Human Genetics
https://doi.org/10.1007/s00439-024-02674-1

ORIGINAL INVESTIGATION

lncRNA CDKN2B-AS1 regulates collagen expression

Weiwei Shi[1] · Jiahui Song[1] · January Mikolaj Weiner 3rd[1] · Avneesh Chopra[1] · Henrik Dommisch[1] · Dieter Beule[2] · Arne S. Schaefer[1]

Received: 14 December 2023 / Accepted: 27 April 2024

Abstract

The long noncoding RNA CDKN2B-AS1 harbors a major coronary artery disease risk haplotype, which is also associated with progressive forms of the oral inflammatory disease periodontitis as well as myocardial infarction (MI). Despite extensive research, there is currently no broad consensus on the function of CDKN2B-AS1 that would explain a common molecular role of this lncRNA in these diseases. Our aim was to investigate the role of CDKN2B-AS1 in gingival cells to better understand the molecular mechanisms underlying the increased risk of progressive periodontitis. We downregulated CDKN2B-AS1 transcript levels in primary gingival fibroblasts with LNA GapmeRs. Following RNA-sequencing, we performed differential expression, gene set enrichment analyses and Western Blotting. Putative causal alleles were searched by analyzing associated DNA sequence variants for changes of predicted transcription factor binding sites. We functionally characterized putative functional alleles using luciferase-reporter and antibody electrophoretic mobility shift assays in gingival fibroblasts and HeLa cells. Of all gene sets analysed, collagen biosynthesis was most significantly upregulated ($P_{adj}=9.7\times10^{-5}$ (AUC > 0.65) with the CAD and MI risk gene *COL4A1* showing strongest upregulation of the enriched gene sets (Fold change = 12.13, $P_{adj} = 4.9\times10^{-25}$). The inflammatory "TNFA signaling via NFKB" gene set was downregulated the most ($P_{adj}=1\times10^{-5}$ (AUC = 0.60). On the single gene level, *CAPNS2*, involved in extracellular matrix organization, was the top upregulated protein coding gene (Fold change = 48.5, $P<9\times10^{-24}$). The risk variant rs10757278 altered a binding site of the pathogen responsive transcription factor STAT1 ($P=5.8\times10^{-6}$). rs10757278-G allele reduced STAT1 binding 14.4% and rs10757278-A decreased luciferase activity in gingival fibroblasts 41.2% ($P=0.0056$), corresponding with GTEx data. CDKN2B-AS1 represses collagen gene expression in gingival fibroblasts. Dysregulated collagen biosynthesis through allele-specific CDKN2B-AS1 expression in response to inflammatory factors may affect collagen synthesis, and in consequence tissue barrier and atherosclerotic plaque stability.

Introduction

CDKN2B-AS1 (CDKN2B antisense RNA 1; ANRIL) encodes a long non-coding RNA (lncRNAs), a class of molecules that are considered critical players of gene regulation in multiple biological processes. In general, lncRNAs act as transcriptional repressors, downregulating gene activity by directly interacting with the chromatin or mRNA of their target genes (Statello et al. 2021). Accordingly, CDKN2B-AS1 has been shown to regulate the neighboring genes *CDKN2A* and *CDKN2B* (cyclin dependent kinase inhibitors 2 A and 2B) (Yap et al. 2010) as well as the expression of distant genes (Holdt et al. 2013). CDKN2B-AS1 expression is particularly strong in gastrointestinal barrier tissues like small intestines and colon found in the Human Protein Atlas project from analysis of 27 different tissues in order to determine tissue-specificity implying a functional role of CDKN2B-AS1 for gene regulation in gastrointestinal tissues. A recent study that investigated the function of CDKN2B-AS1 in the intestines showed that cells with reduced CDKN2B-AS1 activity exhibited enhanced tissue barrier function and showed that here, CDKN2BAS-1 is expressed mainly in epithelial cells (Rankin et al. 2019).

Dysregulation of lncRNAs is associated with many complex diseases (Mercer et al. 2009; Zhao et al. 2016).

