0.2	4°C	hold
H_20 (nuclease-free)	0.8		
Total	**2.0**		

Abbreviations: min, minutes | MM, master mix.

Table S7.4 | PCR amplification: mix and thermocycler program.

		Thermocycler program:		
	MM (µl)	**Temperature**	**Time (min)**	
a) using single-indexing		98°C	00:30	
2x iProof HF Master Mix	12.5	98°C	00:10	for 18-28 cycles*
Primer_FWD (100µM)	0.125	98°C	00:30	
H_20 (nuclease-free)	6.125	60°C	00:30	
Primer_REV with BC (10µM)	1.25	72°C	2:00	
Template smMIPs (exo-treated)	5.0	4°C	hold	
Total	**25.0**			
b) using double-indexing				
2x iProof HF Master Mix	12.5			
H_20 (nuclease-free)	5.0			
Primer_FWD with BC (10µM)	1.25			
Primer_REV with BC (10µM)	1.25			
Template smMIPs (exo-treated)	5.0			
Total	**25.0**			

*Determined individually for each smMIP panel in a series of test PCRs with different numbers of cycles. Abbreviations: BC, barcode | FWD, forward | min, minutes | MM, master mix | REV, reverse | smMIP, single-molecule Molecular Inversion Probe.

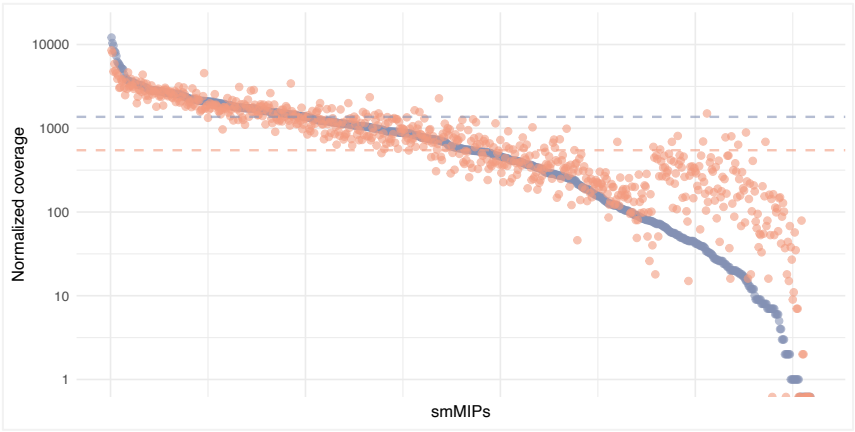

Figure S1 | HAE panel smMIP (re)-balancing.

Figure S1 displays the normalized coverage for each of the 718 designed smMIPs (69.4 kb target region). Shown are the results of the balancing (blue) and rebalancing (orange) run in six and two samples, respectively. For rebalancing the concentration of under-performers (mean coverage < 100) was increased, while over-performers (mean coverage > 5000) were attenuated. The dotted lines represent the mean coverage per run. To generate comparable results, the coverage of each run was normalized (coverage/mean coverage*1000).

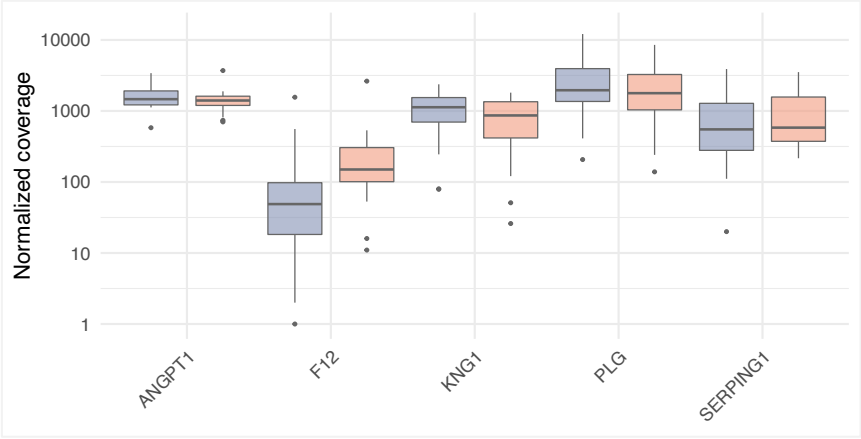

Figure S2 |(Re-)balancing results of the five genes of interest for the candidate gene analysis.

A total of 116 smMIPs (10.1 kb target region), covering the five genes of interest, were drawn from the whole HAE candidate gene panel. In Figure S2 the normalized coverages (coverage/mean coverage*1000) per candidate gene before (balancing; blue) and after the adjustment (rebalancing; orange) are displayed. Adjustments were done according to the (re)-balancing results of the whole gene panel.

APPENDIX A4: Supplementary Methods and Tables for Section 3.7

Details on the externally performed EstBB GWAS:

Samples contained in the EstBB data set were genotyped in the Core Genotyping Lab of the Institute of Genomics, University of Tartu using either the Illumina GSAv1.0, GSAv2.0 or GSAv2.0_EST array. Genotype calling was performed using the Illumina's GenomeStudio software (v.2.0.4) and obtained genotypes were extracted in PLINK format. During sample QC individuals with a call rate < 95% and mismatching sex information (genetically determined sex did not match the reported sex) were excluded. During SNP QC variants were filtered for a call rate > 95%, MAF >1% and deviation from HWE ($P < 10^{-4}$, autosomal variants only). Moreover, indels were removed. Pre-phasing of the data was done using the the Eagle software (v2.3, (Loh et al., 2016). The imputation was carried out using 2297 Estonian population specific WGS samples as a reference (Mitt et al., 2017) and the Beagle (v.28Sep18.793, Browning and Browning, 2007). The GWAS was performed using SAIGE (Zhou et al., 2018) and included sex, birthyear, birthyear squared and the first 10 PCs as covariates.

Table S8 | Overview of the number of excluded individuals during the GWAS QC stratified per cohort.

	Pre-QC		Pre-imputation QC		Subject relatedness		Ancestry outlier		Population outlier		Post-QC	
Cohort	**Patients**	**Controls**	**Patients**	**Controls**	**Patients**	**Controls**	**Patients**	**Controls**	**Patients**	**Controls**	**Patients**	**Controls**
vARIANCE	106	4,249	3	1	0	82	2	8	6	26	95	4,135
Denmark	51	1,628	1	50	0	81	1	3	4	5	45	1,489
Sweden	44	1,033	1	1	0	9	0	33	1	15	42	975
VanMar$_{EUR}$	107	330	0	1	0	1	1	3	0	4	106	321
VanMar$_{AFR}$	65	155	0	0	0	2	2	4	0	0	63	149
UKB	90	360	0	0	0	0	0	0	4	4	86	356

Abbreviation: QC, quality control.

Table S9 | External summary statistics used in the cross-trait LD score regression analyses.

Trait group	Trait	Ncase/ Nctrl or Ntotal	Source
Cardiovascular	Hypertension	93,560/ 267,581	http://www.nealelab.is/uk-biobank/; Phenotype code: 20002_1065
Allergic	Asthma	11,717/ 80,070	http://www.nealelab.is/uk-biobank/; Phenotype code: 22127
Blood clotting	Blood clot in the leg (DVT)	7,386/ 353,141	http://www.nealelab.is/uk-biobank/; Phenotype code: 6152_5
	Blood clot in the lung	2,984/ 357,543	http://www.nealelab.is/uk-biobank/; Phenotype code: 6152_7
Medication use	Intake of agents acting on the renin-angiotensin system	62,752/ 174,778	Wu et al. 2019 (https://doi.org/10.1038/s41467-019-09572-5)
Reported risk factors	Coronary artery disease	22,233/ 64,762	Schunkert et al. 2011 (https://doi.org/10.1038/ng.784)
	Ever smoked	518,633	Karlsson Linnér et al. 2019 (https://doi.org/10.1038/s41588-018-0309-3)
	Hayfever/allergic rhinitis	20,904/ 70,883	http://www.nealelab.is/uk-biobank/; Phenotype code: 22126
Reported protective factor	Type 2 diabetes	80,154/ 853,816	Mahajan et al. 2022 (https://doi.org/10.1038/s41588-022-01058-3)

Abbreviations: DVT, deep vein thrombosis | Ncase, number of cases | Nctrl, number of controls | Ntotal, number of total individuals.

APPENDIX B | Results

APPENDIX B1: Supplementary Figures and Tables for Section 4.1

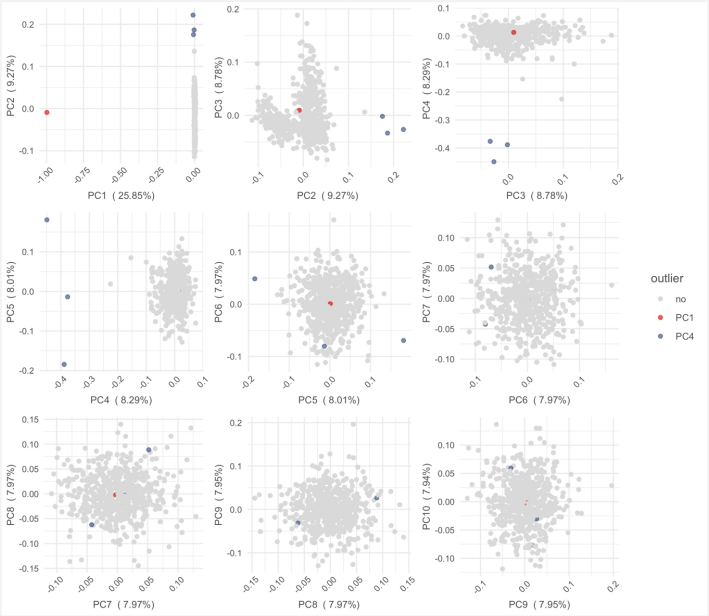

Figure S3 | In-sample principal component analysis results.

