From the Department of Molecular Epidemiology,
Institute of Epidemiology II,
Helmholtz Center Munich, German Research Center for Environmental Health
Head: Prof. Dr. rer. biol. hum. Annette Peters

Is Acute-phase Serum Amyloid A Protein a Risk Factor for Type 2 Diabetes – Epidemiologic Perspective Including a Genetic Approach

Thesis
Submitted for a Doctoral Degree in Human Biology
at the Faculty of Medicine, Ludwig-Maximilians-University Munich, Germany

Carola Marzi
Waiblingen, Germany
2014

With approval of the Faculty of Medicine
Ludwig-Maximilians-University Munich, Germany

First reviewer:	Prof. Dr. rer. nat. Thomas Illig
Second reviewer:	Prof. Dr. Dr. h.c. Christian Haass
	Prof. Dr. Alexander Baethmann
	Prof. Dr. Martin Reincke
	Prof. Dr. Eva Grill, MPH
Thesis committee:	PD Dr. oec. troph. Barbara Thorand, MPH
	Dr. rer. nat. Holger Prokisch
Dean:	Prof. Dr. med. Dr. h.c. M. Reiser, FACR, FRCR

Date of the oral examination: 13.01.2014

Table of contents

Summary

Type 2 diabetes is a metabolic disorder with globally increasing prevalence. Therefore, the identification of etiological factors is of ascending relevance for the understanding, treatment, and prevention of the disease. Levels of the acute-phase serum amyloid A (A-SAA) protein have been found to be elevated in type 2 diabetic subjects, but little is known about their causal implication in the development of type 2 diabetes so far. This doctoral thesis presents an epidemiological perspective on the association between circulating levels of A-SAA and risk of type 2 diabetes and assesses a possible causality in this association using a genetic approach.

Three studies were conducted. In a **prospective cohort study,** A-SAA levels were measured in 836 initially non-diabetic, elderly, Western European subjects without clinically overt inflammation who participated in a seven-year follow-up examination. Results of this study provided first evidence that levels of A-SAA are elevated years before the manifestation of type 2 diabetes independent of other type 2 diabetes risk factors. However, adjustment for parameters related to glucose metabolism, particularly levels of 2h-glucose, attenuated the association suggesting a potential link via post-challenge hyperglycemia in the association between elevated levels of A-SAA and type 2 diabetes or, alternatively, a possible reverse causality between levels of A-SAA and 2h-glucose. In a **meta-analysis of genome-wide association studies** (GWAS) on levels of A-SAA conducted in three population-based studies and one prospective case-cohort study including a total of 4,212 participants of European descent two biologically highly plausible genetic susceptibility loci for A-SAA proteins at chromosome 11p15.5-p13 and chromosome 1p31 were identified. One of these loci represented a suitable candidate for a **Mendelian Randomization study**. In Mendelian Randomization studies, genetic variants are used as proxies for a biomarker. These studies benefit from the fact that genotypes are randomly assorted at meiosis and are largely independent of non-genetic confounding and disease processes. Thus, they constitute a genetic approach to assess whether the association between a biomarker and a disease is causal. The associations between genetic variants of the candidate locus and type 2 diabetes were extracted from the results of a meta-analysis of eight GWAS (8,130 cases, 38,987 controls) published by DIAGRAM, a large diabetes and genetic consortium. In spite of sufficient power, the above mentioned associations were not significant suggesting that there are genetic mechanisms that raise plasma levels of A-SAA without translating into an increase in type 2 diabetes risk.

In conclusion, results of this doctoral thesis indicated that levels of A-SAA are elevated years before the manifestation of type 2 diabetes but could not provide evidence that the association is truly causal using a genetic approach. Rather it seems likely that the association between levels of A-SAA and risk of type 2 diabetes is substantially influenced by post-challenge hyperglycemia. Time-series studies are warranted to elucidate the role of post-challenge hyperglycemia in this association.

Zusammenfassung

Typ 2 Diabetes ist eine Stoffwechselerkrankung mit weltweit zunehmender Prävalenz. Daher ist die Identifizierung ätiologischer Faktoren für Verständnis, Behandlung und Prävention der Erkrankung von steigender Bedeutung. In Typ 2 Diabetikern wurden erhöhte Spiegel des Akut-phase Serum Amyloid A (A-SAA) Proteins festgestellt, bisher ist allerdings wenig über ihre kausale Mitwirkung an der Entstehung von Typ 2 Diabetes bekannt. Diese Dissertation bietet eine epidemiologische Sicht auf den Zusammenhang zwischen zirkulierenden A-SAA Spiegeln und das Typ 2 Diabetesrisiko und untersucht mithilfe eines genetischen Ansatzes eine mögliche Kausalität dieses Zusammenhangs.

Drei Studien wurden durchgeführt. In einer **prospektiven Kohortenstudie** wurden A-SAA Spiegel in 836 anfangs nicht-diabetischen, älteren, westeuropäischen Probanden ohne klinisch offenkundige Inflammation, die an einer sieben Jahre späteren Anschlussuntersuchung teilnahmen, gemessen. Ergebnisse dieser Studie gaben erste Hinweise darauf, dass A-SAA Spiegel unabhängig von anderen Typ 2 Diabetes Risikofaktoren vor der Manifestation von Typ 2 Diabetes erhöht sind. Eine Adjustierung für Parameter des Glukosestoffwechsels, insbesondere für 2h-Glukose Spiegel, schwächten jedoch den Zusammenhang ab, was auf einen potentiellen Link zwischen erhöhten A-SAA Spiegeln und Typ 2 Diabetes über post-challenge Hyperglykämie oder, alternativ, eine mögliche umgekehrte Kausalität zwischen A-SAA-und 2h-Glukose Spiegeln hindeutet. In einer **Metaanalyse genomweiter Assoziationsstudien** (GWAS) zu A-SAA Spiegeln, die in drei populationsbasierten und einer prospektiven Fall-Kontroll-Studie mit insgesamt 4,212 Teilnehmern europäischer Herkunft durchgeführt wurde, wurden zwei biologisch sehr plausible A-SAA Prädispositionsorte auf Chromosom 11p15.5-p13 und Chromosom 1p31 gefunden. Einer davon stellte einen geeigneten Kandidaten für eine **Mendelsche Randomisierungsstudie** dar. In Mendelschen Randomisierungs-studien werden genetische Varianten stellvertretend für Biomarker untersucht und die Tatsache genutzt, dass Genotypen in der Meiose zufällig vererbt werden und weitgehend unabhängig von nicht-genetischen Störfaktoren und Krankheitsprozessen sind. Somit stellen sie einen genetischen Ansatz zur Untersuchen der Kausalität zwischen einem Biomarker und einer Erkrankung dar. Die Assoziationen zwischen genetischen Varianten des Kandidatenortes und Typ 2 Diabetes wurden den Ergebnissen einer von DIAGRAM, einem großen Diabetes- und Genetikkonsortium, veröffentlichten Metaanalyse von acht GWAS (8,130 Fälle, 38,987 Kontrollen) entnommen. Sie waren trotz ausreichender Power nicht signifikant, was darauf hinweist, dass es genetische Mechanismen gibt, die A-SAA Spiegel erhöhen, ohne zu einem Anstieg des Typ 2 Diabetesrisikos zu führen.

Zusammenfassend zeigen die Ergebnisse dieser Dissertation, dass A-SAA Spiegel Jahre vor der Entstehung von Typ 2 Diabetes erhöht sind, dass der Zusammenhang aber tatsächlich kausal ist, konnte mithilfe eines genetischen Ansatzes nicht gezeigt werden. Es scheint vielmehr wahrscheinlich, dass der Zusammenhang zwischen A-SAA Spiegeln und Typ 2 Diabetesrisiko wesentlich durch post-challenge Hyperglykämie beeinflusst wird. Zeit-Reihen Studien sind erforderlich, um die Rolle von post-challenge Hyperglykämie in diesem Zusammenhang zu klären.

1. Scientific Background

1.1. Type 2 diabetes

1.1.1. Overview of type 2 diabetes

Diabetes mellitus

Diabetes mellitus is a heterogeneous group of metabolic disorders characterized by chronic hyperglycemia [1]. First classical symptoms of diabetes mellitus are weight loss, polyuria, polydipsia and blurring of vision [2]. They may be absent or not severe for a long time before diabetes mellitus is diagnosed. Diabetes mellitus may cause severe long-term micro- and macrovascular damage eventually leading to the dysfunction or failure of heart, blood vessels, eyes, kidney and nerves [1, 2]. The clinically most relevant consequences of untreated diabetes mellitus are retinopathy and blindness, diabetic nephropathy and renal failure, neuropathy and foot ulcers, amputations, and autonomic dysfunction, stroke, coronary artery and peripheral vascular disease [1-3].

Types of diabetes mellitus

Based on etio-pathogenesis diabetes mellitus can be classified into different types [1, 2]. Type 1 diabetes is characterized by deficient insulin production primarily caused by an immune-mediated destruction of pancreatic β-cells. It accounts for about 5 % of diabetes mellitus cases. Type 2 diabetes comprises about 90 % of all diabetes mellitus cases thus constituting the most common form of diabetes mellitus. It is mainly the result of a decreased insulin sensitivity which may not be compensated for by an increased insulin secretion at a later stage of disease pathogenesis. Another type of diabetes mellitus is gestational diabetes with a prevalence of 2-5 % among pregnant women in Europe. It is characterized by a carbohydrate intolerance resulting in hyperglycemia with onset or first recognition during pregnancy. Further types of diabetes mellitus include late-onset autoimmunity diabetes in the adult (LADA) which is a slow-onset metabolic disorder with characteristics of type 1 diabetes, other types of diabetes mellitus prone to genetic defects in insulin action or the β-cells such as maturity-onset diabetes in the young (MODY), diseases of the exocrine pancreas, endocrinopathies, and chemically or pharmaceutically induced impairments of glucose homeostasis [1, 2].

Diagnostic criteria for type 2 diabetes

Type 2 diabetes can be diagnosed on the basis of an oral glucose tolerance test (OGTT). In an OGTT, venous plasma glucose levels are measured in the fasting state and 2 hours after the intake of 75 g anhydrous glucose [1, 4]. Additionally, type 2 diabetes can be diagnosed based on the presence of classical symptoms of hyperglycemia (weight loss, polyuria, polydipsia, and blurring of vision) accompanied by occasionally elevated venous plasma glucose concentrations [1, 4]. An impaired glucose homeostasis may be present a long time before the onset of type 2 diabetes. It may appear in

two pre-diabetic conditions of intermediate hyperglycemia, impaired fasting glucose (IFG) and impaired glucose tolerance (IGT). Individuals with IFG or IGT are at high risk for developing type 2 diabetes as well as long-term complications [1, 4, 5]. Therefore, the identification of these individuals is of major importance. Like type 2 diabetes, the two pre-diabetic conditions IFG and IGT can be diagnosed on the basis of an OGTT [6]. Cut-off levels of venous plasma glucose for the diagnosis of type 2 diabetes and intermediate hyperglycemia according to the current recommendations of the World Health Organization (WHO) [4] are presented in Table 1.

Differing from WHO recommendations, the American Diabetes Association (ADA) recommends a lower fasting venous plasma glucose cut-off value of 100 mg/l (5.6 mmol/l) for the diagnosis of IFG to meet a potential clinical relevance in particular for individuals at high risk [6, 7].

Table 1. Cut-off levels of venous plasma glucose for the definition and diagnosis of type 2 diabetes and intermediate hyperglycemia

	Type 2 diabetes	IFG	IGT
Fasting glucose **2h-glucose**	≥ 126 mg/dl (≥ 7.0 mmol/l) or ≥ 200 mg/dl (≥ 11.1 mmol/l)	110-125 mg/dl (6.1-6.9 mmol/l) and < 140 mg/dl (< 7.8 mmol/l)	< 126 mg/dl (< 7 mmol/l) and 140-199 mg/dl (7.8-11.0 mmol/l)
Random glucose*	≥ 200 mg/dl (≥ 11.1 mmol/l)		

IFG= impaired fasting glucose, IGT= impaired glucose tolerance
* In patients with classical symptoms of hyperglycemia or hyperglycemic crisis

HbA1c - another diagnostic marker for type 2 diabetes?

Recently, the ADA also expanded its recommendations for the diagnosis of type 2 diabetes to hemoglobin A1c (HbA1c) and defined HbA1c values ≥ 6.5 % as clinically relevant for the diagnosis of type 2 diabetes [6]. HbA1c is the percentage of the glycated fraction of hemoglobin. It is primarily measured to identify the average plasma glucose concentration over a prolonged period of time (2-3 months). Thus, it constitutes a useful measure for monitoring the progress of diabetes mellitus in therapy [4, 8]. The major advantage of measuring HbA1c is that it does not require fasting blood probes or an OGTT and that it integrates both, fasting blood glucose concentrations as well as peaks of post-prandial blood glucose concentrations. However, the definition of type 2 diabetes on the basis of HbA1c has been discussed controversially [4, 6, 8, 9]. Apart from higher costs, the diagnosis of type 2 diabetes on the basis of HbA1c values leads to a different classification of type 2 diabetes cases compared to the diagnosis based on an OGTT [10] with implications for the individual and population

prevalence estimates. Moreover, it has been shown that the association between levels of blood glucose and HbA1c may be different in different ethical populations [4, 11].

Epidemiology of type 2 diabetes

Before presenting prevalence and incidence estimates for type 2 diabetes, it should be noted that global estimations are based on country reports which may differ in quality. Furthermore, for several countries no population-based data on diabetes prevalence and undetected cases are available making it hard to estimate global prevalence and incidence precisely [12]. However, dimensions and tendencies are without controversy.

According to estimations of the International Diabetes Federation (IDF), the global prevalence for diabetes and IGT in adults aged 20-79 years was 8.3 % which corresponds to some 366 million people in 2011 [13]. These estimations include 50 % undiagnosed or unreported diabetic cases [13]. Country-specific estimates for diabetes prevalence vary with most diabetic people living in the economically less developed regions [13].

Diabetes prevalence has increased dramatically in the past and is expected to rise further in the future. It is assumed that this development is possibly driven by population growth and aging, increasing urbanization, dietary changes, and a reduced physical activity [14, 15]. Data extrapolation resulted in an estimated rise in global prevalence to 9.9 % corresponding to 522 million individuals with manifest diabetes by 2030 [13]. This equates to a cumulative incidence of ten million cases per year [13]. Like prevalence numbers incidence rates are geographically heterogeneous; whereas the increase in type 2 diabetic patients between 2010 and 2030 is estimated to be 20 % in Europe, it is estimated to be 42 % in North America and even 72-98 % in Africa, the Middle East, and South-East Asia [13].

In Europe, the number of diabetic people is thought to be 52.6 million corresponding to 8.1 % of the adult population [13]. Thus, Europeans have a moderate prevalence of diabetes compared to other ethnic groups. Similarly, in Germany the number of diabetic patients was estimated to be about 8.2 % (95 % confidence interval (CI), 7.3-9.2 %) of the adult population [16]. Within Germany, the prevalence of diabetes is geographically heterogeneous and dispays an imbalanced gradient from North to South and from East to West with highest prevalence in North-East Germany [16]. In a population-based study conducted in the region of Augsburg, Southern Germany, about 50 % of type 2 diabetes cases was previously undetected and newly diagnosed in participants aged 35 years and older [17, 18]. However, due to the geographical differences in the German population, these data estimates may not be extrapolated to the whole of Germany. Moreover, the described study did not differentiate between the different types of diabetes making it difficult to define the proportion of estimates attributed to type 2 diabetes.

1.1.2. Type 2 diabetes risk factors

Epidemiological studies have provided evidence for a variety of different type 2 diabetes risk factors [19]. These risk factors can be classified into non-modifiable and modifiable risk factors. The most important non-modifiable risk factors are age and genetic predisposition; adiposity and physical inactivity are the most relevant modifiable risk factors [20].

Age

Globally, the prevalence of type 2 diabetes and glucose intolerance increases with age [13, 21, 22]. Results of time-series analyses conducted in 6,538 participants (35-55 years of age) of a prospective cohort study (the Whitehall-II study) showed a decrease in insulin sensitivity with age during 13 years of follow-up and thus provided evidence for a direct association between age and insulin sensitivity [23]. In some studies, however, the decrease in insulin sensitivity with age was attributed to an increase in intra-abdominal fat [24, 25]. Contrary to the findings on insulin sensitivity, deficiencies in ß-cell function had been consistently demonstrated in aging humans independent of adiposity measures [25, 26].

Genetic determinants

A comparison of type 2 diabetes prevalence across different ethnicities living in a similar environment suggests a possible genetic contribution to the pathogenesis of type 2 diabetes [27, 28]. For example, Asian Americans have been reported a 60 % higher type 2 diabetes prevalence adjusted for body-mass index (BMI) than non-Hispanic white Americans in a great population-based survey conducted in the United States [27]. Moreover, family studies have shown that first degree relatives of individuals with type 2 diabetes are about 3 times more likely to develop the disease than individuals without a positive family history [29]. Up to now, more than 50 genetic susceptibility loci for type 2 diabetes have been identified [30-32]. Of those, common variants of the transcription factor 7-like 2 gene *(TCF7L2)* showed the strongest replicated effect on type 2 diabetes risk across different ethnicities with odds ratios (OR)s varying between 1.3 and 1.6 in an additive genetic model [33-36]. The majority of type 2 diabetes genetic susceptibility loci has been identified in large genome-wide scans, only two, potassium inwardly-rectifying channel, subfamily J, member 11 (*KCNJ11*), and peroxisome proliferator-activated receptor gamma (*PPARγ*), were reproducibly found in candidate gene studies (chapter 1.3.2) [37, 38]. Although the identified type 2 diabetes susceptibility loci explain only about 10 % of the estimated heritability so far [30, 39], some insights have been gained. For one thing, results of genetic studies on type 2 diabetes demonstrated that the primary genetic effect on type 2 diabetes susceptibility seems to be mediated through deficiencies in insulin secretion rather than insulin action [37, 40]. Moreover, some of the identified type 2 diabetes loci also harbor rare

mutations that are causal for monogenic types of diabetes or exert pleiotropic effects indicating an overlap in disease etiology [40].

Adiposity and physical inactivity

The most important modifiable type 2 diabetes risk factor is adiposity. In addition to adiposity per se, lifestyle and environmental conditions such as chronic excess nutrition and physical inactivity contributing to adiposity are of major relevance in the context of type 2 diabetes [41, 42].

Adipose tissue is an endocrine organ and releases amongst others free fatty acids, hormones, pro-inflammatory cytokines such as interleukin-6 (IL-6) and tumor necrosis factor-alpha (TNF-α), and acute-phase reactants like plasminogen activator inhibitor-1 (PAI-1) and serum amyloid A (SAA) [43-46]. These adipokines interfere with different metabolic pathways which might induce, promote or influence adverse metabolic conditions contributing to type 2 diabetes pathogenesis (see also 1.2.1.).

Generally used measures for adiposity include BMI and waist-to-hip ratio (WHR). BMI is calculated as weight [kg] divided by the square of height [m] and is a measure of adiposity regardless of body fat distribution. In contrast, WHR is the ratio between the circumferences of the waist to that of the hip and reflects abdominal fat. The question which adiposity measure is more strongly associated with type 2 diabetes has been discussed controversially [47-49].

Large part of type 2 diabetes risk is attributable to adiposity. For example, in a multi-ethnic cohort study conducted in 74,970 subjects aged between 45 and 75 years of Caucasian, Native Hawaiian, and Japanese American ethnicity about half of the incident type 2 diabetes cases were calculated to be attributed to a BMI $\geq$ 25 kg/m² [50]. Closely related to adiposity is physical inactivity. According to a global estimation of the year 2012, physical inactivity accounted for 7 % (95 % CI, 3.9-9.6) of type 2 diabetes cases [51]. It is thought that physical activity directly impacts on glucose homeostasis via activation of glucose transporters in muscle and adipose tissue as well as enhanced glycogen synthase resulting in an increased non-oxidative disposal of glucose [52-54].

Further type 2 diabetes risk factors

Apart from these risk factors several more type 2 diabetes risk factors have been identified. For example, active and passive smoking [55, 56] and non-moderate alcohol intake [57] constitute two more modifiable lifestyle risk factors.

Moreover, low birth weight reflecting fetal malnutrition has been found to be associated with IGT and type 2 diabetes pointing towards an influence of intrauterine environment on the development of type 2 diabetes later in life [58, 59].

In the last years, evidence arose that also psychosocial factors play a role in the development of type 2 diabetes. It has been shown that job strain, depression, or - closely connected to both - quality of sleep contribute to diabetes risk [60-63]. Furthermore, low education, low income and low employment all

of them representing a low socio-economic status were found to be significantly associated with an increased type 2 diabetes risk in industrialized countries [64-67].
Finally, prospective studies have shown an inverse association between serum 25-hydroxyvitamin D and insulin resistance [68, 69].

Prevention of type 2 diabetes via reduction of modifiable risk factors
Representing targets for primary prevention, type 2 diabetes risk factors that are at least in part modifiable are of major public health relevance [50, 52, 70-72]. Several intervention studies aiming on the reduction of weight, a balanced diet with a low total intake of fat and saturated fat and an increased intake of fiber, as well as an increased physical activity demonstrated a significantly reduced progression of type 2 diabetes in individuals at high risk. For example, an intervention study conducted in 557 Chinese subjects with IGT who were randomly assigned to a diet group, an exercise group, a diet-and-exercise group, and a control group showed a significant more than 30 % reduction in risk for developing type 2 diabetes within 6 years of follow-up in any of the intervention groups compared to the control group [73]. In another intervention study, the Finish Diabetes Prevention Study, conducted in 522 overweight subjects with IGT a randomly assigned intervention group received individualized counselling aimed at reducing weight by diet and physical activity. In this group, the risk for developing type 2 diabetes was reduced by 58 % within 3.2 years compared to a control group [74].

1.1.3. Pathogenic mechanisms of type 2 diabetes

Given the diversity of factors contributing to type 2 diabetes risk the question arises whether all these risk factors are based on certain common underlying concepts [20].

Framework of glucose homeostasis
In order to understand putative concepts, it is necessary to first describe the principal framework of glucose homeostasis. In normoglycaemic subjects, levels of blood glucose are sustained stable within a narrow range between 60-126 mg/dl. Insulin is the key hormone for the regulation of blood glucose and in a healthy organism normoglycaemia is maintained by the balanced interplay between insulin action and insulin secretion [75]. Insulin is synthesized in the ß-cells within the islets of Langerhans, an endocrine region of the pancreas. It is secreted directly in the blood flow and exerts its effect on insulin-dependent cells through binding to insulin receptors on the surface of the cells. The insulin receptor in return phosphorylates itself and several substrates thus starting the downstreaming signaling and promoting the uptake of glucose into peripheral insulin-depending tissue such as muscle and adipose tissue [76, 77]. Furthermore, insulin inhibits hepatic glycogen breakdown and

gluconeogenesis. Insulin resistance is a state in which the effect of insulin on its target cells is less than expected. It can be partly compensated for as healthy pancreatic ß-cells can adapt to changes in the insulin action [78, 79]. This means that a decreased insulin action is compensated by a higher insulin secretion. However, the compensatory concept does not work unlimited. When insulin action decreases, concentrations of blood glucose increase and although this increase may be small, over time it promotes glucose toxicity and eventually leads to ß-cell dysfunction in itself [75]. Time-series analyses in the Whitehall-II study showed that the insulin sensitivity decreased slightly in all participants during the 13 years of follow-up, while ß-cell function displayed a short increase followed by a rapid decrease in incident type 2 diabetes cases four years before diagnosis [23]. This demonstrates that ß-cell dysfunction is a critical feature in type 2 diabetes pathogenesis.

Pathogenic mechanisms

The main mechanisms to explain insulin resistance and ß-cell dysfunction discussed today are glucotoxicity and lipotoxicity. These features describe the pathological consequences of increased glucose and lipid concentrations in blood. On an intracellular level, glucotoxicity and lipotoxicity might possibly lead to mitochondrial dysfunction and an increase of reactive oxidative species (ROS), endoplasmic reticulum stress, and subclinical inflammation [80-82].

Hyperglycemia may affect different metabolic pathways, molecules, and cells, thus increasing the risk for the development of type 2 diabetes as well as micro- and macrovascular complications of the disease. For example, it is thought to activate protein kinase C (PKC) isoforms, increase levels of glucose-derived advanced glycation end products, and to increase glucose flux through the aldose reductase pathway in vascular endothelium and in adipocytes [83]. Beside other effects, ultimately, these features might possibly result in an increase of mitochondrial ROS as well as an activation of pro-inflammatory signaling [83, 84].

Lipotoxicity is caused by an imbalance between the production of lipids and their oxidation or transport. This imbalance may lead to amplification, functional disturbance, and morphological changes in adipose tissue as well as ectopic lipid deposition in the muscle, liver, and pancreas with functional implications for these tissues [85].

Mitochondrial dysfunction is thought to be one of the results of an intracellular increase of triglycerides and concentrations of intermediate metabolites related to lipid metabolism as a consequence of deficient lipid oxidization. Mitochondrial dysfunction is assumed to activate pro-inflammatory signaling and to inhibit insulin signaling resulting in a reduced uptake of glucose in target cells [85-88]. Thus, mitochondrial dysfunction is discussed to represent the linking mechanism between lipotoxicity and insulin resistance.

ROS is assumed to be a by-product of mitochondrial respiration during times of excess intracellular nutrient availability. Amongst others, an increase of ROS possibly causes endoplasmic reticulum stress. Endoplasmic reticulum stress is characterized by slow protein folding and, ultimately, an

unfolded protein response which in return possibly leads to additional production of ROS. This has been studied in different cell lines including adipocytes [84, 89]. Moreover, endoplasmic reticulum stress has also been reported to be present in pancreatic ß-cells, especially in insulin resistant individuals where the flux of proteins through the endoplasmic reticulum is higher compared to normoglycaemic individuals [90]. Endoplasmic reticulum stress is thought to possibly lead to reduced insulin sensitivity as a consequence of a reduced ability to increase glucose uptake in response to insulin [91]. Furthermore, it is assumed to trigger the activation of inflammatory signaling by activating nuclear factor κB (NF-κB) which in return may possibly lead to the induction of several cytokines like IL-6 and the suppression of insulin signaling in insulin-dependent target cells [91].

Subclinical inflammation seems to be a common feature in the concepts described. Moreover, many type 2 diabetes risk factors such as adiposity, physical activity, alcohol intake, and lipids, but also risk factors which are not prone to excess or malnutrition like age, psychosocial stress, and quality of sleep have been reported to be correlated with inflammatory markers [20]. Therefore, it has been postulated that type 2 diabetes might actually be an inflammatory disease [92].

1.2. Acute-phase serum amyloid A – a measure of subclinical inflammation

1.2.1. Subclinical inflammation

Inflammation is the first innate immune response to inflammatory stimuli such as injury, infection, trauma, and stress. The cardinal signs of inflammation are redness, swelling, heat, pain, and loss of function. In healthy individuals, the acute inflammatory process is transient and has mainly beneficial effects in restoring homeostasis. However, inflammation is a "friendly fire" [93]. Persistent, excessive or recurrent inflammatory stimuli may result in chronic subclinical inflammation, a condition of a low-graded increase of immune mediators in blood in individuals without classical symptoms of inflammation [92, 94, 95]. It is thought that subclinical inflammation may affect the body's susceptibility to a number of severe chronic diseases such as type 2 diabetes, atherosclerosis and their clinical complications [93-95].

Evidence for an implication of subclinical inflammation in type 2 diabetes

First evidence of an association between subclinical inflammation and type 2 diabetes was given about 50 years ago in the early 1960s when cross-sectional studies found that immune mediators were up-regulated in type 2 diabetic cases [96]. Later studies presented evidence that a large range of immune mediators are already up-regulated in individuals with IGT and years before the onset of overt type 2 diabetes. These immune mediators include high sensitivity-C reactive protein (hs-CRP), IL-6, TNF-α, leucocyte count, interleukin-1 (IL-1), PAI-1, factor VIII, von Willebrand factor, orosomucoid, fibrinogen, sialic acid, and gamma-globulin [94, 95]. The prospective studies provided a hint that subclinical inflammation might be a cause rather than a consequence of the disease. This might be the case particularly in women: In a case-cohort study conducted in 7,936 subjects aged 35-74 years at baseline who participated in the population-based Monitoring of Trends and Determinants in Cardiovascular Disease/Cooperative Research in the Region of Augsburg (MONICA/ KORA) study, it has been shown that the association between hs-CRP and incident type 2 diabetes was sex specific and more pronounced in women [97]. However, not all cross-sectional and prospective studies stratified their analyses by sex. Furthermore, results of some studies have been partly conflicting regarding confounding by adiposity measures and baseline glycaemia [98].