✉ Arne S. Schaefer
arne.schaefer@charite.de

[1] Dept. of Periodontology, Oral Medicine and Oral Surgery, Institute for Dental and Craniofacial Sciences, Charité - University Medicine Berlin, Berlin, Germany

[2] Core Unit Bioinformatics, Berlin Institute of Health at Charité, Berlin, Germany

Published online: 04 June 2024

Despite the biased expression in the colon, *CDKN2B-AS1* is the major genetic risk locus of coronary artery disease (CAD) (Consortium et al. 2013; WTCCC 2007). The CAD associated haplotype block (tagged by GWAS lead SNP rs1333049) is also associated with progressive early-onset forms of the oral inflammatory disease periodontitis (Schaefer et al. 2009) and myocardial infarction (MI) (Helgadottir et al. 2007b; Myocardial Infarction Genetics et al. 2009; Nikpay et al. 2015) (Munz et al. 2018; Schaefer et al. 2009). Genetic risk variants of this haplotype block have an influence on CDKN2B-AS1 transcript levels (Folkersen et al. 2009), implying a molecular biological link between susceptibility for these diseases and regulation of CDKN2B-AS1 expression.

A link of CDKN2B-AS1 to inflammation has also been established. Reduced CDKN2B-AS1 transcript levels repressed TNFA induced IL6 and IL8 expression (Zhou et al. 2016), whereas upstream in the inflammatory signaling cascade the pro-inflammatory cytokine IFNG stimulated CDKN2B-AS1 expression (Harismendy et al. 2011). Moreover, the CAD and MI GWAS lead SNP rs10757278 (Nikpay et al. 2015; Tcheandjieu et al. 2022), which is in strong linkage disequilibrium (LD, $r2 > 0.9$) with rs1333049, disrupts a binding site for the immune signal transducer STAT1 (Harismendy et al. 2011).

However, in general, genes that had differential expression after increasing or decreasing CDKN2B-AS1 activity showed little overlap between the different studies, possibly due to the heterogeneity of the methods and cell models used.

The aim of the current study was to investigate the role of CDKN2B-AS1 in gingival fibroblasts. To this end, we investigated the cell type-specific downstream regulatory effects of CDKN2B-AS1 on gene expression in gingival fibroblasts. Here, we followed the rational of previous studies (Alfeghaly et al. 2021; Bochenek et al. 2013; Hubberten et al. 2019; Rankin et al. 2019), which suggested that overexpression of CDKN2B-AS1 can induce cellular stress and that in cells where CDKN2B-AS1 is already naturally expressed, overexpression does not lead to further downregulation of the potentially suppressed target genes because they would have already been silenced. In addition, we searched for biologically functional genetic variants in the associated haplotype block in order to obtain information about the upstream signaling events that regulate CDKN2B-AS1 activity.

Here, we show that CDKN2B-AS1 suppresses collagen synthesis in gingival fibroblasts and that the periodontitis and infarction associated susceptibility gene polymorphism rs10757278-G reduced STAT1 binding and increased CDKN2B-AS1 expression.

Materials and methods

Knockdown of CDKN2B-AS1 transcript levels by locked nucleic acids

LNA GapmeRs (single stranded antisense oligos) either targeting unique regions of CDKN2B-AS1 isoforms or non-targeting any region (scrambled, used as a negative control) were designed by QIAGEN as published before (Alfeghaly et al. 2021). A mix of four CDKN2B-AS1 LNA GapmeRs or a scrambled LNA GapmeR was used. The GapmeRs hybridized to unique regions of the main CDKN2B-AS1 isoforms as follows: gene globe ID LG00217779-DFA (exon 1), LG00217777-DFA (exon 7–13, numbered as in transcript EU741058), LG00217784-DFA (exon 12–13, numbered as in transcript DQ485454), LG00217785-DFA (exon 17–18, numbered as in transcript NR_003529). Primary gingival fibroblasts(pGFs) (from 3 different donators) and immortalized Human Gingival Fibroblast cells (iHGF) were sown with a density of 1.3×10^5 cells / well in 6-well tissue culture plates (Techno Plastic Products, Switzerland) one day prior to LNA GapmeR transfection with Lipofectamine 2000 (Thermo Fisher Scientific, Waltham, USA) at a final concentration of 225 µm 48 h after transfection. Total RNA was extracted using the RNeasy Mini Kit (Qiagen, Germany).

Cell culture

Primary gingival fibroblasts (pGFs) were isolated from gingival biopsies and cultured as previously described (Freitag-Wolf et al. 2019). In brief, pGFs cells and immortalized Human Gingival Fibroblast cells (iHGF) (ABM, Canada) were grown in Dulbecco's modified Eagle's medium (DMEM, PAN Biotech, Germany) supplemented with 10% fetal bovine serum (FBS, Gibco, USA), and 1% non-essential amino acids (MEM-NEAA, PAN Biotech). HeLa cells were cultured as recently described (Chopra et al. 2021). Briefly, cells were cultured in Earle's MEM (Bio&Sell, Nuremberg, Germany) supplemented with 10% FBS, 2mM L-Glutamine (Bio&Sell), and 1%NEAA.