Every dot represents a single individual of the study cohort. Percentages in parentheses indicate the variance explained by each PC. Colored dots represent individuals that are outlying in any of the first ten PCs, as depicted by the respective color. Abbreviation: PC, principal component.

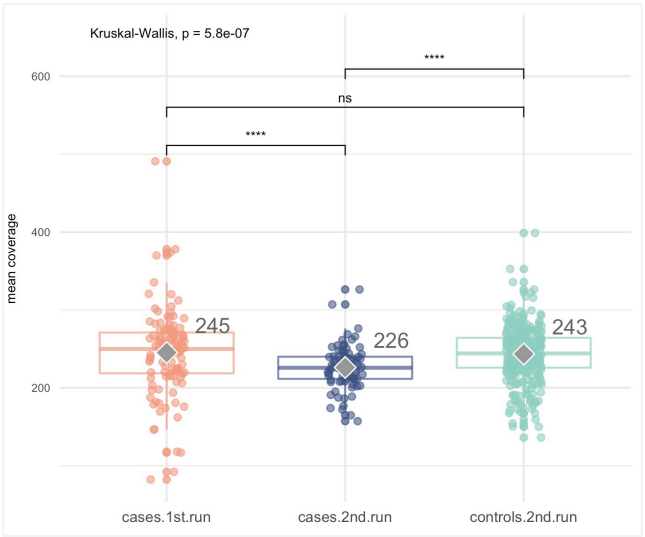

Figure S4 | Mean coverages of the first and second run cases and second run controls.

Displayed are the achieved mean coverages stratified by case-control status and the respective run the individual was sequenced in. The grey diamonds depict the overall mean coverage per group.

APPENDIX B2: Supplementary Figures and Tables for Section 4.2

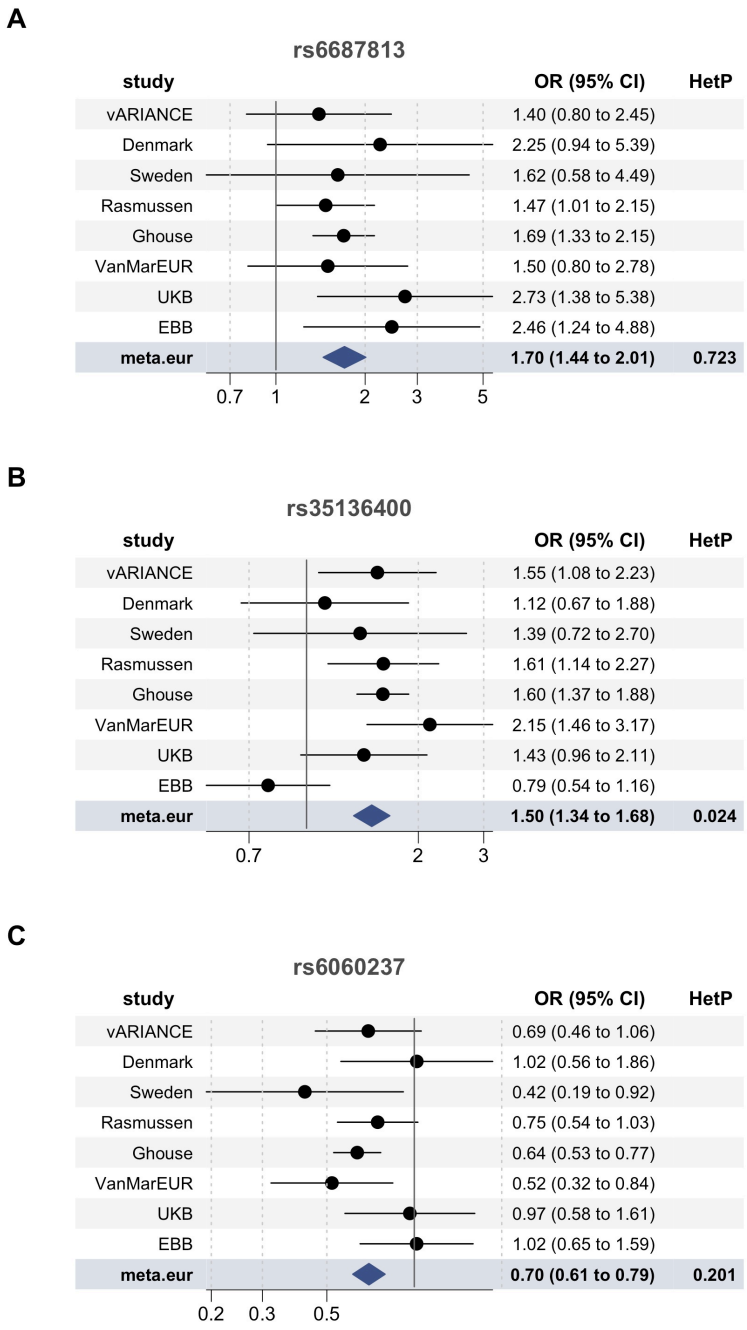

Figure S5 | Forest plots of the meta$_{EUR}$ lead SNPs.

Forest plots of the three top SNPs marking the three genome-wide significant loci on 1q24.2 (A), 14q32.2 (B), and 20q11.22 (C). Each graph illustrates the OR and corresponding 95% CI on a logarithmic scale. In addition, the between-study heterogeneity is reported as a heterogeneity p-value (Het*P*). The results of the individual GWAS studies are shown as black dots, and the result of the meta-analysis is shown as a blue diamond. Abbreviations: CI, confidence interval | OR, odds ratio.

Table S9 | Suggestive loci identified in the meta$_{EUR}$ analysis.

Suggestive locus	Cytoband	SNP	Chr	Pos	Non effect allele	Effect allele	MAF	*P*	Beta	SE	Nearest Gene	Dist	Func
1	2p22.2	2:38180325	2	38180325	C	A	0.012	1.45E-06	-1.037	0.215	RMDN2:RMDN2-AS1	00:00	ncRNA_intronic
2	2q24.1	rs61740878	2	159672252	G	A	0.154	1.93E-06	0.290	0.061	*DAPL1*	0	exonic
3	2q36.1	rs77506056	2	224620594	C	T	0.037	5.04E-06	0.506	0.111	*AP1S3*	0	intronic
4	3p24.1	rs9310952	3	30882346	G	A	0.159	1,01E-06	-0.286	0.059	*GADL1*	0	intronic
5	5p15.2	rs2401902	5	14528587	G	A	0.377	8.28E-08	-0.255	0.048	*TRIO*	0	intronic
6	5q12.3	rs3121690	5	65714450	G	A	0.402	3.40E-06	0.220	0.047	RP11-5P22.3	88921	intergenic
7	5q13.3	rs4704331	5	75841726	G	C	0.172	1.78E-06	0.299	0.063	*IQGAP2*	0	intronic
8	5q23.1	rs6885949	5	117799807	C	T	0.265	2.46E-06	-0.251	0.053	CTD-2281M20.1	4043	intergenic
9	7p15.2	rs731008	7	25560512	G	A	0.093	7.05E-06	0.382	0.085	AC091705.1	26869	intergenic
10	7p15.1	rs41345	7	28715007	C	T	0.472	9.57E-06	-0.207	0.047	*CREB5*	0	intronic
11	7q22.1	rs2690942	7	101403404	G	C	0.342	5.41E-06	-0.216	0.048	*CUX1*	55554	intergenic
12	8p23.3	rs568850179	8	1250886	G	A	0.016	4.45E-06	0.996	0.217	CTD-2281E23.1	56	upstream
13	8q12.2	rs7823926	8	61726970	C	T	0.111	4.93E-06	0.363	0.079	*CHD7*	0	intronic
14	8q24.3	rs4961375	8	142364222	G	A	0.307	8.31E-06	-0.219	0.049	CTD-3064M3.3	0	ncRNA_exonic
15	10p12.1	10:26371171	10	26371171	G	T	0.015	9.50E-06	0.926	0.209	*MYO3A*	0	intronic
16	12q23.3	rs1051924	12	108983977	C	A	0.100	1.15E-06	0.408	0.084	*TMEM119*	0	UTR3
17	19p13.2	rs794454	19	7960005	G	A	0.354	9.18E-06	-0.233	0.053	*LRRC8E*	0	intronic
18	19q13.2	rs12611340	19	41160287	G	A	0.350	5.16E-06	-0.223	0.049	*NUMBL*	12308	intergenic
19	19q13.2	19:41942508	19	41942508	C	T	0.007	2.55E-06	-1.231	0.262	*ATP5SL*	0	intronic
20	21q22.12	rs78203858	21	36445895	C	T	0.046	9.62E-06	-0.519	0.117	*RUNX1*	0	intronic

Abbreviations: Chr, chromosome | Pos, genomic position (hg19) | MAF, minor allele frequency | *P*, *p*-value | SE, standard error | Dist, distance to the nearest gene (SNPs located within a gene or 1kb up- or downstream of a transcription start site or transcription end site have a distance of 0) | Func, functional consequence of the SNP according to ANNOVAR annotations.

Table S10 | Functional annotation of the three lead and all candidate SNPs identified at the risk loci.

