Genetic loci on chromosome 5 are associated with circulating levels of interleukin-5 and eosinophil count in a European population with high risk for cardiovascular disease

Olga McLeod[a], Angela Silveira[a], Elsa Valdes-Marquez[b], Harry Björkbacka[c], Peter Almgren[d] Karl Gertow[a], Jesper R. Gådin[a], Alexandra Bäcklund[a], Bengt Sennblad[a,e], Damiano Baldassarre[f,g], Fabrizio Veglia[g], Steve E. Humphries[h], Elena Tremoli[f,g], Ulf de Faire[i], Jan Nilsson[c], Olle Melander[d], Jemma C Hopewell[b], Robert Clarke[b], Hanna M. Björck[a], Anders Hamsten[a], John Öhrvik[a] and Rona J. Strawbridge[a] on behalf of the IMPROVE study group

[a]Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden.

bCTSU – Nuffield Department of Population Health, University of Oxford, Oxford, UK

[c]Experimental Cardiovascular Research Unit, Dept. of Clincal Sciences Malmö, Lund Univeristy, Malmö, Sweden.

dDepartment of Clinical Sciences, Malmö, Lund University, Malmö, Sweden

[e]Cardiovascular Medicine Unit, Department of Medicine, Science for Life Laboratory, Karolinska Institutet, Stockholm, Sweden.

[f]Dipartimento di Scienze Farmacologiche e Biomolecolari, Università di Milano

[g]Centro Cardiologico Monzino, IRCCS, Milan, Italy.

[h]Centre for Cardiovascular Genetics, University College London, United Kingdom.

[i]Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.

Correspondence to:
Rona J Strawbridge
L8:03, Centre for Molecular Medicine,
Karolinska Universitetsjukhuset
Solna, 171 76
Stockholm

Abstract

IL-5 is a Th2 cytokine which activates eosinophils and is suggested to have an atheroprotective role. Genetic variants in the *IL5* locus have been associated with increased risk of CAD and ischemic stroke. In this study we aimed to identify genetic variants associated with IL-5 concentrations and apply a Mendelian randomisation approach to assess IL-5 levels for causal effect on intima-media thickness in a European population at high risk of coronary artery disease. We analysed SNPs within robustly associated candidate loci for immune, inflammatory, metabolic and cardiovascular traits. We identified 2 genetic loci for IL-5 levels (chromosome 5, rs56183820, BETA= 0.11, P= $6.73E^{-5}$ and chromosome 14, rs4902762, BETA= 0.12 , P= $5.76E^{-6}$) and one for eosinophil count (rs72797327, BETA= -0.10, P= $1.41E^{-6}$). Both chromosome 5 loci were in the vicinity of the *IL5* gene, however the association with IL-5 levels failed to replicate in a meta-analysis of 2 independent cohorts (rs56183820, BETA= 0.04, P= 0.2763, I^2 =24, I^2-P = 0.2516). No significant associations were observed between SNPs associated with IL-5 levels or eosinophil count and IMT measures. Expression quantitative trait analyses indicate effects of the IL-5 and eosinophil-associated SNPs on *RAD50* mRNA expression levels (rs12652920 (r2=0.93 with rs56183820) BETA= -0.10, P=$8.64E^{-6}$ and rs11739623 (r2= 0.96 with rs72797327) BETA= -0.23, P=$1.74E^{-29}$, repsectively). Our data do not support a role for IL-5 levels and eosinophil count in intima-media thickness, however SNPs associated with IL-5 and eosinophils might influence stability of the atherosclerotic plaque via modulation of *RAD50* levels.

Key words

Subclinical atherosclerosis, intima-media thickness, genetics, IL-5, eosinophil count, *RAD50*

Abbreviations

$IMT_{mean-max}$ Mean of the maximum IMT measures along the entire carotid tree

$IMT\text{-}CC_{mean}$ Mean of IMT measures in the common carotid

$IMT\text{-}CC_{max}$ Max of IMT measures in the common carotid

$IMT\text{-}Bif_{mean}$ Mean of IMT measures in the carotid bifurcation

$IMT\text{-}Bif_{max}$ Max of IMT measures in the carotid bifurcation

1. Introduction

Atherosclerosis is a process leading to coronary artery disease (CAD) and ischemic stroke, and is among the leading causes of mortality and morbidity throughout the world [1]. It is thought that a key component of atherosclerosis is lipid accumulation in the arterial wall, leading to leukocyte recruitment, inflammation and participation of both innate and adaptive immune responses. To support the immune responses hypothesis, a number of immune or inflammatory biomarkers for cardiovascular disease (CVD) have been investigated. Some, such as C-reactive protein (CRP), appear to reflect general disease processes [2, 3],whereas a number of cytokines (such as IL-1, IL-2, IL-6, IL-8, IL-10, IL-12, IL-17 and TNFα) are robustly associated with CVD and are thought to be actively involved in the inflammatory processes leading to atherosclerosis [4, 5].

