



Tobacco smoking is associated with sex- and plaque-type specific upregulation of *CRLF1* in atherosclerotic lesions

Tian Lan^{a,b}, Kaylin C.A. Palm^b, Luka Hoeben^a, Ernest Diez Benavente^a, R. Noah Perry^{c,d}, Mete Civelek^{c,d}, Dominique P.V. de Kleijn^e, Hester M. den Ruijter^a, Gerard Pasterkamp^b, Michal Mokry^{a,b,*}

^a Laboratory of Experimental Cardiology, Department of Cardiology, University Medical Center Utrecht, University Utrecht, Utrecht, the Netherlands

^b Central Diagnostics Laboratory, University Medical Center Utrecht, University Utrecht, Utrecht, the Netherlands

^c Center for Public Health Genomics, University of Virginia, Charlottesville, USA

^d Department of Biomedical Engineering, University of Virginia, Charlottesville, USA

^e Department of Vascular Surgery, University Medical Center Utrecht, Utrecht, the Netherlands

ABSTRACT

Background and aims: Tobacco smoking is a known risk factor for atherosclerotic disease, with more elevated risks in women compared to men. We hypothesized that atherosclerotic plaques from smokers show different gene expression patterns compared to non-smokers, in a sex-specific manner.

Methods: Gene expression data of 625 carotid plaques (151 females and 474 males) were analyzed for differential gene expression between current smokers (n = 226) and non-smokers (n = 399). All analyses were stratified by sex and by molecular plaque characteristics. Finally, we projected the activity of gene regulatory networks and utilized single-cell transcriptomics from 38 plaques (26 males and 12 females) to interpret the sex- and plaque-type specific signals.

Results: We observed higher expression levels of *CRLF1* gene in atherosclerotic plaques from smokers compared to non-smokers (log2FC = 0.48, FDR = 0.012). *CRLF1* upregulation was interacting with sex ($p = 0.01$) and was more pronounced in females (log2FC = 0.93, $p = 1.53E-05$) compared to males (log2FC = 0.35, $p = 0.0018$). Through single-cell RNA-seq analysis, we identified the highest *CRLF1* expression within the transitioning and synthetic smooth muscle cell populations. *CRLF1* expression was increased in fibro-inflammatory and fibro-cellular plaque types. Gene annotations pointed to increased expression of *CRLF1* in networks with extracellular matrix related genes.

Conclusions: Atherosclerotic plaques from current smokers show sex-dependent upregulation of smooth muscle cell gene *CRLF1*. This may explain the different contributions of smoking to cardiovascular risk in females.

1. Introduction

Smoking is a leading preventable risk factor for atherosclerosis and cardiovascular disease (CVD), which are major causes of death worldwide [1]. Exposure to thousands of chemicals found in cigarette smoke can trigger the formation of plaques in the arteries by increased oxidative stress, platelet activation, and inflammation [2]. In addition, smoking can accelerate the recruitment of inflammatory cells, degradation of matrix metalloproteinases, and intraplaque hemorrhage, which can destabilise atherosclerotic plaque [3].

Tobacco smoking differentially affects men and women, with females being more susceptible to harmful outcomes, according to epidemiological studies [4–6]. For example, a systematic review including 75 cohort studies showed smoking increases the risk for coronary heart disease more in female smokers compared to male smokers

[7]. Additionally, the amount of consumption of cigarettes may lead to differential mortality risks, with women at a higher risk of death with lower levels of smoking compared to men [8]. Sex hormones, such as estradiol and progesterone, may contribute to women's higher nicotine addiction and difficulty quitting smoking [9,10]. The mechanisms involved in the sex-specific effects of smoking-related elevated risks for CVD remain unclear.

Numerous studies have established that tobacco smoking is associated with changes in DNA methylation and gene expression. For instance, it is well-known that smoking leads to DNA hypomethylation at the aryl hydrocarbon receptor repressor (AHHR) and F2R like thrombin or trypsin receptor loci in blood samples and atherosclerotic plaques possibly contributing to cardiovascular disease [11–15]. Previously, we demonstrated that the CpG sites mapped to the *CRLF1* gene are significantly changed in current smokers of carotid plaques [16].

* Corresponding author. UMC Utrecht, Heidelberglaan 100, 3584CX, Utrecht, the Netherlands.

E-mail address: M.Mokry@umcutrecht.nl (M. Mokry).

<https://doi.org/10.1016/j.atherosclerosis.2024.118554>

Received 9 November 2023; Received in revised form 18 June 2024; Accepted 6 August 2024

Available online 7 August 2024

0021-9150/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

However, our understanding of sex-specific smoking-related gene expression changes in these lesions remains limited.

Here, we examined the differential gene expression between smokers and non-smokers in carotid plaques. Our analysis revealed that the *CRLF1* gene exhibited significantly differential expression between the two groups. Female smokers showed almost three times higher upregulation of *CRLF1* in atherosclerotic plaques compared to male smokers. Interestingly, the *CRLF1* gene is mostly expressed in ACTA2+ smooth muscle cells, and the upregulation of *CRLF1* was also affected by molecular plaque types. These results provide evidence for the sex-specific effect of smoking on gene expression in atherosclerotic plaques and help to improve our understanding of sex specific differences in mechanisms that lead to cardiovascular diseases.

2. Materials and methods

2.1. Study design and subject selection

The Athero-Express Biobank Study (AE-biobank) is an ongoing longitudinal biobank that started in 2002. It involves patients undergoing carotid endarterectomy (CEA) at two Dutch hospitals. This study collects extensive baseline data, blood samples, and plaque specimens from patients, detailed in prior publications. The criteria for CEA in the study align with established clinical trials and are assessed by a multidisciplinary vascular team. Experienced surgeons perform plaque removal following standardized protocols, with pre-operative neurological assessments conducted by neurologists. The cohort study design has been previously described in detail [17–19]. For the present study, subsequent patients who underwent carotid endarterectomy “CEA” between 2002 and 2016 and for whom transcriptomic and clinical data were available. The baseline characteristics were obtained from patients’ records and through standardized questionnaires. This questionnaire contains questions on the history of cardiovascular disease, risk factors (smoking habits, hypertension, diabetes mellitus), and use of medication. Current tobacco smoking (i.e., including hand-rolled cigarettes, cigars, etc.) was defined as smoking within 1 year prior to admission for CEA and was assessed by questionnaire. We estimated the number of pack years smoking based on a categorical question regarding the number of cigarettes smoked and defined the “estimated Pack Years Smoking” = (number of cigarettes smoked per day x number of years smoked)/20; where 1 pack is defined as 20 cigarettes. Symptom categories were ‘asymptomatic’, defined as not having ipsilateral cerebrovascular symptoms in the previous 6 months; ‘ocular’—amaurosis fugax, defined as ipsilateral mono-ocular blindness of acute onset lasting <24 h; cerebral ‘transient ischemic attack’, defined as the ipsilateral focal neurologic deficit of acute onset lasting <24 h; and ipsilateral ‘stroke’. Total cholesterol, triglycerides, high-density lipoprotein cholesterol, glucose, hemoglobin and creatinine were assessed before operation. This study complies with the Declaration of Helsinki and all participants provided informed consent (University Medical Center, Utrecht, The Netherlands, and St. Antonius Hospital, Nieuwegein, The Netherlands). The medical ethical committees of the respective approved these studies.

2.2. Sample handling

Carotid plaque specimens were removed during surgery and immediately processed in the laboratory. An experienced technician identified the culprit lesion, which is defined as the segment with the smallest lumen. Samples were cut transversely into segments of 5 mm. The culprit lesion (the region with the most severe stenosis) was identified, fixed in 4 % formaldehyde, embedded in paraffin, and processed for histological examination. Plaque histological features were routinely scored through chemical- and immunohistochemical techniques as described below. The rest of the segments were stored at –80 °C.