Genetic association studies further supported a pathogenic role of immune mediators in type 2 diabetes as they reported associations between diabetes and polymorphisms in functional regions of genes involved in the immune response such as in the promoter region of the pro-inflammatory cytokines TNF-α and IL-6 [99-102]. Furthermore, the Ala allele of the common Pro12Ala polymorphism in isoform 2 of *PPARγ* has been robustly found to be associated with reduced risk for type 2 diabetes in numerous studies [103]. *PPARγ* is a transcription factor with some anti-inflammatory activities. The reduction in transcriptional activity of *PPARγ* is thought to modulate the

production and release of adipose-derived factors, including the pro-inflammatory cytokine TNF-α as well as resistin which has been shown to increase the expression of several pro-inflammatory cytokines [104].

Moreover, immune gene disruption or transgenic overexpression of immune mediators have been found to be associated with protection from diabetes or higher diabetes prevalence in animal studies [95]. For example, the absence of TNF-α improved insulin sensitivity in obese mice and protected from the obesity-related reduction in insulin receptor signaling in muscle and fat tissue of these animals [105]. In another study, mice lacking the gene encoding PAI-1 were protected from developing obesity and insulin resistance after a high-fat diet [106].

Finally, a number of intervention studies showed a decrease in inflammatory mediators after weight reduction and/ or improved physical activity which was correlated with an improved glucose metabolism [95]. For example, one study conducted in 19 obese and 10 lean men (initial mean BMI 38.6 kg/m², SD 0.6 kg/m² and 23.4 kg/m², 0.4 kg/m², respectively) found a correlation between changes in IL-6 but not interleukin-8 (IL-8) with an improved insulin sensitivity after 16 weeks of energy-restricted diet and a subsequent weight loss in the obese men [107]. Intervention studies also reported on a decrease in inflammatory mediators after the treatment with anti-diabetic agents. In a review on inflammation and type 2 diabetes Kolb and Mandrup-Poulsen (2005) presented evidence for an effect of sulphonylurea on TNF-α, metformin on hs-CRP, glitazones on hs-CRP, SAA, TNF-α, soluble CD40 ligand, and PAI-1, and insulin on hs-CRP, IL-1, TNF-α, soluble intercellular adhesion molecule-1 (ICAM-1), monocyte chemotactic protein-1 (MCP-1), and PAI-1 [95]. However, interventions need to target inflammatory or glucose pathways only and may not have pleiotropic effects to allow deductions on the direction of causation. For example, in a randomized controlled trial undertaken in 27 type 2 diabetic subjects (mean age 63.0 +/- 1.7 years, HbA1c 8.8 +/-0.3 %, BMI 32.7 +/- 0.8 kg/m2, therapy duration 15.2 +/- 1.4 years, insulin dose 73.3 +/-7.0 U/day) receiving daily either 400 mg troglitazone or placebo for 16 weeks glucose metabolism was improved and levels of circulating SAA were decreased in the intervention group [108]. Troglitazone acts as insulin sensitizer and has been found in cell lines and animal models to decrease hepatic glucose output by decreasing the rate of gluconeogenesis in the liver or by increasing glycolysis [109]. Being a ligand particularly to PPARγ it has also been found to be associated with a decrease of NF-κB and an increase of IκB thus modulating the transcription of a number of pro-inflammatory genes [110, 111]. The implications of troglitazone into both, glucose metabolism as well as inflammation, make it difficult to determine cause and consequence.

Inflammatory pathways in insulin signaling

Several inflammatory pathways which are initiated by extracellular mediators like cytokines and lipids as well as by intracellular conditions like endoplasmic reticulum stress and ROS are thought to lead to the attenuation of insulin signaling and insulin action (Figure 1).

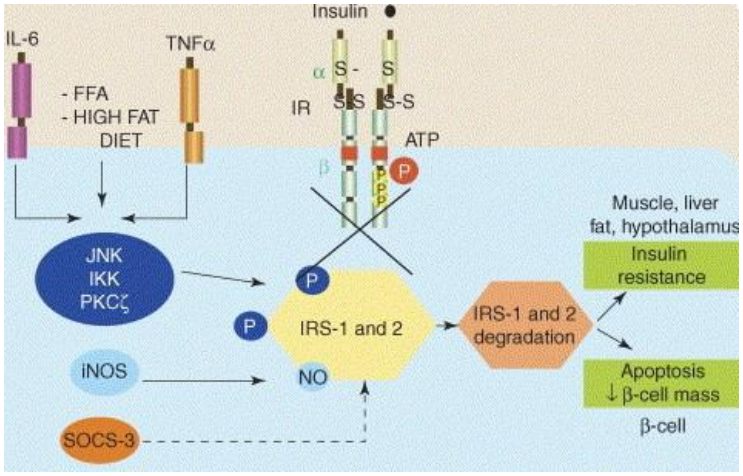

Figure 1: Attenuation of insulin signaling and insulin action by post-translational modification of insulin receptor substrates. *Modified according to Schmid MI et al., 2005 [112].*
FFA - free fatty acids; IL-6 - interleukin-6; TNA-α - tumor necrosis factor-α, JNK – JunkN-terminal kinase, IKK – inhibitor of NF-κ B kinase, PKC-θ -protein kinase C-θ, iNOS - inducible nitric oxide synthase, NO – nitric oxide, SOCS-3 – suppressor of cytokine signaling 3, IRS-1 and 2 – insulin receptor substrate 1 and 2.

It has been shown in vitro that exposure of cells to the pro-inflammatory cytokine TNF-α stimulates the phosphorylation of serine residues of the insulin receptor substrate-1 (IRS-1) which prohibits both tyrosine phosphorylation of IRS-1 in response to insulin and the ability to associate with the insulin receptor [113]. Several kinases including Junk N-terminal kinase (JNK), inhibitor of NF-κB kinase (IκB), and protein kinase C-θ (PKC-θ) are thought to be essential to the inflammatory signaling. These kinases possibly activate inflammatory gene expression as well as directly inhibit insulin signaling [85, 94]. Furthermore, members of the protein suppressor of cytokine signaling (SOCS) family have been found to be implicated in inflammation-induced insulin resistance in cell line studies and animal models [85, 94]. It is hypothesized that they inhibit insulin signaling by interfering with IRS phosphorylation or proteosomal degradation [114]. Moreover, inflammation is discussed to lead to an overproduction of nitric oxide which is thought to contribute to both, reduced insulin sensitivity in muscle cells as well as an impaired ß-cell function [85]. The modulation of these pathways might connect subclinical inflammation with insulin resistance and reduction in β-cell mass.

Inflammation and loss of ß-cell function

Amyloid deposition has been reported to be present in type 2 diabetes pancreatic islet tissue [90]. Furthermore, it has been shown that levels of cytokines and chemokines were up-regulated in pancreatic islets of patients with type 2 diabetes and also animal models consistently revealed an infiltration of immune cells in type 2 diabetes pancreatic islets [90]. All these points towards an implication of inflammation into the loss of ß-cell function. One possible mechanism is the interleukin-1ß (IL-1ß)/ NF-κB pathway [115]. It is assumed that exposure to prolonged or recurrent high glucose levels increases the production of IL-1ß at large extents which in return promotes Fas-triggered apoptosis partly via NF-κB activation [115]. Furthermore, it possibly increases the number of cytokines leading to enhanced infiltration of macrophages, which in return might then further induce IL-1ß expression as well as other cytotoxic factors possibly responsible for ß-cells dysfunction and failure [90]. The process is thought to be amplified by IL-1ß capacity of auto-induction as well as by the presence of free fatty acids which further increase IL-1ß expression and the release of IL-6 and IL-8 in pancreatic islets [90].

1.2.2. Acute-phase serum amyloid A

This chapter is partly extracted from the manuscripts Marci C. et al., ***Acute-phase serum amyloid A protein and its implication in the development of type 2 diabetes in the KORA S4/ F4 study***, Diabetes Care, 2012 Dec 13, epub ahead of print, and Marzi, C. et al., ***Genome-Wide Association Study Identifies Two Novel Regions at 11p15.5-p13 and 1p31 with Major Impact on Acute-Phase Serum Amyloid A,*** PloS Genetics PLoS Genet. 2010 Nov 18;6(11):e1001213.

SAA protein is a sensitive marker of subclinical inflammation. It is an apo-lipoprotein mainly associated with high-density lipoprotein (HDL) in plasma. SAA isoforms are expressed constitutively (C-SAA) and show a rapid increased expression in response to inflammatory stimuli such as infection, tissue injury, trauma, and stress during the acute phase (A-SAA) [116, 117]. The high inductive capacity along with a high conservation of genes and proteins throughout evolution of vertebrates and invertebrates suggests that A-SAA plays a key role in pathogen defense and probably functions as an effector molecule of the immune-system [116].

Functional aspects of A-SAA

The precise role of A-SAA proteins in host defense during inflammation has not been completely elucidated so far. Three potential clinically important functions are proposed. First of all, A-SAA proteins have been found to induce extracellular-matrix degrading enzymes such as collagenase, which are essential in repair processes after tissue damage. Furthermore, A-SAA seems to be involved

in lipid metabolism and transport by reversing cholesterol transport to supply lipid to peripheral cells that may have an increased requirement for tissue regeneration at the inflammatory site or, alternatively, by facilitating the removal of cholesterol liberated at sites of tissue damage during inflammation. Finally, A-SAA is assumed to have chemotactic properties and to recruit inflammatory cells such as monocytes, leukocytes, mast cells and T-lymphocytes to sites of inflammation. Summarized according to Uhlar and Whitehead, 1998 [116].

A-SAA expression and regulation

Serum concentrations of A-SAA increase by up to a 1000-fold within 24h in response to inflammatory stimuli such as infection, injury, trauma, and stress [116]. The expression of A-SAA during the acute-phase is summarized in Figure 2.

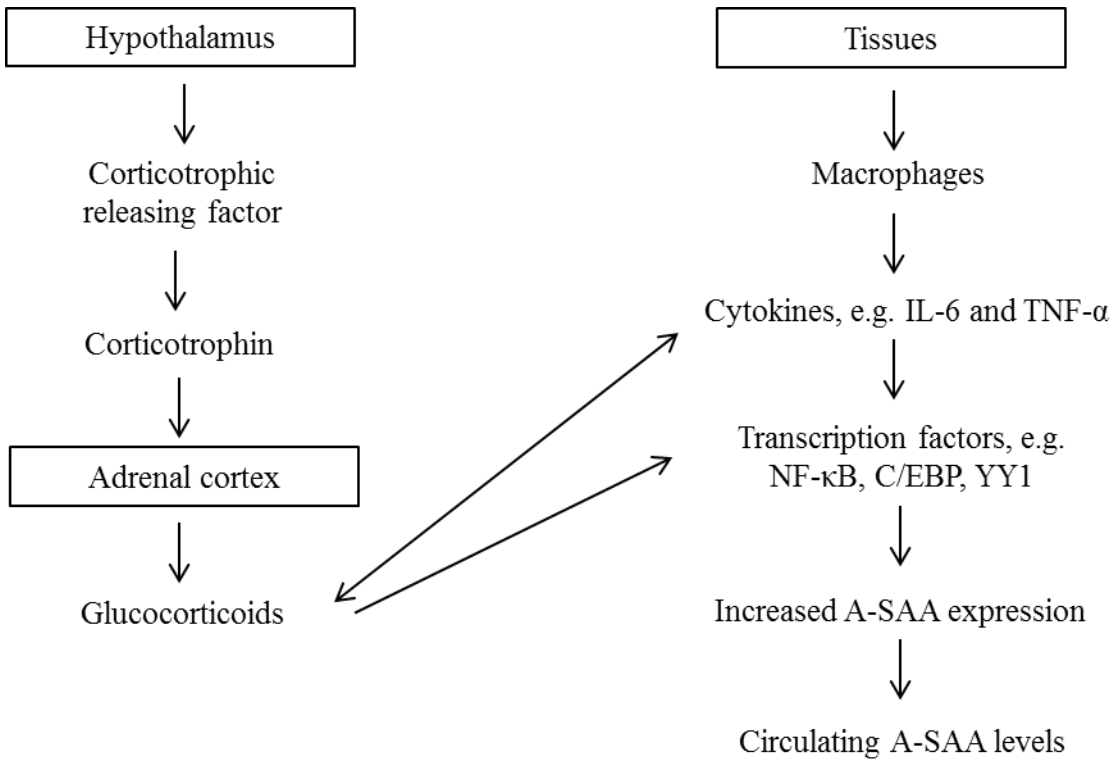

Figure 2: A-SAA expression during the acute-phase. *Modified according to Uhlar, CM and Whitehead, AS, 1999 [116].* IL-6 – interleukin-6, TNF-α – tumor necrosis factor-α, NF-κB – nuclear factor-κB, C/EBP – CCAAT/ enhancer binding protein, YY1 – ying and yang 1

A-SAA expression is induced by pro-inflammatory cytokines including IL-1, TNF-α and IL-6 as well as glucocorticoids [116-119]. The dramatic induction of A-SAA mRNA during the acute-phase is probably mainly attributed to the synergistic effects of cytokine signaling pathways, particularly those of IL-1 and IL-6 type cytokines after the recruitment of macrophages to the inflammatory site [116]. Inflammatory cytokines are also thought to increase glucocorticoid production by the adrenal cortex after corticotrophic releasing factor and corticotrophin have been released from the hypothalamus. Glucocorticoids may possibly enhance A-SAA synthesis as well as down-regulate the systemic acute-phase response [116]. Furthermore, A-SAA is supposed to having cytokine-like properties and enhancing the induction of pro-inflammatory cytokines itself thus augmenting the immune response [116]. Also post-transcriptional mechanisms including changes in mRNA stability and translation efficiency are discussed to be involved in the regulation of A-SAA [117]. Several transcription factors have been found to engage with A-SAA promoters including NF-κB and CCAAT/enhancer binding protein (C/EBP) for the up-regulation and ying and yang 1 (YY1) for the inhibition of A-SAA expression [116].

On the cellular level, pro-inflammatory cytokines bind to their hepatic receptors thus activating signal transduction into the cell. The binding of IL-6 to its receptor is thought to phosphorylate the nuclear factor for IL-6 which is possibly subsequently transferred to the nucleus where it might initiate A-SAA transcription. In addition to this, it might activate the expression of the nuclear factor for IL-6ß which in return might further enhance A-SAA expression. The binding of TNF-α and IL-1 to its receptors is assumed to lead to the phosphorylation of the cytoplasmic NF-κB-IκB complex resulting in the degradation of IκB and the translocation of NF-κB to the nucleus where it might bind to the promoter region of A-SAA. At the later stage of the acute-phase response, the expression of A-SAA is assumed to be down-regulated via an increased production of cytokine antagonists such as interleukin-1 receptor antagonist (IL-1Ra) and soluble cytokine receptors. The A-SAA gene transcription may possibly be controlled by the NF-κB-IκB complex. The activation of NF-κB might result in an increased transcription of the *IκB* gene. IκB might subsequently enter the nucleus and induce the dissociation of NF-κB from its binding site in the A-SAA promoter. Another putative control mechanism for the A-SAA transcription involves the transcription factor YY1 which might also contribute to the dissociation of NF-κB from its binding site. The NF-κB-IκB complex is thought to be subsequently translocated into the cytoplasm. The reactivation of transcriptions factors is thought to be prevented by IL-1Ra which might bind to the IL-1 receptor and thereby stop pro-inflammatory signal transduction. Summarized according to Jensen and Whitehead, 1998 [117].

Sometimes, in patients with chronic or recurrent inflammatory conditions such as present in rheumatoid arthritis and atherosclerosis, control mechanisms seem to be not sufficient and the long-term maintenance of some aspects of the acute-phase response may possibly cause subclinical inflammatory conditions with chronically elevated levels of A-SAA [116]. Chronically elevated levels of A-SAA might also be induced by exposure to allergens or chemicals, physical and mental stress,

certain metabolic conditions such as obesity, as well as genetic immune deficiencies and defects in the normal control mechanisms [117, 120].

Like other acute-phase proteins, A-SAA is expressed primarily by the liver [119, 121]. However, extra-hepatic expression has been reported for different cells including monocyte, epithelial cells, endothelial cells, adipocytes, atherosclerotic lesions, and aortic smooth muscle cells [119, 122, 123]. Furthermore, SAA expression has been detected in the brain of patients with Alzheimer disease [124]. The broad spectrum of extra-hepatic expression provides a potential local source of A-SAA proteins. These proteins may be expressed under conditions independent of the systemic acute-phase response and may thus play a role related to the site of expression [119].

Serum amyloid A gene family

The SAA gene family is located within 150 kb at chromosome 11 and comprises four genes: *SAA1* and *SAA2*, the bona fide acute-phase SAA isoforms which code for A-SAA proteins, *SAA3*, a pseudo-gene in humans, and *SAA4*, a low level expressed gene coding for the constitutively expressed C-SAA proteins [121, 125, 126]. The alignment of SAA genes on the chromosome is shown in Figure 3.

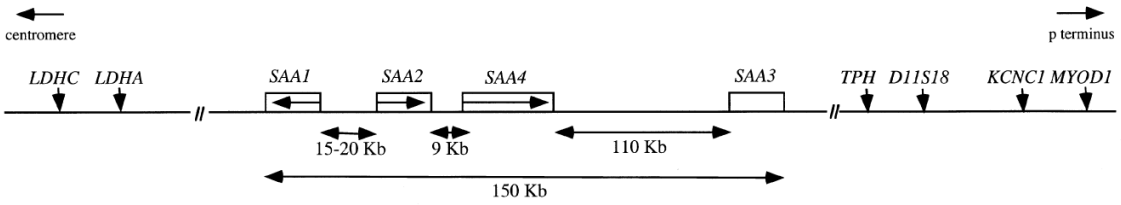

Figure 3: The SAA gene family on chromosome 11. *Modified according to Sellar GC et al., 1994 [126].* Arrows within SAA genes represent 5'to 3' orientation of the gene. The SAA gene family is surrounded by the following genes: lactate dehydrogenase A and C (*LDHA, LDHC*), tryptophan hydroxylase (*TPH*), a member of the potassium channel family (*KCNC1*), and myogenic factor 3 (*MYOD1*). D11S8 is an anonymous human marker.

SAA1 and *SAA2* are both allelic [116, 127]. *SAA1* has at least five alleles, three of which encode distinct proteins (*SAA1.1, SAA1.2*, and *SAA1.3),* and two are neutral polymorphisms of *SAA1.1* [116]. The alleles are based on two polymorphisms (rs1136743, rs1136747) within exon 3 of the *SAA1* gene. *SAA2* has three alleles*: SAA2.1* of which there is a *Hind III* polymorphism of neutral effect and *SAA2.2* [116]. The two genes, *SAA1* and *SAA2*, are concurrently induced during the acute-phase [116] and cluster within 15-20 kb of each other in a head to head arrangement [128]. The third functional gene, *SAA4*, is located 9 kb downstream of *SAA2* in the same direction [129]. Mature protein

sequences of *SAA1* and *SAA2* share greater than 90 % identity [116]. In contrast, the protein sequence of *SAA4* shares only 53 % and 55 % identities with protein sequences of the human *SAA1* and *SAA2* genes, respectively [116].

While *SAA3* expression has been reported in other species such as hamsters and rats [130, 131], the human *SAA3* gene is a not a functional gene due to a single base insertion in exon 3 and a subsequent down-stream stop signal [116].

Twin studies suggest a substantial genetic contribution to baseline A-SAA concentrations with heritability estimates of 59 % (95 % CI, 49–67 %) [132]. However, in the past, association analyses between genetic variants and A-SAA levels were limited and restricted to allelic variants of the SAA gene family [133, 134]. More studies were set out to investigate the association between genetic variants of the *SAA* genes and amyloidosis (see section below) [135, 136]. These studies revealed great ethnic differences suggesting that results of these studies might not be transferable to other ethnic groups [135].

Acute-phase serum amyloid A and disease implications

Numerous studies have shown that prolonged and recurrent chronic infection as well as inflammation is causally involved in the pathogeneses of amyloidosis [116]. Amyloidosis is a heterogeneous group of disorders characterized by an abnormal amyloid deposition eventually leading to dysfunction and failure of the affected tissues and organs. The abnormal amyloid depositions present in amyloidosis are presumed to be primarily a consequence of rheumatoid arthritis and Familial Mediterranean Fever. [116].

Apart from its implication in the etiology of amyloidosis, in recent years, clinical and epidemiological studies have gathered substantial evidence that A-SAA is associated with obesity. In a meta-analysis carried out in 11 cross-sectional studies including a total of 6,993 participants of both genders from Canada, USA, Greece, Australia, New Zealand, Finland, the Netherlands, and Belgium SAA and BMI were moderately, but significantly correlated with a correlation coefficient of 0.230 (95 % CI, 0.160–0.297, $p < 0.0005$). In the same study, the difference in SAA levels before and after weight reduction was significant in 10 prospective studies with a total of 353 participants of both genders from Western Europe and the USA [120].

Furthermore, numerous studies suggested that elevated levels of A-SAA induce, promote or influence the susceptibility to several chronic diseases such as atherosclerosis and its clinical complications [137-145] and various malignancies [146]. For example, in 1,268 initially healthy consecutive patients who underwent serial carotid ultrasound investigations at baseline and after about 7 months the ORs (95 % CI) for atherosclerosis progression (defined as an increase of the degree of stenosis by at least one category with category classes: 0 % to 29 % (carotid plaques), 30 % to 49 % (advanced plaques), 50 % to 69 % (moderate stenosis), 70 % to 89 % (high-grade stenosis), 90 % to 99 % (subocclusive stenosis), and 100 % (occlusion)) for the highest versus the lowest quintile of SAA concentrations

was 2.28 (1.24-4.20) [137]. In a case-control study carried out in 506 men with coronary heart disease and 1,025 male controls, men at the top third of baseline SAA concentrations had an OR (95 % CI) for coronary heart disease of 1.12 (0.71-1.77) compared to those of the lowest third of SAA concentrations adjusted for age, town, smoking, vascular risk factors, and indicators of socioeconomic status [141]. Finally, in a review on SAA and tumor pathogenesis, evidence was presented that SAA was up-regulated around 30-fold in renal cell carcinoma and about two-fold in samples from lung cancer patients [146].

Cross-sectional data have demonstrated an association between elevated systemic A-SAA concentrations and prevalent type 2 diabetes [147, 148] as well as related metabolic parameters [149, 150]. For example, in a baseline examination of 145 type 2 diabetic Finish subjects between 50 and 75 years of age SAA was significantly correlated with HbA1c [149, 150]. In another study, a case-control study (53 type 2 diabetic subjects, 114 controls), the correlations between levels of A-SAA and homeostatic model assessment for insulin resistance (HOMA-IR), HbA1c, and insulin were significant with correlation coefficients (p-values) of 0.24 ($p = 0.002$), 0.19 ($p = 0.02$), and 0.21 ($p = 0.006$) [149]. In this study, most of these correlations were more pronounced in women than in men suggesting that sex differences should be taken into account when analyzing A-SAA concentrations [149]. Furthermore, evidence that circulating levels of A-SAA are not only increased in prevalent type 2 diabetes subjects but also in individuals with IGT was presented by comparing 152 type 2 diabetic subjects, 80 subjects with IGT, and 77 healthy control subjects of the KORA S4 survey [148]. However, one study conducted in 492 Aboriginal Canadian subjects investigated whether baseline levels of SAA are prospectively associated with type 2 diabetes and did not find evidence for a significant association in this study population [151].

1.3. Epidemiological perspective and genetic approach

1.3.1. Epidemiology and causal inferences

Overview of epidemiology

Epidemiology is the study of the distribution and determinants of diseases frequencies in human populations [152]. Briefly, it can be classified into descriptive and analytical epidemiology. Descriptive epidemiology assesses disease prevalence and incidences in a population. The analysis of associations between putative risk factors and diseases is subject to analytical epidemiology. Analytical epidemiology is based on the assumption that diseases do not occur by chance but rather by causes which might be identified and at best eliminated [153].

When interpreting an association between exposure and disease *confounding* by a third variable associated with the exposure and the disease as well as *effect modification* (i.e. the effect of an exposure varies by levels of another variable) have to be considered. In part, this can be accounted for in the statistical analysis by multivariate regression and/ or stratified analyses. This is, however, only feasible if the confounding variable and/ or the effect modifier is known and has been collected in the study investigation. Summarized according to Kreienbrock and Schach, 1995 [154] and Rothman and Greenland, 1998 [155] .

Causal inferences

The hallmark of epidemiological studies is an observational nature. Therefore, causal inferences on disease etiology in epidemiological studies have to be drawn with all due caution and require further considerations to assess evidence for causation [156]. Those proposed by Sir Austin Bradford Hill in 1965 are the most frequently used considerations to deduce on causation in epidemiology [157]. One of these criteria, temporal relationship, i.e. the chronological order of exposure and disease, is implemented in *prospective cohort studies* (follow-up studies, longitudinal studies). In a prospective cohort study, a defined study population during a given period of time (follow-up period) is monitored with respect to certain events such as the onset of disease or death. This study is most reasonable when analyzing diseases with high incidence rates such as type 2 diabetes. For the identification of risk factors, prospective cohort studies compare different risk factor profiles between subjects with regard to disease incidence. The temporal relationship which is incorporated in these study type is not evidence for causality per se but it is an indispensable prerequisite to it.

A more distinct indication for causality between a putative risk factor and a disease can be obtained by analytical-experimental studies, i.e. *clinical trials and intervention studies*. In most of these studies, participants are randomly assigned to either an intervention or a control group. In order to avoid placebo and observation bias participants and study investigators are often blinded, i.e. not knowing which participant is assigned to which group. A significant reduction in disease risk in the intervention group compared to the control group provides evidence for a causal implication of this

risk factor in disease etiology. Summarized according to Kreienbrock and Schach, 1995 [154] and Rothman and Greenland, 1998 [155].

1.3.2. Genome-wide association studies (GWAS)

Aims of genetic association studies

The identification of genetic variants that influence disease susceptibility could help to detect previously unknown etiological pathways and may thus provide clues to better understand disease pathogenesis and to point towards novel potential targets for intervention [158, 159]. The identification of genetic variants involved in a disease might also be of relevance for further sub-grouping diseases with heterogeneous phenotypes, and this might improve therapy towards a more personalized medication approach [160, 161]. In addition, although the genetic predisposition to disease is non-modifiable, an improved risk prediction by genetic testing might encourage the individual at risk to reduce modifiable risk factors such as high-caloric diet and physical inactivity [160]. Finally, the identification of genetic variants that determine biomarker levels is a presumption for a Mendelian Randomization study (chapter 1.3.3.).

From candidate gene approach to high-throughput genotyping methods

In the past, genetic association studies were mainly restricted to certain candidate gene regions which were identified in linkage studies or which were by themselves hypothesized to be of relevance for certain diseases [162]. In any case, some previous knowledge was required. Since about 2005, hypothesis-free testing in genetic association studies has become feasible with the technological development of high-throughput/ low-cost genotyping enabling the conduction of GWAS comprising a dense set of genetic markers across the whole genome [162, 163].

Chances of GWAS

Including genetic variants of the whole genome and being hypothesis-free this approach bears the chance to detect novel genetic loci involved in pathogenic networks, especially of complex and common diseases [164, 165]. It is hypothesized that common diseases in a population (i.e. those with a high prevalence) have to be influenced by genetic determinants that are long-standing and thus common (i.e. have a high frequency within the population). This assumption has become known as the *common-disease/ common-variant* hypothesis [166, 167]. The *common disease/ common variant hypothesis* implies that genetic determinants of common phenotypes should be found in studies of large sample size even if the attributable risk may be small [166, 168]. Indeed, several hundred genetic susceptibility loci for different phenotypes have been identified in the past [169] and the

capacity to undertake GWAS has already largely improved the understanding of the genetic basis of common complex phenotypes [165].