Genetic studies have identified single nucleotide polymorphisms (SNPs) that are associated with increased risk of CAD in the interleukin-5 gene (*IL5*) [6] and increased risk of ischemic stroke in the interleukin-5 receptor alpha subunit gene (*IL5RA*) [7]. As with most genetic variants associated with CVD, the mechanism by which the *IL5* SNP influences disease is unclear. A number of studies (including one from our group) have also proposed an athero-protective role for IL-5 [8-11], which is supported by the fact that it is produced mainly by anti-inflammatory T helper 2 (Th2) cells. In addition to activated Th2 cells, IL-5 can be produced by mast cells and to a lesser extent by eosinophils. IL-5 is a growth and differentiation factor for B-cells, an immunoglobulin-A (IgA)-enhancing factor and a key cytokine in eosinophil maturation, differentiation, activation and survival [12, 13]; thus, it is interesting that the *IL5* locus has also been associated with eosinophil count [14]. Cellular responses are dependent upon IL-5 activating a dimeric receptor consisting of an α and a β-subunit. The β-subunit is shared with IL-3 and granulocyte-macrophage colony-stimulating factor (GM-CSF) [15], however the α-subunit is specific to IL-5. Thus, availability of the α-chain on the cell surface controls responsiveness of the cell to IL-5 [16, 17].

Thickness of the intima-media layer of the carotid artery can be considered a marker of cardiovascular risk. Differences in carotid intima-media thickness (IMT) over time can be indicative of changes in the structure, and possibly functions, of the vessel wall. The aims of this study were to identify genetic variants associated with IL-5 levels and to use these in a Mendelian randomisation approach to assess IL-5 levels for causal effects on carotid IMT. As eosinophil count and IL-5 levels are inter-related, we also investigated the genetic regulation of eosinophil count. We used the prospective IMPROVE study (the Carotid Intima Media

Thickness [IMT] and IMT-Progression as Predictors of Vascular Events in a High Risk European Population study), which is based on 3,703 subjects at high risk of CVD with extensive genetic data, biochemical phenotyping and repeat detailed ultra-sound measurements of the carotid artery IMT [18, 19]. Genetic markers robustly associated with increased levels of a biomarker or constituting a causal component of increased IMT should be useful for risk stratification and early prevention of CVD.

2. Materials and Methods

2.1 Discovery cohort: the IMPROVE cohort

The design features, biobank and dataset of the IMPROVE study have been reported [18, 19]. Of 3,703 participants with at least three cardiovascular risk factors but without symptoms or diagnosed prevalent disease, 3,435 passed phenotypic quality control and were considered for the current study. Ethics committee approval for the study was obtained from each centre and written informed consent was obtained from all participants.

2.2 Phenotypes

A wide range of anthropometric and biochemical measurements were acquired, including fasting concentrations of HDL and LDL cholesterol (by Friedewald's formula), triacylglycerol (TG), high sensitivity CRP (hs-CRP), creatinine and plasma glucose as described [19]. Diabetes was defined as diagnosis of diabetes and/or fasting blood glucose ≥7 mmol/L and/or glucose lowering treatment. Levels of IL-5 were measured in EDTA-plasma samples using the ultra-sensitive kit for human IL-5 (MesoScale Discovery, Gaithersburg, MD, USA) following instructions provided by the manufacturer. Details of the high-resolution ultrasound measurements of carotid IMT at baseline and follow-up have been reported [18, 19]. A number of studies suggested that atherosclerotic changes can be site-specific and have distinct clinical implications [20-22]. Thus one composite and four segment-specific measurements of carotid IMT by were included in the study: $IMT_{mean\text{-}max}$ (mean of the maximum IMT measures along the entire carotid tree), $IMT\text{-}CC_{mean}$ and $IMT\text{-}CC_{max}$ (mean and maximum of common carotid segment), $IMT\text{-}Bif_{mean}$ and $IMT\text{-}Bif_{max}$ (mean and maximum of the bifurcation). The ultrasonographic measurements were obtained at baseline and after 15 and 30 months. These three time points were used in linear regression to assess change over time in IMT [18], referred to here as progression of IMT. Characteristics of the IMPROVE cohort with available genetic and phenotypic data (n=3,435) are presented in Table 1.

2.3 Genetic analysis

Genotyping was performed using Illumina 200K Cardio-Metabochip [23]and Immunochip [24] genotyping arrays (hence referred to as the combined chip). Whilst the combined chips do not have genome-wide coverage, they do have good coverage of candidate or robustly associated loci for immune, inflammatory, metabolic and cardiovascular traits, which allows for a less conservative threshold for significance compared to genome-wide platforms.

Genotyping and calling were performed at the SNP Technology Platform (Uppsala University, Sweden Uppsala). Standard genetic quality control procedures were followed with exclusions for SNPs failing call rate (<95%) and Hardy-Weinberg ($p<5E^{-7}$) checks or with minor allele frequency (MAF) <1% and for subjects with cryptic relatedness or sex check failures. After quality control, 3,433 individuals with genetic and phenotypic data were included in the analysis.

