*This is the accepted manuscript of the article, which has been published in Genetic Testing and Molecular Biomarkers. 2020, 24(1). https://doi.org/10.1089/gtmb.2019.0119*

# Association of collagen type IV alpha 1 chain gene rs3783107 (G>A) major genotype with hypertension, asthma and eczema, the TAMRISK study

Tarja Kunnas[1], Seppo T Nikkari[1*]

[1]Department of Medical Biochemistry, Faculty of Medicine and Health Technology, Tampere University and Fimlab laboratories, Tampere, Finland

* Corresponding author:

Seppo Nikkari, MD, PhD, Department of Medical Biochemistry, Faculty of Medicine and Health Technology, FI33014 Tampere University, Finland. E-mail: seppo.nikkari@uta.fi

Running title: COL4A1 gene variant rs3783107 and disease

**Abstract**

**Aims.** Basement membranes (BMs) provide structural support to the tissue, and also offer functional input to modulate cellular function. Type IV collagen is the most abundant protein in BMs. Collagen type IV alpha 1 chain (COL4A1) gene variant rs3783107 G>A has previously been associated with intracranial aneurysms. We wanted to study whether this polymorphism associates with hypertension; we also examined its association with cerebrovascular events, asthma and long-term eczema in the Tampere adult population cardiovascular risk study (TAMRISK).

**Materials and Methods.** A Finnish periodic health examination (PHE) cohort of 331 subjects with diagnosed hypertension and 440 normotensive controls was analyzed. DNA was extracted from buccal swabs. Genotyping was performed using KASP (competitive allelic specific amplification). Prevalence of hypertension, asthma or long-term eczema were obtained from PHE. Incidence of cerebrovascular diseases and transient cerebral ischemic attacks (TIA) were obtained from the National Hospital Discharge Registry (HILMO).

**Results**. Even after adjusting for BMI-class and gender, subjects with rs3783107 major genotype GG had significantly more hypertension (OR 1.38, CI 1.01-1.87, p=0.043), asthma (OR 2.78, CI 1.25-6.19, p=0.012), and eczema (OR 2.00, CI 1.08-3.70, p=0.027), than those with the A allele. Furthermore, subjects with major genotype GG had significantly higher systolic- (p=0.026) and diastolic blood pressure (p=0.013) compared to those with the A-allele. Variant rs3783107 did not significantly associate with cerebrovascular events.

**Conclusions.** Our findings suggest that there are genotype-phenotype associations between the *COL4A1* gene variant rs3783107 and hypertension, asthma, and eczema.

KEYWORDS: Single Nucleotide Polymorphism, Collagen alpha1(IV), Vascular diseases

## Introduction

Basement membranes (BMs) are abundant components of the extracellular matrix (ECM) in every tissue of the human body. Different collagen types are expressed throughout the human body of which type I is the most abundant; collagen type IV is the predominant type in BM. (LeBleu et al. 2007). BMs are specifically found in the epithelium lining the skin and respiratory tract. The principal cell responsible for arterial structure, function and repair is the smooth muscle cell (SMC). All SMCs in the medial layer of arteries are surrounded by their own BM, which seems to be crucial for the maintenance of SMC quiescence. The disruption of the interaction between SMCs and BM is part of cellular activation after atherogenic or traumatic stimuli (Hedin et al. 1999). Lungs are composed of airways and lung parenchyma, and changes in the ECM in the airways may affect respiratory diseases, including asthma (Burgess et al. 2016). Also, pathological studies have demonstrated that even minor structural differences in collagen IV can lead to distinct, clinically different diseases of the skin, including subepidermal blistering diseases (Abreu-Velez and Howard. 2012). In addition to providing structural support to the tissue, BMs thus also offer functional input to modulate cellular function.

Genetic variations in the collagen type IV alpha-1 gene (*COL4A1*) have been associated with a variety of vascular abnormalities, including angiopathy (Vahedi and Alamowitch. 2011), prevalence of myocardial infarction (Yamada et al. 2008), and differences in pulse wave velocity, the rate at which pressure waves move down the vessel, measuring arterial stiffness (Tarasov et al. 2009). We wanted to study whether the intronic *COL4A1* gene variant rs3783107 associates with hypertension, cerebrovascular events, asthma or long-term eczema in the Tampere adult population cardiovascular risk study (TAMRISK),

because there was a significant relationship between rs3783107 and intracranial aneurysm (IA) in a Dutch population (Ruigrok et al. 2006).