The major challenge of GWAS is the identification of true causative variants. There are three possible reasons for a statistically valid association between a single-nucleotid polymorphism (SNP) and a phenotype: (i) the SNP is causally involved in disease etiology, (ii) the SNP is in linkage disequilibrium (LD) with a causal SNP, or (iii) the association is per chance or a false-positive signal [158]. False positive findings may be the result of (i) genotyping or imputation errors, (ii) population stratification, or (iii) insufficient multiple testing [158].

(i) Genotyping and imputation

In spite of all technological advances in the development of high-throughput genotyping it is currently not possible to cover all known variation in the human genome at reasonable cost. Therefore, it is possible that the causative genetic variants of a disease are not typed on the SNP-chip [169-171]. Genotyping chips differ in the way in which the SNPs are chosen and the total number of SNPs assayed, i.e. in coverage of SNPs [169]. The selection of SNPs for the chip design can be based on chance, on a physically consistent distribution across the whole genome, or based on selection methods accounting for LD structures [158] with implications for the power to detect putative causative variants as well as for the analyses in different ethnical populations [169].

To ensure high *genotyping quality* stringent filtering based on quality scores is necessary [172] and those SNPs with a low *genotyping call rate* which is an indicator for poor DNA quality should be excluded from the data set [173]. Furthermore, those SNPs with deviations from *Hardy-Weinberg Equilibrium* (HWE) are likely to be erroneous and should also be excluded from the data set [173].

Additional information on genetic variation for the genome-wide association analyses can be obtained by integrating information of existing catalogs of genetic variation into the analysis [165]. One convincing approach for the integration of such information is the *imputation* method developed by Marchini et al., 2007 [174]. The core aim of this method is to predict (or "impute") the missing data based upon the observed genotype data using information on the SNPs directly genotyped and combining them with population-based genotype data such as those of the Haplotype Mapping (HapMap) consortium and estimates of the fine-scale recombination map across the genome [174]. The major advantage of this method is that it uses information from all markers in LD with an untyped SNP but in a way that weights genotypes that are consistent with the local patterns of LD more and decreases with genetic distance from the SNP being imputed [174]. Population stratification (see section below) is a particular concern in the context of imputation methods, because allele frequencies and LD patterns between the study sample and the sample used in catalogs of genetic variation may differ and reduce the accuracy of the imputation. However, it is assumed that approaches to adjust for population stratification should also work for imputed SNPs [174].

(ii) Population stratification

An association between exposure and disease can be biased by confounding. In this context, population stratification is of most relevance in GWAS [175-177]. Population stratification is caused by population admixture or subgroups within cases or controls which are systematically heterogeneous with respect to disease/ phenotype prevalence and allele frequencies [158, 173, 178]. The strategy of choice would be to avoid bias induced by population stratification by ensuring that study samples are drawn from homogenous populations [173]. Indeed, the importance of population stratification as a cause for false-positive findings is thought to be modest if cases and controls are well matched for ethnic background and measures are taken to identify and exclude individuals whose genome-wide data reveal substantial differences in genetic background [165, 179]. There are also statistical approaches to detect and correct for population stratification. These approaches are based on the fact that genetic confounding results in deviations from HWE and LD pattern between genetic markers on different chromosomes [158]. The mostly used approach to detect and correct for population stratification is the *genomic control approach* [158]. It requires further "neutral" genetic markers (i.e. not associated with the phenotype of interest) which can be used to estimate substructures. The genomic control approach assesses a scaling factor by which the test statistic is extended if population stratification is present [177]. This extension can be taken into account by adjusting the test statistic by this scaling factor using different statistical methods [175, 180].

(iii) Multiple testing, power and sample sizes

In GWAS, the association between SNPs across the whole genome and the disease/ phenotype of interest is tested. This results in a large number of tests and thus an increased probability of false-positive findings (type I errors) [165]. One commonly used method to correct for multiple testing is the Bonferroni correction where the level of significance is divided by the number of independent tests [181]. However, the Bonferroni correction is over-conservative and leads to an increase in false-negative findings (type II errors), if there are strong correlations between the independent variables to be tested, as is the case for SNPs in LD blocks [165, 181]. As a solution to this problem, information on LD structures collected by the HapMap consortium was integrated in power calculations and the power was corrected for a million independent tests genome-wide in Europeans, and twice that number in Africans [181, 182].

Given the high dimensionality of GWAS and the resulting multiple testing problem the *power* to detect a statistically significant association between a genetic variant and the phenotype of interest is limited and consequently, true positive associations might not be found. In the past, results of GWAS showed only small to moderate effects of genetic variants on various phenotypes [183]. The power to detect associations at a scale of such small effect sizes is only sufficient if the minor allele frequency (MAF) is sufficiently high [173].

Power can be boosted by sample size. This fact has led to the formation of large consortia of research institutes such as the *Diabetes Genetics Replication and Meta-analysis* (DIAGRAM) consortium and the *Meta-Analyses of Glucose and Insulin-related traits Consortium* (MAGIC). These consortia provide a framework for the combination of several GWAS into large *meta-analyses*. Meta-analyses combine study results either with or without weights (e.g. inverse variance) in random or fixed effects models [184]. As the two models are based on the assumption that effects in the single studies are homogenous (fixed effects model) or heterogeneous (random effects model) a major point to be considered when choosing the appropriate model is between-study heterogeneity and population stratification [40, 185, 186]. There are multiple approaches to measure homogeneity between studies and consistency of study results [183]. In one approach provided by Higgins et al. in 2003, a quantity measure (I^2) is assessed that describes the percentage of total variation across studies attributed to heterogeneity rather than chance [187]. I^2 metrics can range between 0 % and 100 % with 0 % indicating homogeneity [187].

Replication and multi-stage analysis

Validation of GWAS has been shown to be essential, firstly, because many initial reports of SNP-phenotype associations have turned out to be false-positive findings [188] and secondly, because almost all initial studies have overestimated the true effect size in the past ("the winners curse") [189, 190]. This overestimation of effect sizes is hypothesized to be attributed to bias and genuine population diversity [190]. Some studies use a multistage approach to provide replication evidence for GWAS findings. At a first stage, SNPs that are associated with a phenotype of interest on a genome-wide significance level are identified. At a second stage, the identified SNPs are tested in an independent replication sample in subsequent analyses [191-193]. A major advantage to this approach is that it retains power while reducing genotyping costs [165].

1.3.3. Mendelian Randomization Study

Concept of Mendelian Randomization studies

Being observational in its nature epidemiological studies are subject to many potential biases, to environmental confounding and to reverse causation which limits their ability to robustly identify causal associations (see also 1.3.1.) [194]. Mendelian Randomization studies are an approach to strengthen causal inferences by using genetic variants as proxy for environmentally modifiable risk factors of a disease [194-196]. The concept of a Mendelian Randomization study is displayed in Figure 4.

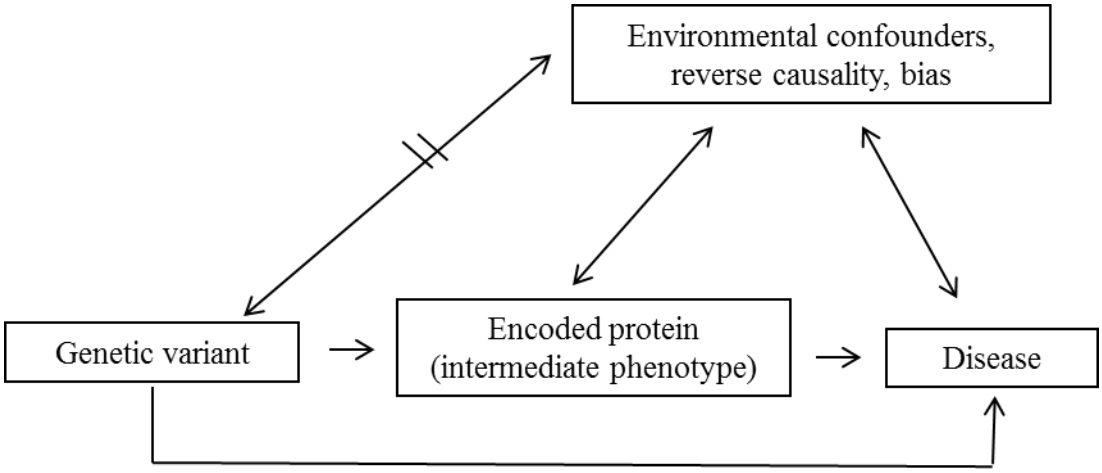

Figure 4*: *Concept of a Mendelian Randomization study

Different alleles of a gene may influence concentrations of the encoded protein via a different degree of gene expression activity. If the encoded protein is found to be associated with a specific genetic variant as well as a specific disease the genetic variant should also be associated with the disease provided that the association between protein and disease is causal [197, 198]. Being based on the assumption of a random assortment of genes from parents to offspring, Mendelian Randomization studies benefit from a "randomization by nature" and combine the concepts of cohort studies with that of randomized controlled trials [199].

Advantages of Mendelian Randomization studies

There are a number of advantages to this study concept. (i) The random assignment of study participants according to their genotypes ensures an equal distribution of potential confounding variables such as smoking in the study groups. (ii) Mendelian Randomization studies are double-

blinded, as neither study participants nor investigators usually know the genotype. This averts effects caused by placebo and/ or observation bias. (iii) The study concept incorporates the chronological order of exposure and disease as the genotype is defined before the onset of the disease so that reverse causality is eliminated.

Generally, Mendelian Randomization studies provide the opportunity to detect associations between genetic variants as proxies for intermediate phenotypes and a disease largely unbiased by potential environmentally modifying factors [196-198, 200]. Thus, the Mendelian Randomization approach might be particularly useful for the identification of causal risk factors for complex diseases characterized by multifactorial environmental contributions to pathogenic mechanisms such as type 2 diabetes.

Limitations

However, besides these major advantages some limitations should be mentioned. These can be summarized to three critical issues [196-198, 200]:

(i) First of all, a prerequisite of the concept is a suitable genetic variant. This means, that the genetic variant within the encoding gene has to be functional otherwise it would not lead to different phenotypes. Furthermore, genetic confounding might arise if the genetic variant or another genetic variant in LD with it exhibits pleiotropic effects on other intermediate phenotypes associated with the disease.

(ii) Moreover, population stratification, canalization (i.e. the developmental adaption to the genetically determined "problem"), selective survival and gene-gene interactions might cause genetic confounding.

(iii) Finally, genetic effects on complex disease phenotypes via intermediate risk factors are usually poor and thus large sample sizes are required in order to provide sufficient power to identify gene-disease associations (section 1.3.2.).

2. Outline of the doctoral thesis

This doctoral thesis aimed to elucidate the association between circulating levels of A-SAA and the development of type 2 diabetes in an epidemiological perspective and to assess a possible causal implication of A-SAA in the etiology of type 2 diabetes by means of genetics. Three studies were conducted: (i) a prospective cohort study entitled *Acute-phase serum amyloid A protein and its implication in the development of the KORA S4/ F4 study* (Marzi et al., Diabetes Care, 2012 Dec 13, epub ahead of print), (ii) a meta-analysis of GWAS entitled *Genome-wide Association Study Identifies Two Novel Regions at 11p15.5-p13 and 1p31 with Major Impact on Acute-Phase Serum Amyloid A* (Marzi et al., PLoS Genet. 2010 Nov 18;6(11):e1001213), and (iii) a Mendelian Randomization study on genetic susceptibility loci for levels of A-SAA and type 2 diabetes (unpublished data).

Prospective cohort study on elevated levels of A-SAA and incident type 2 diabetes

Cross-sectional data demonstrated a significant association between elevated systemic A-SAA concentrations and prevalent type 2 diabetes as well as the related metabolic parameters HOMA-IR, fasting insulin, and HbA1c (chapter 1.2.2.). However, only one study conducted in 492 Aboriginal Canadian subjects had investigated whether baseline levels of SAA were prospectively associated with type 2 diabetes and did not find evidence for a significant association (chapter 1.2.2.). Therefore, the prospective cohort study of this doctoral thesis was set out to investigate whether elevated levels of A-SAA precede the onset of type 2 diabetes during seven years of follow-up in a large population-based study of initially non-diabetic, elderly Western European subjects without clinically overt inflammation. Moreover, it was assessed whether this association is independent of other type 2 diabetes risk factors including parameters of glucose metabolism (i.e. fasting glucose, fasting insulin, HbA1c, and 2h-glucose) as an indicator for early impairments of glucose homeostasis and whether there are differences between men and women as suggested previously (chapter 1.2.1. and 1.2.2.). Finally, the strength of the putative association was compared to that between the most frequently used inflammatory marker, circulating hs-CRP and incident type 2 diabetes.

Meta-analysis of GWAS on levels of A-SAA

The identification of genetic variants that determine levels of A-SAA is a prerequisite for a Mendelian Randomization study. Furthermore, it might provide important clues to the immune response pathways involved in the regulation of A-SAA levels. This might also be of relevance for related clinical entities such as type 2 diabetes. Twin studies suggested a substantial genetic contribution to baseline A-SAA concentrations but in the past, association analyses between genetic variants and A-SAA levels had been limited and had been restricted to allelic variants of *SAA* genes and protein concentrations (chapter 1.2.2.). Therefore, the meta-analysis of GWAS of this doctoral thesis was designed to analyze the genetic basis of A-SAA levels in a more comprehensive way. In a meta-

analysis four genome-wide scans (KORA S4, LURIC, TwinsUK and Sorbs) the association between genetic variants across the whole genome and circulating levels of A-SAA was analyzed in 4,212 participants of European descent. Additionally, in order to account for potential sex-specific differences in the regulation of A-SAA (chapter 1.2.2.) the analysis was stratified by sex.

Mendelian Randomization study on A-SAA susceptibility loci and type 2 diabetes

The identification of risk factors that causally contribute to the development of a disease is of relevance for the prevention as well as the treatment of a disease. Mendelian Randomization studies can be used to assess whether an association between a biomarker and a disease is causal using genetic variants as proxies for this biomarker (chapter 1.3.3.). Therefore, in order to analyze whether elevated levels of A-SAA causally contribute to the development of type 2 diabetes a Mendelian Randomization study was conducted. The associations between genetic variants with the lowest p-values in the regions/ subregions identified in the meta-analysis of GWAS of this doctoral thesis and type 2 diabetes were extracted from a previously published meta-analysis comprising of eight GWAS (8,130 type 2 diabetes cases and 38,987 control subjects of European descent). Furthermore, a power calculation was conducted to ensure high statistical validity.

3. Epidemiological studies

3.1. Acute-phase serum amyloid A protein and its implication in the development of the KORA S4/ F4 study

This chapter is modified according to the manuscript Marci C. et al., ***Acute-phase serum amyloid A protein and its implication in the development of type 2 diabetes in the KORA S4/ F4 study*** Diabetes Care, 2012 Dec 13, epub ahead of print.

3.1.1. Research Design and Methods

Study population

The present study included participants of the population-based KORA Survey S4 (1999-2001, N=4,261) and its seven-year follow-up study F4 (2006-2008, N=3,080) and was conducted in the region of Augsburg, Southern Germany. The study was approved by the ethics committee of the Bavarian Medical Association and informed written consent was obtained from all participants. The selection process of study participants is displayed in Figure 5. The sample included all KORA S4 participants for whom an OGTT was performed at baseline (i.e. only the age group between 55 and 74 years) and who were non-diabetic according to the OGTT (n=1,231). For the present analyses, subjects with elevated levels of hs-CRP at baseline which indicated an acute pro-inflammatory state (hs-CRP > 10 mg/l, n = 44) or with missing A-SAA data (n = 16) were excluded. Furthermore, participants in the follow-up examination (n = 878) with missing data in type 2 diabetes status (n = 20) or in one of the co-variates used in the different models (n = 22) were also excluded. Thus, 836 subjects were included in the present analyses.

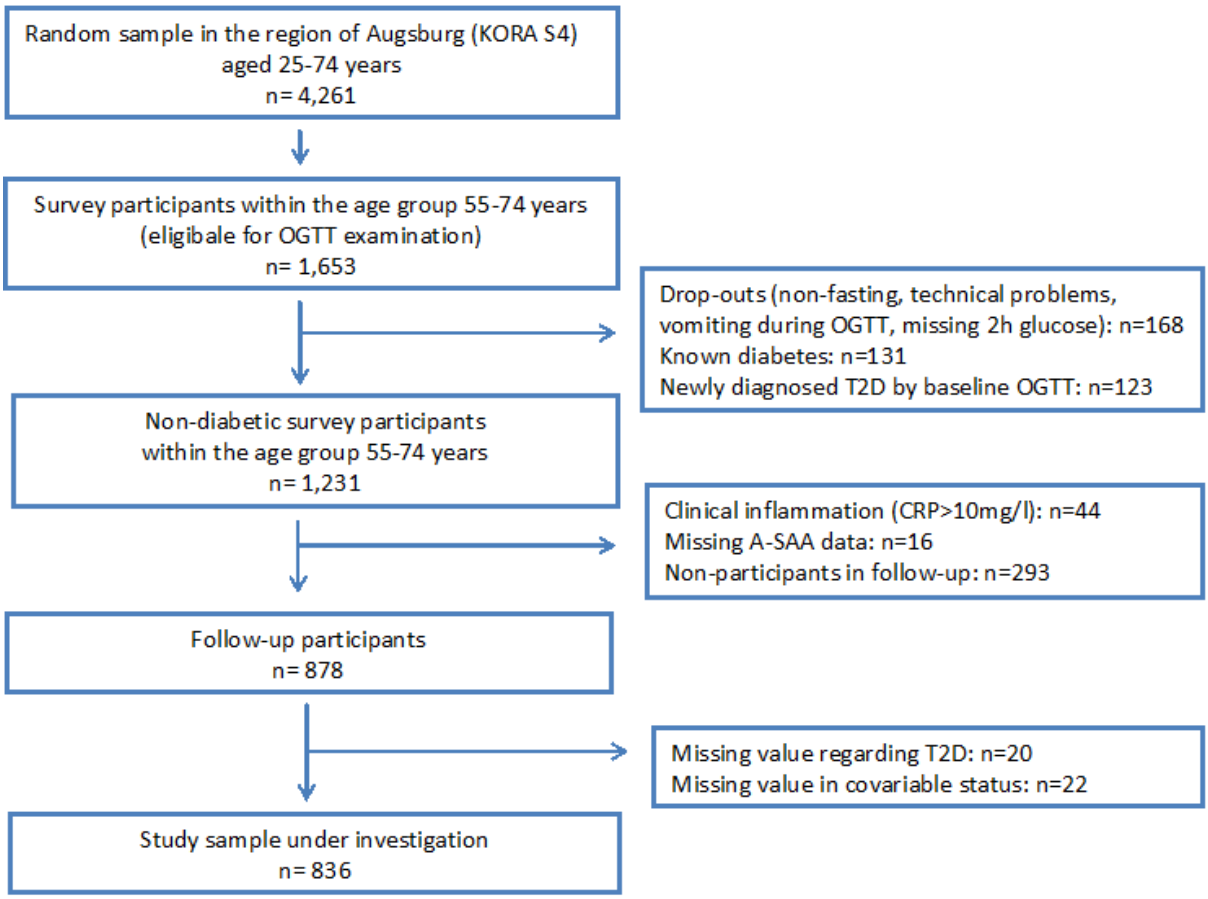

Figure 5: Flowchart of the selection process of study participants for the present analyses

Ascertainment of type 2 diabetes and assessment of independent variables

Type 2 diabetes was defined on the basis of an OGTT according to the 1999 WHO diagnostic criteria [2] or by a self-report which was validated by data on glucose-lowering medication or questioning the treating physician. At baseline, fasting EDTA plasma samples were collected and A-SAA was analyzed by immunonephelometry on a BN II analyzer using an N Latex enhanced test from Siemens, Germany. The inter-assay coefficient of variation was below 7 % [148, 201].

Information regarding the collection of laboratory, anthropometric, and socioeconomic variables, are described in detail elsewhere [17, 97, 202, 203].

Statistical analyses

A-SAA, hs-CRP and co-variables were tested for normal distribution. Furthermore, all numerical co-variables were tested for correlation among each other using the Pearson's coefficient of correlation for normally distributed variables and the Spearman's coefficient of correlation for skewed distributed variables. Baseline characteristics of subjects who developed type 2 diabetes and those who remained diabetes-free were compared using five different tests: i) Pearson's chi-square test to test for

independence of categorical variables between groups, (ii) Fisher's exact test to test for independence between groups with small sample sizes in categories, (iii) Student's two sample t-test for normally distributed numerical variables with equal variances, (iv) Welch's t-test for normally distributed numerical variables with unequal variances, and (v) the Mann-Whitney-U test for skewed distributed numerical variables. Multivariate linear and logistic regression analyses with various degrees of adjustment were performed. Interactions between A-SAA and sex as well as A-SAA and age were assessed by adding an interaction term to the multivariate models. Two-sided p-values that were below 0.05 were considered to be statistically significant. All statistical analyses were conducted using SAS (Version 9.2, SAS Institute, Cary, NC).

3.1.2. Results

Baseline characteristics stratified by incident type 2 diabetes status are provided in Table 2. Seven-year cumulative incidence of type 2 diabetes was 11.1 % (n = 93). Circulating systemic A-SAA levels were significantly higher (i) in diabetic subjects (median of A-SAA levels 4.0 mg/l [IQR, 2.6-6.8]) than in non-diabetic subjects (3.4 mg/l, [IQR, 2.2-5.4]); $p = 0.009$), (ii) in women (median 3.9 mg/l [IQR, 2.6-6.1]) than in men (2.9 mg/l, [IQR, 2.0-4.7]); $p < 0.0001$), and (iii) in subjects between 65 and 74 years of age (median 3.6 mg/l [IQR, 2.3-6.2]) than in subjects between 55 and 64 years of age (3.3 mg/l [IQR, 2.2-5.0]; $p = 0.03$). All co-variables were correlated with each other with $r < 0.5$ except for hs-CRP and A-SAA ($r = 0.50$, Table 1 Appendix).

Table 2. Baseline characteristics of the study participants stratified by incident type 2 diabetes status.

	No diabetes	Incident type 2 diabetes	p
N (%)	743 (88.9)	93 (11.1)	
Age (years)	63.02 (5.35)	65.16 (5.06)	**0,0002**
Male (%)	48,3	63,4	**0.006**
BMI (kg/m²)	27.84 (3.97)	30.30 (3.41)	**<.0001**
Never smoking/ former smoker/ current smoker (%)	54.1/ 34.9/ 11	39.8/ 45.2/ 15,1	**0,03**
Alcohol intake (g/day)*	8.57 (0.89-23.14)	11.43 (1.57-26.14)	0,42
Vigorous/ moderate/ mild/ no leisure time physical activity (%)	20.1/ 27.9/ 17.4/ 34,7	16.1/ 28.0/ 15.1/ 40,9	0.62
No educational degree/ Lower secondary educational degree/ Secondary educational degree/ High secondary educational degree/ University degree (%)	1.4/ 64.1/ 21.1/ 4.7/ 8.8	2.2/ 68.8/ 14.0/ 7.5/ 7.5	0.38
Systolic blood pressure (mmHg)	133 (19)	138 (19)	**0.009**
Lipids			
Non-HDL cholesterol (mmol/l)	4.78 (1.07)	4.90 (1.16)	0.35
Triglycerides (mmol/l)*	1.22 (0.89-1.73)	1.55 (1.21-2.02)	**<.0001**
Glucose homeostasis			
Fasting glucose (mg/dl)	97.78 (8.75)	106.40 (10.49)	**<.0001**
Fasting insulin (µU/ml)*	9.3 (6.75-13.20)	13.5 (10.05-21.15)	**<.0001**
HbA1c (%)*	5.6 (5.4-5.8)	5.8 (5.6-6.0)	**<.0001**
HOMA-IR	3.03 (4.01)	4.88 (4.92)	**<.0001**
2h glucose (mg/dl)*	106 (89-126)	146 (117-176)	**<.0001**
Inflammatory markers			
A-SAA (mg/l)*	3.40 (2.20-5.40)	4.00 (2.60-6.80)	**0.009**
hs-CRP (mg/l)*	1.39 (0.72-2.75)	2.52 (1.25-3.78)	**<.0001**
At least one parent diabetic/ no parent diabetic/ no information (%)	21.9/ 59.9/ 18.2	36.6/ 44.1/ 19.4	**0.004**

Data with normal and skewed distribution (indicated by *) are given as mean (SD) and median (interquartile range) of the variables. P-values: Pearson's chi-square test, Fisher's exact test, Student's two sample t-test, Welch's t-test, and Mann-Whitney-U test. P-values printed in bold letters indicate a statistically significant difference.

Table 3 presents the association between baseline A-SAA levels and incident type 2 diabetes. A-SAA concentrations were significantly associated with the development of type 2 diabetes during the seven-year follow-up period. The OR (adjusted for age and sex, model 1) for a 1-SD (4.16 mg/l) increase in A-SAA was 1.28 [95% CI, 1.08-1.53], p = 0.005. The association remained statistically significant after adjustment for multiple co-variables including BMI (model 2) and further traditional risk factors (smoking, alcohol intake, physical activity, education, parental history of diabetes, non-HDL cholesterol, fasting triglycerides, systolic blood pressure, model 3). To further evaluate the impact of body fat and body composition on the association between A-SAA levels and incident type 2 diabetes BMI was replaced by WHR or, alternatively, body fat mass index, lean body mass index, and appendicular skeletal muscle mass index as proposed by Kyle [204, 205] in model 2 of the association analysis. These analyses yielded results similar to those displayed in Table 3 (Table 2 Appendix). The effect estimates decreased and the association was attenuated after additional adjustment for baseline parameters of glucose metabolism (fasting glucose, fasting insulin, HbA1c and 2h-glucose, model 4). Sensitivity analyses demonstrated that the attenuation was attributable mostly to adjustment for 2h-glucose (model 3 + fasting glucose + fasting insulin: OR = 1.25 [95 % CI, 1.03-1.52], p = 0.02; model 3 + HbA1c: OR = 1.21 [95 % CI, 1.01-1.46], p = 0.04; model 3 + 2h-glucose: OR = 1.17 [95 % CI, 0.97-1.42], p = 0.11). In multivariable linear regression analyses the association between baseline levels of A-SAA and ln-transformed levels of 2h-glucose at follow-up was statistically significant in all four models (model 1: ß = 0.01, se(ß) = 0.003, p < 0.0001, model 2: ß = 0.008, se(ß) = 0.003, p = 0.0009, model 3: ß = 0.007, se(ß) = 0.002, p = 0.004, model 4: ß = 0.006, se(ß) = 0.002, p = 0.008, N = 803).

Table 3. Results of the overall and age-stratified associations between circulating concentrations of A-SAA and incident type 2 diabetes according to four different models of co-variables adjustment.

Stratum (T2D/nondiabetic)		all (93/743)		55-64 years (43/457)		65-74 years (50/286)	
	Covariables	**OR (95% CI)**	**P**	**OR (95% CI)**	**P**	**OR (95% CI)**	**P**
1	age + sex	1.28 [1.08; 1.53]	**0.005**	1.48 [1.17; 1.88]	**0.001**	1.21 [0.94; 1.55]	0.14
2	model 1 + BMI	1.25 [1.05; 1.50]	**0.01**	1.40 [1.09; 1.79]	**0.009**	1.21 [0.94; 1.09]	0.15
3	model 2 + traditional risk factors*	1.22 [1.02; 1.46]	**0.03**	1.38 [1.06; 1.80]	**0.02**	1.15 [0.88; 1.51]	0.3
4	model 3 + glucose related parameters**	1.16 [0.95; 1.42]	0.16	1.24 [0.85; 1.80]	0.27	1.07 [0.80; 1.43]	0.67

* traditional risk factors are smoking, physical activity, alcohol consumption, education, parental history of diabetes, non-HDL cholesterol, fasting triglycerides, and systolic blood pressure
**glucose related parameters include fasting glucose, fasting insulin, HbA1c, and 2h-glucose
Odds ratios (95% CI) are given for a 1-SD increase of A-SAA concentrations (4.16 mg/l for all, 3.34 mg/l for younger, and 5.11 mg/l for older age stratum). Smoking is coded in three classes (never, former and current smoker). Physical activity is coded in four classes (none, mild = irregularly 1h per week, moderate = regularly 1 h per week, and vigorous= regularly 2h or more per week). Education is coded in five classes based on the highest degree of education. Parental history is coded in three classes (at least one parent has/had T2D, none of the parents has/had T2D, no information on parental history).