2.4 Replication cohorts: the PROCARDIS and MDC cohorts

Replication was sought in the PROCARDIS [25] and MDC cohorts [26, 27].

The PROCARDIS cohort consists of cases and controls from the UK, Sweden, Italy and Germany. All cases were diagnosed with CAD before the age of 66 years and most (80%) had a sibling with CAD before 66 years. Controls were recruited from the same populations and did not have a diagnosis of CAD, or a sibling with diagnosis of CAD, before 66 years of age. Plasma levels of IL-5 were measured in duplicate on the same plate by a sandwich immunoassay with electrochemiluminescence detection using a MesoScale Diagnostics platform (Gaithersburg, Maryland, USA). Genomewide genotyping of PROCARDIS was conducting using the Illumina 1M or 610K platforms. Genotype calling was performed at the Centre National du Genotypage, Paris or the Uppsala SNP Genotyping platform, Sweden. Standard quality control was conducted as detailed above.

The Malmö Diet and Cancer (MDC) study is a prospective population-based cohort study including 28,449 subjects recruited during 1991–1996. Subjects aged 45–69 years, living in Malmö, Sweden were eligible for participation. Between November 1991 and February 1994, every second enrolled subject was also invited to take part in a substudy of the epidemiology of carotid artery disease. This MDC-CardioVascular cohort consists of 6103 subjects (60% women), 5540 of whom also agreed to have blood collected under standardized fasting conditions. MDC was genotyped using the Illumina OmniExpressExome chip and standard quality control was conducted as detailed above. The methodology used for measuring IL-5 levels has previously been described [28]. Briefly, levels of cytokines released by mononuclear cells after stimulation (with CD3/CD28 beads) were analysed using a multiplex cytokine immunoassay (MesoScale Discovery, Gaithersburg, MD, USA).

Whilst the genotyping platforms utilised by the discovery and replication cohorts are not the same, all 3 platforms are produced by the same company and based upon the same bead chip technology. In addition, the same quality control parameters are used for the genotype calling, as well as for the downstream analyses (as detailed above).

2.5 Statistical methods

Differences between two groups according to sex were assessed by Wilcoxon Mann-Whitney ranksum test for continuous variables and Fisher's exact test for binary variables (Table 1). Since a sex difference in IL-5 levels has been observed [11], we adjusted for sex. Risk factors and biomarkers significantly associated with plasma levels of IL-5 (by Spearman rank correlation) were considered for inclusion in regression models. Variables with skewed distributions were log-transformed prior to regression analysis. In both the analyses of IL-5 and of eosinophil count, the basic model was adjusted for population structure (as assessed by the first 3 multi-dimensional scaling components), age and sex. For IL-5 analyses, the extended model also included LDL and hsCRP. For eosinophil count, the extended model included BMI, fasting proinsulin concentration and white blood count. Variables for inclusion in the extended model were determined by backward stepwise selection in a multiple regression model. Linear regression analysis was used to assess associations between SNP genotypes and plasma IL-5 concentrations or eosinophil count, assuming an additive genetic model. The threshold for replication was set at $p < 1.00E^{-5}$, whereas array-wide significance was set at $p < 1.92E^{-7}$ (based on the calculation of the number of uncorrelated SNPs on the combined Metabo-Immuno chips as previously described [29]). Linear regression assuming an additive genetic model was also used for analysis of IMT measures, with adjustment for age, sex and population structure. Further adjustment for established CVD risk markers included age, sex, population structure, BMI, SBP, HDL, TG, hsCRP, current smoking and type 2 diabetes.

Replication analyses were conducted in the same manner as the discovery analyses, but the models were amended for PROCARDIS, where country was used instead of MDS components and case-control status was included as a covariate. Statistical analyses were performed in STATA (Statacorp, Houston, TX, USA) except chip-wide genetic analyses which were conducted using PLINK [30]. Because of high correlation between IMT measurements of carotid segments [19], correction for testing multiple phenotypes was not applied.

2.6 Gene expression analysis

Differential gene expression analysis for SNPs robustly associated with plasma IL-5 concentrations was performed. Firstly, in the Advanced Study of Aortic Pathology (ASAP) [31], where aortic adventitia (n=132), aortic media (n=137) and heart (n=127) tissue samples

were analysed. In brief, tissue samples were collected from patients undergoing elective surgery for aortic aneurysm and/or aortic valve repair. RNA was extracted from biopsy material and hybridised to Affymetrix ST 1.0 exon arrays (Santa Clara, CA, USA). DNA was extracted from whole blood and genotyped on the Illumina 610w-Quad Bead Array (San Diego, CA, USA). SNPs with call rate >95% were used for imputation, and imputed SNPs with quality scores of MACH <0.3 were excluded. The ASAP study was approved by the ethics committee at Karolinska Institutet. Associations between SNP genotypes and gene expression were assessed assuming an additive genetic model. Significance was set at $p<0.0056$ (Bonferroni-correction for 3 SNPs in 3 tissues). Secondly the publicly available MuTHER gene expression data was also interrogated [32]. Of the IL-5 or eosinophil-associated SNPs identified in IMPROVE, 5 were available in the MuTHER dataset, thus $p<0.01$ (Bonferroni-correction for 5 SNPs in one tissue) was set for the level of significance. Finally, the publically available rSNPBase [33] was searched effects of SNPs on gene expression. As the data in this resource is qualitative, no threshold for significance was employed.