SNP	Chr	Pos	Non effect allele	Effect allele	MAF	*P*	Beta	SE	r2	IndSigSNP	Nearest Gene	Func	CADD	RDB	minChrState
1:169090748	1	169090748	C	T	0.011	7.14E-07	-0.607	0.122	0.787	rs6025	*ATP1B1*	intronic	0.41	5	2
1:169160458	1	169160458	C	T	0.012	7.44E-08	0.642	0.119	0.859	rs6025	*NME7*	intronic	5.36	5	5
1:169208179	1	169208179	C	T	0.011	2.95E-06	-0.648	0.139	0.787	rs6025	*NME7*	intronic	4.43	NA	5
1:169216412	1	169216412	G	C	0.012	3.84E-08	0.634	0.115	0.859	rs6025	*NME7*	intronic	0.89	5	5
1:169324793	1	169324793	C	T	0.013	5.21E-08	0.626	0.115	0.801	rs6025	*NME7*	intronic	1.75	4	5
1:169435027	1	169435027	C	T	0.007	1.90E-06	-0.608	0.128	0.639	rs6025	*SLC19A2*	UTR3	9.62	NA	4
1:169467654	1	169467654	G	A	0.011	5.51E-07	-0.610	0.122	0.927	rs6025	*AL021068.1*	intergenic	0.69	5	5
rs4264045	1	169470748	G	T	0.062	5.52E-10	0.528	0.085	0.937	rs6687813	*F5*	intergenic	1.37	7	5
rs6670848	1	169472899	G	A	0.066	4.52E-10	0.529	0.085	1.000	rs6687813	*F5*	intergenic	1.68	7	5
rs10737547	1	169476052	G	A	0.066	4.65E-10	0.528	0.085	1.000	rs6687813	*F5*	intergenic	2.96	5	5
rs6025	1	169519049	C	T	0.012	5.81E-09	0.680	0.117	1.000	rs6025	*F5*	exonic	18.92	NA	4
rs970740	1	169479974	C	T	0.062	3.18E-10	-0.533	0.085	0.937	rs6687813	*F5*	intergenic	3.00	7	5
rs6427194	1	169481121	T	A	0.069	2.71E-10	-0.535	0.085	0.955	rs6687813	*F5*	intergenic	4.77	4	1
rs6427195	1	169481176	T	A	0.069	2.71E-10	0.535	0.085	0.955	rs6687813	*F5*	intergenic	1.99	4	1
rs6009	1	169498834	C	T	0.068	3.17E-10	0.533	0.085	0.969	rs6687813	*F5*	intronic	14.67	3a	2
rs9332666	1	169486641	G	C	0.064	3.34E-10	-0.532	0.085	0.907	rs6687813	*F5*	intronic	0.67	7	4
rs2420370	1	169490392	G	C	0.063	3.58E-10	-0.531	0.085	0.922	rs6687813	*F5*	intronic	1.86	7	4
rs6682179	1	169490401	C	T	0.067	2.79E-10	0.534	0.085	0.984	rs6687813	*F5*	intronic	4.01	7	2
rs2420371	1	169491555	G	A	0.064	3.56E-10	-0.530	0.084	0.907	rs6687813	*F5*	intronic	0.90	6	4
rs2420372	1	169498056	G	A	0.063	3.66E-10	0.532	0.085	0.891	rs6687813	*F5*	intronic	6.81	4	1
rs6687813	1	169477574	C	A	0.066	2.67E-10	0.533	0.084	1.000	rs6687813	*F5*	intergenic	6.40	4	2
rs6427197	1	169500590	C	A	0.069	3.00E-10	-0.534	0.085	0.955	rs6687813	*F5*	intronic	0.23	7	4
rs1018827	1	169514006	G	A	0.063	1.84E-09	0.504	0.084	0.830	rs6687813	*F5*	intronic	0.71	6	4
rs6427196	1	169481223	G	C	0.069	2.72E-10	0.535	0.085	0.955	rs6687813	*F5*	intergenic	0.10	3a	2
rs2213868	1	169521553	G	A	0.060	2.59E-08	-0.532	0.096	0.755	rs6687813	*F5*	intronic	8.63	6	4
rs34033283	14	96599656	G	A	0.228	1.65E-11	-0.389	0.058	0.943	rs35136400	*AL137190.1*	intergenic	1.78	7	5
rs59804216	14	96600316	C	T	0.229	5.14E-12	-0.398	0.058	0.949	rs35136400	*AL137190.1*	intergenic	2.77	5	5
rs2369539	14	96600583	C	T	0.228	7.43E-12	-0.394	0.058	0.954	rs35136400	*AL137190.1*	intergenic	0.71	6	5
rs4627266	14	96600596	T	A	0.228	5.60E-12	-0.396	0.058	0.954	rs35136400	*AL137190.1*	intergenic	2.68	6	5
rs2369541	14	96600784	G	T	0.228	5.00E-12	0.397	0.058	0.954	rs35136400	*AL137190.1*	intergenic	3.11	7	5
rs72704813	14	96612609	G	A	0.229	2.48E-12	-0.400	0.057	0.971	rs35136400	*AL137190.1*	intergenic	18.11	2b	2
rs7156430	14	96603267	C	T	0.244	1.10E-09	0.338	0.055	0.881	rs35136400	*AL137190.1*	ncRNA_exonic	0.55	5	7
rs1889372	14	96603815	G	C	0.244	1.09E-09	-0.338	0.055	0.881	rs35136400	*AL137190.1*	downstream	1.55	5	5
rs11160314	14	96604435	C	T	0.244	1.05E-09	0.338	0.055	0.881	rs35136400	*AL137190.1*	intergenic	1.67	7	5
rs3939400	14	96604627	C	T	0.244	1.14E-09	-0.337	0.055	0.881	rs35136400	*AL137190.1*	intergenic	0.47	NA	7

...Table S10 continued...

rs55940712	14	96605573	G	A	0.229	2.78E-12	-0.400	0.057	0.960	rs35136400	*AL137190.1*	intergenic	1.66	5	7
rs60634508	14	96606733	C	T	0.230	1.98E-12	-0.403	0.057	0.954	rs35136400	*AL137190.1*	intergenic	0.48	5	5
rs36024935	14	96608042	C	T	0.228	3.33E-12	-0.398	0.057	0.966	rs35136400	*AL137190.1*	intergenic	0.33	5	5
rs56334881	14	96608221	G	A	0.228	1.69E-12	-0.404	0.057	0.966	rs35136400	*AL137190.1*	intergenic	0.00	7	5
rs2151767	14	96608962	C	T	0.228	2.43E-12	-0.401	0.057	0.966	rs35136400	*AL137190.1*	intergenic	0.93	NA	5
rs7144843	14	96610075	C	T	0.249	3.17E-10	0.348	0.055	0.920	rs35136400	*AL137190.1*	intergenic	0.15	5	5
rs34485356	14	96611271	C	T	0.229	1.71E-12	0.404	0.057	0.971	rs35136400	*AL137190.1*	intergenic	2.37	6	5
rs12888576	14	96611391	G	A	0.229	1.53E-12	0.403	0.057	0.971	rs35136400	*AL137190.1*	intergenic	2.98	5	5
rs68023675	14	96615137	G	A	0.229	2.65E-12	0.400	0.057	0.971	rs35136400	*AL137190.1*	intergenic	3.87	2c	5
rs12894970	14	96612682	C	T	0.250	3.15E-10	0.348	0.055	0.915	rs35136400	*AL137190.1*	intergenic	7.99	4	2
rs12885218	14	96614325	G	A	0.249	3.36E-10	-0.348	0.055	0.920	rs35136400	*AL137190.1*	intergenic	9.59	7	5
rs71415026	14	96615089	G	A	0.229	2,29E-12	0.401	0.057	0.971	rs35136400	*AL137190.1*	intergenic	2.76	4	5
rs35136400	14	96619480	G	A	0.234	1,28E-12	0.404	0.057	1.000	rs35136400	*AL137190.1*	intergenic	3.28	5	5
rs36092996	14	96616039	G	C	0.234	1,42E-12	0.404	0.057	1.000	rs35136400	*AL137190.1*	intergenic	1.47	7	5
rs4905449	14	96616330	G	T	0.250	3,30E-10	0.348	0.055	0.915	rs35136400	*AL137190.1*	intergenic	1.85	7	5
rs10140368	14	96616571	G	A	0.249	2,95E-10	0.349	0.055	0.920	rs35136400	*AL137190.1*	intergenic	7.18	6	5
rs11365128	14	96616606	GA	G	0.229	1,75E-12	-0.403	0.057	0.971	rs35136400	*AL137190.1*	intergenic	8.28	NA	5
rs8022837	14	96616845	C	T	0.234	2,02E-12	-0.401	0.057	1.000	rs35136400	*AL137190.1*	intergenic	0.55	6	5
rs34845487	14	96617046	C	A	0.229	2.57E-12	-0.400	0.057	0.971	rs35136400	*AL137190.1*	intergenic	5.62	5	5
rs35974883	14	96617205	G	A	0.234	2.34E-12	-0.400	0.057	1.000	rs35136400	*AL137190.1*	intergenic	1.62	7	5
rs12883511	14	96617282	C	T	0.234	2.05E-12	0.400	0.057	1.000	rs35136400	*AL137190.1*	intergenic	0.17	7	5
rs1959041	14	96617880	G	A	0.249	3.74E-10	0.346	0.055	0.920	rs35136400	*AL137190.1*	intergenic	2.33	NA	5
rs34985854	14	96617940	C	T	0.229	2.00E-12	-0.402	0.057	0.971	rs35136400	*AL137190.1*	intergenic	1.53	7	5
rs11850248	14	96618312	G	A	0.234	1.61E-12	-0.402	0.057	1.000	rs35136400	*AL137190.1*	intergenic	0.79	7	5
rs11850332	14	96618733	G	C	0.234	1.62E-12	0.402	0.057	1.000	rs35136400	*AL137190.1*	intergenic	8.77	6	5
rs11850303	14	96618740	C	A	0.234	1.62E-12	0.402	0.057	1.000	rs35136400	*AL137190.1*	intergenic	7.71	7	5
rs11850334	14	96618781	T	A	0.234	1.54E-12	0.403	0.057	1.000	rs35136400	*AL137190.1*	intergenic	2.15	7	5
rs55668608	14	96619307	G	T	0.234	1.28E-12	-0.404	0.057	1.000	rs35136400	*AL137190.1*	intergenic	1.40	7	5
rs4905447	14	96602783	G	A	0.244	9.88E-10	0.338	0.055	0.881	rs35136400	*AL137190.1*	upstream	1.11	2b	5
rs12894873	14	96619739	G	T	0.234	1.39E-12	-0.404	0.057	1.000	rs35136400	*AL137190.1*	intergenic	2.31	7	5
rs11846378	14	96620441	G	A	0.234	1.60E-12	0.403	0.057	1.000	rs35136400	*AL137190.1*	intergenic	3.39	7	5
rs11846531	14	96620508	G	A	0.234	1.61E-12	-0.403	0.057	1.000	rs35136400	*AL137190.1*	intergenic	3.55	7	5
rs11846417	14	96620556	G	A	0.234	1.61E-12	0.403	0.057	1.000	rs35136400	*AL137190.1*	intergenic	3.78	6	5
rs11846465	14	96620639	G	A	0.234	1.60E-12	0.403	0.057	1.000	rs35136400	*AL137190.1*	intergenic	5.99	7	5
rs11846550	14	96620826	G	A	0.234	1.60E-12	0.403	0.057	1.000	rs35136400	*AL137190.1*	intergenic	4.67	7	5
rs71415027	14	96620874	C	T	0.234	1.58E-12	-0.403	0.057	1.000	rs35136400	*AL137190.1*	intergenic	3.40	7	5
rs71415028	14	96620955	G	A	0.234	1.60E-12	0.403	0.057	1.000	rs35136400	*AL137190.1*	intergenic	4.22	7	5
rs2369544	14	96621379	G	A	0.234	1.59E-12	0.403	0.057	1.000	rs35136400	*AL137190.1*	intergenic	3.25	6	5
rs28849215	14	96621809	G	A	0.234	1.67E-12	-0.402	0.057	1.000	rs35136400	*AL137190.1*	intergenic	0.97	6	5