2.3. Histology assessment

The assessment was performed according to a standardized protocol, which was described in detail previously [20]. Cross-sections of the culprit lesion are stained and quantified for each patient at × 40 magnification. Hematoxylin and eosin (H&E) staining was performed for the assessment of calcification, plaque hemorrhage and picrosirius red for collagen. Plaque hemorrhage was confirmed by stainings for fibrin and/or glycophorin A. Immunohistochemical staining was performed for assessment of macrophages (CD68), and smooth-muscle cells (alpha-actin). The amount of collagen, calcifications, and plaque hemorrhage were classified as minor or major. The lipid content of the plaque was estimated as a percentage of the total plaque area, with a cutoff at 10 % and 40 % for carotid plaques. Plaques with <10 %, 10–40 % and >40 % fat were categorized as fibrous, fibro-atheromatous and atheromatous, respectively. The criteria for classified Smooth Muscle Cells (SMCs) were defined as follows: (1) no or few scattered cells; (2) minor alpha-actin staining over the entire circumference with absent staining at parts of the circumference of the arterial wall; (3) moderate: positive cells along the circumference of the luminal border, with locally at least few scattering cells; (4) heavy: SMC-dominant plaque with cells within the entire cap and also large clusters deep in the lesion. All histological characteristics were scored and quantified with dedicated intra-observer and inter-observer reproducibility by two independent observers [21]. Associations of current tobacco smoking with pathological features were determined by general linear modelling GLM() in R.

2.4. Bulk RNA sequencing and differential gene expression analysis

A total of 700 atherosclerotic samples were selected in this study. We isolated RNA using in-house standardized protocols. Then we employed the CEL-seq2 method for library preparation. The methodology captures the 3’ end of polyadenylated RNA species and includes unique molecular identifiers (UMIs), which allow direct counting of unique RNA molecules in each sample. Plaque sample processing for RNA isolation and library preparation is described in detail previously [22]. Libraries were sequenced on the Illumina NextSeq 500 platform; a high output paired-end run of 2 × 75 bp was performed (Utrecht Sequencing Facility). The reads were demultiplexed and aligned to human cDNA reference (Ensembl 84) using the Burrows–Wheeler Aligner (BWA) (0.7.13) by calling ‘bwa aln’ with settings -B 6 -q 0 -n 0.00 -k 2 -l 200 -t 6 for R1 and -B 0 -q 0 -n 0.04 -k 2 -l 200 -t 6 for R2, ‘bwa sampe’ with settings -n 100 -N 100. Multiple reads mapping to the same gene with the same UMI (6 bp long) were counted as a single read. The raw read counts were corrected for UMI sampling (corrected_count = $-4096 * (\ln(1 - (\text{raw_count}/4096)))$), normalized for sequencing depth and quantile normalized. We discarded samples (n = 46) with fewer than 9000 detected genes from further analysis. We used DESeq2 [23] version 1.38.2 in R to examine the differences in gene expression related to smoking status, and smoking amount, sex differences, histological features, preoperative clinical symptoms, secondary events, and molecular characteristics of carotid plaques. The interaction analyses and likelihood ratio tests within DESeq2 were used to understand the relationships among these variables. Additionally, we used the Pearson correlation coefficient [cor()] function in R to assess the relationship between *CRLF1* gene expression and expression of other genes. To reduce the effects of confounding due to non-random assignment in our observational study, we employed propensity score adjustment. Propensity scores were calculated using logistic regression, modeling the probability of treatment assignment conditional on observed baseline covariates. The covariates included in the model were Age, Sex, History of Coronary Artery Disease (CAD), Hypertension, Body Mass Index (BMI), Diabetes Mellitus (DM), and Glomerular Filtration Rate estimated (eGFR) by the MDRD study equation.

2.5. Mass spectrometry proteomics analysis

As a part of AE biobank, samples are 200 human carotid plaque. Proteins were extracted using sequential incubation with 0.5 M sodium chloride (NaCl), 0.1 % sodium dodecyl sulfate (SDS), and 4 M guanidine hydrochloride (GuHCl).¹⁰ All extracts were labeled using tandem mass tags and analyzed on an Orbitrap Fusion Lumos Tribrid MS for proteomics (Thermo Scientific). A parallel reaction monitoring method was developed on a Q Exactive HF MS (Thermo Scientific). Protein abundances were normalized based on total protein per sample and scaled using log2 transformation for relative protein quantities. The data was then filtered to only include protein data with less than 30 % missing values and any remaining missing values within the data were imputed using KNN-Impute method, described in detail previously [24]. In the end 1499 proteins was detected from 200 plaque lesions. Linear regression was employed to analyze the association between CRLF1 protein and tobacco smoke.

2.6. Identifying the CRLF1 gene features in single-cell sequencing mapping plaque

The cell population of human atherosclerotic carotid plaque from two scRNA-seq datasets which have been previously generated and published by Slenders et al. and Pan et al. [25,26]. The data can be assessed on the open-source single-cell portal for research website (<http://plaqview.uvadcos.io/>). The segment procedures and data analysis for single-cell sequencing are described in detail in its original publications.

2.7. Gene regulatory networks (GRNs)

To investigate the gene regulatory networks (GRNs) associated with the CRLF1 gene, we utilized the Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task database (STARNET) database (starnet.mssm.edu) [27]. The counted reads from the RNA sequencing

were normalized using counts per million (CPM), and the trimmed mean of M values (TMM) normalized factors. To compare the network activity between different plaque types, we 1) calculated the module score as a proxy to network expression using normalized counts. 2) Gene connectivity was measured as the average pairwise correlation between gene expression values using log2-transformed normalized counts. The multiple comparisons of module score and connectivity in each plaque type by using One-way ANOVA and Tukey's range test after ANOVA test in R.

2.8. Statistical analysis

Statistics were performed using the R statistical software. Data normality was determined using the Shapiro-Wilk tests. For comparison of 2 groups of continuous variables with normal distribution and equal variances, 2-tailed unpaired Student *t* tests were performed with a confidence level of 95 %. For comparison of 2 groups of continuous variables with normal distribution and unequal variances, 2-tailed unpaired Student *t* tests followed by a Welch correction were performed with a confidence level of 95 %. Two-tailed unpaired Mann-Whitney *U* test with a confidence level of 95 % was conducted if data were non-normally distributed.

2.9. Pathway analysis

We performed pathway enrichment analysis by using the "ReactomePA_1.42.0" and "clusterProfiler_4.6.2" packages [28,29].

3. Results

3.1. Study population

In this study, we analyzed a bulk transcriptomic dataset obtained from atherosclerotic plaques of 625 patients from the AE-biobank. A

Table 1
Clinical characteristics of smokers and non-smokers.

Characteristics	Smoker (n = 226)	Non-smoker (n = 399)	<i>p</i>
Age, years (IQR)	65.3(12)	71.1 (13)	<0.001
Males (%)	158(69.9)	316 (79.1)	0.012
SBP, mm Hg (IQR)	154(36)	151 (32)	0.894
DBP, mm Hg (IQR)	80(16)	80 (17)	0.668
eGFR, ml/min Per 1.73 m ² (SD)	74.79(1.0)	71.18 (1.4)	0.040
BMI, kg/m ² (IQR)	25.9(4.9)	26.2 (4.2)	0.080
Pack years (IQR)	10.0(19.3)	0.5 (0.9)	<0.001
Comorbidities			
Diabetes mellitus (%)	28(7.0)	57(13.9)	0.557
Hypertension (%)	181(80.0)	361(90.4)	<0.001
Medication use			
Hypertensive drugs ^a (%)	153(67.6)	329(82.4)	<0.001
Anticoagulants (%)	21(9.2)	55(13.7)	0.127
Antiplatelet drugs (%)	204(90.2)	354(88.7)	0.642
LLDs (%)	6(2.6)	24(6.0)	0.090
Atherosclerotic phenotypes			0.08
Fibrous (%)	84(37.2)	114(31.8)	
Atheromatous (%)	59(26.2)	129(30.7)	
Fibro-atheromatous (%)	82(36.6)	150(37.5)	
Symptoms			0.07
Asymptomatic (%)	29(13.1)	66(16.8)	
Ocular (%)	50(22.6)	57(14.5)	
TIA (%)	90(40.7)	173(44.1)	
Stroke (%)	52(23.5)	96(24.4)	

p values were calculated by using the chi-square test for categorical variables, the T-test for normally distributed variables, and the Mann-Whitney *U* test for non-normally distributed variables. The distributions were significantly non-normal for the variables age, SBP, DBP, BMI, and eGFR according to Shapiro-Wilk tests. Smokers represent current smokers, while non-smokers represent previous or never smokers. SD, standard deviation; IQR, interquartile range; SBP, systolic blood pressure; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; BMI indicates body mass index; LLDs, use of lipid-lowering drugs and TIA, transient ischemic attack.