## Materials and Methods

### Subjects

TAMRISK study data was obtained in 2003 from periodic health examinations (PHE) done by a public health nurse for 50-year-old men and women living in Tampere, a city in southern Finland with 220 000 inhabitants (Maatta et al. 2015). Height (cm) and weight (kg) were recorded from which the body mass index (BMI) was calculated. We further divided the subjects to BMI classes of <25, 25 to 29.9, 30 to 34.9, and >35 kg/$m^2$. Blood pressure (mmHg) was measured using a calibrated mercury sphygmomanometer. For most patients, physicians diagnose hypertension when diastolic blood pressure readings are consistently 90 mmHg and systolic 140 mmHg, or above. Serum total cholesterol (mmol/L) was measured after an overnight fast by standard techniques. The public health nurse conducted an interview using a structured questionnaire about health and health-related behavior. Current and previous diseases were identified based on self-report of diagnosis by a physician, including hypertension. During years 2006–2010, buccal swabs for DNA extraction were collected from the participants by mail. Using the patient's national identity code, data on hospitalizations including ICD-10 codes for discharge diagnoses were obtained from the National Hospital Discharge Registry (HILMO) maintained by the National Institute of Health and Welfare. The incidence of cerebrovascular diseases (I60-I69), and transient cerebral ischemic attacks (TIAs) (G45) was followed up from 2005 to 2014 until the subjects were on average 60 years old (Piesanen et al. 2018). In follow-up of the genotyped subjects, there were 14 who had a diagnosis of cerebrovascular disease and 7 with TIA. The subjects with cerebrovascular disease and TIA were combined for the

group with cerebrovascular events. The Ethics Committees of the Tampere University Hospital and the City of Tampere approved the study.

Cases were subjects who had self-reported hypertension at the age of 50 years (as diagnosed by a physician) and for each case, at least one normotensive control subject with the same age, sex, and similar smoking habits, was chosen in order of admission from the PHE cohort (n=6000).

Genotyping

*COL4A1* gene variant rs3783107 G>A is positioned at chr13:110189755 (GRCh38.p12), with MAF/MinorAlleleCount A=0.3624/1815 (1000 Genomes). More information on this polymorphism can be found from https://www.ncbi.nlm.nih.gov/snp/rs3783107. DNA was extracted from buccal swabs using a commercial kit (Qiagen Inc., Valencia, Calif., USA). Genotyping was performed using KASP (competitive allelic specific amplification) genotyping services at KBioscience Institute, UK. Details of this method can be obtained from https://www.lgcgroup.com/genotyping/.

Statistical analysis

One-way ANOVA or T-test for continuous variables, and Chi-square test or Fisher's exact test for categorical variables were applied for the comparison of cases, controls and genotype groups. Logistic regression was adjusted by BMI-class and gender. Analyses were carried out using SPSS 23.0 for Windows (SPSS Inc., Chicago, Illinois, USA).

## Results

Clinical characteristics of case group of hypertensive subjects and controls at the age of 50 years have been previously described (Kunnas et al. 2012). Samples were available and genotyping for *COL4A1* gene variant rs3783107 was successful for all 771 tested

subjects: 331 cases and 440 controls (304 women and 467 men). The genotype frequencies for the whole tested population were: GG 354 (45.9%), GA 331 (42.9%), and AA 86 (11.2%). G allele frequency was 67.4% and A allele frequency 32.6%. The genotypes were in Hardy-Weinberg equilibrium (Chi-square=0.29, p>0.05). The genotype frequencies in the case group (139 women and 192 men) were: GG 168 (50.8%), GA 131 (39.6%), and AA 32 (9.7 %). The genotype frequencies in the control group (165 women and 275 men) were: GG 186 (42.3%), GA 200 (45.5%), and AA 54 (12.3%).

Association analyses of the different genotypes with background characteristics showed that subjects with major genotype GG had significantly higher systolic- (p=0.026) and diastolic blood pressure (p=0.013) compared to those with the A-allele (Table 1).