3. Results

3.1 Cardiovascular, metabolic, inflammatory or immune-associated SNPs vs IL-5 levels or eosinophil count

In order to assess whether there are SNPs in loci previously associated with immune, inflammatory, metabolic or CVD traits which influence plasma levels of IL-5 or eosinophil count, the combined chip was analysed. When considering IL-5 levels, there were no genetic variants that demonstrated array-wide significant associations (i.e. $p < 1.92E^{-7}$), however 2 loci reached the threshold set for replication ($p<1.00E^{-5}$, Figure 1, Upper panel and Table 2). Further adjustment using the extended model had a negligible effect on the results (Table 2). The chromosome 14 locus consists of a single SNP (Supplementary Figure 1), whilst the chromosome 5 locus (defined by the positions of the first and last SNPs reaching $p<1.00E^{-5}$ in the region) encompasses 79 SNPs extending across and upstream of the *IL5* gene. When considering associations with eosinophil count, one locus, including 14 SNPs, reached the threshold for replication but not array-wide significance (Figure 1, Lower panel and Table 2) in both the basic and extended models (Table 2). This locus is also on chromosome 5 and extends across the *IL5* gene.

3.2 Genetic architecture of IL5 and eosinophil-associated loci

Given that only a single SNP was genotyped in the chromosome 14 locus, it is hard to say anything about the genetic architecture of this locus. More genotyping in this region would be needed to further explore this locus.

The 2 loci on chromosome 5 discovered here are located in physical proximity to each other, however the SNPs in the 2 loci are distinct and appear to be independent from each other by linkage disequilibrium (LD, Figure 2, Upper and Lower panels, Supplementary Figure 2). The 3 genetic markers in the vicinity of the *IL5* gene reported in earlier studies were independent from each other and the loci identified in the current study as assessed by LD (Supplementary Figure 2), however the eosinophil-associated SNP (rs4143832) and the CVD-associated SNP (rs2706399) fell within our eosinophil-associated locus. To further test the independence of the 2 loci, conditional analysis was conducted, whereby the IL-5 analysis was adjusted for the

lead eosinophil-associated SNP and *vice versa*. As demonstrated in Table 3, there is negligible attenuation of the effect of the IL-5-associated SNPs on IL-5 levels when including the lead eosinophil-associated SNP as a covariate. The same is observed if the reverse is attempted, thus we conclude that the loci are independent.

3.3 Replication of IL5-associated loci

Ideally we would have liked to replicate both the IL-5 and eosinophil findings but due to lack of applicable cohorts it was only possible to attempt replication for the IL-5-assocaited SNPs. Characteristics of the replication cohorts are presented in Supplementary Table 1. Replication was attempted in 2 independent European cohorts for the IL-5-associated loci, namely rs56183820 on chromosome 5 and rs4902762 on chromosome 14. The size of effect in the replication meta-analysis (Table 4) was similar to that seen in the discovery analysis, however the direction was inconsistent and the associations were solidly non-significant.

3.4 IL-5 and eosinophil-associated SNPs in relation to IMT

Our aim was to use a Mendelian randomization approach to assess whether the IL-5 or eosinophil-associated loci influenced measures of IMT. As we were not able to replicate the association of the chr 5 locus with IL-5 levels, conducting a Mendelian randomization experiment was unjustified. However, we did assess the effect of these SNPs on IMT measures. Considering 6 independent loci (3 previously reported SNPs, 1 eosinophil-associated locus and 2 IL-5-associated loci), the Bonferroni-corrected significance threshold was set at $p<0.0083$. No associations were observed between the SNPs and baseline measures of IMT (Supplementary Table 2). Eosinophil-associated SNPs demonstrated nominally significant associations with progression of IMT-CC_{max}, but only rs2706399 (the G allele of which was previously associated with increased CVD risk [6]) met the Bonferonni-corrected p for significance (G allele, Beta -0.008, 95%CI (-0.012, -0.004), p = 0.0017, Supplementary Table 3). Further adjustment for emerging or established CVD risk markers had negligible effect on the associations (Supplementary Table 3). Inclusion of eosinophil count in the model rendered the eosinophil-associated SNPs non-significant, but did not have an effect on the assocation of rs2706399 (CVD-associated) with IMT-CC_{max} progression (Supplementary Table 3).