^a Use of one or more antihypertension drugs.

comparison of baseline clinical characteristics, risk factors, and medication was conducted between smokers (226 current smokers) and non-smokers (399 former or never smokers) (Table 1). Our findings indicated that current smokers presented a favourable clinical profiling and demonstrated a trend towards more histological stable plaques when contrasted with non-smokers. This pattern was further substantiated by data from 3939 subjects in AE-biobank with smoking status (Supplementary Table 1). From a histo-pathological perspective, we noted that current smokers had higher odds of calcification (odds ratio [95 % CI]: 1.22 [1.04, 1.43], $p = 0.01$) and collagen presence (odds ratio [95 % CI]: 1.44 [1.16, 1.79], $p < 0.001$) within carotid plaques. Conversely, they had lower odds of fat content (odds ratio [95 % CI]: 0.74 [0.61, 0.90], $p = 0.003$) (Supplementary Table 2).

Upon stratifying clinical profiles by sex, we found significant

differences in clinical symptoms between female smokers and non-smokers ($p = 0.02$). This difference was not present in males ($p = 0.56$). Specifically, female never smokers showed a higher incidence of severe events, with TIA occurring in 46.9 % compared to 32.8 % in smokers, and stroke occurring in 27.1 % compared to 20.8 % in smokers, as shown in Supplementary Table 3.

3.2. Tobacco smoking is associated with differential gene expression of *CRLF1* in carotid plaques

We performed a differential gene expression analysis [23] of carotid atherosclerotic plaques of the smoker group and non-smokers group. We found that the *CRLF1* gene was the only differentially expressed gene ($\log_2FC = 0.50$, $lfcSE = 0.10$, $FDR = 0.008$) in smokers, after adjusting

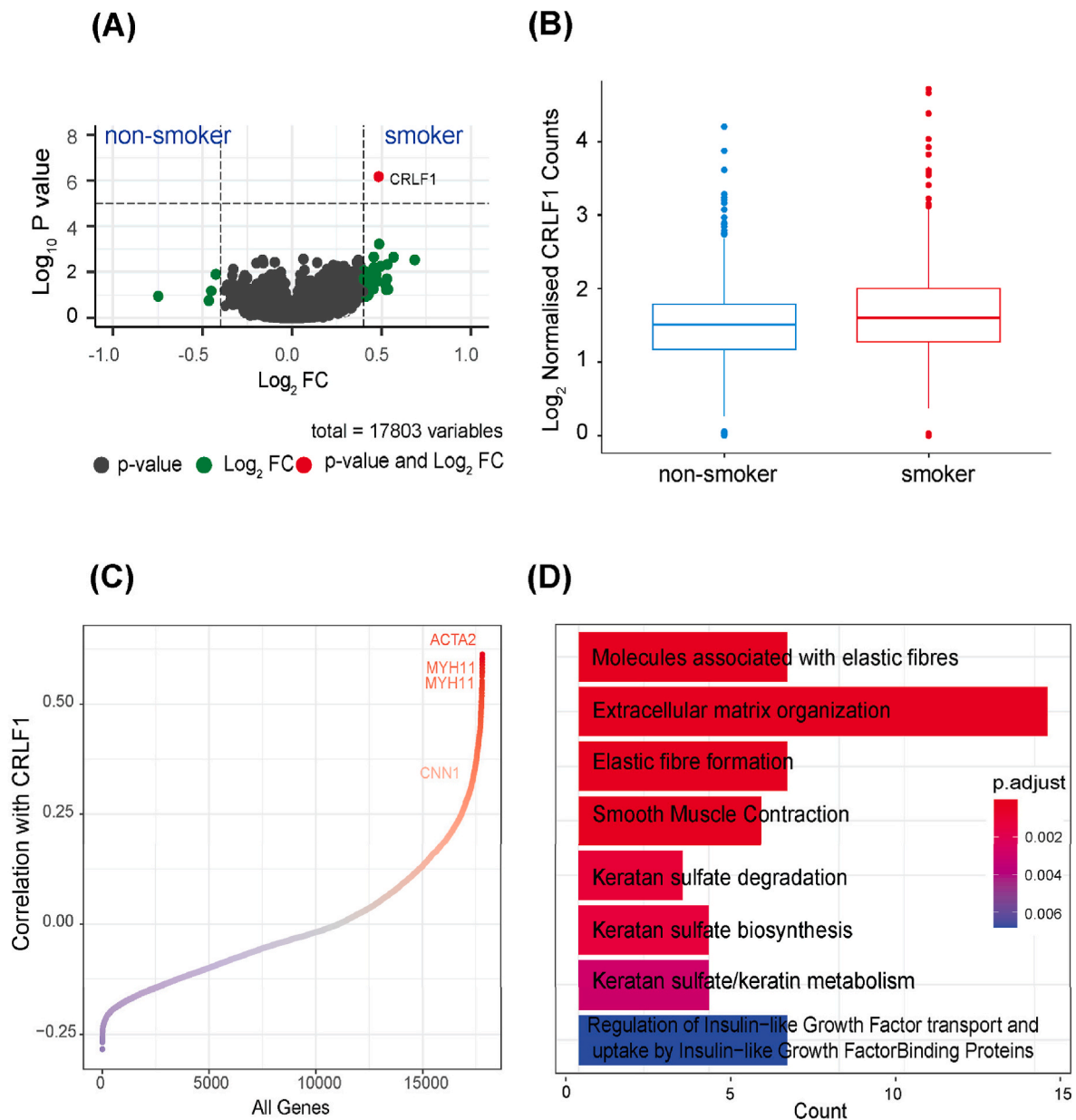


Fig. 1. The association of *CRLF1* gene expression with tobacco smoking in carotid plaque.

(A) Volcano plot of differentially expressed genes between smokers and non-smokers. The X-axis denotes \log_2 Fold Change ($\log_2 FC$) values while the Y-axis shows $-\log_{10}$ p-values. The $\log_2 FC$ and $-\log_{10}$ p-value thresholds are set at 0.4 and 5, respectively. (B) Boxplot showing the *CRLF1* gene expression levels in smokers and non-smokers based on normalized unique molecular identifiers (UMIs) counts. (C) Pairwise Pearson correlation of *CRLF1* gene expression with every other gene. (D) Pathway enrichment analysis of 212 genes, which are the most positively correlated with the *CRLF1* gene (calculated correlation coefficients >0.4). The top 10 enriched pathways are shown by name.

for sex and age (Fig. 1A and B). The observed upregulation remained significant after adjusting for clinical parameters associated with smoking, such as eGFR, hypertension, and hypertensive drug use ($\log_2FC = 0.48$, $lfcSE = 0.09$, $FDR = 0.012$) and after correcting for traditional risk factors for atherosclerosis ($\log_2FC = 0.56$, $lfcSE = 0.09$, $FDR = 0.0003$) (Supplementary Fig. 1). Proteomics analysis also confirmed higher levels of *CRLF1* protein in the plaques of smokers ($p = 0.02$) (Supplementary Fig. 2). We also observed a significant association

between the expression of the *CRLF1* gene and the estimated smoking pack-year (Supplementary Table 4).