Stratification by sex showed that effect sizes in general were stronger in women than in men but the multiplicative interaction term between sex and A-SAA was not statistically significant ($p > 0.1$). Replacement of BMI by WHR, body fat mass index, lean fat mass index, or appendicular skeletal muscle mass index also did not indicate a statistically significant sex effect. However, the multiplicative interaction term between the metric age variable and A-SAA was borderline significant regarding the outcome of incident type 2 diabetes (p-value for interaction adjusted for age and sex = 0.047). Therefore, we also calculated age-stratified ORs using 10-year age strata (Table 3). Stratified analyses showed that the A-SAA variability was higher in the age stratum 65-74 years (SD for age stratum 55-64 years = 3.34 mg/l, SD for age stratum 65-74 years = 5.11 mg/l). The association between A-SAA and incident type 2 diabetes was only significant in younger subjects (OR for the age group 55-64 years adjusted for age and sex = 1.48 [95 % CI, 1.17-1.88], p = 0.001; OR for the age group 65-74 years adjusted for age and sex=1.21 [95 % CI, 0.94-1.55]; p = 0.14). Furthermore, sensitivity analyses demonstrated, that for the age group 55-64 years the effect was mostly attenuated

by adjustment for HbA1c and not 2h-glucose (model 3 + fasting glucose + fasting insulin: OR = 1.44 [95% CI, 1.07-1.92], p = 0.03; model 3 + HbA1c: OR = 1.26 [95 % CI, 0.93-1.70], p = 0.14, model 3 + 2h-glucose: OR = 1.37 [95 % CI, 1.00-1.86], p = 0.05], whereas for the age group 65-74 years, it was mainly affected by adjustment for 2h-glucose (model 3 + fasting glucose + fasting insulin: OR = 1.17 [95 % CI, 0.89-1.55], p = 0.31; model 3 + HbA1c: OR = 1.18 [95 % CI, 0.90-1.56], p = 0.23, model 3 + 2h-glucose: OR = 1.06 [95 % CI, 0.80-1.41], p = 0.69).

The association between baseline levels of hs-CRP and incident type 2 diabetes was analyzed similarly to the baseline levels of A-SAA (Figure 6). The hs-CRP effect estimate was somewhat larger than the effect size for A-SAA when adjusted only for age and sex (OR = 1.39 [95 % CI, 1.10-1.68], p = 0.0006), but similar after full adjustment for all co-variables (OR = 1.13 [95 % CI, 0.88; 1.45], p = 0.34). No evidence for an age-specific effect was found for hs-CRP (p-value for interaction > 0.1).

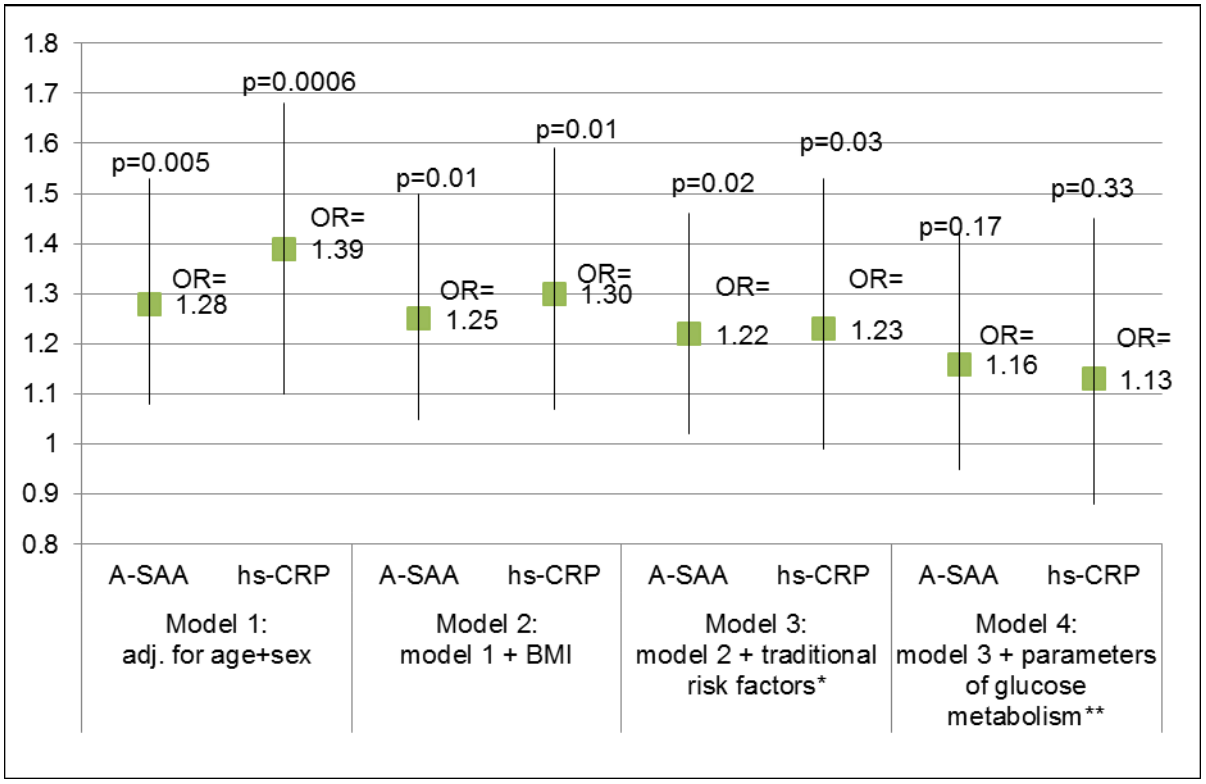

Figure 6*: *OR and 95 % CI for the association between baseline levels of A-SAA and hs-CRP with incident type 2 diabetes.

*Traditional risk factors include smoking, alcohol intake, physical activity, education, parental history of diabetes, non-HDL cholesterol, fasting triglycerides, and systolic blood pressure.

** Parameters of glucose metabolism are baseline levels of fasting glucose and fasting insulin, HbA1c, and 2h-glucose.

Smoking is coded in three classes (never, former and current smoker). Physical activity is coded in four classes (none, mild = irregularly 1h per week, moderate = regularly 1 h per week, and vigorous = regularly 2h or more per week). Education is coded in five classes based on the highest degree of education. Parental history is coded in three classes (at least one parent has/ had type 2 diabetes, none of the parents has/ had type 2 diabetes, no information on parental history).

Sensitivity analyses were conducted to assess the impact of menopausal status, hormone replacement therapy, intake of aspirin, and liver disease in the association between levels of A-SAA and incident type 2 diabetes by adjusting for these variables, but all analyses yielded very similar results as those displayed in Table 3 (Table 3 Appendix).

3.1.3. Discussion

In this study a significant association between elevated systemic levels of A-SAA and the development of type 2 diabetes was observed in a large population-based cohort with long-term follow-up. This finding extends observations from earlier cross-sectional studies that reported a positive association between systemic A-SAA concentrations and prevalent type 2 diabetes and related traits [147-149]. Furthermore, the study showed that the association between A-SAA and incident type 2 diabetes was independent of various established type 2 diabetes risk factors, most of which were correlated with A-SAA cross-sectionally in this study. These risk factors include age and sex and the traditional type 2 diabetes risk factors smoking, alcohol, physical activity, education and parental history of diabetes, total cholesterol, fasting triglycerides and systolic blood pressure. The association of A-SAA with incident type 2 diabetes was also independent of BMI or, alternatively, other measures of body fat and body composition such as WHR, body fat mass index and lean body mass index. This is particularly notable since concentrations of A-SAA are known to be elevated in obese subjects [120] and since the pro-inflammatory cytokines and A-SAA inducers IL-6 and TNF-α as well as A-SAA itself were found to be partly expressed in adipose tissue [45, 46, 206]. Additionally, the administration of recombinant A-SAA in mice adipocytes results in a down-regulation of genes that are critical for insulin sensitivity in the treated cells [207]. Our results suggest that apart from adipose tissue other metabolically critical sites, such as the liver, seem to be involved in the course of the disease. Furthermore, there was no evidence for a role of appendicular skeletal muscle mass index in the association between A-SAA and incident type 2 diabetes which makes confounding effects by myokines unlikely.

Relationship between A-SAA, post-challenge hyperglycemia, and incident type 2 diabetes

The difference in A-SAA levels between individuals with incident type 2 diabetes and study participants who remained diabetes-free was attenuated after adjustment for baseline parameters of glucose metabolism which indicate early impairments of glucose homeostasis. Sensitivity analyses showed that levels of 2h-glucose contributed most strongly to this attenuation. This suggests that elevated levels of A-SAA could be associated with early alterations in glucose control. Thus, elevated levels of circulating A-SAA might be a consequence of a pre-diabetes process, rather than a cause of it. Alternatively, post-challenge hyperglycemia might be an intermediate variable linking A-SAA to the development of type 2 diabetes. To further elucidate the question regarding the direction of causality, baseline levels of A-SAA and levels of 2h-glucose at follow-up were analyzed. These analyses yielded a significant association independent of baseline parameters of glucose metabolism which suggests an independent impact of circulating levels of A-SAA on glucose homeostasis and thus supports the hypothesis of post-challenge hyperglycemia as an intermediate variable. However, the analyses do not allow inferences on reverse causality between levels of 2-glucose and A-SAA.

Consequently, it cannot be ruled out that the relationship between A-SAA and 2h-glucose might be bi-directional.

Comparison of the data with the literature

In accordance with a previous study [149], our study demonstrated that A-SAA concentrations were higher in women than in men but the estimates for A-SAA concentrations regarding risk of incident type 2 diabetes did not differ significantly between men and women, as had been suggested for hs-CRP [97].

The finding of the present study differed from those of the Sandy Lake Health and Diabetes Project [151]. This study had an even broader age range (10-79 years), similar co-variables and a similar sex distribution but did not observe a significant association between SAA and incident type 2 diabetes, although effect estimates pointed into the same direction. The 10-year cumulative incidence was high (17.5%), but fewer subjects (n=492) were analyzed. Thus, results may differ due to a lower statistical power of the Sandy Lake Health and Diabetes Project. In addition, the study was conducted in Aboriginal Canadians and ethnic factors might also be responsible for the different results.

The association between A-SAA and incident type 2 diabetes is possibly modulated by age

In this study, the association between levels of A-SAA and incident type 2 diabetes seemed to be age-dependent. Stratified analyses showed that the variability of A-SAA levels was higher in older subjects and the association of A-SAA with incident type 2 diabetes was only significant in younger subjects. In the age group 55-64 years the effect was attenuated mainly by adjustment for HbA1c but not levels of 2h-glucose. This suggests that elevated levels of A-SAA may play a more important role in the development of type 2 diabetes in younger subjects and that the pathogenic mechanisms may differ in an age-dependent manner. Also, there was no evidence of an influence of hormonal changes during menopause or hormone replacement therapy in women on the observed age effect. Notably, the age-dependent differences were evident in our study population with elderly participants of a relatively small age range (55-74 years). It might be even more pronounced in study populations with younger participants and broader age ranges. However, we only found an effect modification by age for A-SAA, but not for hs-CRP. Therefore, further research is required to confirm and clarify the role of age in the context of innate immunity and the development of type 2 diabetes.

A-SAA versus hs-CRP and incident type 2 diabetes

To compare the strength of the putative association between levels of A-SAA and incident type 2 diabetes with that of another acute-phase protein, also levels of hs-CRP as the best established inflammatory marker were analyzed. Effect estimates for hs-CRP were initially higher than for A-SAA, but were attenuated more strongly when adjusting for the same co-variables. After full adjustment for all co-variables effect sizes for A-SAA and hs-CRP were similar.

Limitations and strengths of the study

Several potential limitations of the study have to be mentioned. Firstly, due to the restrictions in laboratory methods analyses were confined to the A-SAA isoforms and did not capture the constitutively expressed C-SAA isoform which, however, responds only moderately to inflammatory stimuli. Secondly, only one A-SAA measurement was available, although multiple measurements for the determination of inflammatory proteins are preferred [208]. And third, this study did not assess the possible influence of energy intake or liver steatosis on the association between levels of A-SAA and incident type 2 diabetes. Strengths of this study are its prospective design, the well-defined study population sample and the definition of type 2 diabetes based on validated diagnosis and OGTT at both time points. In addition, the present study accounted for baseline parameters of glucose metabolism and thus early impairments of glucose homeostasis.

3.2. Genome-wide association study identifies two novel regions at 11p15.5-p13 and 1p31 with major impact on acute-phase serum amyloid A

This chapter has been published in Marzi, C. et al., ***Genome-Wide Association Study Identifies Two Novel Regions at 11p15.5-p13 and 1p31 with Major Impact on Acute-Phase Serum Amyloid A,*** PloS Genetics PLoS Genet. 2010 Nov 18;6(11):e1001213.

3.2.1. Materials and Methods

Participating studies

This meta-analysis combined data from four genome-wide scans: the KORA S4 study, the Ludwigshafen Risk and Cardiovascular Health study (LURIC), the UK Adult Twin Register (TwinsUK) and a self-contained population from Eastern Germany (Sorbs). Approval was obtained by each of the local Ethic Committees for all studies and written informed consent was given by all study participants. In total, the meta-analysis included 4,212 individuals (1,928 males, 2,284 females) of European ancestry with measured baseline A-SAA concentrations. For validation analyses, data of 2,136 participants of the KORA S4 sample, which were not included in the meta-analysis were used (Text 1 Appendix). Sample sizes and characteristics of the study participants of the four genome-wide scans and the validation sample are displayed in Table 4 Appendix.

Measurement of A-SAA concentrations

In all four studies, study participants were fasting and EDTA plasma samples were analyzed by immunonephelometry on a BN II device from Siemens, Germany, and well-validated automated microparticle capture enzyme immunoassays [148, 209]. The inter-assay coefficients of variation were below 7 % in all four studies.

Genome-wide genotyping and imputation

For genotyping different platforms as the Affymetrix 500K GeneChip array (Sorbs), Affymetrix 6.0 GeneChip array (KORA S4, LURIC, Sorbs), Illumina HumanHap300 BeadChip (317K) (TwinsUK) and Illumina Human 610K BeadChip (TwinsUK) were used. Quality control before imputation was undertaken in each study separately. Detailed information on genotyping and imputation is reported in Table 5 Appendix. Imputation based on the HapMap Phase 2 CEU population was performed using IMPUTE [174] in all studies. After imputation all genotype data had to meet the following quality criteria: a MAF ≥ 0.01, a call rate per SNP ≥ 0.9, and r^2.hat metrics ≥ 0.40 for imputed SNPs. In total, 2,593,456 genotyped or imputed autosomal SNPs were analyzed in the meta-analysis.

For validation and comparison of genotyping platforms, 27 of the SNPs with the lowest p-values in the association analysis were selected. Genotyping of these SNPs was performed with the MassARRAY system using the iPLEX technology (Sequenom, San Diego, CA) in the KORA S4 study. The allele-dependent primer extension products were loaded onto one 384-element chip using a nanoliter pipetting system (SpectroCHIP, Spectro-POINT Spotter; Sequenom), and the samples were analyzed by matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (Bruker Daltonik, Leipzig, Germany). The resulting mass spectra were analyzed for peak identification via the SpectroTYPER RT 3.4 software (Sequenom). To control for reproducibility, 9.8 % of the samples was genotyped in duplicate with a discordance rate of less than 0.5 %.

Genome-wide association analyses and meta-analysis

In each study, linear regression models for all available SNPs were calculated on ln-transformed A-SAA levels in mg/l. The genetic effect was assumed to be additive. Adjustment was made for age, gender, BMI, and study specific covariates, i.e. the Friesinger Score in the LURIC population [210] and a genotyping batch variable in the TwinsUK population. Additionally, this analysis was undertaken stratified by gender. The genome-wide scans were calculated with the analysis software SNPTEST (http://www.stats.ox.ac.uk/~marchini/software/gwas/snptest.html) (KORA S4, LURIC) QUICKTEST (http://toby.freeshell.org/software/quicktest.shtml) (Sorbs) and Merlin (http://www.sph.umich.edu/csg/abecasis/Merlin/) (TwinsUK).

The results of all four genome-wide scans were meta-analyzed using a fixed-effects model applying inverse variance weighting with the METAL software (www.sph.umich.edu/csg/abecasis/metal). Study specific results were corrected for population stratification using the genomic control method. For the overall meta-analysis, the inflation factor was 1.009. No further correction was applied.

P-values below the threshold of $p = 5\times10^{-8}$, which corresponds to a Bonferroni correction for the estimated number of one million tests for independent common variants in the human genome of European individuals [181], were considered to be significant.

As a measure for between study heterogeneity I^2 was calculated [187]. Deviations from Hardy-Weinberg-Equilibrium were tested for all identified SNPs by means of the exact Hardy Weinberg test. For the calculation of explained variances, the multiple R^2 value of the covariate model was subtracted from those of the full model including covariates and top hits of the loci in every single study and the weighted mean (KORA S4, LURIC, and the Sorbs) was assessed. Adjacent regions were tested for independency by analyzing the significance of their top SNPs in a joint model.

Accession numbers

The OMIM (http://www.ncbi.nlm.nih.gov/omim) accession numbers for genes mentioned in this article are 104750 for *SAA1*, 607521 for *HPS5*, 189972 for *GTF2H1*, 150000 for *LDHA*, 150150 for *LDHC*, 601007 for *LEPR*, and 606051 for *SERGEF*.

3.2.2. Results

In this meta-analysis of four genome-wide scans 106 SNPs distributed across two regions showed genome-wide significant associations with p-values below the threshold of 5×10^{-8} (Figure 7, Table 6 Appendix). Table 4 shows study specific results for the top hits within the two regions and three identified subregions (see below) of the meta-analysis as well as an additional region for men in the gender stratified analysis. Genotypic mean levels are provided in Table 7 Appendix. Results of the single genome-wide studies were consistent across all four studies regarding the direction and magnitude of the effects. In addition, results were consistent between different genotyping technologies (Table 8 Appendix). No deviations from the Hardy-Weinberg-Equilibrium were observed. The variable of inter-study heterogeneity (I^2) showed homogeneity at the 1p31 locus. At the 11p15.5-p13 locus we observed I^2 values that indicated a more distinct heterogeneity. This reflects the relatively large and varying beta values and differences in the MAF (Table 6 Appendix). However, taking into account that this locus was clearly significantly associated with A-SAA in all studies included in the meta-analysis, results of the meta-analysis are reported based on a fixed effect model.

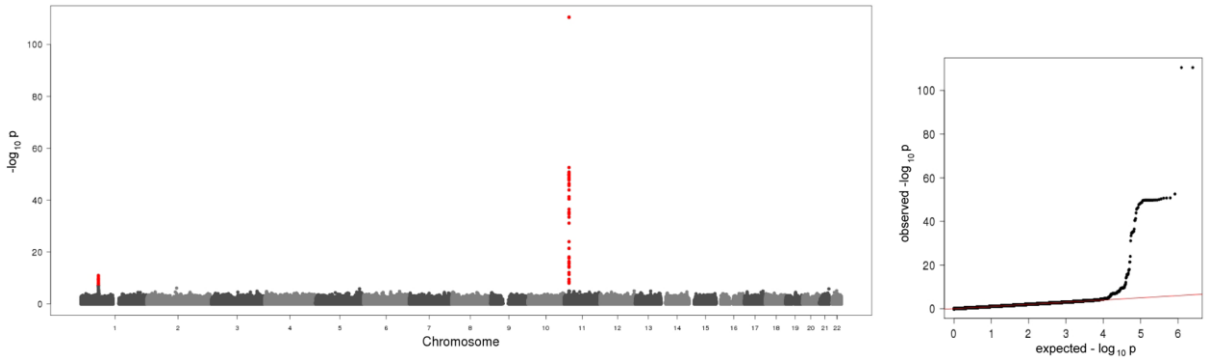

Figure 7: Manhattan plot and quantile-quantile plot of the results of the meta-analysis on baseline A-SAA levels.

The Manhattan plot on the left hand side displays all analyzed SNPs with their calculated p-values (p-values below the threshold of genome-wide significance are colored red). The quantile-quantile plot on the right hand side points out the observed significant associations beyond those expected by chance.

Table 4. Study-specific results for the hits within the regions/ subregions.

chr	SNP region/subregion	study	effect allele	other allele	imp	Effect allele freq	n	beta	se(beta)	p
11	**rs4150642**	KORA S4	G	C	I	0.200	1785	0.529	0.032	1.92E-58
	11p15.5-p13	LURIC	G	C	I	0.212	961	0.569	0.075	6.51E-14
	region	Sorbs	G	C	I	0.212	883	0.524	0.042	1.22E-32
		TwinsUK	G	C	I	0.109	550	0.358	0.050	8.27E-13
		combined	**G**	**C**	**-**	**0.193**	**4179**	**0.501**	**0.022**	**3.20E-111**
		validation	G	C	G	0.217	2082	0.473	0.029	7.41E-58
11	**rs4638289**	KORA S4	A	T	I	0.347	1785	0.323	0.027	4.95E-31
	SAA1	LURIC	A	T	I	0.358	961	0.235	0.063	2.17E-04
	subregion	Sorbs	A	T	I	0.320	883	0.405	0.044	1.71E-19
		TwinsUK	A	T	I	0.214	514	0.206	0.041	5.70E-07
		combined	**A**	**T**	**-**	**0.327**	**4143**	**0.305**	**0.020**	**2.77E-53**
		validation	A	T	G	0.332	2091	0.292	0.025	1.61E-30
11	**rs4353250**	KORA S4	T	C	I	0.339	1785	0.278	0.027	1.06E-24
	HPS5/GTF2H1	LURIC	T	C	I	0.351	961	0.364	0.060	2.21E-09
	subregion	Sorbs	T	C	I	0.413	883	0.310	0.035	3.21E-18
		TwinsUK	T	C	I	0.256	577	0.188	0.035	9.01E-08
		combined	**T**	**C**	**-**	**0.346**	**4206**	**0.272**	**0.018**	**1.68E-51**
		validation	T	C	G	0.352	2125	0.345	0.025	5.82E-42
11	**rs2896526**	KORA S4	G	A	I	0.177	1785	0.265	0.033	3.99E-15
	LDHA/LDHC	LURIC	G	A	I	0.181	961	0.271	0.079	6.58E-04
	subregion	Sorbs	G	A	I	0.163	883	0.226	0.048	3.00E-06
		TwinsUK	G	A	I	0.156	582	0.123	0.042	3.27E-03
		combined	**G**	**A**	**-**	**0.172**	**4211**	**0.221**	**0.023**	**4.12E-22**
		validation	G	A	G	0.187	2097	0.216	0.031	8.18E-12
1	**rs12753193**	KORA S4	A	G	I	0.613	1785	0.103	0.027	1.94E-04
	1p31 (*LEPR)*	LURIC	A	G	I	0.601	961	0.094	0.064	1.44E-01
	region	Sorbs	A	G	I	0.599	883	0.177	0.038	4.15E-06
		TwinsUK	A	G	G	0.628	583	0.133	0.033	5.86E-05
		combined	**A**	**G**	**-**	**0.609**	**4212**	**0.125**	**0.018**	**1.22E-11**
		validation	A	G	G	0.606	2127	0.085	0.025	8.60E-04
11	**rs549485**	KORA S4	T	C	G	0.335	871	0.155	0.040	9.77E-05
	males only	LURIC	T	C	G	0.368	691	0.105	0.076	1.17E-01
	11p14 *(SERGEF)*	Sorbs	T	C	G	0.309	361	0.272	0.144	3.18E-05
	region	**combined**	**T**	**C**	**-**	**0.342**	**1923**	**0.173**	**0.031**	**2.76E-08**
		vlidation*	C	G	G	0.349	1049	0.091	0.035	8.50E-03

*In the validation analysis rs549485 was replaced by rs493767 (r^2=0.961, 3rd lowest p-value within this region in the gender stratified meta-analysis) for technical reasons.

The first region (193.3 kb of length) resides at 11p15.5-p13 and includes *SAA1* one of the structure genes of A-SAA. Within this region the strongest association was found for two highly correlated intronic polymorphisms of the general transcription factor 2 H1 *(GTF2H1)* gene, rs4150642 ($p = 3.20\times10^{-111}$) and rs7103375 ($p = 3.26\times10^{-111}$) (Figure 8A). These two top hits show modest correlation ($r^2 \leq 0.376$) with other significantly associated SNPs within this region. When the structure of correlation and explained variances within the region were analyzed three mostly independent subregions were identified (Table 9 Appendix, Figure 8B – D). The first subregion encloses the 5′ end of *SAA1* (Figure 8B) with strongest association for rs4638289 ($p = 2.77\times10^{-53}$). The other two subregions harbor the genes Hermansky-Pudlak Syndrome 5 *(HPS5)* and *GTF2H1* (Figure 8C) and lactate dehydrogenase A and C (*LDHA* and *LDHC*) (Figure 8D) with strongest associations for rs4353250 ($p = 1.68\times10^{-51}$) and rs2896526 ($p = 4.12\times10^{-22}$), two intronic polymorphisms of *HPS5* and *LDHA*, respectively.

The second region was detected at 1p31 (Figure 8E). All 38 significantly associated variants cluster around the 3′ end of the leptin receptor gene *(LEPR)*. The SNP with the lowest p-value, rs12753193, ($p = 1.22\times10^{-11}$) is located downstream of *LEPR*.

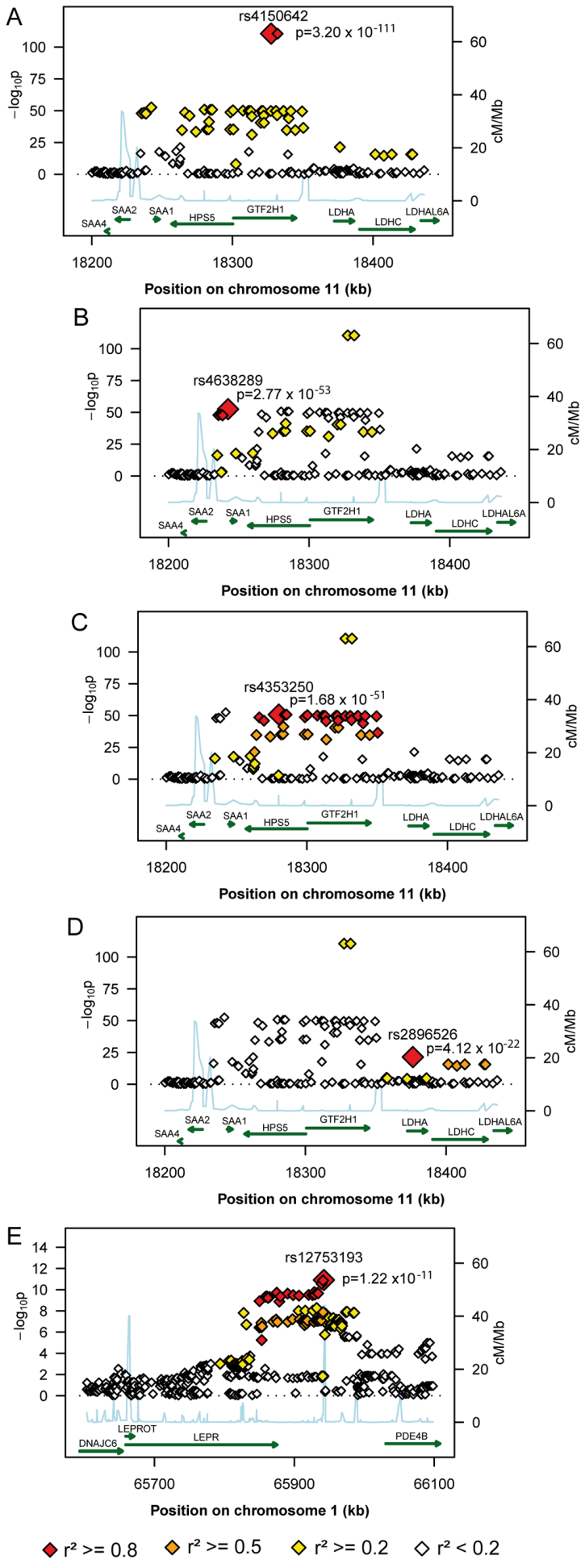
A
rs4150642
p=3.20 x 10-111
B
rs4638289
p=2.77 x 10-53
C
rs4353250
p=1.68 x 10-51
D
rs2896526
p=4.12 x 10-22
E
rs12753193
p=1.22 x10-11
−log10p
cM/Mb
SAA4
SAA2
SAA1
HPS5
GTF2H1
LDHA
LDHC
LDHAL6A
LEPROT
LEPR
DNAJC6
PDE4B
Position on chromosome 11 (kb)
Position on chromosome 1 (kb)
r² >= 0.8
r² >= 0.5
r² >= 0.2
r² < 0.2

Figure 8: ***Regional plots of the genetic susceptibility regions/ subregions.***
Regional plots present gene regions and block structures of the region at 11p15.5-p13 (A), the *SAA1* subregion (B), the *HPS5/GTF2H1* subregion (C), the *LDHA/LDHC* subregion (D), and the region at 1p31 (E) and picture the probability values of the significantly associated SNPs, the colour representing the degree of correlation with the top hit of the respective region/subregion.