3.5 Expression QTL

To identify functional effects of the IL-5 or eosinophil-associated loci, SNPs which met the threshold for p<1.00E^{-5} were examined for influence on expression of genes at the locus. In the ASAP study, aorta (adventitia and media) and heart tissues were examined. A region 250kb up and downstream of the lead SNP was defined for each locus, and genes within this region were assessed for genotype-specific expression. Bonferroni correction was made for 3 tissues and 4 loci (one on chromosome 14, 3 on chromosome 5, p for significance <0.0042, (Supplementary Figure 2)). The chromosome 14 region included 2 genes, *CTAGE5* and *FBXO33* and rs4902762 was associated with *CTAGE5* expression levels in heart tissue (p=0.0073), but this did not reach statistical significance. The chromosome 5 eosinophil- and IL-5-associated loci overlapped to a large degree, and included many genes (*P4HA2* and *PDLIM4* (eosinophil count locus), *SLC22A4, SLC22A5, IRF1, IL5, RAD50, IL13, IL4, KIF3A* and *SEPT8* (both loci)*, SHROOM1, GDF9, UQCRQ, LEAP2* and *AFF4* (IL-5 locus). Rs12652920 was associated with *RAD50* expression in aortic media (p=0.0093), but this did not reach the level of statistical significance. The MuTHER lymphocyte data [32] supports a positive effect of the rs4705959 T allele (associated with increased eosinophil count) on mRNA expression of *RAD50* and *SLC22A5*, but a negative effect on expression of *IL5* (Table 5). The C allele of rs12652920 (associated with increased IL-5 levels) was also associated with increased mRNA expression of *RAD50* (Table 5). The rSNPBase was queried to identify proxies (r2> 0.8) for the lead SNPs and whether they have been reported to influence expression of genes. For the IL-5-associated locus (rs56183420) 63 proxies were identified, of which 47 reported eQTL effects on *RAD50* but only 24 for *IL5*. Of note, 35 of these SNPs were available in the IMPROVE cohort, all showed associations (p<1.00E^{-4}) with IL-5 levels, but not eosinophil count. For the eosinophil-associated loucs (rs72797327) 4 proxies were identified, of which 2 each were reported for effects on *IL-5* and *RAD50* expression, but all 4 reported effects on *SEPT8* expression. Of these SNPS, all 3 that were available in IMPROVE showed strong (p<1.00E^{-5}) effects on eosinophil but not IL-5 levels. Considering the chromosome 14 locus, 23 proxies were identified with 5 influencing *CTAGE5* expresison and 7 influencing *FBXO33* levels. Only one proxy was available in IMPROVE and was associated with IL-5 levels but not eosinophil count.

4. Discussion

This study identified genetic regulators of plasma IL-5 levels and eosinophil count, however the lack of replication of the genetic associations precludes the ability to formally assess the causality of IL-5 levels on IMT. To the best of our knowledge this is the largest study with measurement of IL-5 levels and the first to investigate genetic regulation of IL-5 levels.

The genetic loci associated with IL-5 levels and eosinophil count are in close proximity to each other and in the vicinity of the *IL5* gene on chromosome 5, but act independently. Earlier studies identified SNPs in the region of the *IL5* and *IL5a* genes as being associated with inflammatory bowel disease (rs2188962 [34]), eosinophil count (rs4143832 [14]) and coronary artery disease (rs2706399 [6]). These variants were present on the combined chip and thus were also analysed. No association was observed between rs2706399 and either IL-5 levels or eosinophil count (Supplementary Table 5). In contrast, rs2188962 demonstrated nominal associations and rs4143832 demonstrated convincing associations with IL-5 levels and eosinophil count (Supplementary Table 5). In agreement with Gudbjartsson et al, the minor allele of rs4143832 was associated with eosinophil count [14] as well as IL-5 levels. Of note, whilst these SNPs demonstrate some significance in relation to levels of IL-5 or eosinophil count, they do not reach the preset level of significance for being forwarded to replication analysis. Rs4143832 and rs2706399 are located within the eosinophil-associated locus reported here but are independent from the SNPs studied here (by LD, Supplementary Table 2).

Whilst SNPs at the loci identifed here do show some associations with *IL5* mRNA levels, expression of the *RAD50* gene demonstrates stronger and more consistent associations across the loci. The results from rSNPBase are less clear, and whilst they do support *RAD50* as a possible candidate, they suggest that *SEPT8* (chromosome 5) and *FBXO33* (chromosome 14)

might also be relevant. Recently, up-regulation of RAD50 and other double-strand repair proteins was reported in vascular smooth muscle cells from atherosclerotic plaque [35]. The authors suggest that these proteins are associated with plaque stability rather than atherosclerosis. Thus it is plausible that the *IL5* locus influences plaque stability via RAD50-associated mechanisms rather than IL-5-associated mechanisms. This speculative mechanism fits with the *IL5* locus being associated with MI [6] and unstable angina [36], typically resulting from plaque rupture, but not with IMT, which is indicative of vessel wall remodelling.