In our investigation of the relationship between *CRLF1* gene expression and various histological characteristics, preoperative clinical symptoms, and subsequent cardiovascular events, we found a significant overexpression of *CRLF1* in the presence of macrophages ($\log Fold\text{-}Change = -0.29$, $p = 0.001$). Although we observed a trend towards downregulation of *CRLF1* expression in fat content, plaque hemorrhage,

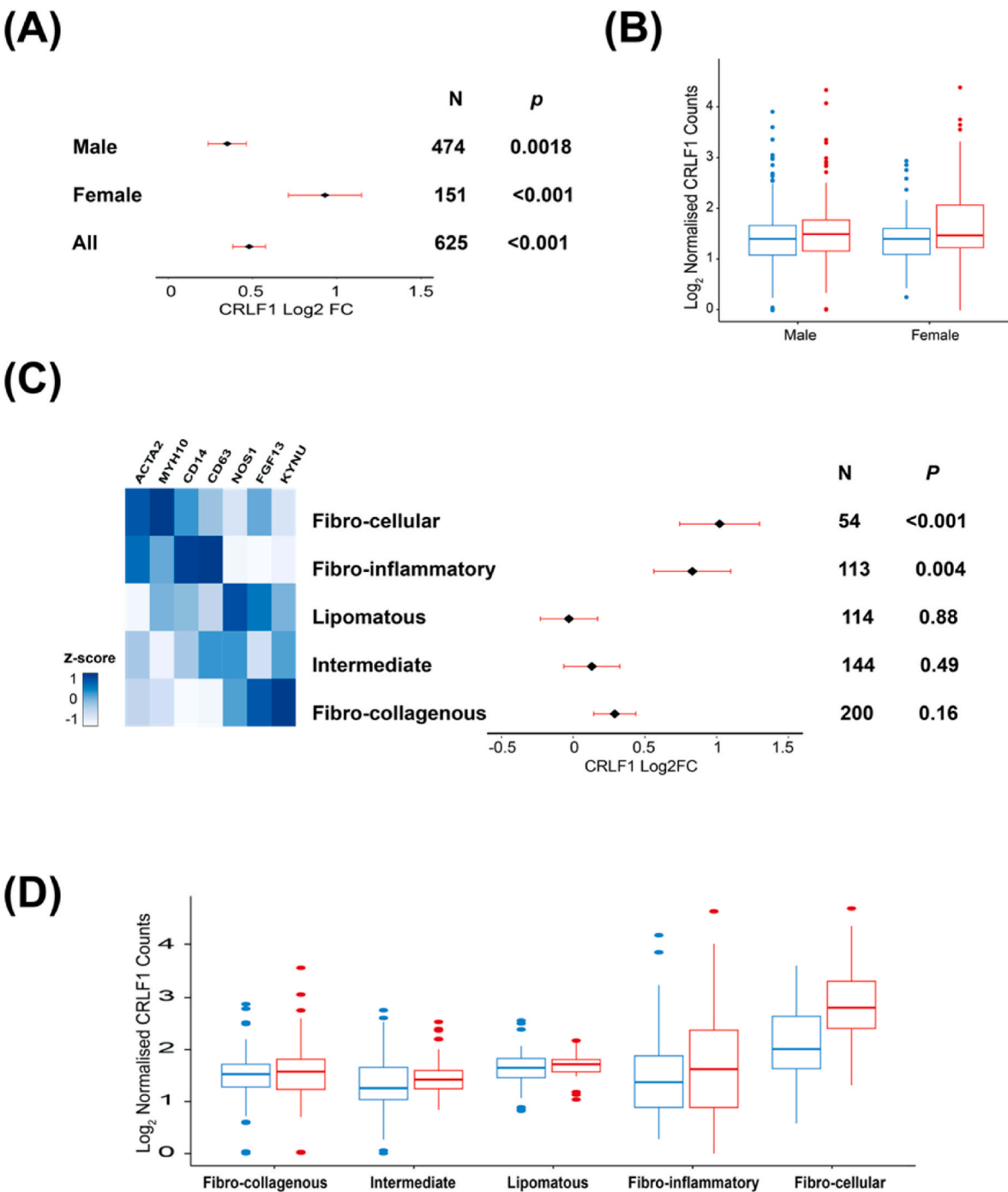


Fig. 2. The smoking related *CRLF1* gene expression stratified for sex and plaque. (A) and (C) The “forest” plot, presenting the Log2 Fold Changes (Log2 FC) of the expression of the *CRLF1* gene in smokers within the entire cohort and each subgroup. The P represents the nominal p-value associated with the Log2 FC of the *CRLF1* gene in smokers compared to non-smokers. The red line denotes the Log2 FC \pm standard error (SE). The heatmap shows the mean expression of selected marker genes in transcriptomic-defined clusters of carotid plaque by Mokry et al. NCVR 2022. (B) and (D) Each boxplot shows the normalized expression level of the *CRLF1* gene between smokers (red) and non-smokers (blue) in males and females, and across molecular clusters. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

and major adverse cardiovascular events (MACE), and upregulation related to calcification, smooth muscle cells (SMCs), and collagen content, these changes did not reach statistical significance. Further details are provided in [Supplementary Table 5](#). Furthermore, *CRLF1* expression was positively correlated with smooth muscle cell marker genes. These co-expressed genes are involved with the extracellular matrix and smooth muscle activity, as shown by Gene Ontology (GO) term enrichments ([Fig. 1C and D](#)). Gene set enrichment analysis (GSEA) identified

top significant pathways ([Supplementary Figs. 3A and 3B](#)).

3.3. The impact of sex on *CRLF1* upregulation in smokers

To investigate if the effect of tobacco smoking was similar in males and females, we analyzed the smoking-associated *CRLF1* gene expression changes in a sex-stratified manner. We found that the increase in *CRLF1* expression in smokers is higher in females ($\text{Log2FC} = 0.85, \text{IcfSE}$