Quantitative real-time PCR

Complementary DNA (cDNA) was synthesized from 100 ng total RNA, using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Thermo Fisher Scientific). Quantitative real-time PCR (qRT-PCR) was performed using SYBR Select Master Mix (Applied Biosystems) to validate downregulation of CDKN2B-AS1 mRNA levels. The results were analyzed by using the $2-\Delta CT$ or $2-\Delta\Delta CT$ method and normalized to GAPDH as an internal

control. The primer sequences were described in (Supplementary Materials Table. S1).

RNA-sequencing

Total RNA was extracted from pGFs and iHGF cells using the RNeasy Mini Kit. 500 ng total RNA of transfected cell cultures were sequenced with 16 million reads (75 bp single end) on a NextSeq 500 using the NextSeq 500/550 High Output Kit v2.5 (75 cycles). RNA-Seq was performed at the Berlin Institute of Health, Core Facility Genomics. Reads were aligned to the human genome sequences (build GRCh38.p7) using the STAR aligner v. 2.7.8a (Dobin et al. 2013). Quality control (QC) of the reads was inspected using the multiqc reporting tool (Ewels et al. 2016) summarizing a number of approaches, including fastqc (available online at http://www.bioinformatics.babraham.ac.uk/projects/fastqc), dupradar (Sayols et al. 2016), qualimap (Garcia-Alcalde et al. 2012), and RNA-SeqC (DeLuca et al. 2012). Raw counts were extracted using the STAR program. For differential gene expression, the R package DESeq2 (Love et al. 2014), version 1.30 was used. Gene set enrichment was performed using the CERNO test from the tmod package (Zyla et al. 2019), version 0.50.07, using the gene expression profiling-based gene set included in the package, as well as the MSigDB (Liberzon et al. 2015), v.7.4.1. For the hypergeometric test and the Gene Ontology gene sets, the goseq package, version 1.38 (Young et al. 2010) was used. The *P* values of the differently expressed genes were corrected for multiple testing using Benjamini-Hochberg correction. The corrected *P* values are given as q values (false discovery rate [FDR]).

Western blotting

To validate the expression of iHGF after knockdown the CDKN2B-AS1, the total protein of iHGF was extracted using RIPA with I.P. for 30 min on ice. The lysed samples were separated on polyacrylamide gels and transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore, USA). Then the PVDF membranes were incubated with the primary antibody COL4A1(1:1000, Cell Signaling), Col 6A1 (1:500, Santa Cruz), CAPNS2(1:1000, Invitrogen) and β-actin (1:2000, Santa Cruz) overnight at 4 °C. Subsequently, the membranes were incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies at room temperature for 1 h. The signal was acquired by using chemiluminescence detection (Chemostar Touch, INTAS, Indian). Then, ImageJ software was used to calculate the band intensities, and β-actin antibody was used as an internal control for signal normalization.

Table 1 TFs predicted to bind at the SNP sequences with $p < 0.01$

SNP	Name	*P*-value (common allele)	*P*-value (rare allele)	*p*-value (combined)
rs10757278	V$STAT_01	0.0080	0.077	0.0052
rs7859727	V$GATA6_01	< 0.00057	0.824	0.0041

Calculated with the Transcription factor Affinity Prediction (TRAP) Web Tool TRAP (multiple sequences), using the matrix file 'transfac_2010.1 vertebrates', the background model 'human_promoters' and 10 bp +/- the SNP alleles:

rs10757278: CATTCCGGTA[A/G]GCAGCGATGC,

rs7859727: ATCTGAATGA[T/C]AGGCATTCCT

Screening for functional periodontitis associated variants

For screening functional periodontitis associated variants, LD between the lead SNP rs1333049 and other common SNPs of this haplotype block was assessed using LDproxy Tool (Machiela and Chanock 2015) with population groups CEU (Utah Residents from North and West Europe) and GBR (British in England and Scotland). We assessed LD using r^2 as a measure of correlation of alleles for 2 genetic variants (Supplementary Materials Fig. S2 and Table 1). We analyzed whether these SNPs located to chromatin elements that correlate with regulatory functions of gene expression provided from ENCODE (ENCODE-Project- Consortium 2012) (Supplementary Materials Fig. S3). To annotate eQTL effects of the associated SNPs, we used the software tool QTLizer (Munz et al. 2020). To investigate whether these SNPs changed predicted TFBSs, we used the TF databases Transfac (Thomas-Chollier et al. 2011) and the open-access database Jaspar (Sandelin et al. 2004). If Transfac as well as Jasper TF matrix files predicted a TF binding affinity at a SNP, with a stronger binding affinity at the common allele compared to the alternative allele, we selected the SNP for functional follow-up experiments. This conservative selection criterion was preassigned to avoid choosing false positive TFBS predictions for functional follow-up. TF binding motives were confirmed using the web interface for Position Weight Matrix (PWM) model generation and evaluation, PWMTools (Ambrosini et al. 2018).