Eosinophil count and IL-5 levels are closely inter-connected: IL-5 is essential for eosinophil maturation and differentiation [37]. Conversely, eosinophil-specific granules store IL-5 [38], indicating that these cells might also contribute to increasing circulating IL-5 levels. In IMPROVE, there is a strong correlation between IL-5 concentrations and eosinophil count (Spearmans rho=0.560, $p<0.001$) [11]. Thus it is interesting that these 2 variables are regulated by independent loci in the same genetic region. One possibility is that a common factor, potentially RAD50, regulates both IL-5 levels and eosinophil count. An alternative scenario is that, given the roles of IL-5 and eosinophils in inflammation, upregulation of *RAD50* might be a DNA-protective stress response mechanism. This would explain the puzzling finding that the allele associated with increased eosinophil count shows increased *RAD50* and the inverse effect on *IL5* gene expression, whilst the allele associated with increased IL-5 levels also shows increased *RAD50* gene expression. It is of note that 2 of the top eQTL SNPs for *RAD50* [39] are those with the strongest associations with eosinophils (rs11739623 and rs4705959) in IMPROVE. Previous reports have demonstrated the use of studying specific immune cell types to explore immune or inflammation-related diseases [40, 41]. Whilst the relationship between the SNPs, *RAD50* expression, IL-5 levels and eosinphil counts requires clarification, these findings are of potential relevance for eosinophil-associated disorders.

Whilst this is the largest study with detailed IMT measures, plasma IL-5 concentrations and extensive genetic data for all subjects, there are limitations to the study. This study was conducted in subjects with at least 3 classical CVD risk factors, thus findings might not pertain to other populations. A sex difference in IL-5 levels has been reported [11], thus sex-stratified analyses would be of interest. However, this approach has not been used here, for two reasons: the distribution of outliers (≥1.5 interquartile range above the upper quartile)

between men and women might explain the sex difference in levels. Because of this and the loss of power when stratifying the analyses, we believe that the chance of detecting false positive signals is too high for this analysis to be robust. The lack of Th2 count or measurements of Th2 cytokines mean that we cannot conclusively determine the origin of IL-5 and whether Th2 responses are activated in these subjects. However, it has previously been shown that the percentage of Th2 cells has only a very weak (non significant) correlation with IL-5 levels released by activated leukocytes [28]. The genotyping chips employed in this study are based on a large scale candidate gene/loci experiment, which allows a less stringent significance threshold for the discovery analysis, however the whole genome is not represented. Thus we cannot comprehensibly describe the genetic architecture of the IL-5-associated locus on chromosome 14, nor can we exclude the possibility that true IL-5-influencing genetic variants exist which lie within areas of the genome not examined in this study.

5. Conclusions

This study identified genetic loci on chromosome 5 associated with IL-5 levels and eosinophil count. Whilst these loci are close together in the vicinity of the *IL5* gene, they are independent. Furthermore, expression analysis suggests *RAD50*, rather than *IL5*, as a prime candidate gene for these associations. An additional locus associated with IL-5 levels was observed on chromosome 14. This report suggests that these loci have negligible influence on IMT measures.

6. Acknowledgements

The IMPROVE study was supported by the European Commission (LSHM-CT- 2007-037273), the Swedish Heart-Lung Foundation, the Swedish Research Council (8311, 8691), the Knut and Alice Wallenberg Foundation, the Foundation for Strategic Research, the Torsten and Ragnar Söderberg Foundation, the Strategic Cardiovascular Programme of Karolinska Institutet and the Stockholm County Council and the Stockholm County Council (560183). PROCARDIS was supported by Supported by the British Heart Foundation, the European Community Sixth Framework Program (LSHM-CT-2007-037273), AstraZeneca, the Wellcome Trust, the United Kingdom Medical Research Council, the Swedish Heart–Lung Foundation, the Swedish Medical Research Council, the Knut and Alice Wallenberg

Foundation, and the Karolinska Institutet. MDC was supported by grants from the Swedish Research Council, the Swedish Heart–Lung foundation, the Knut and Alice Wallenberg foundation, VINNOVA, Swedish Foundation for Strategic Research. KG acknowledges support from the Swedish Heart-Lung Foundation and Stiftelsen för Gamla Tjänarinnor. BS acknowledge support from the Magnus Bergvall Foundation and the Foundation for Old Servants. SEH is a British Heart Foundation Professor and is supported by the British Heart Foundation (RG008/08) and by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. JCH acknowledges support from the British Heart Foundation (FS/14/55/30806). RJS is supported by the SRP Diabetes program at Karolinska Institutet.