…Table S10 continued…

rs28823359	14	96621833	G	C	0.234	1.59E-12	0.403	0.057	1.000	rs35136400	*AL137190.1*	intergenic	1.02	7	5
rs12432014	14	96622427	C	T	0.249	4.00E-10	-0.346	0.055	0.920	rs35136400	*AL137190.1*	intergenic	1.18	6	5
rs12881275	14	96622754	G	A	0.251	2.81E-12	-0.403	0.058	0.910	rs35136400	*AL137190.1*	intergenic	1.33	6	5
rs7492727	14	96622909	C	A	0.250	4.25E-10	-0.345	0.055	0.915	rs35136400	*AL137190.1*	intergenic	5.10	6	5
rs34393530	14	96624517	G	A	0.231	2.54E-12	0.405	0.058	0.960	rs35136400	*AL137190.1*	intergenic	3.25	7	5
rs35526305	14	96624550	C	T	0.230	2.54E-12	-0.405	0.058	0.966	rs35136400	*AL137190.1*	intergenic	0.07	7	5
rs34870532	14	96624719	G	T	0.230	3.02E-12	-0.403	0.058	0.966	rs35136400	*AL137190.1*	intergenic	1.40	7	5
rs72704824	14	96624929	C	T	0.230	2.73E-12	-0.404	0.058	0.966	rs35136400	*AL137190.1*	intergenic	0.26	6	5
rs112558727	14	96625519	C	T	0.235	1.63E-12	-0.407	0.058	0.994	rs35136400	*AL137190.1*	intergenic	0.11	6	5
rs66680728	14	96625772	C	T	0.235	1.88E-12	0.406	0.058	0.994	rs35136400	*AL137190.1*	intergenic	3.27	7	5
rs35291022	20	33537157	G	A	0.088	3.59E-06	-0.337	0.073	0.656	rs6060237	*GSS*	intronic	1.69	NA	2
rs80109502	20	33587596	G	A	0.087	4.48E-06	0.334	0.073	0.665	rs6060237	*MYH7B*	exonic	17.02	5	4
rs34174778	20	33544277	G	C	0.087	3.07E-06	-0.340	0.073	0.649	rs6060237	*GSS*	upstream	5.87	5	1
rs35552264	20	33544973	T	A	0.088	2.78E-06	-0.340	0.073	0.656	rs6060237	*GSS*	intergenic	2.21	4	5
rs55909363	20	33549407	C	T	0.088	2.82E-06	0.340	0.073	0.656	rs6060237	*GSS*	intergenic	5.62	5	5
rs55696836	20	33549887	G	A	0.088	2.86E-06	0.340	0.073	0.656	rs6060237	*GSS*	intergenic	2.01	7	5
rs80170004	20	33551462	G	A	0.088	2.85E-06	0.340	0.073	0.656	rs6060237	*GSS*	intergenic	0.74	7	5
rs76110461	20	33552305	C	T	0.088	2.71E-06	0.341	0.073	0.656	rs6060237	*GSS*	intergenic	0.65	7	5
rs17401737	20	33552642	C	A	0.088	3.46E-06	0.337	0.073	0.656	rs6060237	*GSS*	intergenic	2.96	6	5
rs75635914	20	33555815	C	T	0.088	3.38E-06	0.338	0.073	0.656	rs6060237	*MYH7B*	intergenic	5.83	5	5
rs147927753	20	33557550	G	A	0.088	2.76E-06	0.340	0.073	0.656	rs6060237	*MYH7B*	intergenic	1.45	6	5
rs75866240	20	33558839	C	T	0.088	3.32E-06	0.338	0.073	0.656	rs6060237	*MYH7B*	intergenic	5.89	5	5
rs73905019	20	33560314	C	A	0.088	2.75E-06	0.341	0.073	0.656	rs6060237	*MYH7B*	intergenic	0.38	5	5
rs74599371	20	33561495	G	A	0.088	5.07E-06	0.331	0.073	0.656	rs6060237	*MYH7B*	intergenic	1.77	5	5
rs79197732	20	33562476	C	T	0.088	5.17E-06	0.331	0.073	0.656	rs6060237	*MYH7B*	upstream	1.20	5	2
rs76191812	20	33563911	G	A	0.088	5.34E-06	0.331	0.073	0.656	rs6060237	*MYH7B*	intronic	4.33	5	5
rs77437249	20	33564738	G	A	0.088	5.39E-06	0.330	0.073	0.656	rs6060237	*MYH7B*	intronic	2.46	5	1
rs867186	20	33764554	G	A	0.087	1.22E-06	-0.348	0.072	0.618	rs6060237	*EDEM2:PROCR*	exonic	16.65	1f	4
rs55738930	20	33569515	C	T	0.088	6.09E-06	0.328	0.073	0.656	rs6060237	*MYH7B*	intronic	0.59	6	2
rs55641088	20	33569619	C	T	0.088	5.45E-06	0.331	0.073	0.656	rs6060237	*MYH7B*	intronic	3.32	6	3
rs7269138	20	33570007	C	T	0.088	3.84E-06	-0.335	0.073	0.656	rs6060237	*MYH7B*	intronic	3.90	5	2
rs8118978	20	33574458	C	A	0.087	4.36E-06	-0.333	0.073	0.665	rs6060237	*MYH7B*	intronic	3.62	5	3
rs143368271	20	33576651	C	T	0.087	4.37E-06	0.334	0.073	0.665	rs6060237	*MYH7B*	intronic	0.11	7	4
rs6060238	20	33694540	G	A	0.123	3.49E-08	0.360	0.065	1.000	rs6060237	*EDEM2*	intergenic	15.49	7	5
rs7263203	20	33773375	C	A	0.087	1.30E-06	-0.349	0.072	0.618	rs6060237	*EDEM2*	intronic	13.96	6	5
rs2295700	20	33591627	G	A	0.087	4.92E-06	0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	0.00	5	4
rs73905041	20	33592148	G	A	0.087	3.86E-06	0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	3.26	5	4
rs75537616	20	33592588	C	T	0.087	4.84E-06	0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	4.43	5	4
rs17317888	20	33594959	G	T	0.087	4.93E-06	0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	7.85	5	4