Association analyses of *COL4A1* gene variant rs3783107 genotypes with hypertension, asthma and long-term eczema for are given in Table 1. Even after adjusting for BMI-class and gender, subjects with major genotype GG had significantly more hypertension (OR 1.38, CI 1.01-1.87, p=0.043), asthma (OR 2.78, CI 1.25-6.19, p=0.012), and eczema (OR 2.00, CI 1.08-3.70, p=0.027), than those with the A allele. The *COL4A1* gene variant rs3783107 did not significantly associate with cerebrovascular events.

**Discussion**

The intronic rs3783107 minor allele A in *COL4A1* gene has previously associated with IA in a Dutch population (Ruigrok et al. 2006), and the risk of minor allele carriers in the Dutch population was significantly higher than that in the Japanese population (Meng et al. 2017, Ruigrok et al. 2006). We report that the major allele of this polymorphic site associates with hypertension, self-reported asthma and eczema in the TAMRISK study. *COL4A1* rs3783107 does not lead to amino acid change. However, it may be in linkage

disequilibrium with a functional polymorphism that could lead to pathological changes at the type IV collagen protein level.

The SMC basement membrane is considered to be crucial for the maintenance of cell quiescence (Hedin et al. 1999) and its alterations, also those possibly contributed by *COL4A1* variation, could lead to SMC activation. There may be a link to hypertension that is characterized by arterial stiffness with vascular SMC hypertrophy (Hixon and Gualberto. 2003).

The ECM has been regarded as an inert structure in the lung. However, there is also evidence that the ECM is a bioactive environment, and it might be that ECM changes in lung diseases do not solely result from a disease process, but may themselves cause disease, such as asthma (Burgess et al. 2016). Airway wall modeling in asthma (Wiggs et al. 1992) may thus associate with the investigated type IV collagen gene variant.

Finally, there was an association of *COL4A1* rs3783107 with self-reported long-term eczema. Several skin conditions fall under the heading of eczema, and differential diagnosis could not be made on the basis of the PHE. However, it is known that structural differences in collagen IV can lead to diseases of the skin, since type IV collagen is a type of collagen found primarily in the skin within the basement membrane zone (Abreu-Velez and Howard. 2012).

This was an exploratory study, and the population was relatively small. Nevertheless, it is noteworthy that all three findings for disease were in the similar direction, where subjects with major rs3783107 genotype GG had significantly more hypertension, asthma, and long-term eczema, than those with the A allele. We did not find a significant association of rs3783107 genotype with cerebrovascular disease, although there was a trend of more such cases in the GG genotype group. These observations for involvement of the GG

genotype are somewhat in contrast to previous observations, where there was evidence of the minor A allele for contributing to development of intracranial aneurysms (Ruigrok et al. 2006). Our findings however suggest that the GG genotype of *COL4A1* gene variant rs3783107 may be involved in disease.

Strengths of the study include that all participating subjects were 50 years old when the PHE examination took place. Since the study subjects are from a restricted genetic pool of a large city in Finland (Finnish Caucasian), the findings might not be extrapolated to different genetic populations.

## Acknowledgments

The expert technical assistance by Mirka Pietiläinen and Nina Peltonen is gratefully acknowledged. This study was supported by grants from Competitive research funding of the Pirkanmaa Hospital District.

## Author Disclosure Statement

No competing financial interests exist.

## References

Abreu-Velez AM, Howard MS (2012) Collagen IV in Normal Skin and in Pathological Processes. N Am J Med Sci 4:1-8.

Burgess JK, Mauad T, Tjin G*, et al* (2016) The extracellular matrix - the under-recognized element in lung disease? J Pathol 240:397-409.

Hedin U, Roy J, Tran PK*, et al* (1999) Control of smooth muscle cell proliferation--the role of the basement membrane. Thromb Haemost 82 Suppl 1:23-26.

Hixon ML, Gualberto A (2003) Vascular smooth muscle polyploidization--from mitotic checkpoints to hypertension. Cell Cycle 2:105-110.