7. References

[1] Mortality GBD, Causes of Death C. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015;385:117-71.
[2] Kannel WB. Overview of hemostatic factors involved in atherosclerotic cardiovascular disease. Lipids. 2005;40:1215-20.
[3] Ridker PM. C-reactive protein and the prediction of cardiovascular events among those at intermediate risk: moving an inflammatory hypothesis toward consensus. Journal of the American College of Cardiology. 2007;49:2129-38.
[4] Ranjbaran H, Sokol SI, Gallo A, Eid RE, Iakimov AO, D'Alessio A, et al. An inflammatory pathway of IFN-gamma production in coronary atherosclerosis. Journal of immunology. 2007;178:592-604.
[5] Signorelli SS, Fiore V, Malaponte G. Inflammation and peripheral arterial disease: the value of circulating biomarkers (Review). International journal of molecular medicine. 2014;33:777-83.
[6] Consortium IKC. Large-scale gene-centric analysis identifies novel variants for coronary artery disease. PLoS genetics. 2011;7:e1002260.
[7] Ingelsson E, Yin L, Back M. Nationwide cohort study of the leukotriene receptor antagonist montelukast and incident or recurrent cardiovascular disease. The Journal of allergy and clinical immunology. 2012;129:702-7 e2.
[8] Binder CJ, Hartvigsen K, Chang MK, Miller M, Broide D, Palinski W, et al. IL-5 links adaptive and natural immunity specific for epitopes of oxidized LDL and protects from atherosclerosis. The Journal of clinical investigation. 2004;114:427-37.
[9] Sampi M, Ukkola O, Paivansalo M, Kesaniemi YA, Binder CJ, Horkko S. Plasma interleukin-5 levels are related to antibodies binding to oxidized low-density lipoprotein and to decreased subclinical atherosclerosis. Journal of the American College of Cardiology. 2008;52:1370-8.
[10] Taleb S, Tedgui A, Mallat Z. Adaptive T cell immune responses and atherogenesis. Current opinion in pharmacology. 2010;10:197-202.

[11] Silveira A, McLeod O, **Strawbridge RJ**, Gertow K, Sennblad B, Baldassarre D, et al. Plasma IL-5 concentration and subclinical carotid atherosclerosis. Atherosclerosis. 2015;239:125-30.
[12] Ikutani M, Yanagibashi T, Ogasawara M, Tsuneyama K, Yamamoto S, Hattori Y, et al. Identification of innate IL-5-producing cells and their role in lung eosinophil regulation and antitumor immunity. Journal of immunology. 2012;188:703-13.
[13] Kouro T, Takatsu K. IL-5- and eosinophil-mediated inflammation: from discovery to therapy. International immunology. 2009;21:1303-9.
[14] Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. Nature genetics. 2009;41:342-7.
[15] Asquith KL, Ramshaw HS, Hansbro PM, Beagley KW, Lopez AF, Foster PS. The IL-3/IL-5/GM-CSF common receptor plays a pivotal role in the regulation of Th2 immunity and allergic airway inflammation. Journal of immunology. 2008;180:1199-206.
[16] Kitamura T, Sato N, Arai K, Miyajima A. Expression cloning of the human IL-3 receptor cDNA reveals a shared beta subunit for the human IL-3 and GM-CSF receptors. Cell. 1991;66:1165-74.
[17] Tavernier J, Devos R, Cornelis S, Tuypens T, Van der Heyden J, Fiers W, et al. A human high affinity interleukin-5 receptor (IL5R) is composed of an IL5-specific alpha chain and a beta chain shared with the receptor for GM-CSF. Cell. 1991;66:1175-84.
[18] Baldassarre D, Hamsten A, Veglia F, de Faire U, Humphries SE, Smit AJ, et al. Measurements of carotid intima-media thickness and of interadventitia common carotid diameter improve prediction of cardiovascular events: results of the IMPROVE (Carotid Intima Media Thickness [IMT] and IMT-Progression as Predictors of Vascular Events in a High Risk European Population) study. Journal of the American College of Cardiology. 2012;60:1489-99.
[19] Baldassarre D, Nyyssonen K, Rauramaa R, de Faire U, Hamsten A, Smit AJ, et al. Cross-sectional analysis of baseline data to identify the major determinants of carotid intima-media thickness in a European population: the IMPROVE study. European heart journal. 2010;31:614-22.
[20] Ebrahim S, Papacosta O, Whincup P, Wannamethee G, Walker M, Nicolaides AN, et al. Carotid plaque, intima media thickness, cardiovascular risk factors, and prevalent cardiovascular disease in men and women: the British Regional Heart Study. Stroke; a journal of cerebral circulation. 1999;30:841-50.
[21] Finn AV, Kolodgie FD, Nakano M, Virmani R. The differences between neovascularization of chronic total occlusion and intraplaque angiogenesis. JACC Cardiovascular imaging. 2010;3:806-10.
[22] Johnsen SH, Mathiesen EB, Joakimsen O, Stensland E, Wilsgaard T, Lochen ML, et al. Carotid atherosclerosis is a stronger predictor of myocardial infarction in women than in men: a 6-year follow-up study of 6226 persons: the Tromso Study. Stroke; a journal of cerebral circulation. 2007;38:2873-80.
[23] Voight BF, Kang HM, Ding J, Palmer CD, Sidore C, Chines PS, et al. The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS genetics. 2012;8:e1002793.
[24] Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. Nature genetics. 2011;43:1193-201.
[25] Farrall M, Green FR, Peden JF, Olsson PG, Clarke R, Hellenius ML, et al. Genome-wide mapping of susceptibility to coronary artery disease identifies a novel replicated locus on chromosome 17. PLoS genetics. 2006;2:e72.
[26] Berglund G, Elmstahl S, Janzon L, Larsson SA. The Malmo Diet and Cancer Study. Design and feasibility. Journal of internal medicine. 1993;233:45-51.
[27] Hedblad B, Nilsson P, Janzon L, Berglund G. Relation between insulin resistance and carotid intima-media thickness and stenosis in non-diabetic subjects. Results from a cross-sectional study in Malmo, Sweden. Diabetic medicine : a journal of the British Diabetic Association. 2000;17:299-307.