…Table S10 continued…

rs6060230	20	33689308	C	T	0.089	1.94E-06	0.344	0.072	0.648	rs6060237	*TRPC4AP*	intergenic	11.96	6	5
rs6579208	20	33598612	C	T	0.087	3.85E-06	0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	1.27	4	4
rs8123978	20	33598789	C	T	0.087	4.94E-06	0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	3.03	7	4
rs76507298	20	33623522	T	A	0.087	4.84E-06	-0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	11.23	5	4
rs75535620	20	33601498	C	T	0.087	4.94E-06	-0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	2.37	5	4
rs10485508	20	33605098	C	T	0.087	4.98E-06	0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	0.98	5	4
rs6579210	20	33605802	C	T	0.087	3.73E-06	0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	0.07	4	4
rs6579211	20	33605857	G	A	0.087	4.94E-06	0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	3.77	4	4
rs7274866	20	33608616	G	A	0.087	3.88E-06	-0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	3.87	5	4
rs8121710	20	33610118	C	T	0.087	5.84E-06	0.329	0.073	0.665	rs6060237	*TRPC4AP*	intronic	0.90	6	4
rs7271729	20	33610992	C	T	0.087	5.08E-06	0.331	0.073	0.665	rs6060237	*TRPC4AP*	intronic	3.54	5	4
rs75383229	20	33613651	C	T	0.087	3.93E-06	-0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	0.31	6	4
rs117320301	20	33614137	C	A	0.087	5.09E-06	0.331	0.073	0.665	rs6060237	*TRPC4AP*	intronic	0.72	7	4
rs76223987	20	33618472	C	T	0.087	8.59E-06	-0.322	0.072	0.665	rs6060237	*TRPC4AP*	intronic	2.06	7	4
rs142275707	20	33620457	G	A	0.087	4.79E-06	-0.332	0,073	0.665	rs6060237	*TRPC4AP*	intronic	2.43	6	4
rs7263253	20	33649376	C	T	0.087	4.95E-06	-0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	10.44	5	4
rs56363533	20	33626005	G	C	0.088	4.35E-06	0.333	0.073	0.657	rs6060237	*TRPC4AP*	intronic	0.43	7	4
rs55946144	20	33629610	C	T	0.087	3.86E-06	-0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	4.34	7	4
rs8118005	20	33636219	G	C	0.088	3.85E-06	0.335	0.073	0.656	rs6060237	*TRPC4AP*	intronic	0.24	6	4
rs8124662	20	33639256	G	A	0.087	7.33E-06	0.323	0.072	0.665	rs6060237	*TRPC4AP*	intronic	0.74	7	4
rs11905081	20	33640920	C	T	0.087	3.82E-06	0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	1.83	4	4
rs11907438	20	33641220	T	A	0.088	4.45E-06	0.333	0.073	0.656	rs6060237	*TRPC4AP*	intronic	2.41	6	4
rs143373163	20	33752110	G	T	0.088	6.75E-07	0.361	0.073	0.641	rs6060237	*EDEM2*	intronic	10.28	5	5
rs9941751	20	33645709	G	A	0.087	3.89E-06	-0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	7.01	7	4
rs11427024	20	33768439	C	CT	0.087	1.27E-06	0.348	0.072	0.618	rs6060237	*EDEM2*	intronic	10.27	NA	4
rs7268447	20	33649593	G	C	0.087	4.90E-06	0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	5.93	6	4
rs2145558	20	33650069	G	A	0.087	5.08E-06	-0.331	0.073	0.665	rs6060237	*TRPC4AP*	intronic	6.44	5	4
rs75165171	20	33651453	G	A	0.087	4.59E-06	0.333	0.073	0.665	rs6060237	*TRPC4AP*	intronic	0.33	7	4
rs149906242	20	33652362	C	CCA	0.087	4.61E-06	-0.333	0.073	0.665	rs6060237	*TRPC4AP*	intronic	1.23	NA	4
rs717593	20	33652964	G	C	0.087	4.62E-06	0.333	0.073	0.665	rs6060237	*TRPC4AP*	intronic	2.24	6	4
rs11167254	20	33654584	C	T	0.087	3.94E-06	0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	3.95	6	4
rs11167255	20	33656603	G	A	0.087	4.03E-06	0.334	0.073	0.665	rs6060237	*TRPC4AP*	intronic	2.28	7	4
rs74543591	20	33658658	C	T	0.087	4.77E-06	-0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	4.21	6	2
rs144797168	20	33659312	G	C	0.087	4.79E-06	-0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	1.77	6	5
rs139403823	20	33659750	C	T	0.087	4.79E-06	-0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	3.82	6	5
rs140622086	20	33660442	G	A	0.087	4.85E-06	-0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	5.05	7	5
rs142304991	20	33663515	C	T	0.087	4.99E-06	-0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	4.55	7	4
rs8116257	20	33664583	G	C	0.087	4.02E-06	-0.334	0.073	0.665	rs6060237	*TRPC4AP*	intronic	3.49	7	4
rs78704804	20	33665831	C	A	0.087	4.79E-06	-0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	4.84	5	4

…Table S10 continued…

rs78202808	20	33667424	C	T	0.087	4.04E-06	-0.334	0.073	0.665	rs6060237	*TRPC4AP*	intronic	4.91	5	4
rs117236853	20	33667783	C	A	0.086	5.02E-06	0.332	0.073	0.658	rs6060237	*TRPC4AP*	intronic	9.48	5	4
rs6579215	20	33668260	C	T	0.086	3.93E-06	0.335	0.073	0.658	rs6060237	*TRPC4AP*	intronic	0.69	5	4
rs6579216	20	33668297	C	T	0.086	3.93E-06	-0.335	0.073	0.658	rs6060237	*TRPC4AP*	intronic	6.84	5	4
rs147614901	20	33671947	G	A	0.087	5.25E-06	-0.331	0.073	0.665	rs6060237	*TRPC4AP*	intronic	0.82	6	4
rs55993524	20	33674004	C	T	0.087	5.06E-06	0.331	0.073	0.665	rs6060237	*TRPC4AP*	intronic	2.79	7	4
rs8121957	20	33676109	C	T	0.087	3.86E-06	-0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	0.04	5	1
rs11905354	20	33677164	G	A	0.087	3.84E-06	0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	6.47	5	4
rs192024492	20	33678368	G	A	0.087	4.90E-06	0.332	0.073	0.665	rs6060237	*TRPC4AP*	intronic	6.24	5	3
rs73903009	20	33678732	G	A	0.087	7.60E-06	0.323	0.072	0.665	rs6060237	*TRPC4AP*	intronic	9.32	4	1
rs6060222	20	33682570	C	A	0.088	2.96E-06	-0.337	0.072	0.657	rs6060237	*TRPC4AP*	intergenic	2.73	5	5
rs11908681	20	33682906	C	T	0.088	2.66E-06	0.339	0.072	0.657	rs6060237	*TRPC4AP*	intergenic	5.17	6	5
rs145850164	20	33683058	G	A	0.087	2.76E-06	0.340	0.072	0.665	rs6060237	*TRPC4AP*	intergenic	2.43	7	5
rs78517073	20	33683430	G	A	0.087	2.79E-06	0.339	0.072	0.665	rs6060237	*TRPC4AP*	intergenic	1.70	5	5
rs7261312	20	33684909	T	A	0.088	2.24E-06	-0.341	0.072	0.657	rs6060237	*TRPC4AP*	intergenic	1.72	7	5
rs6060225	20	33686404	G	C	0.088	2.26E-06	-0.341	0.072	0.657	rs6060237	*TRPC4AP*	intergenic	1.64	6	5
rs6060244	20	33699435	G	A	0.087	2.52E-06	0.341	0.072	0.665	rs6060237	*EDEM2*	intergenic	8.15	1f	4
rs6060235	20	33691652	G	A	0.087	3.19E-06	-0.336	0.072	0.665	rs6060237	*TRPC4AP*	intergenic	0.37	7	5
rs8117100	20	33692261	C	T	0.087	2.75E-06	-0.339	0.072	0.665	rs6060237	*EDEM2*	intergenic	7.10	5	5
rs6060236	20	33692618	C	A	0.087	2.11E-06	0.343	0.072	0.665	rs6060237	*EDEM2*	intergenic	9,53	4	5
rs17319967	20	33693650	G	C	0.087	2.47E-06	0.341	0.072	0.665	rs6060237	*EDEM2*	intergenic	0.56	7	5
rs11167260	20	33775200	G	A	0.087	1.36E-06	0.348	0.072	0.618	rs6060237	*EDEM2*	intronic	7.20	1f	5
rs2295888	20	33722863	G	A	0,088	4.30E-07	-0.367	0.073	0.641	rs6060237	*EDEM2*	intronic	6.89	2b	2
rs6060239	20	33694580	C	T	0.123	4.33E-08	-0.357	0.065	1.000	rs6060237	*EDEM2*	intergenic	5.26	6	5
rs6058182	20	33696486	C	T	0.087	3.15E-06	0.338	0.073	0.665	rs6060237	*EDEM2*	intergenic	1.08	7	5
rs6060241	20	33696495	G	A	0.087	3.15E-06	0.338	0.073	0.665	rs6060237	*EDEM2*	intergenic	0.12	7	5
rs6060242	20	33698016	G	T	0.088	5.19E-07	-0.380	0.076	0.657	rs6060237	*EDEM2*	intergenic	7.08	6	4
rs6060245	20	33699625	G	A	0.087	2.43E-06	-0.342	0.073	0.665	rs6060237	*EDEM2*	intergenic	5.08	5	4
rs28469723	20	33700717	G	A	0.087	2.49E-06	-0.341	0.073	0.665	rs6060237	*EDEM2*	intergenic	7.95	7	4
rs8117847	20	33642480	C	T	0.087	3.84E-06	0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	5.52	1f	4
rs6058185	20	33701652	G	A	0.088	2.67E-06	-0.339	0.072	0.657	rs6060237	*EDEM2*	intergenic	4.54	4	4
rs79341738	20	33702104	C	T	0.087	2.50E-06	-0.341	0.072	0.665	rs6060237	*EDEM2*	intergenic	2.94	6	4
rs79438986	20	33702280	G	T	0.087	3.14E-06	-0.338	0.073	0.665	rs6060237	*EDEM2*	downstream	7.44	7	4
rs112318873	20	33702831	C	T	0.087	2.49E-06	-0.341	0.073	0.665	rs6060237	*EDEM2*	downstream	2.23	7	4
rs75888794	20	33702998	G	C	0.089	2.50E-06	0.341	0.073	0.647	rs6060237	*EDEM2*	downstream	4.81	5	4
rs17092297	20	33703134	G	A	0.089	2.75E-06	0.339	0.072	0.648	rs6060237	*EDEM2*	downstream	3.85	6	4
rs111641740	20	33707177	C	T	0.088	2.83E-06	0.340	0.073	0.656	rs6060237	*EDEM2*	intronic	0.07	7	4
rs11908647	20	33720592	C	T	0.088	4.42E-07	-0.367	0.073	0.641	rs6060237	*EDEM2*	intronic	3.65	7	4
rs11908683	20	33720920	C	T	0.088	4.52E-07	-0.367	0.073	0.641	rs6060237	*EDEM2*	intronic	1.71	7	4