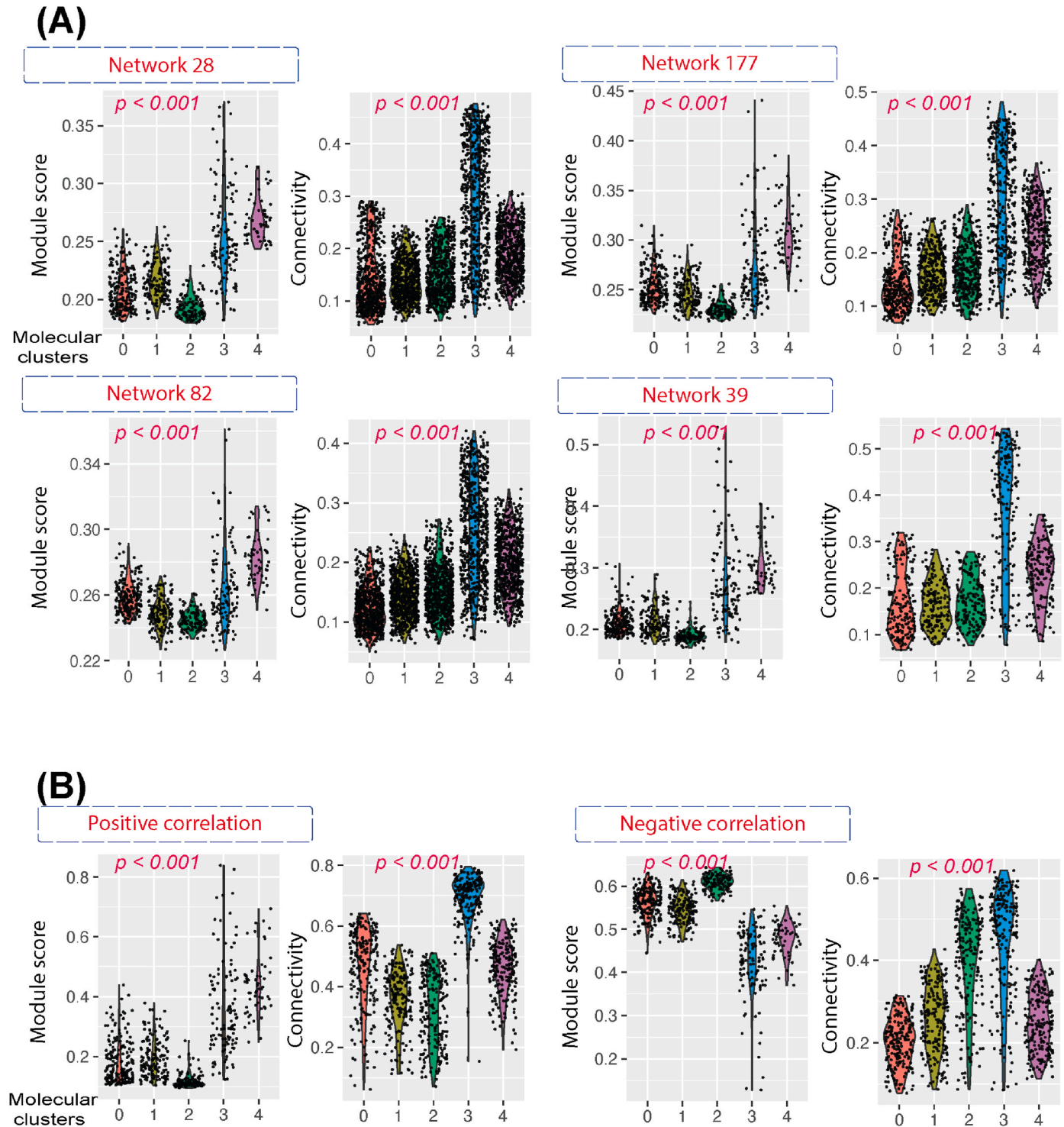


Fig. 3. The overall expression (module score) and connectivity of genes from regulatory networks (GRNs) in different molecular plaque types. (A) Module scores and connectivity of genes annotated to *CRLF1* associated GRNs. (B) Module scores and Connectivity of genes with positive and negative correlation with *CRLF1* expression in different plaque types.

= 0.18, $p < 0.001$) compared to male smokers ($\text{Log2FC} = 0.39$, $\text{lfcSE} = 0.10$, $p = 0.001$). GSEA results revealed that the extra cellular matrix pathway was enriched in both female and male smokers; and the cellular response to lipids was enriched in male smokers (Supplementary Figs. 4A and 4B). After adjustment for confounders, the association remained significant only in female smokers (Supplementary Table 6). The association between *CRLF1* expression and smoking quantity (packyears) was significant in female smokers ($\text{Log2FC} = 0.47$, $p < 0.001$), and not present in males ($\text{Log2FC} = 0.03$, $p = 0.55$) (Supplementary Table 7).

3.4. Smoking-related *CRLF1* upregulation depends on the molecular phenotype of carotid plaque

Atherosclerotic plaque can vastly differ in the activity of different molecular pathways. We delved into whether molecular composition within the plaques influenced smoking induced *CRLF1* expression. To study this, we classified plaques into the five transcriptomic-defined [22] plaque types: Fibro-collagenous (plaque type 0), Intermediate (plaque type 1), Lipomatous (plaque type 2), Fibro-inflammatory (plaque type 3) and Fibro-cellular (plaque type 4), as illustrated by the expression of various marker genes (Fig. 2C). These plaque types differ in the expression of genes involved in numerous molecular mechanisms, including extracellular matrix organization, immune response, and metabolism. This study found no statistically significant difference in smokers' presence among different plaque types ($p = 0.35$).

Building upon these findings, our analysis revealed that the influence of smoking on the *CRLF1* gene expression is statistically significant ($p = 0.003$) across different molecular plaque types, as determined by the likelihood ratio test. The pronounced upregulation of *CRLF1* is observed in fibro-inflammatory and fibro cellular plaques ($\text{Log2FC} = 0.61$, $\text{lfcSE} = 0.268$, $p = 0.022$, and $\text{Log2FC} = 1.03$, $\text{lfcSE} = 0.278$, $p < 0.001$, respectively) in smokers compared to non-smokers. These plaque types are characterized by higher expression of smooth muscle cell marker genes, such as *MYH11*, *MYH10*, and *ACTA2* (Fig. 2C and D).

We investigated the activity of gene regulatory networks (GRNs) associated with the *CRLF1* gene. We first utilized the STARNET database [27] to identify the GRNs enriched for genes that functionally annotated with the *CRLF1* (Fig. 1C). We identified 30 significantly enriched (FDR < 0.05) GRN modules from all network modules (see Supplementary

Table 8). The fibro inflammatory and fibro-cellular plaque types displayed significantly increased expression and connectivity (ANOVA $p < 0.001$) in the top four enriched GRNs: GRN28, GRN82, GRN177, and GRN39, as shown in Fig. 3A. Remarkably, these GRNs (GRN28, GRN82, and GRN39) were found to be active in biological processes implicated in atherosclerosis development, such as extracellular matrix organization and vasculature development. Additionally, we observed higher expression of genes positively correlated to *CRLF1* in the fibro-inflammatory and fibro-cellular plaque types, while the negatively correlated genes had higher average expression in plaque the other plaque types (Fig. 3B). These findings suggest that the specific upregulation of the *CRLF1* gene may be mediated via the activity of these GRNs.

3.5. *CRLF1* gene expression in plaques is cell-type-specific

To understand the source of *CRLF1* expression in atherosclerotic plaques, we have utilized single-cell transcriptomics datasets derived from 38 patients from the AE biobank cohort [25]. Among the 20 distinct cell populations derived presently in carotid plaques, the *CRLF1* gene was predominantly expressed in transitioning and synthetic smooth muscle cells. In contrast, this gene was minimally expressed in contractile smooth muscle cells. Further pathway analysis suggested that SMCs with elevated *CRLF1* expression were involved in biological processes such as extracellular matrix organization and coagulation (Fig. 4A and B). This observation was confirmed in a publicly available dataset [26] where *CRLF1* gene was primarily overexpressed in smooth muscle cells (SMCs) and their transdifferentiated cell types, like SMC-derived intermediate cell states (ICS) and fibrochondrocytes (Fig. 4C).

To investigate the potential function of the *CRLF1* gene in SMCs, we analyzed the expression levels of *CRLF1* genes in relation to 12 phenotypes and assessed a large collection of primary SMCs derived from ascending aorta [30]. Our results showed a negative correlation between *CRLF1* gene expression and the calcification phenotype of SMCs (Pearson correlation coefficient = -0.18 , $p = 0.04$).

4. Discussion

Our study aimed to investigate the impact of tobacco smoking on gene expression in atherosclerotic carotid lesions. Our findings indicate

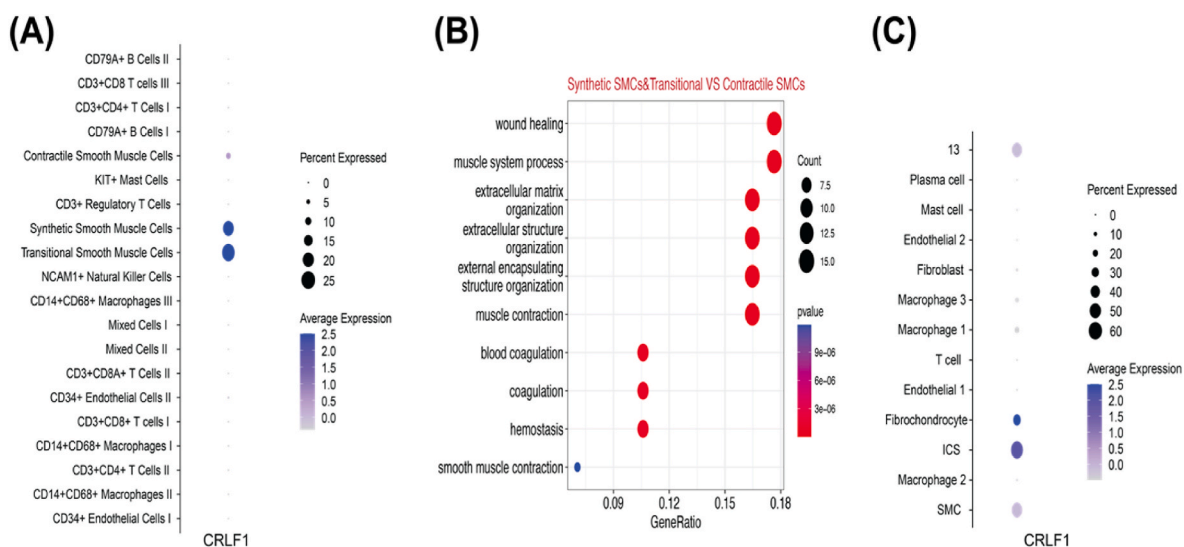


Fig. 4. *CRLF1* expression pattern in single-cell transcriptomic cell populations.