[28] Engelbertsen D, Andersson L, Ljungcrantz I, Wigren M, Hedblad B, Nilsson J, et al. T-helper 2 immunity is associated with reduced risk of myocardial infarction and stroke. Arteriosclerosis, thrombosis, and vascular biology. 2013;33:637-44.
[29] Gertow K, Sennblad B, **Strawbridge RJ**, Ohrvik J, Zabaneh D, Shah S, et al. Identification of the BCAR1-CFDP1-TMEM170A Locus as a Determinant of Carotid Intima-Media Thickness and Coronary Artery Disease Risk. Circulation-Cardiovascular Genetics. 2012;5:656-65.
[30] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics. 2007;81:559-75.
[31] Folkersen L, van't Hooft F, Chernogubova E, Agardh HE, Hansson GK, Hedin U, et al. Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease. Circulation Cardiovascular genetics. 2010;3:365-73.
[32] Grundberg E, Small KS, Hedman AK, Nica AC, Buil A, Keildson S, et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nature genetics. 2012;44:1084-9.
[33] Guo L, Du Y, Chang S, Zhang K, Wang J. rSNPBase: a database for curated regulatory SNPs. Nucleic acids research. 2014;42:D1033-9.
[34] Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature. 2012;491:119-24.
[35] Gray KL, Kumar SV, Figg N, Harrison J, Baker L, Mercer JR, et al. Effects of DNA Damage in Smooth Muscle Cells in Atherosclerosis. Circulation research. 2014.
[36] Avramakis G, Papadimitraki E, Papakonstandinou D, Liakou K, Zidianakis M, Dermitzakis A, et al. Platelets and white blood cell subpopulations among patients with myocardial infarction and unstable angina. Platelets. 2007;18:16-23.
[37] Blanchard C, Rothenberg ME. Biology of the eosinophil. Advances in immunology. 2009;101:81-121.
[38] Melo RC, Spencer LA, Dvorak AM, Weller PF. Mechanisms of eosinophil secretion: large vesiculotubular carriers mediate transport and release of granule-derived cytokines and other proteins. Journal of leukocyte biology. 2008;83:229-36.
[39] Westra H-J, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nature genetics. 2013;45:1238-43.
[40] Orru V, Steri M, Sole G, Sidore C, Virdis F, Dei M, et al. Genetic variants regulating immune cell levels in health and disease. Cell. 2013;155:242-56.
[41] Roederer M, Quaye L, Mangino M, Beddall MH, Mahnke Y, Chattopadhyay P, et al. The genetic architecture of the human immune system: a bioresource for autoimmunity and disease pathogenesis. Cell. 2015;161:387-403.

Figure Legends:

Figure 1: Manhattan plots of associations between SNPs on the combined chip with A) plasma IL-5 levels and B) eosinophil count. Associations were adjusted for age, sex and population stratification. The horizontal line at -log10 p value = 5 indicates the threshold for replication.

Figure 2: Regional plot of SNPs in the chromosome 5 loci which are associated with A) circulating IL-5 concentrations and B) eosinophil count. The purple diamond indicates rs12652920 (the lead SNP in the IL-5-associated locus), and thus demonstrates that the lead SNP in the eosinophil-associated locus (rs57896917) is independent by linkage disequilibrium.

Supplementary Figure 1. Regional plot for a locus on chromosome 14 associated with plasma IL-5 concentrations.

Supplementary Figure 2. Linkage disequilibrium (LD) between candidate SNPs and those associated with IL5 levels or eosinophil count in the IMPROVE study (at $p<1.00E^{-5}$). Block 1 consists of eosinophil-associated SNPs and blocks 2 and 3 are IL5-associated SNPs. The candidate SNP rs2188962 (Irritable bowel disease-associated [34]) falls outside of the 2 novel loci described in this report, whilst the eosinophil count and CVD-associated SNPs (rs4143832 [14] and rs2706399 [6] respectively) are located within the novel eosinophil count-associated locus. Colours and values reflect R2 measures of LD.