Electrophoretic mobility shift assay (EMSA)

To characterize allele-specific DNA-protein interaction and TF STAT1 binding at rs10757278, we performed EMSAs with the Gelshift Chemiluminescent EMSA Kit (Active Motif, Germany) as recently described (Chopra et al. 2021). In brief, allele-specific oligonucleotide probes were synthesized (Metabiom International, Supplementary Materials Table S2). Nuclear protein extract was prepared from iHGF cells using the NE-PER Nuclear and Cytoplasmic Extraction Kit (Thermo Fisher). The

double-stranded oligonucleotides corresponding to both alleles of rs10757278 flanked by 21 bp in both cold and 3′-biotinylated form were obtained by annealing with their respective complementary primers. For supershift EMSA, 20 fmol biotin-labeled, double-stranded oligonucleotides were incubated for 20 min with nuclear extract (5 µg) in 1x binding buffer and 2 µL of a specific monoclonal antibody (STAT1, 10 µg/50 µL each (Santa Cruz Biotechnology, California, USA) at room temperature. For competition assay, 4 pmol unlabeled double-stranded oligonucleotides were added to the binding reaction. The DNA-protein complexes were electrophoresed in a 5% native polyacrylamide gel in 0.5x TBE buffer at 100 V for 1 h. After electric transfer of the products on a nylon membrane and cross-linking, the biotinylated probes were visualized by chemiluminescence detection (Chemostar Touch, INTAS, Indian). Band intensities were quantified by the absolute value area of the shifted antibody bands using the software ImageJ (Rueden et al. 2017).

Luciferase reporter gene assays

The putative regulatory DNA sequences (total length 539 bp) spanning 269 bp up- and downstream of the individual alleles of SNP rs10757278 were cloned into the firefly luciferase vector pGL4.24 (Promega, Madison, USA). The Further details were described in **Supplementary Materials**. iHGF cells were seeded at a density of 330,000 cells per 6-well before transfection with Lipofectamine 2000. HeLa cells were seeded at a density of 80,000 cells per well in 6-well plates and cultured until reaching 50–60% confluence. Transfection of HeLa cells was performed using the jetPEI transfection reagent (Polyplus transfection, France) following the manufacturer's instructions. Cells were co-transfected in triplicates with 2.7 µg firefly luciferase reporter plasmid containing the putative regulatory sequence together with 0.3 µg renilla luciferase reporter vector (phRL-SV40, Promega) in 6-well plates for 24 h. In parallel, cells were transfected with the empty pGL4.24 plasmid and 0.3 µg phRL-SV40 as control. Firefly and renilla luciferase activities were quantified using the Dual-Luciferase Stop & Glo Reporter Assay System (Promega) with the Orion II Microplate Luminometer (Berthold Technologies). Relative fold changes (FC) in activities were normalized according to the manufacturer's instructions (Promega) and differences of transcript levels were calculated with a T-Test using the software GraphPad Prism 9.