(A) Dot plots showing expression of *CRLF1* gene in cell populations identified in the scRNA-seq of 38 atherosclerotic plaques of carotid endarterectomies donors, previously generated by Slenders et al. (B) Top10 pathways upregulated in *CRLF1* in contractile smooth muscle cells (SMCs) (cells with low *CRLF1* expression) and synthetic & translational SMCs (cells with high *CRLF1* expression). (C) Dot plots showing expression of *CRLF1* gene in different cell types identified in the scRNA-seq of 3 atherosclerotic plaques from carotid endarterectomies, generated by Pan et al.

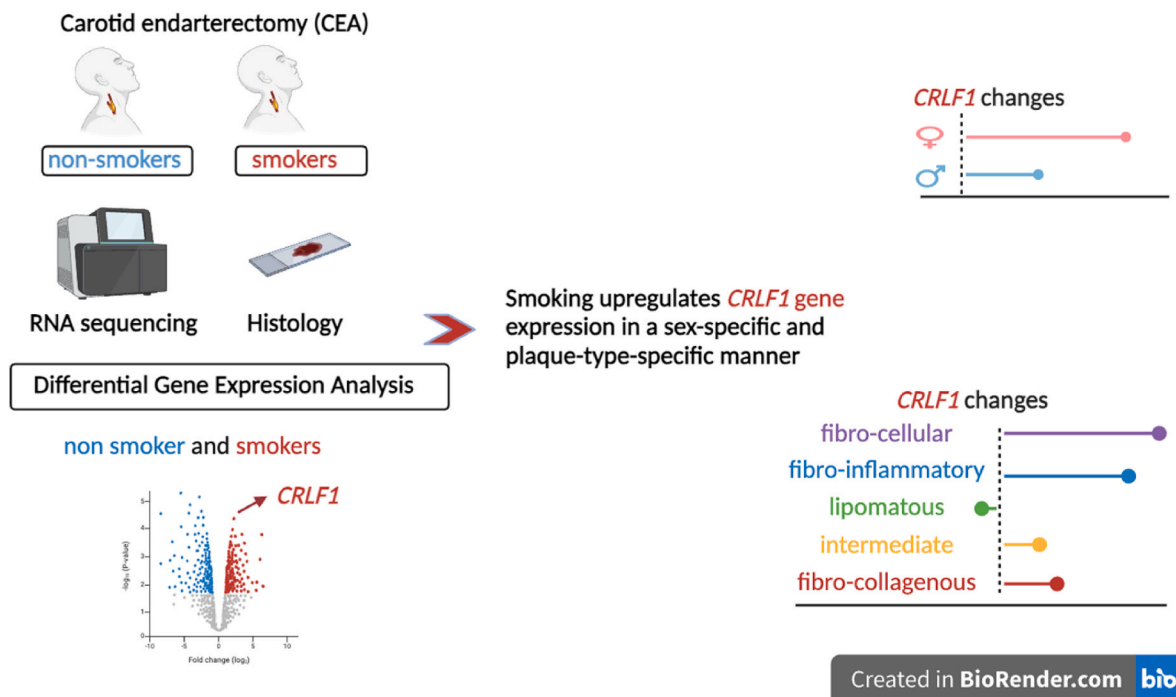


Fig. 5. Graphical abstract.

a statistically significant relationship between smoking status and *CRLF1* gene expression. Furthermore, we observed that female smokers had higher smoker-associated upregulation of the *CRLF1* gene compared to male smokers. Additionally, smoking-induced upregulation of *CRLF1* expression was observed in plaque types with a higher expression of smooth muscle cell markers and in plaques with high levels of ACTA2+ staining. Furthermore, our study observed that the genes positively correlated with *CRLF1* expression in smokers were enriched in biological functions, such as mitochondrial protein-containing complexes and extracellular matrix organization (Fig. 5).

These results are consistent with previous research indicating a dichotomy in biological responses to smoking between male and female smokers [31,32]. For example, a genome-wide expression study showed that sex hormones (progesterone and estrogen) display a major interaction with some of the smoking-responsive immune and inflammatory genes, which suggests those hormones play a potential role in response to smoking, and oncogenic pathway gene sets were significantly more active in female smokers compared to male smokers based on gene expression profiling [33]. Additionally, these sex-related differences in response to smoking were observed in the airway epithelium as well, which showed that female smokers exhibit higher autophagy and lower viral response compared to male smokers [34].

In this study, we observed that smoking status is associated with overexpression of the *CRLF1* gene and CRLF1 protein in human carotid atherosclerotic lesions. We also observed that the *CRLF1* gene is highly overexpressed in ACTA2+ cells based on single-cell transcriptomic analyses of carotid lesions [25,26,35]. *CRLF1* encodes cytokine receptor-like factor-1, a member of the cytokine type I receptor family, is involved in various organs in development and disease. Mutations in *CRLF1* have been associated with Crisponi/cold-induced sweating syndrome (CS/CISS) [36–38]. Many studies attempted to elucidate the bio-mechanism induced by smoking through the analysis of DNAm and gene expression of the impact of smoking in multiple tissues. One previous study showed that the methylation status of CpG sites in carotid plaques mapped to the *CRLF1* gene is associated with smoking [16] and that the *CRLF1* gene is upregulated in adipose tissue of smokers [39]. Brendan et al. identified *CRLF1* as a potential mediator of oxidative stress resistance through a comparative analysis of gene expression

between undifferentiated and differentiated neuroblastoma cell modules after 6-OHDA treatment [40]. Our investigation into the potential role of *CRLF1* gene expression in atherosclerotic lesions revealed that *CRLF1* expression is higher in regions rich in calcification and may influence calcification processes in vascular smooth muscle cells (SMCs). Existing studies [41] suggest that osteochondrogenic changes in SMCs can be initiated by exosomes from synthetic SMCs with elevated *CRLF1* levels. These findings indicate a complex interplay in which *CRLF1* may modulate vascular calcification, potentially acting as a protective feedback mechanism against smoking-induced vascular damage.

The transcriptome profiling improved our understanding of mechanisms that accelerate atherosclerosis [42–44]. In our previous study, we identified five molecular types of carotid plaques in advanced atherosclerosis patients [22]. Building on this work, we investigated the impact of smoking on the expression of the *CRLF1* gene in these molecularly defined plaque clusters. Our findings revealed that molecular plaque types modulate the overexpression of the smoking related *CRLF1* gene in atherosclerotic lesions. Notably, we observed an upregulation of *CRLF1* in Fibro-inflammatory and Fibro-cellular plaques, types that show gene expression that points to smooth muscle cell presence. This upregulation could potentially be orchestrated by gene regulatory networks (GRNs) linked to *CRLF1*. Our results demonstrated that both the expression levels of the *CRLF1* gene and its connectivity to other genes are significantly higher in these two molecular plaque types when compared to other clusters. This observation underscores the pivotal association of molecular plaque-specific characteristics and *CRLF1* expression, particularly in smokers.

Heterogeneous smooth muscle cell (SMC) populations in atherosclerotic plaque contribute to plaque formation and lesion development [45,46]. Single-cell transcriptomics revealed the overexpression of the *CRLF1* gene in ACTA2+ cells, primarily composed of synthetic SMCs within the carotid plaque. This suggests a possible association between *CRLF1* and changes in SMC phenotype. The higher expression of classical SMC gene markers in plaque type 3 and 4 potentially explains the plaque type-specific upregulation of *CRLF1*, given its prominent expression in ACTA2+ cells. Similarly, the relevance of the overexpression of the *CRLF1* gene in female smokers can be explained by previous studies indicating higher activity of SMC-related GRNs in

females compared to males and highlights the importance of phenotypically modulated SMCs in female atherosclerosis [47,48]. However, further research is required to elucidate the precise role of *CRLF1* in SMC in response to smoking.