All associations were consistent in the KORA S4 validation analyses (Table 4 for the top hits, Table 10 Appendix for all SNPs selected for validation).

The entire regression model (including the top SNPs of the two genomic regions (rs4150642 and rs12753193), age, gender and BMI) explains 19.32 % of the total variation of A-SAA in our data. With an explained variance of 10.84 % for the top SNP (rs4150642) of the 11p15.5-p13 locus (5.57 % for rs4638289 of the *SAA1* subregion, 5.34 % for rs4353250 of the *HPS5/GTF2H1* subregion, and 2.37 % for rs2896526 of the *LDHA/LDHC* subregion; Table 9 Appendix) and 0.93 % for the top SNP (rs12753193) of the 1p31 locus the identified genomic regions account for a major part of such variance.

When the analysis was stratified by gender, one additional SNP (rs549485) located about 350 kb apart from the *SAA1* subregion at 11p14 in the secretion regulating guanine nucleotide exchange factor (*SERGEF*) gene showed a borderline significant association with A-SAA levels in men ($p = 2.76\times10^{-8}$) in the meta-analysis. In the validation analysis the association between two highly correlated SNPs within this region (rs493767 and rs550659, $r^2 = 0.961$) and A-SAA levels was also borderline significant ($p = 8.50\times10^{-3}$ and $p = 1.65\times10^{-2}$, respectively). No significant differences between men and women were found within the regions identified in the overall meta-analysis (data not shown).

3.2.3. Discussion

Genetic susceptibility loci for circulating levels of A-SAA

Based on a meta-analysis of four genome-wide association studies including 4,212 participants of European descent two novel genetic susceptibility regions were identified to be associated with baseline A-SAA concentrations in this study. With 11.68 % explained variance in this data, which makes up 19.76 % of the total estimated heritability of 59 %, these two regions seem to have a major impact on baseline A-SAA concentrations.

A-SAA susceptibility region 11p15.5-p13

The region at 11p15.5-p13 accounts for most of the explained variance. Its *SAA1* subregion contains part of a highly conserved region between the two *bona fide* acute-phase structure genes *SAA1* and *SAA2*, which consist of at least 5 and 2 allelic variants, respectively [116, 127]. These two genes are concurrently induced during the acute-phase [116], and cluster within 18 kb of each other in a head to head arrangement [128]. This study is the first presenting the complex genetic architecture of A-SAA levels at this locus. In the identified region, there has been evidence of regulatory elements like C/EBPalpha and C/EBPbeta (http://genome.ucsc.edu), which are necessary for the full responsiveness to IL-1β and IL-6 either alone or in combination [116]. This finding underlines the high functional potential for this region.

The adjacent *GTF2H1* is a basal transcription factor involved in nucleotide excision repair of DNA and RNA transcription by RNA polymerase II [211]. *HPS5* encodes a protein, which is probably involved in organelle biogenesis associated with melanosomes and platelet dense granule, its mutations lead to a homonymous clinical entity [212]. And *LDHA* and *LDHC*, which are expressed in muscle tissue and in testes, respectively, encode for lactate dehydrogenase, an enzyme that catalyzes the interconversion of lactate and pyruvate [213].

Variants of the *GTF2H1* gene have been recently found to be associated with lung cancer in a Chinese population [214]. Furthermore, it was demonstrated that *LDHA* is involved in tumour genecity and its reduction causes bioenergetic and oxidative stress leading to cell death [215-217]. Finally, Kosolowski et al. [218] found *LDHC* to be expressed in several types of tumour cell lines. It is thought, that recurrent or persistent chronic inflammation may play a role in carcinogenesis by causing DNA damage, inciting tissue reparative proliferation and/ or by creating an environment that is enriched with tumour-promoting cytokines and growth factors [146]. Furthermore, SAA synthesis has been found in human carcinoma metastases and cancer cell lines [119].

As the approach taken in this study is observational in nature it is not possible to draw causal inferences. For that reason, it could be possible that not genes, but small regulatory elements may be responsible for the findings. This is most likely the case as the identified region contains one structure gene and the adjacent region. In any case, the major impact on baseline A-SAA concentrations

demonstrates a key role of the 11p15.5-p13 region in the regulation of inflammation. Therefore, the identification of causal variants and their impact on diseases related to elevated baseline A-SAA concentrations represent promising targets for future functional and epidemiological studies.

A-SAA susceptibility region 1p31

The second region was found on chromosome 1p31, harboring the *LEPR* gene locus. Leptin, an important circulating signal for the regulation of body weight, was found to be correlated with SAA concentrations independently of BMI, and both were expressed in adipose tissue [219]. In the KORA F3 study (Text 1 Appendix) a moderate but significant correlation was found between circulating A-SAA and leptin concentrations in blood in 181 participants with measurements of both proteins (Spearman correlation = 0.25, $p < 0.001$). So far it is unclear whether leptin influences SAA expression directly or via the leptin stimulated cytokines, IL-6 and TNF-α [219]. *LEPR* is a single transmembrane receptor of the cytokine receptor family most related to the gp130 signal-transducing component of the IL-6 receptor, the granulocyte colony-stimulating factor (GCSF) receptor, and the Leukaemia Inhibitory Factor (LIF) receptor, all of which are thought to play an essential role in the inflammatory process [220, 221]. Previous studies have provided evidence of an association of the *LEPR* gene locus with hs-CRP and fibrinogen [222-225], which were both correlated with A-SAA in the KORA S4 study (Text 1 Appendix) (hs-CRP: Spearman correlation = 0.58, $p < 0.001$, and fibrinogen: Spearman correlation = 0.31, $p < 0.001$; $N = 1734$). The finding gives confirmative evidence of the importance of the *LEPR* gene locus for inflammatory processes and the close relationship between leptin, A-SAA, hs-CRP and fibrinogen.

Gender specific finding

Furthermore, in the gender stratified analysis one region containing *SERGEF* was identified to be presumably associated with A-SAA in men. The adjacency of this identified region to the *SAA* gene family suggests that regulatory elements may be responsible for this signal. However, the association with A-SAA levels was only borderline significant in this study and therefore awaits replication.

Limitations and strengths

Two limitations of the study have to be mentioned. Firstly, like in the first study of this doctoral thesis due to the restrictions in laboratory methods analyses were confined to the A-SAA isoforms and did not capture the constitutively expressed C-SAA isoform which might also be of interest, especially when analyzing baseline SAA levels. Secondly, the number of studies with genome-wide data and measured A-SAA levels was limited compared to other genome-wide association studies. Nevertheless, the study had enough power to detect two novel genetic susceptibility regions for circulating A-SAA levels explaining 19.76 % of the total estimated heritability already. Furthermore, results were consistent across all four studies and within different genotyping platforms, the regions

are biologically highly plausible, and the results may contribute to future research on the regulation of inflammatory response and its role in related clinical entities.

3.3. Mendelian Randomization study on genetic susceptibility loci for A-SAA and type 2 diabetes

3.3.1. Materials and Methods

The Mendelian Randomization study was conducted within the framework of a DIAGRAM meta-analysis of eight GWAs on type 2 diabetes by Voight, B. et al., 2010 [39]. DIAGRAM is a large consortium which combines the efforts of about 40 scientific institutions working on the genetics of type 2 diabetes, particularly, in samples of European descent. The association between genetic variants which predispose to levels of A-SAA and type 2 diabetes was extracted from the results of this DIAGRAM meta-analysis. The study population and the statistical analysis are described in detail in this study [39] and are only summarized here.

Study population of the published DIAGRAM meta-analysis

The meta-analysis combined results of eight genome-wide scans comprising 8,130 type 2 diabetes cases and 38,987 controls of European descent: DGDG (Diabetes Gene Discovery Group), deCODE, DGI (Diabetes Genetics Initiative), Rotterdam, EUROSPAN (European Special Population Network), FUSION (Finland-United States Investigation of NIDDM Genetics), KORAgen (Cooperative Health Research in the region of Augsburg, Southern Germany) and WTCCC/ UKT2D (Welcome Trust Case Control Consortium/ United Kingdom Type 2 Diabetes Genetics Consortium).

Statistical analysis of the published DIAGRAM meta-analysis

Less than 1,000 SNPs, where the allele frequencies were < 0.40 in at least one study and > 0.60 in others, were excluded to minimize possible difficulties due to mislabeled alleles. Genomic control correction of autosomal data from each individual study (separately for directly genotyped and for imputed data) was performed by inflating the standard error of the estimated effect so the significance of the SNP matched that of the genomic control value. The overall genomic control inflation factor was estimated to be 1.074 for $\geq$ 17,000 effective total samples indicating low population stratification. The genome-wide association analysis in the individual studies was conducted using an additive model, different covariates including age, sex, BMI, study center and/or birth province, and different analysis software as described in the online supplement of the DIAGRAM study [39]. Association data of the individual studies comprising 2,426,886 imputed and genotyped autosomal SNPs were combined in a meta-analysis using a fixed effect model and the inverse variance-weighted approach. Heterogeneity between study results was assessed by calculating I^2.

Genetic susceptibility regions for the Mendelian Randomization study

For the Mendelian Randomization study, the SNPs with the smallest p-values of the two A-SAA susceptibility regions as well as the SNPs with the smallest p-value of each subregion which had been identified in the second study of this doctoral thesis were selected.

Power calculation for the Mendelian Randomization study

A power calculation was conducted to investigate whether the power to detect a significant association between the selected A-SAA loci and type 2 diabetes was sufficient. Therefore, effect estimates for the associations of the five selected SNPs and type 2 diabetes were calculated by combining results of the two studies presented prior to this study in the present doctoral thesis as following: the beta value of the meta-analysis of GWAS for the KORA study was divided by the standard deviation of ln-transformed A-SAA levels (0.66 mg/l) and multiplied by the natural logarithm of the OR for 1-SD increase of ln-transformed A-SAA levels in type 2 diabetes risk adjusted for age and sex (OR = 1.42, ln(OR) = 0.35). The result was subsequently raised to the power of e. To assess the power of the Mendelian Ranodmization study given the calculated combined effect estimates, allele frequencies as reported in the meta-analysis of genome-wide scans on levels of A-SAA for the KORA study, and an additive genetic model, a power calculation was performed using Genetic Power Calculator by S. Purcell and P. Sham, 2001-2009 [226]. The effective sample size was calculated in the DIAGRAM meta-analysis to be 21,063 [39]. Diabetes prevalence was assumed to be 8.1 % [13] and the threshold of significance was considered to be 0.05.

3.3.2. Results

Results of the DIAGRAM association analyses between the five selected SNPs and type 2 diabetes adjusted for age and sex are presented in Table 5. There was no significant association between any of the selected A-SAA susceptibility loci and type 2 diabetes.

Table 6 shows the combined effect estimates for the association between the five selected SNPs and type 2 diabetes based on the results of the two previous studies of this doctoral thesis (beta values of the meta-analysis of GWAS on A-SAA levels and ORs for a 1-SD increase of ln-transformed A-SAA levels adjusted for age and sex of the prospective cohort study on A-SAA levels and incident type 2 diabetes) as well as the power to detect a significant association between the five selected SNPs and type 2 diabetes.

Table 5. Results of the association analysis between five selected A-SAA susceptibility SNPs and type 2 diabetes in the DIAGRAM consortium (8,130 type 2 diabetic cases, 38,987 non-diabetic controls).

SNP	Chr	Region/ subregion	Study*	Effect allele	Other allele	OR [95% CI]	p-value
rs4638289	11	*SAA1*	+-+++--+	A	T	1.04 [0.99-1.09]	0.163
rs2896526	11	*LDHA/LDHC*	+-+++-++	G	A	0.98 [0.93-1.04]	0.569
rs4150642	11	*11p15.5-p13*	--++-++-	G	C	0.99 [0.94-1.06]	0.854
rs4353250	11	*HPS5/GTF2H1*	--++-+++	T	C	1.00 [0.96-1.05]	0.906
rs12753193	1	*1p31 (LEPR)*	+++--+--	A	G	1.00 [0.96-1.04]	0.932

(+) = SNP was analyzed in the study, (-) = SNP was not analyzed in the study. *The order of the studies is DGDG, deCode, DGI, Rotterdam, EUROSPAN, FUSION, KORAgen, and WTCCC/ UKT2D.

Table 6. Effect estimates for the association between the selected SNPs and type 2 diabetes and the estimated power to detect these effect estimates in the DIAGRAM meta-analysis.

		Study 1	Study 2	Combined	Power calculation
	Effect allele frequency	Beta	OR	OR [95% CI]	Power
rs4638289	0.347	0.323	1.42	1.19 [1.12-1.24]	1
rs2896526	0.177	0.265	1.42	1.15 [1.09-1.20]	1
rs4150642	0.200	0.529	1.42	1.32 [1.24-1.37]	1
rs4353250	0.339	0.278	1.42	1.16 [1.10-1.21]	1
rs12753193	0.613	0.103	1.42	1.05 [1.00-1.10]	0.95

ORs adjusted for age and sex were calculated for 1-SD of ln-transformed A-SAA levels.

The power to detect a significant association between the five selected SNPs and type 2 diabetes was > 0.8 for all SNPs. Likewise, the number of cases needed to detect a significant association between the five selected SNPs and type 2 diabetes given a power of 0.8 was far below the effective sample size of 21,063 for all five selected SNPs (585, 972, 1536, 2340, and 12,554, respectively). Sensitivity analysis showed that also for an assumed OR of 1.22 for study 2 which corresponds to an OR for a 1-SD increase of ln-transformed A-SAA levels adjusted for age, sex, BMI, smoking, alcohol intake, physical activity, education, parental history of diabetes, non-HDL cholesterol, fasting triglycerides, systolic blood pressure, fasting glucose, fasting insulin, HbA1c and 2h-glucose the power to detect a significant association between all selected SNPs and type 2 diabetes was > 0.8 (data not shown).

3.3.3. Discussion

This Mendelian Randomization was designed to assess whether elevated levels of A-SAA are causally involved in the development of type 2 diabetes. However, in this Mendelian Randomization study conducted in 8,130 type 2 diabetes cases and 38,987 controls subjects of European descent no significant association between susceptibility loci for circulating levels of A-SAA and type 2 diabetes was found. There may be three possible explanations for this finding: (i) genetic confounding might have biased the association between the selected genetic variants and type 2 diabetes, (ii) the power to detect a significant association was not sufficient, and (iii) the selected SNPs that predispose to plasma levels of A-SAA do not increase type 2 diabetes risk.

Genetic confounding

The association between the A-SAA susceptibility loci and type 2 diabetes might have been biased by genetic confounding evoked by pleiotropic effects or population stratification.

No evidence for pleiotropic effects had been found in electronic searches of Medline and PubMed, supplemented by hand searches of reference lists for the selected SNPs of the 11p15.5-p13 region. Therefore, although it is not possible to totally exclude pleiotropic effects they do not seem likely for the 11p15.5-p13 region. Contrary to the 11p15.5-p13 locus, the A-SAA susceptibility region at 1p31 harboring the LEPR gene had been found to be associated with further intermediate phenotypes of type 2 diabetes, namely leptin, fibrinogen, and hs-CRP [222-225]. All studies report on the same risk allele for the intermediate phenotypes. However, the associations between these intermediate phenotypes and type 2 diabetes might be oppositional as higher levels of fibrinogen and hs-CRP have been found to be associated with type 2 diabetes [95] whereas high levels of leptin have been found to be both, a risk factor and a protective factor for incident type 2 diabetes [227, 228]. Thus, the association between the 1p31 locus and type 2 diabetes might be biased by pleiotropic effects and the direction of this bias cannot be clearly determined.

Furthermore, a negative finding might be prone to population stratification. This Mendelian Randomization study was accomplished within the framework of a meta-analysis of eight different studies. Participants of all studies were of European descent and study specific results were corrected for population stratification using the genomic control method in all individual studies. The overall inflation factor was close to 1 indication low population stratification.

Power calculation

Finally, the absence of association of the selected SNPs with type 2 diabetes could be due to low statistical power. A valid power calculation for Mendelian Randomization studies requires the robust evaluation of the association between the genetic variant and the intermediate phenotype as well as

the valid assessment of effect estimates between the intermediate phenotype and the disease. Being the first genome-wide scan for A-SAA susceptibility loci ascertainment bias known as “the winner’s curse” and thus an overestimation of the effect size for the gene-intermediate phenotype may be present in the meta-analysis of GWAS on circulating A-SAA levels. Yet, results were consistent across all studies and within different genotyping platforms. In addition, the effect estimates that were assessed in the validation sample of the meta-analysis of GWAS on A-SAA levels displayed only slight differences from those assessed in the KORA study and the meta-analysis. Furthermore, the findings of the prospective cohort study indicated a possible reverse causality between post-challenge hyperglycemia and levels of A-SAA in the A-SAA-incident type 2 diabetes association. In this case, effect estimates and thereby combined effect estimates and the power to detect a significant gene-disease association might also be overestimated. However, sensitivity analyses using effect estimates adjusted for parameters of glucose metabolism in the power calculation for the association between A-SAA levels and type 2 diabetes still displayed a sufficient power to detect a significant association between all selected SNPs and type 2 diabetes. Finally, minor differences might arise out of different adjustments for covariates in the studies of the DIAGRAM meta-analysis as well as in the power calculation of this Mendelian Randomization study. Yet, altogether, power calculations showed that the power to detect a significant association between all selected SNPs of the 11p15.5-p13 was by far sufficiently high and the absence of a significant association between the selected SNPs and type 2 diabetes due to low statistical power seems unlikely.

To sum up, for reasons of potential pleiotropic effects, the 1p31 susceptibility region was not qualified for the Mendelian Randomization study. In contrast, for the 11p15.5-p13 region no pleiotropic effects had been reported. Furthermore, population stratification was low and also the power to detect a significant association with type 2 diabetes for the 11p15.5-p13 region was reasonably high. Therefore, for the 11p15.5-p13 region, the negative results of the study suggest that there are genetic mechanisms that determine levels of A-SAA without affecting type 2 diabetes risk. Thus, although a presumed causality driven by other genetic and/ or non-genetic mechanisms might still be possible, evidence for a causal implication of elevated levels of A-SAA into type 2 diabetes pathogenesis on the basis of the results of this Mendelian Randomization study cannot be provided.

4. Conclusions

This chapter is partly extracted from the manuscripts Marci C. et al., ***Acute-phase serum amyloid A protein and its implication in the development of type 2 diabetes in the KORA S4/ F4 study***, Diabetes Care, 2012 Dec 13, epub ahead of print, and Marzi, C. et al., ***Genome-Wide Association Study Identifies Two Novel Regions at 11p15.5-p13 and 1p31 with Major Impact on Acute-Phase Serum Amyloid A,*** PloS Genetics PLoS Genet. 2010 Nov 18;6(11):e1001213.

Type 2 diabetes is an increasing public health problem and the identification of etiological risk factors is of ascending relevance for the prevention as well as the treatment of the disease. There had been emerging evidence from epidemiological, genetic and functional studies that subclinical inflammation is causally involved in the development of type 2 diabetes (chapter 1.2.1.). However, little was known about the inflammatory protein A-SAA, one of the key molecules in host defense, and its implication in type 2 diabetes pathogenesis. It had been shown in cross-sectional studies that elevated levels of A-SAA are correlated with prevalent type 2 diabetes, and related metabolic traits (chapter 1.2.2.) but these studies do not allow inferences on cause and effect (chapter 1.3.1.). Causal inferences in epidemiological studies are difficult to assess because epidemiology is an observational science. One first indication for causality is a significant association between risk factor and disease in a prospective cohort study. These study type incorporates the time sequence between risk factor and disease, an essential prerequisite for causality (chapter 1.3.1). A prospective cohort study on circulating A-SAA levels and incident type 2 diabetes had previously been conducted in 492 Aboriginal Canadians. However, in this study the association between levels of A-SAA and incident type 2 diabetes was not significant. This might be due to low statistical power. Furthermore, there might be ethnic differences, and thus, deductions on an implication of elevated levels of A-SAA in the development of type 2 diabetes in Western Europeans cannot be drawn on the basis of these findings.

Therefore, a **prospective cohort study** was performed to analyze the association between circulating levels of A-SAA and the development of type 2 diabetes in Western Europeans (chapter 3.1.). In this study, A-SAA concentrations were measured in 836 initially non-diabetic elderly Western European subjects, 55-74 years of age, without clinically overt inflammation who participated in a seven-year follow-up examination including an OGTT. Results of the study provided novel evidence that elevated levels of A-SAA are present years before the manifestation of type 2 diabetes independent of a variety of other type 2 diabetes risk factors but not parameters related to glucose metabolism.

As the next step, the causality in the association between levels of A-SAA and type 2 diabetes was to be assessed. Clinical trials and intervention studies my assess causality in an association by analyzing the difference between two groups with initially equal risk factor profiles after the interesting risk factor has been changed or eliminated in one of the two groups (chapter 1.3.1.). This concept has

recently been adapted in genetic epidemiology and named Mendelian Randomization study (chapter 1.3.3.). Mendelian Randomization studies provide the opportunity to detect associations between genetic variants as proxies for intermediate phenotypes and a disease largely unbiased by potential environmentally modifying factors. Thereby, they might be particularly useful for the identification of causal risk factors for complex diseases characterized by multifactorial environmental contributions to pathogenic mechanisms such as type 2 diabetes (chapter 1.3.3.).

The identification of genetic variants that determine biomarker levels is a prerequisite for a Mendelian Randomization study. Apart from this, in the case of A-SAA levels, it could also provide important clues to the immune response pathways involved in the regulation of A-SAA levels. This might also be of relevance for related clinical entities such as type 2 diabetes. Twin studies suggested a substantial genetic contribution of about 60 % to baseline A-SAA concentrations but in the past association analyses between genetic variants and A-SAA levels had been limited to allelic variants of the SAA genes and protein concentrations (chapter 1.2.2.). The second study of this doctoral thesis was the first whole genome approach to identify genetic variants that determine baseline A-SAA concentrations (chapter 3.2.). In a **meta-analysis of four genome-wide scans** (KORA S4, LURIC, TwinsUK, and Sorbs) the association between 2,593,456 genetic variants across the whole genome and circulating levels of A-SAA was analyzed in 4,212 participants of European descent. Two novel genetic susceptibility regions at chromosome 11p15.5-p13 and chromosome 1p31 were identified to be associated with baseline A-SAA concentrations. The findings demonstrated a major impact of the 11p15.5-p13 region on the regulation of inflammation and suggested a close interplay between leptin, A-SAA, and other acute-phase proteins, as well as a larger role of the *LEPR* gene locus in inflammatory processes as it had been assumed in the past. Both loci are biologically highly plausible A-SAA susceptibility regions and might represent promising targets for future functional and epidemiological studies to elucidate the up-regulation of A-SAA and its differential contribution to related diseases.

The **Mendelian Randomization study** was accomplished within the framework of a published meta-analysis of eight GWAS (8,130 type 2 diabetic cases, 38,987 non-diabetic controls) on type 2 diabetes (chapter 2.3.). This meta-analysis was conducted by scientists of the DIAGRAM consortium, a consortium of about 40 research institutes working on the genetics of type 2 diabetes, mainly, in samples of European descent. For the Mendelian Randomization study, those SNPs with the lowest p-values in each region/ subregion identified in the meta-analysis of GWAS on A-SAA levels were selected as proxies for the Mendelian Randomization study and the association results were extracted from the results of the DIAGRAM meta-analysis. For the 11p15.5-p13 region no pleiotropic effects had been reported. In contrast, for reasons of potential pleiotropic effects and thus genetic confounding, the 1p31 susceptibility region was not qualified for the Mendelian Randomization study.

In spite of a reasonably high power the association between the selected SNPs of the 11p15.5-p13 A-SAA susceptibility region and type 2 diabetes was not statistically significant. This suggests that there are genetic mechanisms that raise plasma levels of A-SAA without translating into an increase in type 2 diabetes risk. Thus, the Mendelian Randomization study could not provide evidence that elevated levels of A-SAA are causally involved in type 2 diabetes pathogenesis.

However, a putative causality between elevated levels of A-SAA and type 2 diabetes driven by other genetic and/ or non-genetic mechanisms might still be present. One possibility to further investigate this subject might be animal studies with knockouts or transgenic overexpression of A-SAA. However, although SAA genes and proteins are highly conserved throughout the evolution of vertebrates and invertebrates there are substantial differences between humans and mice, e.g. concerning the functionality of *SAA3* (chapter 1.2.2.). Furthermore, the manipulation of the genome might constitute too gross disturbances to allow a transfer of results from animal models to the complex and delicate human immune and metabolic system [96]. Thus, animal studies might corroborate or falsify the hypothesis of a causal implication of A-SAA levels in the development of type 2 diabetes, but firm conclusions on human type 2 diabetes pathogenesis may not be drawn on the basis of such studies.

Another approach to investigate causality in the association between circulating A-SAA levels and type 2 diabetes are pharmacological or lifestyle and behavioral intervention studies in humans. However, pharmacological interventions need to target inflammatory pathways only and may not be interwoven into glucose metabolism to allow deductions on the direction of causation. Furthermore, this approach requires monoclonal antibodies for A-SAA proteins in order to avoid pleiotropic effects, which is hardly feasible in practical research. In lifestyle and behavioral intervention studies, at first glance, the reduction of weight and at least moderate physical activity might be of particular interest in this context as A-SAA is highly correlated with BMI and partly expressed in adipose and muscle tissues (chapter 1.2.2.). Yet, in the prospective study on levels of A-SAA and incident type 2 diabetes, effect estimates decreased only slightly when adjusting for BMI or, alternatively, other measures of obesity or body composition as well as appendicular skeletal muscle mass index. This implies an only moderate effect of intervention studies targeting adiposity and physical inactivity only via a reduction of A-SAA concentrations. Furthermore, results of the prospective study of circulating A-SAA levels and incident type 2 diabetes indicated that the relevance of high circulating levels of A-SAA in the pathogenesis of type 2 diabetes might be age-dependent. Just as well as a possible difference between men and women (chapter 1.2.1.) this age effect, once confirmed, needs to be taken into account in intervention studies. However, like in pharmacological intervention studies, pleiotropic effects make it hard to draw deductions on a causal implication of levels of A-SAA in type 2 diabetes pathogenesis also in lifestyle and behavioral intervention studies.

In the prospective cohort study, the association between levels of A-SAA and incident type 2 diabetes attenuated when adjusting for parameters of glucose metabolism, particularly 2h-glucose. In contrast, baseline levels of A-SAA were significantly associated with levels of 2h-glucose at the time of follow-up even after adjusting for baseline parameters of glucose metabolism (chapter 3.1.). This indicates that elevated levels of A-SAA contribute to post-challenge hyperglycemia suggesting a possible link with type 2 diabetes pathogenesis via elevated levels of 2h-glucose. However, also a possible reverse impact of high levels of 2h-glucose on circulating levels of A-SAA cannot be ruled out. Thus, the association between elevated levels of A-SAA and 2h-glucose might also be bi-directional. Biologically, a mutual impact between A-SAA and 2h-glucose might reflect a positive feed-back loop in the context of endoplasmic reticulum stress, mitochondrial dysfunction, and subclinical inflammation making it difficult to distinguish between inflammatory and metabolic consequence and condition. Time-series studies as performed for other markers of subclinical inflammation [229] might further elucidate the relationship between levels of A-SAA, post-challenge hyperglycemia, and type 2 diabetes.