Fig. 1 Differentially expressed genes and enriched gene sets in gingival fibroblasts after CDKN2B-AS1 knockdown. **(A)** Volcano plot of LNA GapmeRs transfected gingival fibroblasts showing differential expression of protein coding genes and numerous lncRNAs and pseudogenes. The names of the most significant differentially expressed protein coding genes are shown. The names of the most significant differentially expressed lncRNAs and pseudogenes are not shown to highlight the observed prominent role of CDKN2B-AS1 interaction with non-protein coding genes. **(B)** Transfection of primary gingival fibroblasts with LNA GapmeRs induced significant reduction of CDKN2B-AS1 transcript levels (qRT-PCR). **(C-D)** Western blot analysis validated that reduced CDKN2B-AS1 transcript levels correlated with significantly reduced CAPNS2 protein levels. Western Blot band intensities are normalized to *ACTB* (*$p < 0.05$; **$p < 0.01$). **(E-G)** Gene set enrichment analysis of GapmeR transfected gingival fibroblasts. Shown are evidence plots (receiver operator characteristic curves) for the significant gene sets with an area under the curve (AUC) ≥ 0.6. (C) From Reactome database, REACTOME_COLLAGEN_BIOSYNTHESIS_AND_ MODIFYING_ENZYMES _ M26999, enriched 62 genes. (D) From Reactome database, REACTOME_COLLAGEN_ CHAIN_TRIMERIZATION_ M27812, enriched 40 genes; (E) From Hallmark database, HALLMARK_TNFA_SIGNALING_VIA_ NFKB_ M5890, enriched 191 genes; (F&G) From Tmod database, EXTRACELLULAR MATRIX (I)_ LI.M2.0 enriched 30 genes and COLLAGEN, TGFB FAMILY ET AL_ LI.M77 enriched 31 genes, respectively. The gray rug plot underneath each curve corresponds to genes sorted by *P* value, with the genes belonging to the corresponding gene sets highlighted in red (upregulated genes) or blue (downregulated genes). Bright red or bright blue indicates that the genes are significantly regulated. **(H)** Western blotting validation of LNA GapmeRs transfected gingival fibroblasts showing *COL4A1* and *COL6A1* upregulation after CDKN2B-AS1 knockdown (*COL4A1* fold change = 1.74, $P = 0.0004$; *COL6A1* fold change = 1.57, $P = 0.0155$ respectively)

Results

CDKN2B-AS1 modulates the expression of collagen genes in gingival fibroblasts

In a first step, we quantified the expression of CDKN2B-AS1 in the gingiva to compare it with the expression in colon tissue, where this lncRNA is relatively highly expressed. We quantified the expression levels of CDKN2B-AS1 poly-A transcripts that terminate with exon-13 and − 19 (numbered as in transcript EU741058 and numbered as in transcript NR_003529, respectively) between healthy gingiva and colon biopsies using qRT-PCR. We found an equal expression in the gingiva compared to colon and similar expression of 3'-exon 13 and 3'-exon 19 transcripts between gingival and colon biopsies (Fig. S1A&B).

Silencing CDKN2B-AS1 expression in primary gingival fibroblasts using a mix of four LNA GapmeRs resulted in 84% reduction of CDKN2B-AS1 expression (Fig. 1B). Genomewide expression profiling by RNA-Seq following CDKN2B-AS1 knockdown revealed 1,167 upregulated genes with $\log_2$ Fold change ($\log_2$FC) > 2) and 2,829 downregulated genes with $\log_2$FC < -1 ($P_{adj} < 0.05$; Fig. 1A, Supplementary Materials Table S4). The most upregulated protein-coding gene with the lowest *p*-value was the gene