…Table S10 continued…

rs139791629	20	33721333	G	A	0.088	5.39E-07	0.364	0.073	0.641	rs6060237	*EDEM2*	intronic	0.29	7	4
rs11904893	20	33721356	G	A	0.088	4.33E-07	-0.367	0.073	0.641	rs6060237	*EDEM2*	intronic	0.39	7	4
rs145497211	20	33721426	C	T	0.088	4.58E-07	0.367	0.073	0.641	rs6060237	*EDEM2*	intronic	7.68	7	4
rs8119351	20	33754405	G	A	0.087	5.25E-07	0.365	0.073	0.634	rs6060237	*EDEM2*	intronic	5.48	2b	5
rs141521143	20	33723455	AAATAAT	A	0.088	6.79E-08	-0.411	0.076	0.641	rs6060237	*EDEM2*	intronic	2.10	NA	4
rs75648520	20	33724221	C	T	0.088	4.72E-07	0.366	0.073	0.641	rs6060237	*EDEM2*	intronic	4.45	2b	2
rs73903017	20	33724758	G	A	0.088	4.48E-07	-0.367	0.073	0.641	rs6060237	*EDEM2*	intronic	9.83	4	4
rs11907574	20	33725207	G	A	0.088	4.51E-07	-0.366	0.073	0.641	rs6060237	*EDEM2*	intronic	4.26	7	4
rs60866116	20	33726150	C	T	0.088	4.56E-07	-0.366	0.073	0.641	rs6060237	*EDEM2*	intronic	2.29	5	4
rs55921558	20	33726536	C	T	0.088	4.44E-07	0.367	0.073	0.641	rs6060237	*EDEM2*	intronic	2.84	7	4
rs55750106	20	33726631	C	T	0.088	5.68E-07	0.363	0.073	0.641	rs6060237	*EDEM2*	intronic	3.77	6	4
rs8126407	20	33727023	G	A	0.088	4.66E-07	-0.366	0.073	0.641	rs6060237	*EDEM2*	intronic	0.36	7	4
rs144439724	20	33728235	C	T	0.088	4.33E-07	0.368	0.073	0.641	rs6060237	*EDEM2*	intronic	2.94	6	4
rs11906318	20	33729442	C	A	0.088	4.48E-07	-0.367	0.073	0.641	rs6060237	*EDEM2*	intronic	1.49	5	4
rs11908232	20	33729477	G	C	0.088	4.98E-07	-0.365	0.073	0.641	rs6060237	*EDEM2*	intronic	4.46	5	4
rs11908100	20	33729479	G	A	0.088	4.98E-07	0.365	0.073	0.641	rs6060237	*EDEM2*	intronic	7.00	5	4
rs11167258	20	33730644	C	T	0.088	4.56E-07	0.366	0.073	0.641	rs6060237	*EDEM2*	intronic	4.40	7	4
rs35072131	20	33730691	C	CCT	0.088	4.56E-07	0.366	0.073	0.641	rs6060237	*EDEM2*	intronic	8.26	NA	4
rs12105996	20	33731010	G	A	0.088	4.48E-07	0.366	0.073	0.641	rs6060237	*EDEM2*	intronic	3.55	5	4
rs57690120	20	33731437	T	A	0.088	4.74E-07	-0.366	0.073	0.641	rs6060237	*EDEM2*	intronic	8.35	5	4
rs79048371	20	33731484	T	A	0.088	5.00E-07	0.365	0.073	0.641	rs6060237	*EDEM2*	intronic	2.22	5	4
rs12106264	20	33732369	G	T	0.088	4.90E-07	-0.365	0.073	0.641	rs6060237	*EDEM2*	intronic	3.25	5	4
rs114948279	20	33733641	G	A	0.088	5.10E-07	0.365	0.073	0.641	rs6060237	*EDEM2*	intronic	0.00	7	1
rs11907010	20	33737661	C	T	0.088	6.92E-07	0.361	0.073	0.641	rs6060237	*EDEM2*	intronic	0.23	5	5
rs141474375	20	33745891	G	A	0.088	4.32E-07	-0.368	0.073	0.641	rs6060237	*EDEM2*	intronic	1.02	7	5
rs74626382	20	33746668	G	T	0.088	4.16E-07	0.368	0.073	0.641	rs6060237	*EDEM2*	intronic	0.44	6	5
rs144917890	20	33746789	C	A	0.088	5.12E-07	0.365	0.073	0.641	rs6060237	*EDEM2*	intronic	3.51	7	5
rs6060246	20	33701107	G	A	0.087	2.76E-06	-0.340	0.072	0.665	rs6060237	*EDEM2*	intergenic	2.88	1f	4
rs55734215	20	33585437	C	T	0.087	4.71E-06	0.333	0.073	0.665	rs6060237	*MYH7B*	exonic	1.27	2b	2
rs2069940	20	33759272	G	C	0.087	9.98E-07	-0.352	0.072	0.618	rs6060237	*EDEM2*	intronic	1.10	6	1
rs17092215	20	33595913	C	T	0.087	3.85E-06	0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	1.13	1f	4
rs11907011	20	33767770	C	T	0.087	1.16E-06	0.350	0.072	0.618	rs6060237	*EDEM2*	intronic	2.38	4	2
rs7273734	20	33599403	G	C	0.087	3.85E-06	0.335	0.073	0.665	rs6060237	*TRPC4AP*	intronic	1.01	1f	4
rs7265317	20	33768523	C	T	0.087	1.21E-06	-0.349	0.072	0.618	rs6060237	*EDEM2*	intronic	5.07	4	4
rs945961	20	33769926	G	A	0.087	1.21E-06	-0.349	0.072	0.618	rs6060237	*EDEM2*	intronic	5.35	5	4
rs6060237	20	33694210	G	A	0.123	3.47E-08	-0.360	0.065	1.000	rs6060237	*EDEM2*	intergenic	0.79	7	5
rs117249133	20	33773630	G	A	0.087	1.17E-06	-0.351	0.072	0.618	rs6060237	*EDEM2*	intronic	9.97	7	5
rs56400038	20	33537671	G	A	0,088	3.08E-06	0.340	0.073	0.656	rs6060237	*GSS*	intronic	0.04	2b	4
rs11904888	20	33778866	T	A	0.087	1.33E-06	0.349	0.072	0.618	rs6060237	*EDEM2*	intronic	3.28	7	4

…Table S10 continued…

rs141932846	20	33783805	G	A	0.086	1.23E-06	0.350	0.072	0.611	rs6060237	*EDEM2*	intronic	9.20	7	5
rs11906148	20	33784021	C	A	0.086	1.22E-06	0.350	0.072	0.611	rs6060237	*EDEM2*	intronic	9.61	5	5
rs117802529	20	33786677	C	T	0.086	1.19E-06	0.351	0.072	0.611	rs6060237	*EDEM2*	intronic	7.63	7	5
rs139189391	20	33792559	C	CA	0.086	4.44E-07	0.383	0.076	0.611	rs6060237	*EDEM2*	intronic	6.59	NA	4

The lead SNP of the independent risk loci are highlighted in blue. Abbreviations: SNP, single nucleotide polymorphism | Chr, chromosome | Pos, genomic position (hg19) | MAF, minor allele frequency | *P*, *p*-value | SE, standard error | r^2, highest measured LD between the given SNP and one of the independent significant SNPs | IndSigSNP, rsID of the independent significant SNP that has the maximum r^2 with the SNP | Func, functional consequence of the SNP according to ANNOVAR annotations | CADD, combined annotation-dependent depletion score | RDB, Regulome DB score | minChrState, the minimum obtained 15-core chromatin state across the 127 tissue/cell types | NA, not available.

Table S11 | Genes prioritized by at least two gene mapping methods.

Gene	Chr	Start	End	pLI	posMap SNPs	posMap MaxCADD	eqtlMap SNPs	eqtlMap minP	eqtlMap minQ	eqtlMap tissues	eqtlDir	ciMap	ciMap tissues
ATP1B1	1	169074935	169101960	0.996	1	0.414	0	NA	NA	NA	NA	Yes	IMR90
NME7	1	169101769	169337205	0.004	4	5.355	20	3.64E-06	2.08E-78	Artery_Aorta; Cells_Cultured_fibroblasts; Skin_Sun_Exposed_Lower_leg; Thyroid	-	Yes	Promoter_anchored_loops; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC
CCDC181	1	169364108	169429907	0.000	1	9.616	0	NA	NA	NA	NA	Yes	Liver; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC
SLC19A2	1	169433147	169455241	0.004	1	9.616	23	4.64E-12	3.45E-08	Whole_Blood; Cells_Cultured_fibroblasts; Skin_Not_Sun_Exposed_Suprapubic	+	Yes	Aorta; Left_Ventricle; Liver; GM12878; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Neural_Progenitor_Cell; Trophoblast-like_Cell; hESC
F5	1	169483404	169555826	0.000	16	18.92	0	NA	NA	NA	NA	Yes	Liver; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC
C1orf112	1	169631245	169823221	0.000	0	0	4	1.55E-04	4.14E-04	Muscle_Skeletal	-	Yes	Left_Ventricle; Mesenchymal_Stem_Cell; Mesendoderm; hESC
SELE	1	169691781	169733846	0.000	0	0	2	7.49E-05	2.05E-02	Colon_Transverse	+	Yes	Aorta; Left_Ventricle; Liver; GM12878; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC
METTL18	1	169761670	169764107	0.029	0	0	17	1.77E-05	1.05E-70	Cells_Cultured_fibroblasts; Thyroid	-	Yes	Left_Ventricle; Liver; GM12878; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC
BDKRB2	14	96671016	96710666	0.270	0	0	64	8.85E-11	2.96E-13	Brain_Putamen_basal_ganglia; Lung	-	Yes	IMR90
BDKRB1	14	96722161	96735304	0.070	0	0	64	3.58E-09	5.42E-05	Brain_Putamen_basal_ganglia	+	Yes	Dorsolateral_Prefrontal_Cortex; Hippocampus; Left_Ventricle; Liver; Spleen; GM12878; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Neural_Progenitor_Cell; Trophoblast-like_Cell; hESC
GGT7	20	33432523	33460663	0.998	0	0	169	1.56E-05	1.41E-07	Whole_Blood;Artery_Tibial; Nerve_Tibial; Stomach	-	Yes	EP_links_oneway; Promoter_anchored_loops; Adult_Cortex; Left_Ventricle; Liver; Right_Ventricle; GM12878; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC

…Table S11 continued…

ACSS2	20	33459949	33515769	0.000	0	0	5	7.29E-05	7.91E-17	Nerve_Tibial	+	Yes	EP_links_oneway; Promoter_anchored_loops; Adult_Cortex; Left_Ventricle; Liver; Right_Ventricle; GM12878; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC
GSS	20	33516236	33543620	0.023	10	5.871	0	NA	NA	NA	NA	Yes	Promoter_anchored_loops; IMR90; Mesenchymal_Stem_Cell
MYH7B	20	33563206	33590240	0.000	30	17.02	33	2.23E-06	1.92E-19	Adipose_Subcutaneous; Artery_Aorta; Nerve_Tibial; Skin_Sun_Exposed_Lower_leg	+	Yes	Promoter_anchored_loops; IMR90
TRPC4AP	20	33590207	33680674	1.000	81	17.02	174	6.70E-25	2.64E-19	Adipose_Subcutaneous; Whole_Blood; Brain_Cerebellum; Breast_Mammary_Tissue; Colon_Sigmoid; Esophagus_ Gastroesophageal_Junction; Esophagus_Muscularis; Lung; Nerve_Tibial; Ovary; Pancreas; Pituitary; Skin_Sun_Exposed_ Lower_leg; Small_Intestine_ Terminal_Ileum; Spleen; Stomach; Testis; Thyroid	-	Yes	Promoter_anchored_loops; Adult_Cortex; Left_Ventricle; Liver; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC
EDEM2	20	33703167	33865928	0.000	82	16.65	174	7.40E-27	1.00E-27	Adipose_Subcutaneous; Artery_Tibial; Breast_Mammary_Tissue; Colon_Sigmoid; Heart_Atrial_Appendage; Heart_Left_Ventricle; Nerve_Tibial; Thyroid	+	Yes	IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC
PROCR	20	33759876	33765165	0.010	15	16.65	174	7.45E-19	2.46E-29	Adipose_Subcutaneous; Cells_EBV_ transformed_lymphocytes; Artery_Tibial; Breast_Mammary_ Tissue; Heart_Atrial_Appendage; Heart_Left_Ventricle; Liver; Lung; Muscle_Skeletal; Nerve_Tibial; Pancreas; Cells_Cultured_fibroblasts; Skin_Not_Sun_Exposed_Suprapubic; Skin_Sun_Exposed_Lower_leg	+	Yes	Promoter_anchored_loops; Left_Ventricle; Liver; Right_Ventricle; GM12878; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC
MMP24	20	33814457	33864801	0.029	0	0	110	4.89E-05	2.26E-06	Skin_Not_Sun_Exposed_Suprapubic; Skin_Sun_Exposed_Lower_leg	-	Yes	IMR90; Mesenchymal_Stem_Cell; Mesendoderm; hESC
EIF6	20	33866714	33872788	0.782	0	0	174	6.34E-16	6.06E-67	Adipose_Subcutaneous; Whole_Blood; Heart_Left_Ventricle; Lung; Muscle_Skeletal; Nerve_Tibial; Skin_Not_Sun_ Exposed_Suprapubic; Skin_Sun_ Exposed_Lower_leg; Thyroid	+	Yes	Promoter_anchored_loops; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC

…Table S11 continued…

FAM83C	20	33873534	33880204	0.000	0	0	145	3.90E-06	5.73E-32	Esophagus_Mucosa; Skin_Not_Sun_Exposed_Suprapubic; Skin_Sun_Exposed_Lower_leg	-	Yes	Promoter_anchored_loops; IMR90; Mesenchymal_Stem_Cell; Mesendoderm; Trophoblast-like_Cell; hESC
GDF5	20	34021145	34042568	0.923	0	0	0	NA	NA	NA	NA	Yes	Promoter_anchored_loops; Adult_Cortex; Fetal_Cortex; IMR90; Mesenchymal_Stem_Cell; Trophoblast-like_Cell

Abbreviations: Chr, chromosome | Start/End, start/end position of the gene (hg19) | pLI, probability of being loss-of-function intolerant | posMapSNPs, number of SNPs mapped to genes based on positional mapping and CADD filtering | posMapMaxCADD, maximum CADD score of SNPs mapped by positional mapping | eqtlMapSNPs, number of SNPs mapped to genes based on eQTL mapping | eqtlMapMinP, minimum eQTL p-value of mapped SNPs | eqtlMapminQ, minimum eQTL FDR of mapped SNPs | eqtlMap tissues, tissue types of mapped eQTL SNPs (all GTEx v8) | eqtlDir, consequential direction of mapped eQTL SNPs (after aligning risk increasing alleles of the meta-analysis and the tested alleles in GTEx eQTLs) | ciMap, “yes” if gene is mapped by chromatin interactions, “no” if otherwise | ciMap tissues, tissue/cell typs of mapped chromatin interactions | NA, not available.

Table S12 | The top 50 genes identified in the gene-based analyses.

Gene	*P*	Chr	Start	Stop	NSNPs	N	Zstat
TMEM119	**7.66E-08**	**12**	**108973622**	**109027096**	**136**	**77865**	**5.25**
EDEM2	**2.39E-06**	**20**	**33693167**	**33900928**	**434**	**77036**	**4.57**
PROCR	5.43E-06	20	33724876	33775165	120	78243	4.40
MYH7B	1.54E-05	20	33528206	33600240	153	78126	4.17
TRPC4AP	3.85E-05	20	33580207	33715674	329	78045	3.95
LIMCH1	3.97E-05	4	41326624	41712061	887	77564	3.95
GSS	4.76E-05	20	33506236	33578620	127	78375	3.90
KIAA0020	5.20E-05	9	2710469	2879241	378	77406	3.88
BDKRB2	8.74E-05	14	96636016	96720666	298	77746	3.75
SLC19A2	9.44E-05	1	169423147	169490241	170	77408	3.73
IQGAP2	1.22E-04	5	75664074	76013957	958	77915	3.67
RP11-404P21.8	1.40E-04	14	96636181	96740266	378	77764	3.63
SELPLG	1.60E-04	12	109005686	109062735	105	77998	3.60
LRMP	1.70E-04	12	25138936	25271268	456	77564	3.58
SART3	2.92E-04	12	108906357	108990176	254	78151	3.44
SIX6	2.94E-04	14	60940669	60989568	103	75562	3.44
F5	3.08E-04	1	169473404	169590826	418	78066	3.42
SNAPC2	3.39E-04	19	7950201	7998135	115	77882	3.40
TGFBR3L	4.55E-04	19	7946030	7993982	102	77762	3.32
C14orf39	4.64E-04	14	60853187	61017261	331	75766	3.31
NUMBL	4.90E-04	19	41162596	41231877	180	77913	3.30
C7orf50	4.94E-04	7	1026622	1212896	823	77469	3.29
KIAA1324L	5.77E-04	7	86496222	86724015	594	77890	3.25
DDX47	6.23E-04	12	12931250	12992915	194	77253	3.23
SIX1	6.47E-04	14	61100133	61159977	89	74291	3.22
SIX4	7.38E-04	14	61166246	61226066	46	76123	3.18
MAP2K7	7.48E-04	19	7933728	7989363	115	77075	3.18
CYP2A6	7.57E-04	19	41339443	41391352	209	73869	3.17
INPP4A	7.95E-04	2	99026317	99220853	265	77725	3.16
ITPKC	8.27E-04	19	41188008	41256765	162	77912	3.15
TTC17	8.33E-04	11	43345482	43526483	339	77491	3.14
ADCK4	8.39E-04	19	41187434	41259112	173	77837	3.14
MNAT1	8.67E-04	14	61166460	61446671	266	75183	3.13
CTD-3193O13.9	8.79E-04	19	7923605	7974326	122	77039	3.13
APOLD1	9.13E-04	12	12843851	12992909	320	77266	3.12
ADCK1	1.05E-03	14	78231426	78411355	740	78031	3.08
UNCX	1.12E-03	7	1237543	1286954	53	77089	3.06
ZNF90	1.13E-03	19	20153803	20247885	307	75221	3.05
LRRC8E	1.13E-03	19	7918390	7976901	139	77165	3.05
FAM96B	1.24E-03	16	66955959	67003326	59	77660	3.03
DOCK9	1.28E-03	13	99435741	99773879	1006	78033	3.02
SOCS6	1.28E-03	18	67921137	68007436	300	77516	3.02
UAP1	1.40E-03	1	162496323	162579627	129	77451	2.99
AC010336.1	1.42E-03	19	7953689	8003427	130	77612	2.99
GPR20	1.44E-03	8	142356600	142412367	136	77813	2.98
KNG1	1.65E-03	3	186400065	186471743	303	77638	2.94
C7orf76	1.65E-03	7	96100938	96167835	142	78052	2.94
RRAD	1.81E-03	16	66945582	66994547	66	77909	2.91
COX16	1.88E-03	14	70781798	70861448	249	78223	2.90
SPRR4	1.89E-03	1	152908142	152955050	90	76829	2.90

Genome-wide significantly associated genes ($P < 2.63 \times 10^{-6}$) are written in bold. Abbreviations: Chr, chromosome | N, sample size | NSNPs, number of SNPs annotated to the gene | *P*, *p*-value |Start/Stop, gene annotation boundaries (hg19) | Zstat, gene z-value.