To summarize, results of this doctoral thesis indicated that levels of A-SAA are elevated years before the manifestation of type 2 diabetes. Furthermore, two highly plausible genetic susceptibility loci for circulating A-SAA levels were identified. One of these loci was a suitable candidate for a genetic approach to assess the causality in the association between elevated levels of A-SAA and type 2 diabetes. However, evidence that the association between A-SAA concentrations and type 2 diabetes is truly causal could not be provided using this genetic approach. The prospective cohort study of this doctoral thesis suggested that the relation between elevated levels of A-SAA, post-challenge hyperglycemia and type 2 diabetes might be mutual and of a more complex nature. Time-series studies are warranted to further clarify this relation.

5. Appendix

Table 1 Appendix. Correlation matrix of co-variables

	A-SAA	Age	BMI	Alcohol intake	Non-HDL	TG	SBP	Fasting insulin	HbA1c	Fasting glucose	2h-glucose	hs-CRP
A-SAA	1	**0.1**	**0.19**	**-0.09**	0.02	0.06	**0.08**	**0.14**	0.06	-0.01	**0.13**	**0.5**
Age		1	**0.09**	-0.04	0.02	-0.01	**0.17**	0.03	**0.15**	0.05	**0.19**	**0.15**
BMI			1	-0.02	**0.08**	**0.22**	**0.12**	**0.48**	0.03	**0.23**	**0.27**	**0.32**
Alcohol intake					0.002	**0.008**	**0.12**	-0.06	**-0.08**	**0.17**	-0.04	-0.02
Non-HDL					1	**0.46**	**0.08**	**0.07**	0.05	0.05	0.05	**0.07**
TG						1	**0.19**	**0.36**	0.05	**0.2**	**0.22**	**0.19**
SBP							1	**0.15**	0.03	**0.21**	**0.19**	**0.16**
Fasting insulin								1	0.05	**0.31**	**0.29**	**0.23**
HbA1c									1	**0.2**	**0.14**	**0.11**
Fasting glucose										1	**0.31**	**0.09**
2h-glucose											1	**0.21**
hs-CRP												1

Non-HDL = non-HDL cholesterol; TG = triglycerides; SBP = systolic blood pressure. Pearson's and Spearman's correlation coefficients for normally and not normally distributed variables, significant correlations ($p<0.05$) are marked in bold. Not normally distributed variables: A-SAA, alcohol intake, triglycerides, fasting insulin, HbA1c, 2h-glucose, and hs-CRP.

Table 2 Appendix. Association between circulating concentrations of A-SAA and incident T2DM adjusted for age, sex and different measures of obesity and body composition

	A-SAA	
Covariables	**OR (95% CI)**	**P**
age + sex + BMI (n=836)	1.25 [1.05-1.50]	**0.01**
age + sex + WHR (n=836)	1.26 [1.05-1.51]	**0.01**
age + sex + fat mass index (n=832)	1.26 [1.05-1.51]	**0.01**
age + sex + lean body mass index (n=832)	1.30 [1.08-1.56]	**0.005**
age + sex + appendicular skeletal muscle mass index (n=832)	1.31 [1.09-1.57]	**0.005**

Odds ratios (95% CI) are given for a 1-SD increase of A-SAA concentrations (4.16mg/l). Fat mass index, lean body mass index and appendicular skeletal muscle mass index have been calculated in kg/m^2 as proposed by Kyle [204, 205].

Table 3 Appendix. Association between elevated levels of A-SAA and incident type 2 diabetes adjusted for age, sex and, alternatively, menopausal status, hormone replacement therapy, liver disease and the intake of aspirin

	A-SAA	
Covariates	**OR (95% CI)**	**P**
age + sex + menopausal status (n=836)	1.28 [1.07-1.53]	**0.006**
age + sex + HRT (n=835)	1.28 [1.08-1.53]	**0.005**
age + sex + liver disease (n=836)	1.28 [1.08-1.53]	**0.005**
age + sex + intake of aspirin (n=836)	1.28 [1.08-1.53]	**0.006**

HRT=hormone replacement therapy
Odds ratios (95% CI) are given for a 1-SD increase of A-SAA concentrations (4.16mg/l). Menopausal status is coded in eight categories (ovarectomy and hysterectomy, only ovarectomy, only hysterectomy, either ovarectomy or hysterectomy, and no information on the other variable, pre-menopausal, post-menopausal, sound uterus and hormone therapy, and male). HRT is coded in three categories (yes, no, male). Liver disease within the last 12 months is coded in three categories (yes, no, no information). Intake of aspirin is coded in yes/no.

Text 1 Appendix. Study descriptions

KORA studies

The Cooperative Health Research in the Region of Augsburg (KORA) study is a series of independent population-based epidemiological surveys and follow-up studies of participants living in the region of Augsburg, Southern Germany. All participants are residents of German nationality identified through the registration and informed consent has been given by all participants [230]. The study has been approved by the local ethics committee. The present study includes data of the KORA S4 survey (1999/2000), the follow-up study F4 (2006-2008) and the follow-up study of KORA S3 survey (1994/1995) F3 (2004/2005). For the genome-wide association study we genotyped 1,814 randomly selected participants of KORA F4 with Affymetrix 6.0. A-SAA measurements were available for 1,785 of these participants. Additionally, 27 SNPs, which were identified in the meta-analysis, were genotyped with the MassARRAY system using the iPLEX technology (Sequenom) in all participants with A-SAA measurements of the KORA S4 study (N=3,677) thus providing one sample for validation (N=2,136) and one sample for the comparison of the genotyping technologies (N=1,541, corresponding to the sample genotyped with Affymetrix 6.0). For the comparison of the genotyping technologies, we used an additional subsample of 717 KORA S4 participants with genotyping data available from the Illumina 550K platform. Correlation analysis was conducted in all participants of

S4 (N=4,261) and F3 (N=3,006) where measurements of SAA, hs-CRP, fibrinogen and leptin were available. Methods for measurement have been described in detail elsewhere [148, 231].

LURIC study

The **Lu**dwigshafen **Ri**sk and **C**ardiovascular Health (LURIC) study includes consecutive white patients of Caucasian origin (17 to 92 years of age) hospitalized for coronary angiography between June 1997 and May 2001 [232]. The study was approved by the ethics review committee at the "Landesärztekammer Rheinland-Pfalz". Written informed consent was obtained from each of the participants. Clinical indications for angiography were chest pain or non-invasive tests consistent with myocardial ischemia. To limit clinical heterogeneity, individuals suffering from acute illness other than acute coronary syndromes, chronic non-cardiac diseases and a history of malignancy within the five past years were excluded. The present study includes data of 1002 participants for whom Affymetrix 6.0 genotyping data was available.

Sorbs study

All subjects are part of a sample from an extensively phenotyped self-contained population from Eastern Germany, the Sorbs (PMID: 19584900; PMID: 19729412). The Sorbs are of Slavonic origin, and lived in ethnic isolation among the Germanic majority during the past 1100 years. At present, about 1000 Sorbian individuals are enrolled in the study. Sampling comprised unrelated subjects as well as families. Mean IBD sharing in the pairwise comparison was 0.008, median $<10^{-6}$ (25% percentile $<10^{-6}$, 75% percentile: 0.012). Extensive phenotyping included standardised questionnaires for past medical history and family history, collection of anthropometric data and a 75g-Glucose-tolerance-test. The study was approved by the ethics committee of the University of Leipzig and all subjects gave written informed consent before taking part in the study.

TwinsUK study

The TwinsUK adult twin registry based at St Thomas' Hospital in London is a volunteer cohort of over 10,000 twins recruited among the general population. Twins largely volunteered unaware of any their phenotypic status in relationship to SAA or any other phenotypic trait of interest to the authors and they gave fully informed consent under a protocol reviewed by the St Thomas' Hospital Local Research Ethics Committee. All participants in the TwinsUK study were female and SAA measurements were obtained as described in the main Methods section.

Table 4 Appendix. Study characteristics of the four studies of the meta-analysis and the validation sample

Study	**KORA S4**	**LURIC**	**Sorbs**	**TwinsUK**	**validation sample**
N all (males/females)	1785 (874/911)	961 (693/268)	883 (361/522)	583 (0/583)	2136 (1069/1067)
Age mean [years](sd)	53.91 (8.87)	64.14 (10.80)	47.89 (16.31)	56.74 (10.39)	45.66 (15.76)
BMI mean [kg/m²](sd)	27.72 (4.55)	26.53 (4.52)	27.05 (4.94)	24.67 (4.03)	26.87 (4.75)
SAA geometric mean [mg/L]	3.35	6.61	3.09	5.61	3.10

Table 5 Appendix. Study specific information on genotyping and imputation

study		**KORA S4**	**LURIC**	**Sorbs**	**TwinsUK**
genotyping platform		Affymetrix 6.0	Affymetrix 6.0	Affymetrix 500K Affymetrix 6.0	Illumina 317k Illumina 610k
calling algorithm		Birdseed	Birdseed	BRLMM Birdseed	Illuminus
QC filters before imputation	individual callrate	≥ 0.93	≥ 0.96	none	≥ 0.95
	SNP callrate	none	none	≥ 0.95	≥ 0.95
	HWE	none	none	< 0.0001	< 0.0001
	MAF	none	none	> 0.01	> 0.01
imputation software		IMPUTE v0.4.2	IMPUTE v0.4.2	IMPUTE v1.0.0	IMPUTE v1.0.0
number of SNPs in meta-analysis		2,498,895	2,515,584	2,363,460	2,416,789
inflation factor		0.999	1.025	1.184	1.136
statistical software		SNPTEST v2.1.0	SNPTEST v2.1.0	QUICKTEST v0.95	Merlin v1.1.2

Table 6 Appendix. List of all significantly associated SNPs of the meta-analysis

SNP	chr	pos	effect allele	other allele	ß	se(ß)	p	n	I^2
rs4150642	11	18327474	G	C	0.5012	0.0224	3.20E-111	4179	70.25
rs7103375	11	18332109	G	A	0.5012	0.0224	3.26E-111	4179	70.25
rs4638289	11	18242350	A	T	0.3045	0.0198	2.77E-53	4143	76.45
rs4353250	11	18280093	T	C	0.2719	0.0180	1.68E-51	4206	67.88
rs2061164	11	18285582	T	C	0.2722	0.0180	1.92E-51	4206	67.53
rs10741739	11	18284260	G	T	0.2719	0.0180	2.27E-51	4210	68.05
rs3825025	11	18300398	C	G	0.2711	0.0180	5.32E-51	4210	68.49
rs4150550	11	18307474	C	G	0.2704	0.0180	9.77E-51	4210	68.34
rs4150561	11	18311453	T	C	0.2697	0.0180	1.54E-50	4210	68.18
rs4150562	11	18311571	A	G	0.2697	0.0180	1.58E-50	4210	68.16
rs4150564	11	18311884	G	A	0.2696	0.0180	1.66E-50	4210	68.13
rs4150616	11	18321690	A	T	0.2695	0.0180	1.79E-50	4210	68.02
rs4150641	11	18327315	C	A	0.2695	0.0180	1.79E-50	4210	68.02
rs4150650	11	18330855	G	A	0.2695	0.0180	1.79E-50	4210	68.02
rs4150651	11	18331001	T	C	0.2695	0.0180	1.79E-50	4210	68.02
rs4150661	11	18336205	C	A	0.2694	0.0180	1.80E-50	4210	68.01
rs3740711	11	18336415	T	C	0.2694	0.0180	1.81E-50	4210	68.01
rs4150612	11	18320934	G	T	0.2695	0.0180	1.82E-50	4210	68.02
rs4150610	11	18320716	A	G	0.2694	0.0180	1.83E-50	4210	68.01
rs4150575	11	18313042	C	T	0.2693	0.0180	2.06E-50	4210	67.99
rs9783347	11	18341514	A	G	0.2692	0.0180	2.08E-50	4210	67.98
rs9988866	11	18349559	T	A	0.2673	0.0180	3.68E-50	4210	68.38
rs10741741	11	18298657	T	C	0.2648	0.0179	1.96E-49	4197	63.24
rs1520884	11	18266276	A	G	0.2687	0.0182	3.03E-49	4166	67.39
rs1993373	11	18237211	A	G	0.2860	0.0194	4.21E-49	4206	84.34

rs4150615	11	18321604	C	T	0.2642	0.0180	8.49E-49	4208	65.86
rs7950019	11	18235488	G	A	0.2857	0.0195	1.40E-48	4211	85.10
rs7112278	11	18238492	C	T	0.2828	0.0193	2.09E-48	4157	81.62
rs4150655	11	18331740	A	T	0.2669	0.0185	3.99E-47	4083	67.37
rs4150622	11	18322286	A	G	0.2626	0.0183	8.67E-47	4187	69.91
rs4757638	11	18269529	G	A	0.2619	0.0183	1.12E-46	4127	62.08
rs4150579	11	18313756	A	G	0.2598	0.0182	2.63E-46	4185	67.39
rs4150673	11	18340025	T	G	0.2528	0.0180	1.26E-44	4191	66.21
rs11024614	11	18283334	C	T	0.2338	0.0172	5.05E-42	4205	59.09
rs4150628	11	18322444	G	A	0.2314	0.0172	3.44E-41	4210	58.74
rs4150606	11	18320145	C	A	0.2314	0.0172	3.54E-41	4210	58.73
rs4757645	11	18350421	C	G	0.2284	0.0179	3.26E-37	4092	64.16
rs11024613	11	18283118	T	C	0.2159	0.0172	3.74E-36	4178	55.69
rs3802967	11	18300640	T	C	0.2173	0.0173	5.18E-36	4181	61.73
rs2403254	11	18281722	C	T	0.2150	0.0172	6.07E-36	4168	54.32
rs10832920	11	18298445	A	G	0.2154	0.0172	6.35E-36	4183	55.67
rs4757637	11	18264151	C	A	0.2231	0.0180	2.10E-35	4177	72.63
rs4596	11	18344704	G	C	0.2138	0.0172	2.17E-35	4182	58.32
rs2037867	11	18338597	A	G	0.2137	0.0172	2.17E-35	4182	57.92
rs2305564	11	18274034	T	A	0.2114	0.0173	3.76E-34	4152	60.63
rs4150581	11	18313846	A	G	0.2027	0.0173	7.68E-32	4160	50.29
rs10832918	11	18292260	G	A	0.2437	0.0238	1.09E-24	2722*	0.00
rs3781945	11	18262848	G	A	0.1879	0.0194	3.89E-22	3874	78.09
rs2896526	11	18376382	G	A	0.2211	0.0229	4.12E-22	4211	60.51
rs2271997	11	18259980	C	T	0.1870	0.0211	8.67E-19	3213**	84.34
rs4150563	11	18311678	T	G	0.2284	0.0261	2.15E-18	4024	84.42
rs12218	11	18247897	T	C	0.1650	0.0189	2.18E-18	4104	76.82
rs2045272	11	18234674	T	G	0.1582	0.0188	4.46E-17	4210	55.78

rs2056781	11	18401565	A	G	0.2191	0.0266	1.98E-16	4210	77.72
rs16935424	11	18413313	T	A	0.2191	0.0267	2.13E-16	4210	77.83
rs16935432	11	18427063	A	G	0.2192	0.0267	2.28E-16	4210	77.93
rs12289603	11	18427946	C	T	0.2193	0.0267	2.29E-16	4210	77.98
rs4150667	11	18339487	C	T	0.2011	0.0245	2.31E-16	4015	82.93
rs3740713	11	18407672	C	A	0.2118	0.0268	2.84E-15	4176	77.08
rs7131332	11	18252099	A	G	0.1493	0.0189	3.22E-15	4115	39.79
rs11024600	11	18252386	T	C	0.1468	0.0189	8.22E-15	4210	49.84
rs11024603	11	18262975	G	A	0.1543	0.0215	8.05E-13	4143	43.30
rs35593189	11	18424954	A	G	0.2379	0.0344	4.66E-12	2725*	33.34
rs12291480	11	18262333	G	C	0.1446	0.0231	3.57E-10	4107	34.21
rs1046615	11	18257530	G	A	0.1343	0.0226	2.95E-09	3244**	17.62
rs12416821	11	18257705	A	G	0.1341	0.0226	3.07E-09	3244**	18.03
rs12419588	11	18257706	G	A	0.1341	0.0226	3.07E-09	3244**	18.03
rs2049129	11	18261909	T	C	0.1340	0.0235	1.12E-08	4026	21.75
rs12753193	1	65942267	A	G	0.1248	0.0184	1.22E-11	4212	0.00
rs7541434	1	65941087	C	A	0.1243	0.0184	1.35E-11	4202	0.00
rs7524581	1	65939087	C	T	0.1217	0.0183	2.82E-11	4206	0.00
rs2211651	1	65928609	G	T	0.1132	0.0177	1.66E-10	4210	0.00
rs1805096	1	65874845	G	A	0.1127	0.0177	1.92E-10	4207	0.00
rs6588158	1	65933209	C	T	0.1135	0.0178	1.99E-10	4211	0.00
rs4420065	1	65934049	C	T	0.1133	0.0179	2.23E-10	4212	0.00
rs4655584	1	65928103	T	G	0.1123	0.0177	2.29E-10	4211	0.00
rs4655585	1	65928215	T	C	0.1123	0.0177	2.30E-10	4211	0.00
rs1938492	1	65890417	A	C	0.1122	0.0177	2.33E-10	4210	0.00
rs2889195	1	65929318	C	T	0.1123	0.0178	3.00E-10	4181	0.00
rs11208711	1	65920434	G	T	0.1114	0.0177	3.12E-10	4210	0.00
rs4655582	1	65925951	C	G	0.1114	0.0177	3.20E-10	4211	0.00

rs10789192	1	65898358	G	A	0.1113	0.0177	3.27E-10	4211	0.00
rs12042779	1	65906538	T	C	0.1110	0.0177	3.65E-10	4210	0.00
rs12042807	1	65906640	T	C	0.1110	0.0177	3.65E-10	4210	0.00
rs7516341	1	65860731	T	C	0.1125	0.0180	4.27E-10	4181	18.00
rs7531867	1	65880134	G	A	0.1107	0.0178	4.44E-10	4187	0.00
rs10889569	1	65858782	A	T	0.1122	0.0180	4.71E-10	4183	20.80
rs6588153	1	65864605	T	A	0.1181	0.0190	4.87E-10	4089	0.00
rs6700896	1	65862370	C	T	0.1116	0.0180	5.93E-10	4177	4.96
rs6678033	1	65850212	G	A	0.1097	0.0180	1.21E-09	4178	27.85
rs1892534	1	65878532	C	T	0.1074	0.0177	1.33E-09	4206	0.00
rs4655794	1	65931889	T	C	0.1018	0.0174	5.29E-09	4210	0.00
rs12022410	1	65926521	A	G	0.1001	0.0174	9.51E-09	4211	0.00
rs4655783	1	65915632	T	C	0.1000	0.0174	9.57E-09	4208	0.00
rs11208700	1	65906031	A	T	0.0999	0.0174	9.84E-09	4208	0.00
rs10789198	1	65975343	G	A	0.0991	0.0174	1.20E-08	4183	0.00
rs10889576	1	65941154	C	T	0.1256	0.0221	1.31E-08	4197	0.00
rs4288572	1	65973610	G	A	0.0988	0.0174	1.33E-08	4184	0.00
rs11208728	1	65985631	A	G	0.0985	0.0174	1.48E-08	4182	0.00
rs4655800	1	65984279	A	G	0.0991	0.0175	1.54E-08	4167	0.00
rs6696954	1	65827583	T	G	0.1057	0.0187	1.56E-08	4192	0.00
rs17407727	1	65944594	A	C	0.1038	0.0184	1.80E-08	4197	50.67
rs4425959	1	65940407	A	G	0.1242	0.0221	2.02E-08	4194	0.00
rs17416194	1	65943780	A	G	0.1035	0.0186	2.50E-08	4164	48.62
rs4278348	1	65947983	T	C	0.0947	0.0172	3.54E-08	4209	34.48
rs7532805	1	65955718	G	A	0.0939	0.0172	4.64E-08	4208	35.69

* only KORA S4 and LURIC, ** only KORA S4, LURIC, and TwinsUK

Table 7 Appendix. Genotypic mean levels

		Geometric means (sd) of A-SAA levels				
		KORA	**LURIC**	**Sorbs**	**TwinsUK**	**Validation sample**
rs4150642 11p15.5-p14 locus	GG	7.296 (2.367)	13.052 (3.939)	5.815 (2.114)	7.657 (0.256)	5.933 (2.395)
	GC	4.450 (2.207)	9.573 (3.891)	4.279 (2.036)	7.787 (2.790)	4.192 (2.294)
	CC	2.769 (2.080)	5.151 (3.455)	2.488 (2.061)	5.111(0.471)	2.518 (2.255)
rs4638289 *SAA1* subregion	AA	4.730 (2.383)	7.594 (3.676)	4.398 (2.182)	7.306 (0.618)	4.469 (2.333)
	AT	3.739 (2.188)	7.599 (3.843)	3.507 (2.073)	6.196 (0.599)	3.371 (2.349)
	TT	2.688 (2.107)	5.430 (3.623)	2.393 (2.112)	5.018 (2.240)	2.511 (2.303)
rs4353250 *HPS5/ GTF2H1* subregion	TT	5.091 (2.492)	9.969 (3.837)	4.532 (2.361)	7.511 (2.740)	4.805 (2.674)
	TC	3.534 (2.190)	7.843 (3.762)	3.259 (2.102)	5.989 (2.447)	3.456 (2.254)
	CC	2.845 (2.103)	4.914 (3.517)	2.385 (1.981)	5.178 (0.477)	2.463 (2.264)
rs2896526 *LDHA/LDHC subregion*	GG	4.788 (2.051)	12.469 (4.545)	4.455 (2.383)	6.660 (2.581)	4.227 (2.566)
	GA	4.005 (2.400)	7.793 (3.835)	3.642 (2.142)	6.123 (2.474)	3.607 (2.441)
	AA	3.050 (2.126)	5.974 (3.656)	2.867 (2.146)	5.393 (0.488)	2.861 (2.327)
rs12753193 1p31 locus (*LEPR*)	AA	3.673 (2.209)	7.120 (3.710)	3.484 (2.269)	6.221 (2.494)	3.410 (2.468)
	AG	3.241 (2.234)	6.570 (3.758)	2.981 (2.112)	5.398 (0.508)	2.983 (2.342)
	GG	2.961 (2.223)	5.725 (3.888)	2.614 (2.023)	4.742 (0.364)	2.789 (2.240)

rs549485 11p14 locus *(SERGEF)* (males only)	TT	3.779 (2.595)	8.583 (4.342)	3.320 (2.925)	NA	3.148 (2.419)*
	TC	3.132 (2.202)	6.119 (3.810)	3.072 (2.103)	NA	2.723 (2.217)*
	CC	2.728 (2.202)	6.093 (4.117)	2.239 (2.200)	NA	2.611 (2.455)*

***** For technical reasons rs549485 was replaced by rs493767 (r^2=0.961, 3rd lowest p-value within this region in the gender stratified meta-analysis) in the validation analysis

Table 8 Appendix. Comparison between different genotyping technologies in the KORA study

Region	SNP	Genotyping technology	Effect allele	Other allele	imp	n	beta	se(beta)	p
11p15.5-p14 locus	rs4150642	Affymetrix 6.0	G	C	I	1785	0.529	0.032	1.92E-58
		Sequenom	G	C	G	1494	0.470	0.032	1.03E-44
	rs7103375	Affymetrix 6.0	G	A	I	1785	0.529	0.032	1.93E-58
		Sequenom	G	A	G	1538	0.473	0.032	1.27E-46
		Illumina 550k	G	A	G	707	0.522	0.050	3.61E-24
SAA1 subregion	rs4638289	Affymetrix 6.0	A	T	I	1785	0.323	0.027	4.95E-31
		Sequenom	A	T	G	1516	0.344	0.028	1.46E-32
	rs1993373	Affymetrix 6.0	A	G	I	1785	0.313	0.027	2.07E-29
		Sequenom	A	G	G	1397	0.348	0.029	1.06E-30
	rs2045272	Affymetrix 6.0	T	G	I	1785	0.199	0.028	2.33E-12
		Sequenom	T	G	G	1526	0.202	0.028	2.12E-12
		Illumina 550k	T	G	G	712	0.115	0.042	5.79E-03
HPS5/ GTF2H1 subregion	rs4353250	Affymetrix 6.0	T	C	I	1785	0.278	0.027	1.06E-24
		Sequenom	T	C	G	1534	0.274	0.029	3.23E-21
	rs9988866	Affymetrix	T	A	G	1785	0.270	0.027	9.24E-

		6.0							24
		Sequenom	T	A	G	1534	0.275	0.028	1.26E-21
	rs1520884	Affymetrix 6.0	A	G	G	1780	0.271	0.027	1.00E-23
		Sequenom	T	C	G	1459	0.250	0.028	3.24E-19
	rs4150655	Affymetrix 6.0	A	T	G	1659	0.272	0.028	4.51E-21
		Sequenom	A	T	G	1530	0.266	0.029	5.18E-20
	rs4757637	Affymetrix 6.0	C	A	G	1751	0.213	0.026	4.56E-16
		Sequenom	C	A	G	1526	0.232	0.028	6.93E-17
		Illumina 550k	C	A	G	705	0.176	0.041	2.19E-05
	rs4150581	Affymetrix 6.0	A	G	G	1768	0.203	0.026	8.18E-15
		Sequenom	A	G	G	1499	0.210	0.028	3.93E-14
		Illumina 550k	A	G	G	714	0.191	0.040	1.79E-06
	rs11024603	Affymetrix 6.0	G	A	G	1782	0.188	0.032	5.76E-09
		Sequenom	G	A	G	1540	0.181	0.035	1.90E-07
		Illumina 550k	G	A	G	715	0.204	0.051	6.26E-05
LDHA/LDHC subregion	rs2896526	Affymetrix 6.0	G	A	I	1785	0.265	0.033	3.99E-15
		Sequenom	G	A	G	1511	0.252	0.035	1.46E-12
		Illumina	G	A	G	716	0.304	0.053	1.15E-

		550k							08
	rs12289603	Affymetrix 6.0	A	G	I	1785	0.269	0.039	5.81E-12
		Sequenom	T	C	G	1390	0.212	0.038	1.95E-08
	rs12289603	Affymetrix 6.0	C	T	I	1785	0.269	0.039	6.15E-12
		Sequenom	C	T	G	1535	0.260	0.040	8.56E-11
	rs3740713	Affymetrix 6.0	C	A	G	1749	0.257	0.039	8.35E-11
		Sequenom	G	T	G	1532	0.237	0.041	6.32E-09
		Illumina 550k	G	T	G	702	0.274	0.061	7.54E-06
	rs35593189	Affymetrix 6.0	A	G	G	1764	0.255	0.037	1.07E-11
		Sequenom	A	G	G	1526	0.255	0.040	2.66E-10
1p31 locus (*LEPR*)	rs12753193	Affymetrix 6.0	A	G	I	1785	0.103	0.027	1.94E-04
		Sequenom	A	G	G	1536	0.062	0.029	3.10E-02
		Illumina 550k	A	G	G	715	0.117	0.043	6.02E-03
	rs7524581	Affymetrix 6.0	C	T	I	1785	0.101	0.027	2.45E-04
		Sequenom	C	T	G	1534	0.053	0.029	6.67E-02
		Illumina 550k	C	T	G	715	0.107	0.043	1.20E-02
	rs2211651	Affymetrix 6.0	G	T	I	1785	0.093	0.026	4.17E-04
		Sequenom	G	T	G	1499	0.071	0.029	1.53E-

									02
		Illumina 550k	G	T	G	715	0.113	0.043	8.17E-03
	rs2889195	Affymetrix 6.0	C	T	G	1754	0.092	0.027	5.90E-04
		Sequenom	C	T	G	1519	0.066	0.029	2.16E-02
		Illumina 550k	C	T	G	716	0.113	0.043	7.76E-03
	rs1892534	Affymetrix 6.0	C	T	G	1779	0.092	0.026	5.27E-04
		Sequenom	G	A	G	1511	0.069	0.028	1.42E-02
		Illumina 550k	G	A	G	716	0.104	0.043	1.45E-02
	rs12022410	Affymetrix 6.0	A	G	I	1785	0.093	0.027	5.66E-04
		Sequenom	A	G	G	1498	0.026	0.029	3.79E-01
		Illumina 550k	A	G	G	715	0.097	0.041	1.84E-02
	rs17407727	Affymetrix 6.0	A	C	G	1773	0.106	0.028	1.45E-04
		Sequenom	A	C	G	1509	0.075	0.030	1.29E-02
		Illumina 550k	A	C	G	716	0.129	0.045	3.75E-03
	rs17416194	Affymetrix 6.0	A	G	G	1744	0.105	0.028	2.04E-04
		Sequenom	A	G	G	1491	0.076	0.030	1.17E-02
		Illumina 550k	A	G	G	716	0.124	0.045	5.60E-03
11p14 locus	rs493767	Affymetrix	C	G	G	860	0.140	0.040	5.66E-

(SERGEF) (males only)		6.0							04
		Sequenom	C	G	G	749	0.142	0.041	5.75E-04
	rs550659	Affymetrix 6.0	G	A	G	873	0.131	0.040	1.16E-03
		Sequenom	G	A	G	742	0.143	0.041	5.97E-04

Table 9 Appendix. Analysis of the structure of the chromosome 11 region

		Single SNP model		Multiple SNP model	
rs4150642	**beta**	0.501			
11p15.5-p13	**se**	0.022			
	p	3.20E-111	$\Delta R^2 = 0.1084$*		
rs4638289	**beta**	0.305		0.254*	
SAA1 subregion	**se**	0.020		0.024*	
	p	2.77E-53	$\Delta R^2 = 0.0557$*	7.61E-27*	
rs4353250	**beta**	0.272		0.212*	
HPS5/GTF2H1 subregion	**se**	0.018		0.022*	
	p	1.68E-51	$\Delta R^2 = 0.0534$*	3.64E-22*	
rs2896526	**beta**	0.221		0.106*	
LDHA/LDHC subregion	**se**	0.023		0.025*	
	p	4.12E-22	$\Delta R^2 = 0.0237$*	1.81E-05*	$\Delta R^2 = 0.0910$*

* Values calculated for KORA S4, LURIC, and the Sorbs.