Kunnas T, Maatta K, Palmroos P*, et al* (2012) Periodic cohort health examinations in the TAMRISK study show untoward increases in body mass index and blood pressure during 15 years of follow-up. BMC Public Health 12:654-2458-12-654.

LeBleu VS, Macdonald B, Kalluri R (2007) Structure and function of basement membranes. Exp Biol Med (Maywood) 232:1121-1129.

Maatta KM, Nikkari ST, Kunnas TA (2015) Genetic variant coding for iron regulatory protein HFE contributes to hypertension, the TAMRISK study. Medicine (Baltimore) 94:e464.

Meng Q, Hao Q, Zhao C (2017) The association between collagen gene polymorphisms and intracranial aneurysms: a meta-analysis. Neurosurg Rev. https://doi.org/10.1007/s10143-017-0925-x.

Piesanen J, Kunnas T, Nikkari ST (2018) Association of Desmin Gene Variant rs1058261 with Cardiovascular Disease, the TAMRISK Study. Genet Test Mol Biomarkers 22:574-576.

Ruigrok YM, Rinkel GJ, van't Slot R*, et al* (2006) Evidence in favor of the contribution of genes involved in the maintenance of the extracellular matrix of the arterial wall to the development of intracranial aneurysms. Hum Mol Genet 15:3361-3368.

Tarasov KV, Sanna S, Scuteri A*, et al* (2009) COL4A1 is associated with arterial stiffness by genome-wide association scan. Circ Cardiovasc Genet 2:151-158.

Vahedi K, Alamowitch S (2011) Clinical spectrum of type IV collagen (COL4A1) mutations: a novel genetic multisystem disease. Curr Opin Neurol 24: 63-68.

Wiggs BR, Bosken C, Pare PD*, et al* (1992) A model of airway narrowing in asthma and in chronic obstructive pulmonary disease. Am Rev Respir Dis 145:1251-1258.

Yamada Y, Kato K, Oguri M*, et al* (2008) Genetic risk for myocardial infarction determined by polymorphisms of candidate genes in a Japanese population. J Med Genet 45:216-221.

**Table 1.** Clinical characteristics (means ± SD) of the study population stratified according to *COL4A1* gene variant rs3783107 genotypes. P values <0.05 are in bold.

| | GG (354) | GA (331) | AA (86) | P value * GG vs. GA vs. AA | P value * AA vs. (GA+GG) | P value * GG vs. (GA+AA) | OR ** (95% CI) GG vs. (GA+AA) | P value ** GG vs. (GA+AA) |
|---|---|---|---|---|---|---|---|---|
| Hypertension % (n) | 47.5 (168) | 39.6 (131) | 37.2 (32) | 0.060 | 0.255 | **0.019** | 1.38 (1.01-1.87) | **0.043** |
| Asthma % (n) | 6.0 (21) | 2.5 (8) | 1.2 (1) | **0.024** | 0.238 | **0.007** | 2.78 (1.25-6.19) | **0.012** |
| Eczema % (n) | 9.9 (30) | 4.4 (13) | 6.7 (5) | **0.031** | 0.868 | **0.011** | 2.00 (1.08-3.70) | **0.027** |
| Cerebrovascular events % (n) | 3.7 (13) | 2.1 (7) | 1.2 (1) | 0.293 | 0.498 | 0.136 | 1.84 (0.75-4.50) | 0.185 |
| | | | | | | | | |
| Body mass index kg/m$^2$ (SD) | 27.2 (4.8) | 26.6 (4.6) | 26.6 (4.2) | 0.206 | 0.512 | 0.076 | | |
| Cholesterol, mmol/L (SD) | 5.38(0.92) | 5.39 (0.91) | 5.33 (0.91) | 0.870 | 0.607 | 0.960 | | |
| Systolic blood pressure, mmHg (SD) | 136.5 (16.1) | 134.1 (17.7) | 132.7 (16.4) | 0.068 | 0.187 | **0.026** | | |
| Diastolic blood pressure, mmHg (SD) | 89.0 (9.8) | 87.2 (9.9) | 87.5 (9.4) | **0.044** | 0.616 | **0.013** | | |

SD, standard deviation. * Chi-square test, Fisher's exact test, one-way ANOVA or T-test. ** Logistic regression adjusted by BMI-class and gender