Table S13 | Gene-sets associated at $P < 10^{-3}$.

Gene Set	Ngenes	Beta	Beta_STD	SE	*P*
GO_cc:go_early_endosome	327	0.1721	0.0224	0.0482	1.79E-04
GO_bp:go_endothelial_cell_activation	10	1.1751	0.0270	0.3414	2.89E-04
Curated_gene_sets:ho_liver_cancer_vascular_invasion	12	0.7265	0.0183	0.2198	4.76E-04
GO_bp:go_righting_reflex	8	0.9486	0.0195	0.2891	5.19E-04
Curated_gene_sets:nikolsky_breast_cancer_20p13_amplicon	8	1.8013	0.0370	0.5533	5.68E-04
Curated_gene_sets:biocarta_agr_pathway	28	0.5564	0.0214	0.1737	6.79E-04
GO_bp:go_macromolecule_depalmitoylation	10	1.0357	0.0238	0.3255	7.34E-04
GO_mf:go_palmitoyl_protein_hydrolase_activity	10	1.0357	0.0238	0.3255	7.34E-04
GO_bp:go_positive_regulation_of_metanephric_glomerulus_development	4	1.1603	0.0168	0.3724	9.19E-04

Abbreviations: Ngenes, number of genes contained in the gene-set | Beta, regression coefficient of the gene-set | Beta_STD, semi-standardized regression coefficient | SE, standard error of the regression coefficient | *P*, *p*-value.

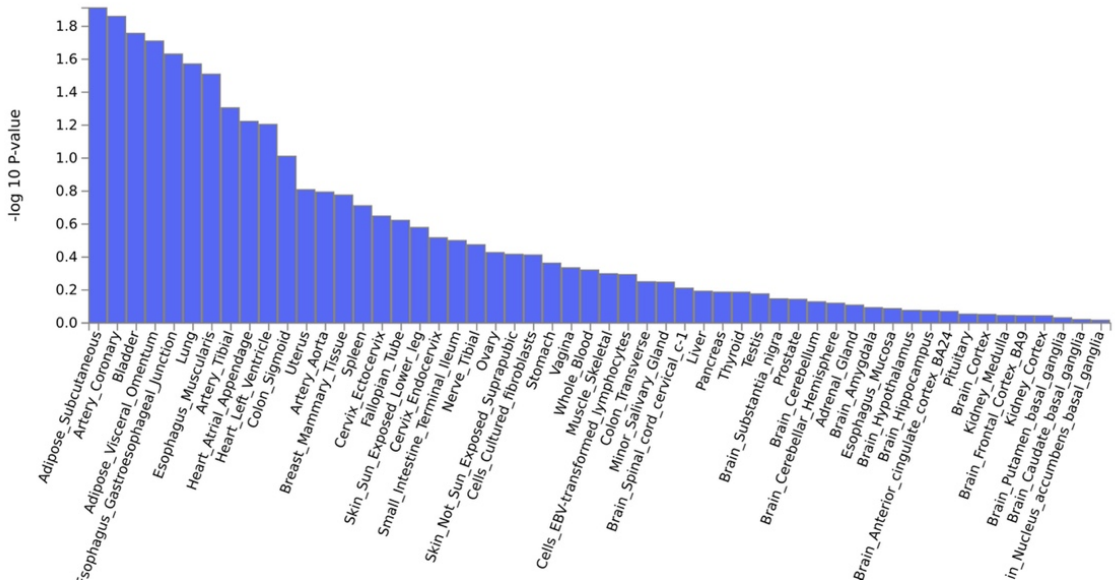

Figure S6 | MAGMA tissue expression analysis results.

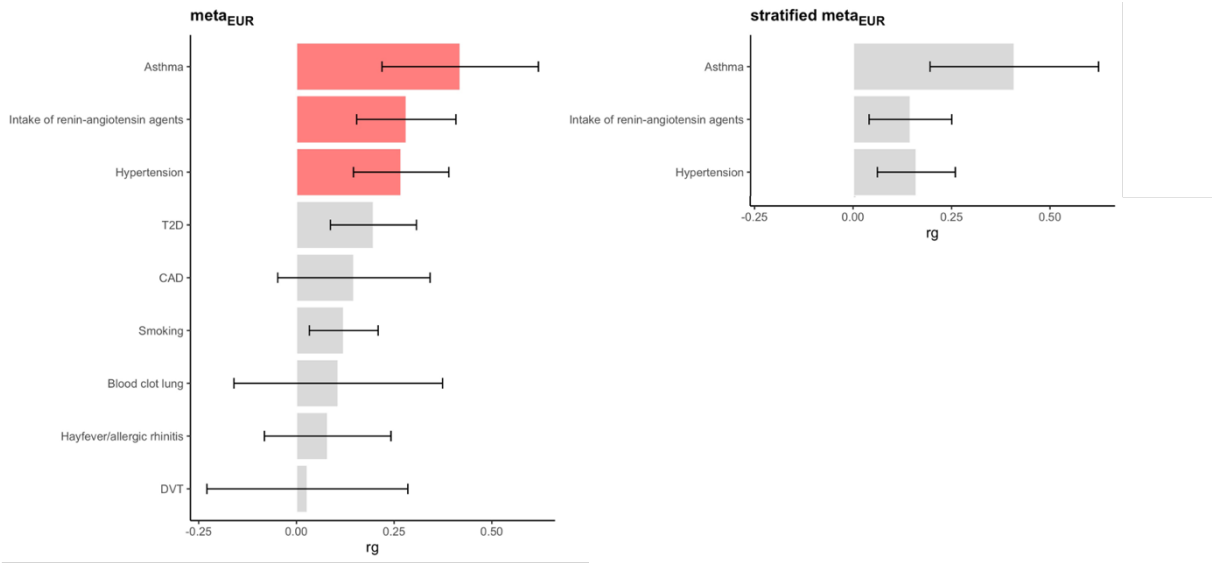

Figure S7 | Genetic correlation results.

Red bars indicate nominal significant ($P < 0.05$) results, while grey bars represent non-significant genetic correlations. The left panel displays the correlation results based on the $meta_{EUR}$ data. The right panel illustrates the re-analysis of the three traits that showed a nominal significant correlation with a stratified $meta_{EUR}$ analysis which contained only treatment-matched controls. Abbreviations: CAD, coronary artery disease | DVT, deep vein thrombosis | T2D, type 2 diabetes | r_g, genetic correlation.

Acknowledgements

This work would not have been possible without the contribution of several people, whom I would like to thank in the following.

First of all, I would like to express my sincere gratitude to my PhD supervisor Prof. Dr. Markus M. Nöthen for providing this exciting research topic, expert guidance, discussion and support, which contributed significantly to the success of the projects and this thesis. A very special thanks also goes to Prof. Dr. Andreas Forstner for stepping in as my mentor and fourth member of my thesis committee. In particular, I would like to thank you, Andreas, for engaging discussions, your constructive criticism and advice, which have contributed significantly to my professional and personal development. Many thanks for your always open ear and your great support during the last years.

Further I would like to thank Prof. Dr. Sven Cichon, as my second reviewer, and Prof. Dr. Julia Stingl, as third member of my thesis committee, for their time and consideration.

My warmest thanks go to Dr. Kerstin Ludwig, PD Dr. Stefanie Heilmann-Heimbach and Dr. Per Hoffmann for their various support during the preparation of this work. Furthermore, I would like to thank Dr. Christiane Stieber for her support during the initial phase of my PhD.

Thanks to all patients who participated in the vARIANCE study. Also, many thanks to all the clinicians who were involved in the recruitment for the vARIANCE study. I am very grateful for the collaborative effort that was essential to the success of the vARIANCE project. Moreover, I would like to thank Prof. Dr. Mia Wadlius, Dr. Eva Rye Rasmussen, Dr. Jonas Ghouse, and Dr. Lili Milani for the fruitful collaboration over the past years, which has contributed significantly to the success of this work.

A big thank you to Dr. Carlo Maj for the QC and imputation of the GWAS data and his support in the following analyses. In addition, a big thanks goes to the both of you, Julia Westmeier and Friederike David, for being my "bioinformatics helpdesk". Moreover, a special thanks to Annika Scheer for her support in the vARIANCE study office.

My warmest thanks go to all my current and former colleagues at the Life&Brain, the entire Institute of Human Genetics and the BfArM. Thank you for the very positive working atmosphere, the collegial and friendly cooperation and support from the very beginning.

A very special and heartfelt thanks goes to my docu room colleagues, especially Anna Si, Julia, Anna So, Friederike, Charlotte and Nora. Thank you for your friendship, your always open ear, encouragement, moral support and the many funny times we shared during the last years. I will be forever grateful that I got to share this PhD journey with all of you!

My deepest gratitude goes to my family, especially my parents and sister, and my friends for their love, unconditional support and encouragement during this journey. Thanks for enabling me to follow my own path and for always being there for me. Without your support I would not have made it to this day.