Table 10 Appendix. Results of the validation analyses

Region	SNP	Effect allele	Other allele	Effect allele freq	n	beta	se(beta)	p
11p15.5-p14	rs4150642	G	C	0.217	2082	0.473	0.029	7.41E-58
locus	rs7103375	G	A	0.215	2130	0.460	0.028	2.27E-56
SAA1 subregion	rs4638289	A	T	0.332	2091	0.292	0.025	1.61E-30
	rs1993373	A	G	0.327	1939	0.297	0.026	2.04E-29
	rs2045272	T	G	0.581	2118	0.207	0.024	4.09E-17
HPS5/ GTF2H1 subregion	rs4353250	T	C	0.352	2125	0.345	0.025	5.82E-42
	rs9988866	T	A	0.359	2125	0.355	0.025	6.62E-45
	rs1520884	T	C	0.356	2034	0.318	0.024	9.14E-38
	rs4150655	A	T	0.323	2123	0.343	0.025	1.52E-41
	rs4757637	C	A	0.406	2109	0.286	0.025	1.48E-29
	rs4150581	A	G	0.486	2072	0.250	0.024	9.31E-25
	rs11024603	G	A	0.828	2127	0.058	0.031	5.74E-02
LDHA/LDHC subregion	rs2896526	G	A	0.187	2097	0.216	0.031	8.18E-12
	rs2056781	T	C	0.170	1942	0.157	0.032	1.17E-06
	rs12289603	C	T	0.120	2131	0.219	0.034	1.69E-10
	rs3740713	G	T	0.109	2117	0.188	0.036	1.67E-07
	rs35593189	A	G	0.116	2113	0.213	0.035	1.11E-09
1p31 locus (*LEPR*)	rs12753193	A	G	0.606	2127	0.085	0.025	8.60E-04
	rs7524581	C	T	0.624	2123	0.080	0.026	1.85E-03
	rs2211651	G	T	0.591	2079	0.069	0.025	6.89E-03
	rs2889195	C	T	0.603	2116	0.075	0.025	2.94E-03
	rs1892534	G	A	0.627	2090	0.070	0.025	5.39E-03
	rs12022410	A	G	0.410	2089	0.072	0.026	5.32E-03

	rs17407727	A	C	0.288	2085	0.067	0.027	1.55E-02
	rs17416194	A	G	0.278	2063	0.055	0.027	4.20E-02
11p14 locus	rs493767	C	G	0.349	1049	0.091	0.035	8.50E-03
(*SERGEF)males*	rs550659	G	A	0.373	1041	0.084	0.035	1.65E-02

6. References

1. American Diabetes Association, *Diagnosis and classification of diabetes mellitus.* Diabetes Care, 2009. 32 Suppl 1: p. S62-7.

2. World Health Organization, *Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications. Report of a WHO Consultation. Part 1: Diagnosis and Classification of Diabetes Mellitus*, 1999: Geneva, Switzerland.

3. World Health Organization, *Diabetes*, 2012, World Health Organization: Geneve, Switzerland.

4. World Health Organization, *Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycaemia. Report of a WHO/ IDF consultation*, 2006, World Health Organization: Geneva, Switzerland.

5. Shaw, J.E., P.Z. Zimmet, and K.G. Alberti, *Point: impaired fasting glucose: The case for the new American Diabetes Association criterion.* Diabetes Care, 2006. 29(5): p. 1170-2.

6. American Diabetes Association, *Standards of medical care in diabetes--2012.* Diabetes Care, 2012. 35 Suppl 1: p. S11-63.

7. Buysschaert, M. and M. Bergman, *Definition of prediabetes.* Med Clin North Am, 2011. 95(2): p. 289-97, vii.

8. World Health Organization, *Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus. Abbreviated Report of a WHO Consultation*, 2011, World Health Organization: Geneva, Switzerland.

9. Sorkin, J.D., et al., *The relation of fasting and 2-h postchallenge plasma glucose concentrations to mortality: data from the Baltimore Longitudinal Study of Aging with a critical review of the literature.* Diabetes Care, 2005. 28(11): p. 2626-32.

10. Rathmann, W., et al., *Hemoglobin A1c and glucose criteria identify different subjects as having type 2 diabetes in middle-aged and older populations: the KORA S4/F4 Study.* Ann Med, 2012. 44(2): p. 170-7.

11. Jorgensen, M.E., et al., *New diagnostic criteria for diabetes: is the change from glucose to HbA1c possible in all populations?* J Clin Endocrinol Metab, 2010. 95(11): p. E333-6.

12. Shaw, J.E., R.A. Sicree, and P.Z. Zimmet, *Global estimates of the prevalence of diabetes for 2010 and 2030.* Diabetes Res Clin Pract, 2010. 87(1): p. 4-14.

13. International Diabetes Federation, *IDF Diabetes Atlas*. 2011, Brussels: International Diabetes Federation.

14. Danaei, G., et al., *National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants.* Lancet, 2011. 378(9785): p. 31-40.

15. World Health Organization, *Prevention of diabetes mellitus. Report of a WHO Study Group.* Vol. No. 844. 1994, Geneva, Switzerland: World Health Organization.

16. Schipf, S., et al., *Regional differences in the prevalence of known Type 2 diabetes mellitus in 45-74 years old individuals: results from six population-based studies in Germany (DIAB-CORE Consortium).* Diabet Med, 2012. 29(7): p. e88-95.

17. Rathmann, W., et al., *High prevalence of undiagnosed diabetes mellitus in Southern Germany: target populations for efficient screening. The KORA survey 2000.* Diabetologia, 2003. 46(2): p. 182-9.

18. Meisinger, C., et al., *Prevalence of undiagnosed diabetes and impaired glucose regulation in 35-59-year-old individuals in Southern Germany: the KORA F4 Study.* Diabet Med, 2010. 27(3): p. 360-2.

19. World Health Organization, *Burden: mortality, morbidity and risk factors* in *Noncommunicable diseases country profiles 2011. WHO global report.* , W.H. Organization, Editor. 2011: Geneva, Switzerland.

20. Herder, C., *Marker der subklinischen Inflammation als Risikofaktoren des Typ 2 Diabetes: Ergebnisse aus epidemiologischen Kohortenstudien*, 2012, Heinrich-Heine-Universität Düsseldorf: Düsseldorf.

21. Meigs, J.B., et al., *The natural history of progression from normal glucose tolerance to type 2 diabetes in the Baltimore Longitudinal Study of Aging.* Diabetes, 2003. 52(6): p. 1475-84.

22. Wild, S., et al., *Global prevalence of diabetes: estimates for the year 2000 and projections for 2030.* Diabetes Care, 2004. 27(5): p. 1047-53.

23. Tabak, A.G., et al., *Trajectories of glycaemia, insulin sensitivity, and insulin secretion before diagnosis of type 2 diabetes: an analysis from the Whitehall II study.* Lancet, 2009. 373(9682): p. 2215-21.

24. Basu, R., et al., *Mechanisms of the age-associated deterioration in glucose tolerance: contribution of alterations in insulin secretion, action, and clearance.* Diabetes, 2003. 52(7): p. 1738-48.

25. Utzschneider, K.M., et al., *Impact of intra-abdominal fat and age on insulin sensitivity and beta-cell function.* Diabetes, 2004. 53(11): p. 2867-72.

26. Chang, A.M. and J.B. Halter, *Aging and insulin secretion.* Am J Physiol Endocrinol Metab, 2003. 284(1): p. E7-12.

27. McNeely, M.J. and E.J. Boyko, *Type 2 diabetes prevalence in Asian Americans: results of a national health survey.* Diabetes Care, 2004. 27(1): p. 66-9.

28. Chandler-Laney, P.C., et al., *Age-related changes in insulin sensitivity and beta-cell function among European-American and African-American women.* Obesity (Silver Spring), 2011. 19(3): p. 528-35.

29. World Health Organization, *Genetics and Diabetes*, World Health Organization: Geneva, Switzerland.

30. Wheeler, E. and I. Barroso, *Genome-wide association studies and type 2 diabetes.* Brief Funct Genomics, 2011. 10(2): p. 52-60.

31. McCarthy, M.I., *Genomics, type 2 diabetes, and obesity.* N Engl J Med, 2010. 363(24): p. 2339-50.

32. Morris, A.P., et al., *Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes.* Nat Genet, 2012. 44(9): p. 981-90.

33. Grant, S.F., et al., *Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes.* Nat Genet, 2006. 38(3): p. 320-3.

34. Marzi, C., et al., *Variants of the transcription factor 7-like 2 gene (TCF7L2) are strongly associated with type 2 diabetes but not with the metabolic syndrome in the MONICA/KORA surveys.* Horm Metab Res, 2007. 39(1): p. 46-52.

35. Humphries, S.E., et al., *Common variants in the TCF7L2 gene and predisposition to type 2 diabetes in UK European Whites, Indian Asians and Afro-Caribbean men and women.* J Mol Med (Berl), 2006. 84(12): p. 1005-14.

36. Damcott, C.M., et al., *Polymorphisms in the transcription factor 7-like 2 (TCF7L2) gene are associated with type 2 diabetes in the Amish: replication and evidence for a role in both insulin secretion and insulin resistance.* Diabetes, 2006. 55(9): p. 2654-9.

37. Prokopenko, I., M.I. McCarthy, and C.M. Lindgren, *Type 2 diabetes: new genes, new understanding.* Trends Genet, 2008. 24(12): p. 613-21.

38. McCarthy, M.I., *Growing evidence for diabetes susceptibility genes from genome scan data.* Curr Diab Rep, 2003. 3(2): p. 159-67.

39. Voight, B.F., et al., *Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis.* Nat Genet, 2010. 42(7): p. 579-89.

40. Prokopenko, I., et al., *Variants in MTNR1B influence fasting glucose levels.* Nat Genet, 2009. 41(1): p. 77-81.

41. van Dam, R.M., *The epidemiology of lifestyle and risk for type 2 diabetes.* Eur J Epidemiol, 2003. 18(12): p. 1115-25.

42. Shaw, J.E. and D.J. Chisholm, *1: Epidemiology and prevention of type 2 diabetes and the metabolic syndrome.* Med J Aust, 2003. 179(7): p. 379-83.

43. Trayhurn, P., *Endocrine and signalling role of adipose tissue: new perspectives on fat.* Acta Physiol Scand, 2005. 184(4): p. 285-93.

44. Trayhurn, P. and I.S. Wood, *Adipokines: inflammation and the pleiotropic role of white adipose tissue.* Br J Nutr, 2004. 92(3): p. 347-55.

45. Sjoholm, K., et al., *A microarray search for genes predominantly expressed in human omental adipocytes: adipose tissue as a major production site of serum amyloid A.* J Clin Endocrinol Metab, 2005. 90(4): p. 2233-9.

46. Yang, R.Z., et al., *Acute-phase serum amyloid A: an inflammatory adipokine and potential link between obesity and its metabolic complications.* PLoS Med, 2006. 3(6): p. e287.

47. Lopatynski, J., G. Mardarowicz, and G. Szczesniak, *A comparative evaluation of waist circumference, waist-to-hip ratio, waist-to-height ratio and body mass index as indicators of impaired glucose tolerance and as risk factors for type-2 diabetes mellitus.* Ann Univ Mariae Curie Sklodowska Med, 2003. 58(1): p. 413-9.

48. Qiao, Q. and R. Nyamdorj, *Is the association of type II diabetes with waist circumference or waist-to-hip ratio stronger than that with body mass index?* Eur J Clin Nutr, 2010. 64(1): p. 30-4.

49. Hadaegh, F., et al., *Waist/height ratio as a better predictor of type 2 diabetes compared to body mass index in Tehranian adult men--a 3.6-year prospective study.* Exp Clin Endocrinol Diabetes, 2006. 114(6): p. 310-5.

50. Steinbrecher, A., et al., *The preventable proportion of type 2 diabetes by ethnicity: the multiethnic cohort.* Ann Epidemiol, 2011. 21(7): p. 526-35.

51. Lee, I.M., et al., *Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy.* Lancet, 2012. 380(9838): p. 219-29.

52. Hu, F.B., et al., *Diet, lifestyle, and the risk of type 2 diabetes mellitus in women.* N Engl J Med, 2001. 345(11): p. 790-7.

53. Goodyear, L.J., et al., *Glucose transporter number, function, and subcellular distribution in rat skeletal muscle after exercise training.* Diabetes, 1992. 41(9): p. 1091-9.

54. Devlin, J.T., *Effects of exercise on insulin sensitivity in humans.* Diabetes Care, 1992. 15(11): p. 1690-3.

55. Willi, C., et al., *Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis.* JAMA, 2007. 298(22): p. 2654-64.

56. Houston, T.K., et al., *Active and passive smoking and development of glucose intolerance among young adults in a prospective cohort: CARDIA study.* BMJ, 2006. 332(7549): p. 1064-9.

57. Pietraszek, A., S. Gregersen, and K. Hermansen, *Alcohol and type 2 diabetes. A review.* Nutr Metab Cardiovasc Dis, 2010. 20(5): p. 366-75.

58. Frayling, T.M. and A.T. Hattersley, *The role of genetic susceptibility in the association of low birth weight with type 2 diabetes.* Br Med Bull, 2001. 60: p. 89-101.

59. Sobngwi, E., et al., *Effect of a diabetic environment in utero on predisposition to type 2 diabetes.* Lancet, 2003. 361(9372): p. 1861-5.

60. Lou, P., et al., *Relation of sleep quality and sleep duration to type 2 diabetes: a population-based cross-sectional survey.* BMJ Open, 2012. 2(4).

61. Agardh, E.E., et al., *Work stress and low sense of coherence is associated with type 2 diabetes in middle-aged Swedish women.* Diabetes Care, 2003. 26(3): p. 719-24.

62. Mezuk, B., et al., *Depression and type 2 diabetes over the lifespan: a meta-analysis.* Diabetes Care, 2008. 31(12): p. 2383-90.

63. Heraclides, A., et al., *Psychosocial stress at work doubles the risk of type 2 diabetes in middle-aged women: evidence from the Whitehall II study.* Diabetes Care, 2009. 32(12): p. 2230-5.

64. Sacerdote, C., et al., *Lower educational level is a predictor of incident type 2 diabetes in European countries: The EPIC-InterAct study.* Int J Epidemiol, 2012. 41(4): p. 1162-1173.

65. Agardh, E., et al., *Type 2 diabetes incidence and socio-economic position: a systematic review and meta-analysis.* Int J Epidemiol, 2011. 40(3): p. 804-18.

66. Agardh, E.E., et al., *Burden of type 2 diabetes attributed to lower educational levels in Sweden.* Popul Health Metr, 2011. 9: p. 60.

67. Maty, S.C., et al., *Education, income, occupation, and the 34-year incidence (1965-99) of Type 2 diabetes in the Alameda County Study.* Int J Epidemiol, 2005. 34(6): p. 1274-81.

68. Thorand, B., et al., *Effect of serum 25-hydroxyvitamin D on risk for type 2 diabetes may be partially mediated by subclinical inflammation: results from the MONICA/KORA Augsburg study.* Diabetes Care, 2011. 34(10): p. 2320-2.

69. Forouhi, N.G., et al., *Baseline serum 25-hydroxy vitamin d is predictive of future glycemic status and insulin resistance: the Medical Research Council Ely Prospective Study 1990-2000.* Diabetes, 2008. 57(10): p. 2619-25.

70. Psaltopoulou, T., I. Ilias, and M. Alevizaki, *The role of diet and lifestyle in primary, secondary, and tertiary diabetes prevention: a review of meta-analyses.* Rev Diabet Stud, 2010. 7(1): p. 26-35.

71. Astrup, A., *Healthy lifestyles in Europe: prevention of obesity and type II diabetes by diet and physical activity.* Public Health Nutr, 2001. 4(2B): p. 499-515.

72. Schulze, M.B. and F.B. Hu, *Primary prevention of diabetes: what can be done and how much can be prevented?* Annu Rev Public Health, 2005. 26: p. 445-67.

73. Pan, X.R., et al., *Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and Diabetes Study.* Diabetes Care, 1997. 20(4): p. 537-44.

74. Tuomilehto, J., et al., *Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance.* N Engl J Med, 2001. 344(18): p. 1343-50.

75. Stumvoll, M., B.J. Goldstein, and T.W. van Haeften, *Type 2 diabetes: principles of pathogenesis and therapy.* Lancet, 2005. 365(9467): p. 1333-46.

76. Myers, M.G., Jr. and M.F. White, *Insulin signal transduction and the IRS proteins.* Annu Rev Pharmacol Toxicol, 1996. 36: p. 615-58.

77. White, M.F., *The insulin signalling system and the IRS proteins.* Diabetologia, 1997. 40 Suppl 2: p. S2-17.

78. Stumvoll, M., et al., *Glucose allostasis.* Diabetes, 2003. 52(4): p. 903-9.

79. Kahn, B.B., *Type 2 diabetes: when insulin secretion fails to compensate for insulin resistance.* Cell, 1998. 92(5): p. 593-6.

80. Donath, M.Y. and S.E. Shoelson, *Type 2 diabetes as an inflammatory disease.* Nat Rev Immunol, 2011. 11(2): p. 98-107.

81. Maedler, K., *Beta cells in type 2 diabetes - a crucial contribution to pathogenesis.* Diabetes Obes Metab, 2008. 10(5): p. 408-20.

82. Zhang, K., *Integration of ER stress, oxidative stress and the inflammatory response in health and disease.* Int J Clin Exp Med, 2010. 3(1): p. 33-40.

83. Nishikawa, T., et al., *Normalizing mitochondrial superoxide production blocks three pathways of hyperglycaemic damage.* Nature, 2000. 404(6779): p. 787-90.

84. Lin, Y., et al., *The hyperglycemia-induced inflammatory response in adipocytes: the role of reactive oxygen species.* J Biol Chem, 2005. 280(6): p. 4617-26.

85. Wellen, K.E. and G.S. Hotamisligil, *Inflammation, stress, and diabetes.* J Clin Invest, 2005. 115(5): p. 1111-9.

86. Rabol, R., R. Boushel, and F. Dela, *Mitochondrial oxidative function and type 2 diabetes.* Appl Physiol Nutr Metab, 2006. 31(6): p. 675-83.

87. Szendroedi, J., E. Phielix, and M. Roden, *The role of mitochondria in insulin resistance and type 2 diabetes mellitus.* Nat Rev Endocrinol, 2012. 8(2): p. 92-103.

88. Parish, R. and K.F. Petersen, *Mitochondrial dysfunction and type 2 diabetes.* Curr Diab Rep, 2005. 5(3): p. 177-83.

89. Hummasti, S. and G.S. Hotamisligil, *Endoplasmic reticulum stress and inflammation in obesity and diabetes.* Circ Res, 2010. 107(5): p. 579-91.

90. Donath, M.Y., et al., *Islet inflammation impairs the pancreatic beta-cell in type 2 diabetes.* Physiology (Bethesda), 2009. 24: p. 325-31.

91. Hotamisligil, G.S., *Inflammation and metabolic disorders.* Nature, 2006. 444(7121): p. 860-7.

92. Pickup, J.C. and M.A. Crook, *Is type II diabetes mellitus a disease of the innate immune system?* Diabetologia, 1998. 41(10): p. 1241-8.

93. Cone, J.B., *Inflammation.* Am J Surg, 2001. 182(6): p. 558-62.

94. Pickup, J.C., *Inflammation and activated innate immunity in the pathogenesis of type 2 diabetes.* Diabetes Care, 2004. 27(3): p. 813-23.

95. Kolb, H. and T. Mandrup-Poulsen, *An immune origin of type 2 diabetes?* Diabetologia, 2005. 48(6): p. 1038-50.

96. Herder, C., et al., *Inflammation and type 2 diabetes: results from KORA Augsburg.* Gesundheitswesen, 2005. 67 Suppl 1: p. S115-21.

97. Thorand, B., et al., *Sex differences in the prediction of type 2 diabetes by inflammatory markers: results from the MONICA/KORA Augsburg case-cohort study, 1984-2002.* Diabetes Care, 2007. 30(4): p. 854-60.

98. Lee, C.C., et al., *Association of C-reactive protein with type 2 diabetes: prospective analysis and meta-analysis.* Diabetologia, 2009. 52(6): p. 1040-7.

99. Costa, A., et al., *Lower rate of tumor necrosis factor-alpha -863A allele and higher concentration of tumor necrosis factor-alpha receptor 2 in first-degree relatives of subjects with type 2 diabetes.* Metabolism, 2003. 52(8): p. 1068-71.

100. Vozarova, B., et al., *The interleukin-6 (-174) G/C promoter polymorphism is associated with type-2 diabetes mellitus in Native Americans and Caucasians.* Hum Genet, 2003. 112(4): p. 409-13.

101. Furuta, M., et al., *Relationship of the tumor necrosis factor-alpha -308 A/G promoter polymorphism with insulin sensitivity and abdominal fat distribution in Japanese patients with type 2 diabetes mellitus.* Diabetes Res Clin Pract, 2002. 56(2): p. 141-5.

102. Dalziel, B., et al., *Association of the TNF-alpha -308 G/A promoter polymorphism with insulin resistance in obesity.* Obes Res, 2002. 10(5): p. 401-7.

103. Gouda, H.N., et al., *The association between the peroxisome proliferator-activated receptor-gamma2 (PPARG2) Pro12Ala gene variant and type 2 diabetes mellitus: a HuGE review and meta-analysis.* Am J Epidemiol, 2010. 171(6): p. 645-55.

104. Stumvoll, M. and H. Haring, *The peroxisome proliferator-activated receptor-gamma2 Pro12Ala polymorphism.* Diabetes, 2002. 51(8): p. 2341-7.

105. Uysal, K.T., et al., *Protection from obesity-induced insulin resistance in mice lacking TNF-alpha function.* Nature, 1997. 389(6651): p. 610-4.

106. Ma, L.J., et al., *Prevention of obesity and insulin resistance in mice lacking plasminogen activator inhibitor 1.* Diabetes, 2004. 53(2): p. 336-46.

107. Bruun, J.M., et al., *Association between measures of insulin sensitivity and circulating levels of interleukin-8, interleukin-6 and tumor necrosis factor-alpha. Effect of weight loss in obese men.* Eur J Endocrinol, 2003. 148(5): p. 535-42.

108. Ebeling, P., et al., *Troglitazone reduces hyperglycaemia and selectively acute-phase serum proteins in patients with Type II diabetes.* Diabetologia, 1999. 42(12): p. 1433-8.

109. Chen, C., *Troglitazone: an antidiabetic agent.* Am J Health Syst Pharm, 1998. 55(9): p. 905-25.

110. Aljada, A., et al., *Nuclear factor-kappaB suppressive and inhibitor-kappaB stimulatory effects of troglitazone in obese patients with type 2 diabetes: evidence of an antiinflammatory action?* J Clin Endocrinol Metab, 2001. 86(7): p. 3250-6.

111. Aljada, A., et al., *Troglitazone reduces the expression of PPARgamma while stimulating that of PPARalpha in mononuclear cells in obese subjects.* J Clin Endocrinol Metab, 2001. 86(7): p. 3130-3.

112. Schmidt, M.I., *Subclinical inflammation and obesity, diabetes and related disorders.* Drug Discovery Tody: Diseaes Mechanisms, 2005. 2(3): p. 307-3012.

113. Hotamisligil, G.S., et al., *IRS-1-mediated inhibition of insulin receptor tyrosine kinase activity in TNF-alpha- and obesity-induced insulin resistance.* Science, 1996. 271(5249): p. 665-8.

114. Lebrun, P. and E. Van Obberghen, *SOCS proteins causing trouble in insulin action.* Acta Physiol (Oxf), 2008. 192(1): p. 29-36.

115. Maedler, K., et al., *Glucose-induced beta cell production of IL-1beta contributes to glucotoxicity in human pancreatic islets.* J Clin Invest, 2002. 110(6): p. 851-60.

116. Uhlar, C.M. and A.S. Whitehead, *Serum amyloid A, the major vertebrate acute-phase reactant.* Eur J Biochem, 1999. 265(2): p. 501-23.

117. Jensen, L.E. and A.S. Whitehead, *Regulation of serum amyloid A protein expression during the acute-phase response.* Biochem J, 1998. 334 (Pt 3): p. 489-503.

118. Thorn, C.F., Z.Y. Lu, and A.S. Whitehead, *Regulation of the human acute phase serum amyloid A genes by tumour necrosis factor-alpha, interleukin-6 and glucocorticoids in hepatic and epithelial cell lines.* Scand J Immunol, 2004. 59(2): p. 152-8.

119. Urieli-Shoval, S., R.P. Linke, and Y. Matzner, *Expression and function of serum amyloid A, a major acute-phase protein, in normal and disease states.* Curr Opin Hematol, 2000. 7(1): p. 64-9.

120. Zhao, Y., et al., *Association between serum amyloid A and obesity: a meta-analysis and systematic review.* Inflamm Res, 2010. 59(5): p. 323-34.

121. Betts, J.C., et al., *The human acute-phase serum amyloid A gene family: structure, evolution and expression in hepatoma cells.* Scand J Immunol, 1991. 34(4): p. 471-82.

122. Meek, R.L., S. Urieli-Shoval, and E.P. Benditt, *Expression of apolipoprotein serum amyloid A mRNA in human atherosclerotic lesions and cultured vascular cells: implications for serum amyloid A function.* Proc Natl Acad Sci U S A, 1994. 91(8): p. 3186-90.

123. Kumon, Y., et al., *Dexamethasone, but not IL-1 alone, upregulates acute-phase serum amyloid A gene expression and production by cultured human aortic smooth muscle cells.* Scand J Immunol, 2001. 53(1): p. 7-12.

124. Liang, J.S., et al., *Evidence for local production of acute phase response apolipoprotein serum amyloid A in Alzheimer's disease brain.* Neurosci Lett, 1997. 225(2): p. 73-6.

125. Sellar, G.C., et al., *The human serum amyloid A protein (SAA) superfamily gene cluster: mapping to chromosome 11p15.1 by physical and genetic linkage analysis.* Genomics, 1994. 19(2): p. 221-7.

126. Sellar, G.C., et al., *Organization of the region encompassing the human serum amyloid A (SAA) gene family on chromosome 11p15.1.* Genomics, 1994. 23(2): p. 492-5.

127. Sipe, J., *Revised nomenclature for serum amyloid A (SAA). Nomenclature Committee of the International Society of Amyloidosis. Part 2.* Amyloid, 1999. 6(1): p. 67-70.

128. Kluve-Beckerman, B. and M. Song, *Genes encoding human serum amyloid A proteins SAA1 and SAA2 are located 18 kb apart in opposite transcriptional orientations.* Gene, 1995. 159(2): p. 289-90.