It is important to note that our study was limited by several factors. First, our cohort consisted of patients with advanced atherosclerotic disease, so our findings may need to be confirmed in earlier stages of the disease. This patient selection may also explain the inverse associations with risk factors and symptoms, which may seem counterintuitive. Additionally, gene expression and regulation are complex and dynamic processes, and our results should be interpreted with caution. Although we accounted for the time-dependent effects of smoking on the plaque, residual confounding or misclassification due to self-reported smoking history cannot be entirely excluded. Finally, we had no validation cohort with extensive phenotypic and exposure information. Further studies are necessary to confirm and expand upon our findings.

Overall, our study presents insights into the transcriptomic responses to tobacco smoking in 625 carotid lesions. Our findings indicate that current smoking is significantly associated with upregulation of *CRLF1* in atherosclerotic plaques. The upregulation of the *CRLF1* gene in the carotid plaque of smokers is sex-specific and is influenced by plaque composition and molecular type.

CRedit authorship contribution statement

Tian Lan: Formal analysis, interpreted data, Writing – original draft. **Kaylin C.A. Palm:** Formal analysis. **Luka Hoebe:** Formal analysis. **Ernest Diez Benavente:** Conceptualization. **R. Noah Perry:** Investigation. **Mete Civelek:** Investigation. **Dominique P.V. de Kleijn:** Conceptualization, All authors have read and agreed to the published version of the manuscript. **Hester M. den Ruijter:** Conceptualization, Supervision. **Gerard Pasterkamp:** Conceptualization, Supervision. **Michal Mokry:** interpreted data, Conceptualization, Supervision, the finalization of the article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Tian Lan is supported by a Chinese Scholarship Council (CSC) fellowship program (NO. 202208500038). We are thankful for the support of the DDutch Heart Foundation, project TRANS (project 01-003-2022-0417 2022), VAR (TKI-2003), Leducq Fondation 'PlaQomics' (18CDV-02), EU H2020 TO_AITION (grant number: 848146), EU HORIZON NextGen (grant number: 101136962), and EU HORIZON MIRACLE (grant number: 101115381).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.atherosclerosis.2024.118554>.

References

- [1] T. Kondo, Y. Nakano, S. Adachi, T. Murohara, Effects of tobacco smoking on cardiovascular disease, *Circ. J.* 83 (10) (2019) 1980–1985, <https://doi.org/10.1253/circj.CJ-19-0323>.
- [2] T. Münzel, O. Hahad, M. Kuntic, J.F. Keaney, J.E. Deanfield, A. Daiber, Effects of tobacco cigarettes, e-cigarettes, and waterpipe smoking on endothelial function and clinical outcomes, *Eur. Heart J.* 41 (41) (2020) 4057–4070, <https://doi.org/10.1093/eurheartj/ehaa460>.
- [3] A. Csordas, D. Bernhard, The biology behind the atherothrombotic effects of cigarette smoke, *Nat. Rev. Cardiol.* 10 (4) (2013) 219–230, <https://doi.org/10.1038/nrcardio.2013.8>.
- [4] J. Mackay, A. Amos, Women and tobacco, *Respirology* 8 (2) (2003) 123–130, <https://doi.org/10.1046/j.1440-1843.2003.00464.x>.
- [5] International Early Lung Cancer Action Program Investigators, C.I. Henschke, R. Yip, O.S. Miettinen, Women's susceptibility to tobacco carcinogens and survival after diagnosis of lung cancer [published correction appears in JAMA.2008 Apr 16; 299(15):1775], *JAMA* 296 (2) (2006) 180–184, <https://doi.org/10.1001/jama.296.2.180>.
- [6] C.I. Henschke, O.S. Miettinen, Women's susceptibility to tobacco carcinogens, *Lung Cancer* 43 (1) (2004) 1–5, <https://doi.org/10.1016/j.lungcan.2003.08.024>.
- [7] R.R. Huxley, M. Woodward, Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies, *Lancet* 378 (9799) (2011) 1297–1305, [https://doi.org/10.1016/S0140-6736\(11\)60781-2](https://doi.org/10.1016/S0140-6736(11)60781-2).
- [8] M.A. Hurley, Light smoking at base-line predicts a higher mortality risk to women than to men; evidence from a cohort with long follow-up, *BMC Publ. Health* 14 (1) (2014) 95, <https://doi.org/10.1186/1471-2458-14-95>.
- [9] E. Jassem, The many faces of tobacco use among women, *Med. Sci. Mon. Int. Med. J. Exp. Clin. Res.* 20 (2014) 153–162, <https://doi.org/10.12659/MSM.889796>.
- [10] S. Pogun, G. Yazarbas, T. Nesil, L. Kanit, Sex differences in nicotine preference, *J. Neurosci. Res.* 95 (1–2) (2017) 148–162, <https://doi.org/10.1002/jnr.23858>.
- [11] L.J. Corbin, S.J. White, A.E. Taylor, et al., Epigenetic regulation of *F2RL3* associates with myocardial infarction and platelet function, *Circ. Res.* 130 (3) (2022) 384–400, <https://doi.org/10.1161/CIRCRESAHA.121.318836>.
- [12] V. Barcelona, Y. Huang, K. Brown, et al., Novel DNA methylation sites associated with cigarette smoking among African Americans, *Epigenetics* 14 (4) (2019) 383–391, <https://doi.org/10.1080/15592294.2019.1588683>.
- [13] L.M. Reynolds, M. Wan, J. Ding, et al., DNA methylation of the aryl hydrocarbon receptor repressor associations with cigarette smoking and subclinical atherosclerosis, *Circ Cardiovasc Genet* 8 (5) (2015) 707–716, <https://doi.org/10.1161/CIRCGENETICS.115.001097>.
- [14] L.P. Breitling, K. Salzmann, D. Rothenbacher, B. Burwinkel, H. Brenner, Smoking, *F2RL3* methylation, and prognosis in stable coronary heart disease, *Eur. Heart J.* 33 (22) (2012) 2841–2848, <https://doi.org/10.1093/eurheartj/ehs091>.
- [15] C. Cao, J. Chen, C. Lyu, et al., Bioinformatics analysis of the effects of tobacco smoke on gene expression, *PLoS One* 10 (12) (2015) e0143377, <https://doi.org/10.1371/journal.pone.0143377>.
- [16] M.A. Siemielink, S.W. van der Laan, S. Haitjema, et al., Smoking is associated to DNA methylation in atherosclerotic carotid lesions, *Circ Genom Precis Med* 11 (9) (2018) e002030, <https://doi.org/10.1161/CIRCGEN.117.002030>.
- [17] W.E. Hellings, Atherosclerotic plaque composition and occurrence of restenosis after carotid endarterectomy, *JAMA* 299 (5) (2008) 547, <https://doi.org/10.1001/jama.299.5.547>.
- [18] B.A. Verhoeven, E. Velema, A.H. Schoneveld, et al., Athero-express: differential atherosclerotic plaque expression of mRNA and protein in relation to cardiovascular events and patient characteristics. Rationale and design, *Eur. J. Epidemiol.* 19 (12) (2004) 1127–1133, <https://doi.org/10.1007/s10564-004-2304-6>.
- [19] W.E. Hellings, W. Peeters, F.L. Moll, et al., Composition of carotid atherosclerotic plaque is associated with cardiovascular outcome: a prognostic study, *Circulation* 121 (17) (2010) 1941–1950, <https://doi.org/10.1161/CIRCULATIONAHA.109.887497>.
- [20] G.W. Van Lammeren, H.M. Den Ruijter, J.E.P. Vrijenhoek, et al., Time-dependent changes in atherosclerotic plaque composition in patients undergoing carotid surgery, *Circulation* 129 (22) (2014) 2269–2276, <https://doi.org/10.1161/CIRCULATIONAHA.113.007603>.
- [21] C.C. Laurie, K.F. Doheny, D.B. Mirel, et al., Quality control and quality assurance in genotypic data for genome-wide association studies, *Genet. Epidemiol.* 34 (6) (2010) 591–602, <https://doi.org/10.1002/gepi.20516>.
- [22] M. Mokry, A. Boltjes, L. Slenders, et al., Transcriptomic-based clustering of human atherosclerotic plaques identifies subgroups with different underlying biology and clinical presentation, *Nat Cardiovasc Res* 1 (12) (2022) 1140–1155, <https://doi.org/10.1038/s44161-022-00171-0>.
- [23] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.* 15 (12) (2014) 550, <https://doi.org/10.1186/s13059-014-0550-8>.
- [24] K. Theofilatos, S. Stojkovic, M. Hasman, et al., Proteomic atlas of atherosclerosis: the contribution of proteoglycans to sex differences, plaque phenotypes, and outcomes, *Circ. Res.* 133 (7) (2023) 542–558, <https://doi.org/10.1161/CIRCRESAHA.123.322590>.
- [25] L. Slenders, L.P.L. Landsmeer, K. Cui, et al., Intersecting single-cell transcriptomics and genome-wide association studies identifies crucial cell populations and candidate genes for atherosclerosis, *Eur Heart J Open* 2 (1) (2022) oeab043, <https://doi.org/10.1093/ehjopen/oeab043>.
- [26] H. Pan, C. Xue, B.J. Auerbach, et al., Single-cell genomics reveals a novel cell state during smooth muscle cell phenotypic switching and potential therapeutic targets for atherosclerosis in mouse and human, *Circulation* 142 (21) (2020) 2060–2075, <https://doi.org/10.1161/CIRCULATIONAHA.120.048378>.
- [27] S. Koplev, M. Seldin, K. Sukhvasi, et al., A mechanistic framework for cardiometabolic and coronary artery diseases, *Nat Cardiovasc Res* 1 (1) (2022) 85–100, <https://doi.org/10.1038/s44161-021-00009-1>.
- [28] G. Yu, Q.Y. He, ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization, *Mol. Biosyst.* 12 (2) (2016) 477–479, <https://doi.org/10.1039/C5MB00663E>.
- [29] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R Package for comparing biological themes among gene clusters, *OMICS A J. Integr. Biol.* 16 (5) (2012) 284–287, <https://doi.org/10.1089/omi.2011.0118>.

- [30] R. Aherrahrou, L. Guo, V.P. Nagraj, et al., Genetic regulation of atherosclerosis-relevant phenotypes in human vascular smooth muscle cells, *Circ. Res.* 127 (12) (2020) 1552–1565, <https://doi.org/10.1161/CIRCRESAHA.120.317415>.
- [31] H.K. Koo, J. Morrow, P. Kachroo, et al., Sex-specific associations with DNA methylation in lung tissue demonstrate smoking interactions, *Epigenetics* 16 (6) (2021) 692–703, <https://doi.org/10.1080/15592294.2020.1819662>.
- [32] C. Bolego, A. Poli, R. Paoletti, Smoking and gender, *Cardiovasc. Res.* 53 (3) (2002) 568–576, [https://doi.org/10.1016/s0008-6363\(01\)00520-x](https://doi.org/10.1016/s0008-6363(01)00520-x).
- [33] S. Paul, Differential effect of active smoking on gene expression in male and female smokers, *J. Carcinog. Mutagen.* 5 (6) (2014), <https://doi.org/10.4172/2157-2518.1000198>.
- [34] C.X. Yang, H. Shi, I. Ding, et al., Widespread sexual dimorphism in the transcriptome of human airway epithelium in response to smoking, *Sci. Rep.* 9 (1) (2019) 17600, <https://doi.org/10.1038/s41598-019-54051-y>.
- [35] M.A.C. Depuydt, K.H.M. Prange, L. Slenders, et al., Microanatomy of the human atherosclerotic plaque by single-cell transcriptomics, *Circ. Res.* 127 (11) (2020) 1437–1455, <https://doi.org/10.1161/CIRCRESAHA.120.316770>.
- [36] L. Crisponi, G. Crisponi, A. Meloni, et al., Crisponi syndrome is caused by mutations in the CRLF1 gene and is allelic to cold-induced sweating syndrome type 1, *Am. J. Hum. Genet.* 80 (5) (2007) 971–981, <https://doi.org/10.1086/516843>.
- [37] L. Crisponi, I. Buers, F. Rutsch, CRLF1 and CLCF1 in development, health and disease, *Int. J. Mol. Sci.* 23 (2) (2022) 992, <https://doi.org/10.3390/ijms23020992>.
- [38] N. Dagoneau, S. Bellais, P. Blanchet, et al., Mutations in cytokine receptor-like factor 1 (CRLF1) account for both Crisponi and cold-induced sweating syndromes, *Am. J. Hum. Genet.* 80 (5) (2007) 966–970, <https://doi.org/10.1086/513608>.
- [39] P.C. Tsai, C.A. Glastonbury, M.N. Eliot, et al., Smoking induces coordinated DNA methylation and gene expression changes in adipose tissue with consequences for metabolic health, *Clin. Epigenet.* 10 (1) (2018) 126, <https://doi.org/10.1186/s13148-018-0558-0>.
- [40] B.D. Looyenga, J. Resau, J.P. MacKeigan, Cytokine receptor-like factor 1 (CRLF1) protects against 6-hydroxydopamine toxicity independent of the gp130/JAK signaling pathway, *PLoS One* 8 (6) (2013) e66548, <https://doi.org/10.1371/journal.pone.0066548>.
- [41] U. Förstermann, N. Xia, H. Li, Roles of vascular oxidative stress and nitric oxide in the pathogenesis of atherosclerosis, *Circ. Res.* 120 (4) (2017) 713–735, <https://doi.org/10.1161/CIRCRESAHA.116.309326>.
- [42] P. Libby, J.E. Buring, L. Badimon, et al., Atherosclerosis, *Nat. Rev. Dis. Prim.* 5 (1) (2019) 56, <https://doi.org/10.1038/s41572-019-0106-z>.
- [43] D.T. Miller, P.M. Ridker, P. Libby, D.J. Kwiatkowski, Atherosclerosis, *J. Am. Coll. Cardiol.* 49 (15) (2007) 1589–1599, <https://doi.org/10.1016/j.jacc.2006.12.045>.
- [44] J.L.M. Björkegren, A.J. Lusis, Atherosclerosis: recent developments, *Cell* 185 (10) (2022) 1630–1645, <https://doi.org/10.1016/j.cell.2022.04.004>.
- [45] K. Pakh, C. Joung, S.M. Jung, et al., Visualization of synthetic vascular smooth muscle cells in atherosclerotic carotid rat arteries by F-18 FDG PET, *Sci. Rep.* 7 (1) (2017) 6989, <https://doi.org/10.1038/s41598-017-07073-3>.
- [46] M.R. Bennett, S. Sinha, G.K. Owens, Vascular smooth muscle cells in atherosclerosis, *Circ. Res.* 118 (4) (2016) 692–702, <https://doi.org/10.1161/CIRCRESAHA.115.306361>.
- [47] R.J.G. Hartman, K. Owsiany, L. Ma, et al., Sex-stratified gene regulatory networks reveal female key driver genes of atherosclerosis involved in smooth muscle cell phenotype switching, *Circulation* 143 (7) (2021) 713–726, <https://doi.org/10.1161/CIRCULATIONAHA.120.051231>.
- [48] E. Diez Benavente, S. Karnewar, M. Buono, et al., Female gene networks are expressed in myofibroblast-like smooth muscle cells in vulnerable atherosclerotic plaques, *Arterioscler. Thromb. Biol.* 43 (10) (2023) 1836–1850, <https://doi.org/10.1161/ATVBAHA.123.319325>.