129. Steel, D.M., et al., *A constitutively expressed serum amyloid A protein gene (SAA4) is closely linked to, and shares structural similarities with, an acute-phase serum amyloid A protein gene (SAA2).* Genomics, 1993. 16(2): p. 447-54.

130. Webb, C.F., P.W. Tucker, and S.B. Dowton, *Expression and sequence analyses of serum amyloid A in the Syrian hamster.* Biochemistry, 1989. 28(11): p. 4785-90.

131. Meek, R.L. and E.P. Benditt, *Rat tissues express serum amyloid A protein-related mRNAs.* Proc Natl Acad Sci U S A, 1989. 86(6): p. 1890-4.

132. MacGregor, A.J., et al., *Genetic effects on baseline values of C-reactive protein and serum amyloid a protein: a comparison of monozygotic and dizygotic twins.* Clin Chem, 2004. 50(1): p. 130-4.

133. Yamada, T., Y. Okuda, and Y. Itoh, *The frequency of serum amyloid A2 alleles in the Japanese population.* Amyloid, 1998. 5(3): p. 208-11.

134. Yamada, T., et al., *Serum amyloid A1 alleles and plasma concentrations of serum amyloid A.* Amyloid, 1999. 6(3): p. 199-204.

135. Booth, D.R., et al., *SAA1 alleles as risk factors in reactive systemic AA amyloidosis.* Amyloid, 1998. 5(4): p. 262-5.

136. Bakkaloglu, A., et al., *Influence of Serum Amyloid A (SAA1) and SAA2 gene polymorphisms on renal amyloidosis, and on SAA/C-reactive protein values in patients with familial mediterranean fever in the Turkish population.* J Rheumatol, 2004. 31(6): p. 1139-42.

137. Schillinger, M., et al., *Inflammation and Carotid Artery--Risk for Atherosclerosis Study (ICARAS).* Circulation, 2005. 111(17): p. 2203-9.

138. Johnson, B.D., et al., *Serum amyloid A as a predictor of coronary artery disease and cardiovascular outcome in women: the National Heart, Lung, and Blood Institute-Sponsored Women's Ischemia Syndrome Evaluation (WISE).* Circulation, 2004. 109(6): p. 726-32.

139. Maier, W., et al., *Inflammatory markers at the site of ruptured plaque in acute myocardial infarction: locally increased interleukin-6 and serum amyloid A but decreased C-reactive protein.* Circulation, 2005. 111(11): p. 1355-61.

140. Fyfe, A.I., et al., *Association between serum amyloid A proteins and coronary artery disease: evidence from two distinct arteriosclerotic processes.* Circulation, 1997. 96(9): p. 2914-9.

141. Danesh, J., et al., *Low grade inflammation and coronary heart disease: prospective study and updated meta-analyses.* BMJ, 2000. 321(7255): p. 199-204.

142. Ridker, P.M., et al., *Inflammation, pravastatin, and the risk of coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events (CARE) Investigators.* Circulation, 1998. 98(9): p. 839-44.

143. Kinlay, S., et al., *Inflammation, statin therapy, and risk of stroke after an acute coronary syndrome in the MIRACL study.* Arterioscler Thromb Vasc Biol, 2008. 28(1): p. 142-7.

144. Chait, A., et al., *Thematic review series: The immune system and atherogenesis. Lipoprotein-associated inflammatory proteins: markers or mediators of cardiovascular disease?* J Lipid Res, 2005. 46(3): p. 389-403.

145. Jousilahti, P., et al., *The association of c-reactive protein, serum amyloid a and fibrinogen with prevalent coronary heart disease--baseline findings of the PAIS project.* Atherosclerosis, 2001. 156(2): p. 451-6.

146. Malle, E., S. Sodin-Semrl, and A. Kovacevic, *Serum amyloid A: an acute-phase protein involved in tumour pathogenesis.* Cell Mol Life Sci, 2009. 66(1): p. 9-26.

147. Pickup, J.C., et al., *NIDDM as a disease of the innate immune system: association of acute-phase reactants and interleukin-6 with metabolic syndrome X.* Diabetologia, 1997. 40(11): p. 1286-92.

148. Müller, S., et al., *Impaired glucose tolerance is associated with increased serum concentrations of interleukin 6 and co-regulated acute-phase proteins but not TNF-alpha or its receptors.* Diabetologia, 2002. 45(6): p. 805-12.

149. Sjoholm, K., et al., *Association of serum amyloid A levels with adipocyte size and serum levels of adipokines: differences between men and women.* Cytokine, 2009. 48(3): p. 260-6.

150. Leinonen, E., et al., *Insulin resistance and adiposity correlate with acute-phase reaction and soluble cell adhesion molecules in type 2 diabetes.* Atherosclerosis, 2003. 166(2): p. 387-94.

151. Ley, S.H., et al., *Adipokines and incident type 2 diabetes in an Aboriginal Canadian [corrected] population: the Sandy Lake Health and Diabetes Project.* Diabetes Care, 2008. 31(7): p. 1410-5.

152. MacMahon B, P.T., *Epidemiology: principles and methods*. 1970, Boston: Little Brown.

153. Beaglehole R, B.R., Kjellström T, *Basic epidemiology*, ed. W.H. Organization. 1993, Geneva, Swizerland.

154. Kreienbrock, L., Schach, S., *Epidemiologische Methoden*. 1995, Stuttgart, Jena, New York G. Fischer Verlag.

155. Rothman, K., Greenland, S, *Modern Epidemiology*. 1998, Philadelphia Lippincott Williams & Wilkins.

156. Rothman, K.J. and S. Greenland, *Causation and causal inference in epidemiology.* Am J Public Health, 2005. 95 Suppl 1: p. S144-50.

157. Hill, A.B., *The Environment and Disease: Association or Causation?* Proc R Soc Med, 1965. 58: p. 295-300.

158. Bickeböller, H., Fischer, C *Einführung in die genetische Epidemiologie*. 2007, Heidelberg: Springer.

159. Ziegler, A., König, IR, *A Statistical Approach to Genetic Epidemiology. Concepts and Applications*. 2nd Edition ed. 2010, Weinheim.

160. Collins, F.S. and V.A. McKusick, *Implications of the Human Genome Project for medical science.* JAMA, 2001. 285(5): p. 540-4.

161. Chan, I.S. and G.S. Ginsburg, *Personalized medicine: progress and promise.* Annu Rev Genomics Hum Genet, 2011. 12: p. 217-44.

162. Bochud, M., *Genetics for clinicians: from candidate genes to whole genome scans (technological advances).* Best Pract Res Clin Endocrinol Metab, 2012. 26(2): p. 119-32.

163. Naidoo, N., et al., *Human genetics and genomics a decade after the release of the draft sequence of the human genome.* Hum Genomics, 2011. 5(6): p. 577-622.

164. Hirschhorn, J.N. and M.J. Daly, *Genome-wide association studies for common diseases and complex traits.* Nat Rev Genet, 2005. 6(2): p. 95-108.

165. McCarthy, M.I., et al., *Genome-wide association studies for complex traits: consensus, uncertainty and challenges.* Nat Rev Genet, 2008. 9(5): p. 356-69.

166. Marian, A.J., *Molecular genetic studies of complex phenotypes.* Transl Res, 2012. 159(2): p. 64-79.

167. Reich, D.E. and E.S. Lander, *On the allelic spectrum of human disease.* Trends Genet, 2001. 17(9): p. 502-10.

168. Christensen, K. and J.C. Murray, *What genome-wide association studies can do for medicine.* N Engl J Med, 2007. 356(11): p. 1094-7.

169. Spencer, C.C., et al., *Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip.* PLoS Genet, 2009. 5(5): p. e1000477.

170. Barrett, J.C. and L.R. Cardon, *Evaluating coverage of genome-wide association studies.* Nat Genet, 2006. 38(6): p. 659-62.

171. Pe'er, I., et al., *Evaluating and improving power in whole-genome association studies using fixed marker sets.* Nat Genet, 2006. 38(6): p. 663-7.

172. Nielsen, R., et al., *SNP calling, genotype calling, and sample allele frequency estimation from New-Generation Sequencing data.* PLoS One, 2012. 7(7): p. e37558.

173. Turner, S., et al., *Quality control procedures for genome-wide association studies.* Curr Protoc Hum Genet, 2011. Chapter 1: p. Unit1 19.

174. Marchini, J., et al., *A new multipoint method for genome-wide association studies by imputation of genotypes.* Nat Genet, 2007. 39(7): p. 906-13.

175. Reich, D.E. and D.B. Goldstein, *Detecting association in a case-control study while correcting for population stratification.* Genet Epidemiol, 2001. 20(1): p. 4-16.

176. Bouaziz, M., C. Ambroise, and M. Guedj, *Accounting for population stratification in practice: a comparison of the main strategies dedicated to genome-wide association studies.* PLoS One, 2011. 6(12): p. e28845.

177. Sillanpaa, M.J., *Overview of techniques to account for confounding due to population stratification and cryptic relatedness in genomic data association analyses.* Heredity (Edinb), 2011. 106(4): p. 511-9.

178. Rodriguez-Murillo, L. and D.A. Greenberg, *Genetic association analysis: a primer on how it works, its strengths and its weaknesses.* Int J Androl, 2008. 31(6): p. 546-56.

179. Thomas, D.C. and J.S. Witte, *Point: population stratification: a problem for case-control studies of candidate-gene associations?* Cancer Epidemiol Biomarkers Prev, 2002. 11(6): p. 505-12.

180. Devlin, B. and K. Roeder, *Genomic control for association studies.* Biometrics, 1999. 55(4): p. 997-1004.

181. Pe'er, I., et al., *Estimation of the multiple testing burden for genomewide association studies of nearly all common variants.* Genet Epidemiol, 2008. 32(4): p. 381-5.

182. International HapMap Consortium, *A haplotype map of the human genome.* Nature, 2005. 437(7063): p. 1299-320.

183. Zeggini, E., et al., *Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes.* Nat Genet, 2008. 40(5): p. 638-45.

184. Begum, F., et al., *Comprehensive literature review and statistical considerations for GWAS meta-analysis.* Nucleic Acids Res, 2012. 40(9): p. 3777-84.

185. Ioannidis, J.P., N.A. Patsopoulos, and E. Evangelou, *Heterogeneity in meta-analyses of genome-wide association investigations.* PLoS One, 2007. 2(9): p. e841.

186. Nakaoka, H. and I. Inoue, *Meta-analysis of genetic association studies: methodologies, between-study heterogeneity and winner's curse.* J Hum Genet, 2009. 54(11): p. 615-23.

187. Higgins, J.P., et al., *Measuring inconsistency in meta-analyses.* BMJ, 2003. 327(7414): p. 557-60.

188. Hirschhorn, J.N., et al., *A comprehensive review of genetic association studies.* Genet Med, 2002. 4(2): p. 45-61.

189. Zöllner, S. and J.K. Pritchard, *Overcoming the winner's curse: estimating penetrance parameters from case-control data.* Am J Hum Genet, 2007. 80(4): p. 605-15.

190. Ioannidis, J.P., *Non-replication and inconsistency in the genome-wide association setting.* Hum Hered, 2007. 64(4): p. 203-13.

191. Skol, A.D., et al., *Optimal designs for two-stage genome-wide association studies.* Genet Epidemiol, 2007. 31(7): p. 776-88.

192. Wang, H., et al., *Optimal two-stage genotyping designs for genome-wide association scans.* Genet Epidemiol, 2006. 30(4): p. 356-68.

193. Muller, H.H., R. Pahl, and H. Schafer, *Including sampling and phenotyping costs into the optimization of two stage designs for genomewide association studies.* Genet Epidemiol, 2007. 31(8): p. 844-52.

194. Lawlor, D.A., et al., *Mendelian randomization: using genes as instruments for making causal inferences in epidemiology.* Stat Med, 2008. 27(8): p. 1133-63.

195. Smith, G.D., N. Timpson, and S. Ebrahim, *Strengthening causal inference in cardiovascular epidemiology through Mendelian randomization.* Ann Med, 2008. 40(7): p. 524-41.

196. Ebrahim, S. and G. Davey Smith, *Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology?* Hum Genet, 2008. 123(1): p. 15-33.

197. Nitsch, D., et al., *Limits to causal inference based on Mendelian randomization: a comparison with randomized controlled trials.* Am J Epidemiol, 2006. 163(5): p. 397-403.

198. Smith, G.D. and S. Ebrahim, *Mendelian randomization: prospects, potentials, and limitations.* Int J Epidemiol, 2004. 33(1): p. 30-42.

199. Hingorani, A. and S. Humphries, *Nature's randomised trials.* Lancet, 2005. 366(9501): p. 1906-8.

200. Smith, G.D. and S. Ebrahim, *'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?* Int J Epidemiol, 2003. 32(1): p. 1-22.

201. Marzi, C., et al., *Genome-wide association study identifies two novel regions at 11p15.5-p13 and 1p31 with major impact on acute-phase serum amyloid A.* PLoS Genet, 2010. 6(11): p. e1001213.

202. Rathmann, W., et al., *Sex differences in the associations of socioeconomic status with undiagnosed diabetes mellitus and impaired glucose tolerance in the elderly population: the KORA Survey 2000.* Eur J Public Health, 2005. 15(6): p. 627-33.

203. Herder, C., et al., *Association of systemic concentrations of macrophage migration inhibitory factor with impaired glucose tolerance and type 2 diabetes: results from the Cooperative Health Research in the Region of Augsburg, Survey 4 (KORA S4).* Diabetes Care, 2006. 29(2): p. 368-71.

204. Kyle, U.G., et al., *Validation of a bioelectrical impedance analysis equation to predict appendicular skeletal muscle mass (ASMM).* Clin Nutr, 2003. 22(6): p. 537-43.

205. Kyle, U.G., et al., *Single prediction equation for bioelectrical impedance analysis in adults aged 20--94 years.* Nutrition, 2001. 17(3): p. 248-53.

206. Trayhurn, P. and J.H. Beattie, *Physiological role of adipose tissue: white adipose tissue as an endocrine and secretory organ.* Proc Nutr Soc, 2001. 60(3): p. 329-39.

207. Scheja, L., et al., *Acute-phase serum amyloid A as a marker of insulin resistance in mice.* Exp Diabetes Res, 2008. 2008: p. 230837.

208. Danesh, J., et al., *Long-term interleukin-6 levels and subsequent risk of coronary heart disease: two new prospective studies and a systematic review.* PLoS Med, 2008. 5(4): p. e78.

209. Wilkins, J., et al., *Rapid automated enzyme immunoassay of serum amyloid A.* Clin Chem, 1994. 40(7 Pt 1): p. 1284-90.

210. Friesinger, G.C., E.E. Page, and R.S. Ross, *Prognostic significance of coronary arteriography.* Trans Assoc Am Physicians, 1970. 83: p. 78-92.

211. Orphanides, G., T. Lagrange, and D. Reinberg, *The general transcription factors of RNA polymerase II.* Genes Dev, 1996. 10(21): p. 2657-83.

212. Huizing, M., et al., *Cellular, molecular and clinical characterization of patients with Hermansky-Pudlak syndrome type 5.* Traffic, 2004. 5(9): p. 711-22.

213. Cori, C.F., *The glucose-lactic acid cycle and gluconeogenesis.* Curr Top Cell Regul, 1981. 18: p. 377-87.

214. Wu, W., et al., *Genetic variants in GTF2H1 and risk of lung cancer: a case-control analysis in a Chinese population.* Lung Cancer, 2009. 63(2): p. 180-6.

215. Fantin, V.R., J. St-Pierre, and P. Leder, *Attenuation of LDH-A expression uncovers a link between glycolysis, mitochondrial physiology, and tumor maintenance.* Cancer Cell, 2006. 9(6): p. 425-34.

216. Xie, H., et al., *LDH-A inhibition, a therapeutic strategy for treatment of hereditary leiomyomatosis and renal cell cancer.* Mol Cancer Ther, 2009. 8(3): p. 626-35.

217. Le, A., et al., *Inhibition of lactate dehydrogenase A induces oxidative stress and inhibits tumor progression.* Proc Natl Acad Sci U S A, 2010. 107(5): p. 2037-42.

218. Koslowski, M., et al., *Multiple splice variants of lactate dehydrogenase C selectively expressed in human cancer.* Cancer Res, 2002. 62(22): p. 6750-5.

219. Lappalainen, T., et al., *Serum concentrations and expressions of serum amyloid A and leptin in adipose tissue are interrelated: the Genobin Study.* Eur J Endocrinol, 2008. 158(3): p. 333-41.

220. Tartaglia, L.A., *The leptin receptor.* J Biol Chem, 1997. 272(10): p. 6093-6.

221. Tartaglia, L.A., et al., *Identification and expression cloning of a leptin receptor, OB-R.* Cell, 1995. 83(7): p. 1263-71.

222. Elliot, G.B. and G. Horner, *False-Positive C-Reactive Protein Reactions Due to Serum Lipoproteins.* Am Rev Respir Dis, 1964. 90: p. 453-4.

223. Zhang, Y.Y., et al., *Genetic variability at the leptin receptor (LEPR) locus is a determinant of plasma fibrinogen and C-reactive protein levels.* Atherosclerosis, 2007. 191(1): p. 121-7.

224. Ridker, P.M., et al., *Loci related to metabolic-syndrome pathways including LEPR,HNF1A, IL6R, and GCKR associate with plasma C-reactive protein: the Women's Genome Health Study.* Am J Hum Genet, 2008. 82(5): p. 1185-92.

225. Sabatti, C., et al., *Genome-wide association analysis of metabolic traits in a birth cohort from a founder population.* Nat Genet, 2009. 41(1): p. 35-46.

226. Purcell, S., S.S. Cherny, and P.C. Sham, *Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits.* Bioinformatics, 2003. 19(1): p. 149-50.

227. Lilja, M., et al., *The impact of leptin and adiponectin on incident type 2 diabetes is modified by sex and insulin resistance.* Metab Syndr Relat Disord, 2012. 10(2): p. 143-51.

228. Schmidt, M.I., et al., *Leptin and incident type 2 diabetes: risk or protection?* Diabetologia, 2006. 49(9): p. 2086-96.

229. Carstensen, M., et al., *Accelerated increase in serum interleukin-1 receptor antagonist starts 6 years before diagnosis of type 2 diabetes: Whitehall II prospective cohort study.* Diabetes, 2010. 59(5): p. 1222-7.

230. Wichmann, H.E., et al., *KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes.* Gesundheitswesen, 2005. 67 Suppl 1: p. S26-30.

231. Karakas, M., et al., *Leptin, adiponectin, their ratio and risk of Coronary Heart Disease: results from the MONICA/KORA Augsburg Study 1984-2002.* Atherosclerosis, 2010. 209(1): p. 220-5.

232. Winkelmann, B.R., et al., *Rationale and design of the LURIC study--a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease.* Pharmacogenomics, 2001. 2(1 Suppl 1): p. S1-73.

7. Guide to abbreviations

A-SAA	acute-phase serum amyloid A
ADA	American Diabetes Association
BMI	body mass index
C-SAA	constitutive serum amyloid A
C/EBP	CCAAT/enhancer binding protein
CI	confidence interval
DGDG	Diabetes Gene Discovery Group
DGI	Diabetes Genetics Initiative
DIAGRAM	Diabetes Genetics Replication and Meta-analysis Consortium
EUROSPAN	European Special Population Network
FFA	free-fatty acids
FUSION	Finland-United States Investigation of Non-Insulin Dependent Diabetes Genetics
GCSF	granulocyte colony-stimulating factor
GTF2H1	general transcription factor 2 H1
GWAS	genome-wide association study
HapMap	Haplotype Mapping
HbA1c	hemoglobin A1c
HDL	high-density lipoprotein
HOMA-IR	homeostasis model assessment for insulin resistance
HPS5	Hermansky-Pudlak Syndrome 5
hs-CRP	high sensitivity-C reactive protein
HWE	Hard-Weinberg equilibrium
ICAM-1	intercellular adhesion molecule-1
IDF	International Diabetes Federation
IFG	impaired fasting glucose
IGT	impaired glucose tolerance
IκK	inhibitor of NF-κ B kinase
IL (-1, -1ß, -6, -8, -18)	interleukin (-1, -1ß, -6, -8, -18)
IL-1Ra	interleukin 1 receptor antagonist
iNOS	inducible nitric oxide synthase
IRS (-1, -2)	insulin receptor substrate (-1, -2)
JNK	JunkN terminal kinase
KCNC1	potassium voltage-gated channel subfamily C member 1
KCNJ11	potassium inwardly-rectifying channel, subfamily J, member 11
KORA	Cooperative Research in the Region of Augsburg

LADA	late-onset autoimmunity diabetes in the adult
LD	linkage disequilibrium
LDHA	lactate dehydrogenase A
LDHC	lactate dehydrogenase C
LEPR	leptin receptor
LIF	leukemia inhibitory factor
LURIC	Ludwigshafen Risk and Cardiovascular Health study
MAF	Minor allele frequency
MAGIC	Meta-Analyses of Glucose and Insulin-related traits Consortium
MCP-1	monocyte chemotactic protein-1
MODY	maturity-onset diabetes in the young
MONICA	Monitoring of Trends and Determinants in Cardiovascular Disease
MYOD1	myogenic differentiation 1
NF-κB	nuclear factor κB
NO	nitric oxide
OGTT	oral glucose tolerance test
OR	odds ratio
PAI-1	plasminogen activator inhibitor-1
PKC (-θ)	protein kinase C (-θ)
ROS	reactive oxidative species
PPAR (α, γ)	peroxisome proliferator-activated receptor (α, γ)
SAA	serum amyloid A
SERGEF	secretion regulating guanine nucleotide exchange factor
SNP	single nucleotide polymorphism
SOCS (-3)	suppressor of cytokine signaling (-3)
TCF7L2	transcription factor 7-like 2
TNF-α	tumor necrosis factor-α
TPH	tryptophan hydroxylase
TWINS	UK adult twin register
WHO	World Health Organization
WHR	waist-to-hip ratio
WTCCC/ UKT2D	Welcome Trust Case Control Consortium/ United Kingdom Type 2 Diabetes Genetics Consortium
YY1	ying and yang 1

8. List of tables

9. List of figures

10. Description of own contribution

This doctoral thesis comprises three studies which are all the result of joint efforts of several co-authors. Therefore, my contribution to the three studies is explained in detail in the following paragraphs.

I initiated all three studies. Furthermore, for all three studies, I worked out the study designs and analysis concepts, performed an extensive literature search for published studies providing information on the scientific background for each of the studies, interpreted all results and wrote the complete manuscripts for all studies except for the paragraphs *Genome-wide genotyping and imputation* and *Genome-wide association analyses and meta-analysis* of the Materials and Methods section of the second study of this doctoral thesis.

In addition to this, for the manuscript ***Acute-phase serum amyloid A protein and its implication in the development of type 2 diabetes in the KORA S4/ F4 study*** the study design and the interpretation of the results were discussed with Dr. Cornelia Huth, PD Dr. Christian Herder, and PD Dr. Barbara Thorand. I created the analysis datasets after having performed data transformations. Furthermore, I performed all analyses and created all corresponding tables and figures. Dr. Jens Baumert gave statistical advice. Prof. Dr. Thomas Illig and Prof. Dr. Wolfgang Koenig supervised the study and many of the co-authors revised the manuscript draft I wrote and gave advice on the analysis concept.

For the study ***Genome-Wide Association Study Identifies Two Novel Regions at 11p15.5-p13 and 1p31 with Major Impact on Acute-Phase Serum Amyloid A*** participating studies were recruited by Prof. Dr. Thomas Illig. I corresponded with the Principal Investigators of each of the participating studies and coordinated the study partners. Harald Grallert performed the laboratory analysis for the validation study. I conducted the correlation analyses. Genome-wide association analyses were undertaken by Dr. Pirro G. Hysi for the TWINS UK, Vasiliki Lagou and Dr. Inga Prokopenko for the Sorbs study, and Eva Albrecht for the KORA the LURIC study. Eva Albrecht also conducted the meta-analysis and created the tables and figures. PD Dr. Christian Gieger gave statistical advice. Together with Prof. Dr. Thomas Illig he also supervised the study. The manuscript was revised by some of the co-authors.

Results of the Mendelian Randomization study were extracted from an analysis conducted by Dr. Benjamin F. Voight who also assisted in questions regarding the power calculation for this study and the interpretation of the results. I performed the power calculation and discussed the procedure also with Dr. Cornelia Huth. The study was revised by PD Dr. Christian Herder and PD Dr. Barbara Thorand.

12. Publications

Marzi C, Huth C, Herder C, Baumert J, Thorand B, Rathmann W, Meisinger C, Wichmann HE, Roden M, Peters A, Grallert H, Koenig W, Illig T. Acute-Phase Serum Amyloid A Protein and Its Implication in the Development of Type 2 Diabetes in the KORA S4/F4 Study. Diabetes Care. 2012 Dec 13. [Epub ahead of print] PubMed PMID: 23238662.

Breitfeld J, Tönjes A, Böttcher Y, Schleinitz D, Wiele N, Marzi C, Brockhaus C, Rathmann W, Huth C, Grallert H, Illig T, Blüher M, Kovacs P, Stumvoll M. Genetic variation in the vaspin gene affects circulating serum vaspin concentrations. Int J Obes (Lond). 2012 Aug 21. doi: 10.1038/ijo.2012.133. [Epub ahead of print] PubMed PMID: 22907691.

Marzi C, Albrecht E, Hysi PG, Lagou V, Waldenberger M, Tönjes A, Prokopenko I, Heim K, Blackburn H, Ried JS, Kleber ME, Mangino M, Thorand B, Peters A, Hammond CJ, Grallert H, Boehm BO, Kovacs P, Geistlinger L, Prokisch H, Winkelmann BR, Spector TD, Wichmann HE, Stumvoll M, Soranzo N, März W, Koenig W, Illig T, Gieger C. Genome-wide association study identifies two novel regions at 11p15.5-p13 and 1p31 with major impact on acute-phase serum amyloid A. PLoS Genet. 2010 Nov 18;6(11):e1001213. PubMed PMID: 21124955; PubMed Central PMCID: PMC2987930.

Grallert H, Herder C, Marzi C, Meisinger C, Wichmann HE, Rathmann W, Illig T. Association of genetic variation in KCNQ1 with type 2 diabetes in the KORA surveys. Horm Metab Res. 2010 Feb;42(2):149-51. Epub 2009 Oct 1. PubMed PMID: 19798621.

Marzi C, Huth C, Kolz M, Grallert H, Meisinger C, Wichmann HE, Rathmann W, Herder C, Illig T. Variants of the transcription factor 7-like 2 gene (TCF7L2) are strongly associated with type 2 diabetes but not with the metabolic syndrome in the MONICA/KORA surveys. Horm Metab Res. 2007 Jan;39(1):46-52. PubMed PMID: 17226113.

13. Acknowledgment

I wish to thank many people who have supported my work as PhD student at the Helmholtz Center Munich.

First off all, I sincerely thank my doctoral thesis supervisor, Prof. Dr. Thomas Illig. He has inspired and constantly encouraged my work in the field of genetic epidemiology and type 2 diabetes. I also thank him for great confidence in my work, excellent scientific advice and enjoyable working conditions during the entire term.

Moreover, I wish to thank my group leader Dr. Harald Grallert as well as PD Dr. Christian Herder, Dr. Cornelia Huth, PD Dr. Barbara Thorand, and Dr. Holger Prokisch for their openness to all kinds of emerging questions, and their invaluable scientific advice and supervision of this doctoral thesis. Likewise, I thank Dr. Jens Baumert, Dr. Melanie Waldenberger, Dr. Eva Reischel, Jennifer Kriebel, Simone Wahl, and Barbara Nitz for great help in different aspects of this doctoral thesis.

I am also grateful to all scientists from the KORA, the Sorbs, the LURIC, and the TWINS UK study as well as the DIAGRAM consortium who were involved in the studies of this doctoral thesis for helpful comments on the manuscripts and a fruitful, straightforward collaboration. In particular, I thank Prof. Dr. Wolfgang Koenig and PD Dr. Christian Gieger for their competent supervision of the manuscripts as well as Eva Albrecht for a delightful cooperation.

I also wish to thank all staff members who were involved in in the planning, organization and conduct of the studies included in this doctoral thesis as well as all staff members involved in the generation of SNP data. Finally, I express my appreciation to all study participants.

I gratefully dedicate this thesis to my parents and my husband who always support me unconditionally and to my two little boys, the sunshine of my life.

14. Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.