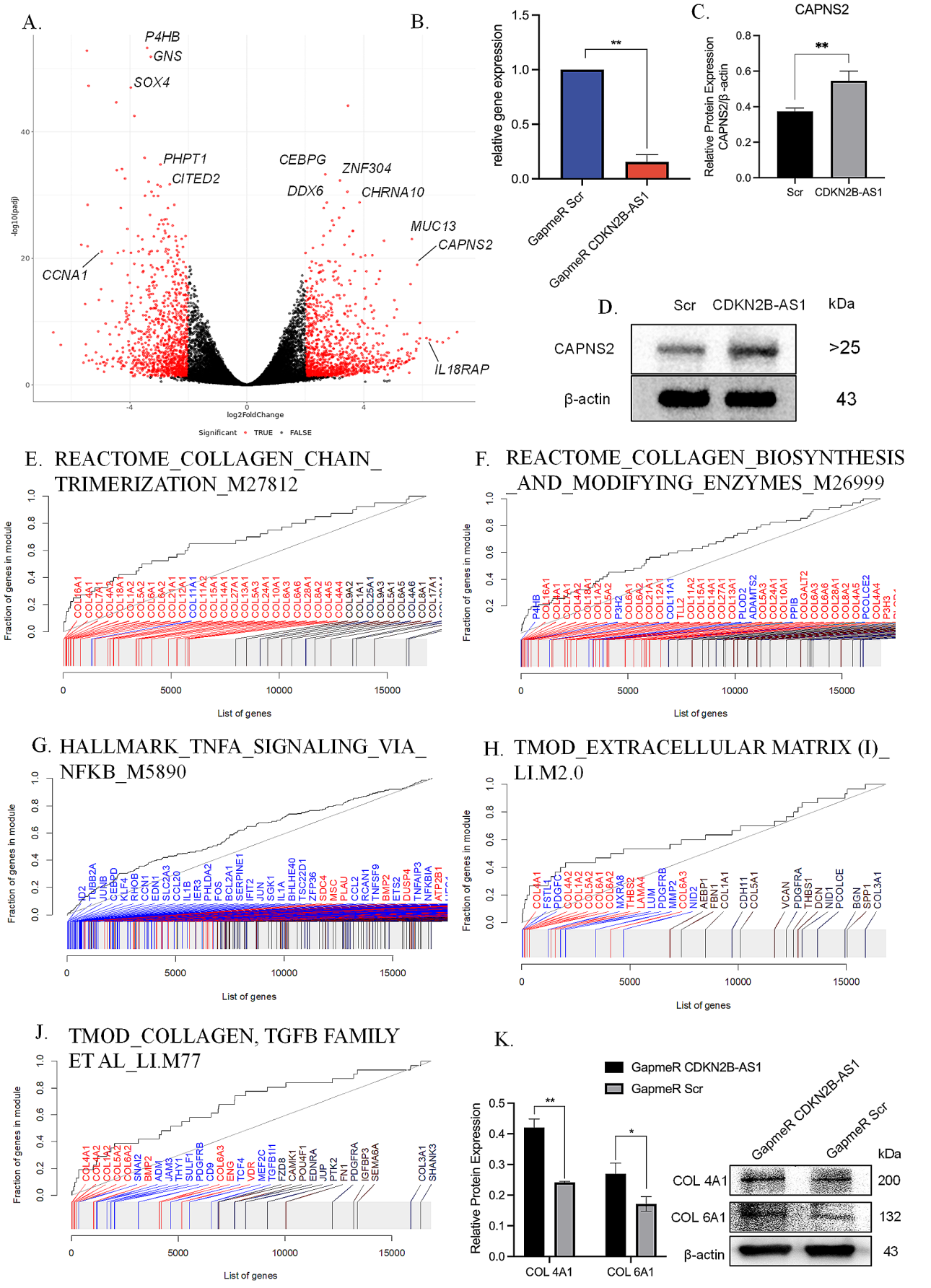

A.
P4HB
GNS
SOX4
PHPT1
CITED2
CEBPG
ZNF304
DDX6
CHRNA10
MUC13
CAPNS2
CCNA1
IL18RAP
-log10(padj)
log2FoldChange
Significant TRUE FALSE
B.
relative gene expression
GapmeR Scr
GapmeR CDKN2B-AS1
C.
CAPNS2
Relative Protein Expression CAPNS2/β-actin
Scr
CDKN2B-AS1
D.
Scr
CDKN2B-AS1
kDa
CAPNS2
>25
β-actin
43
E. REACTOME_COLLAGEN_CHAIN_TRIMERIZATION_M27812
F. REACTOME_COLLAGEN_BIOSYNTHESIS_AND_MODIFYING_ENZYMES_M26999
G. HALLMARK_TNFA_SIGNALING_VIA_NFKB_M5890
H. TMOD_EXTRACELLULAR MATRIX (I)_LI.M2.0
J. TMOD_COLLAGEN, TGFB FAMILY ET AL_LI.M77
Fraction of genes in module
List of genes
K.
GapmeR CDKN2B-AS1
GapmeR Scr
Relative Protein Expression
COL 4A1
COL 6A1
β-actin
kDa
200
132
43

Table 2 Top 10 differentially expressed protein coding genes in gingival fibroblasts after CDKN2B-AS1 knockdown

	Gene	Log_2 Fold change	IfcSE†	P_{adj}(<0.005)
Top Up-regulated	*IL18RAP*	6.261	1.065	7.20E-08
	ATP13A5	5.878	1.045	2.80E-07
	CAPNS2	5.645	0.532	9.00E-24
	KERA	5.497	1.113	0.000008
	CMTM2	5.486	1.15	0.000017
	GSDMC	5.441	1.329	0.00027
	SLC5A12	5.302	1.008	0.0000018
	ASCL3	5.298	1.263	0.00018
	INSL5	5.227	1.177	0.000069
	GOLT1A	5.194	1.475	0.002
Top Down-regulated	*CCNA1*	-4.961	0.489	8.30E-22
	UBE2L5	-4.561	1.012	0.000052
	WDR87	-4.481	0.671	6.70E-10
	KLHDC7B	-4.406	1.057	0.0002
	TRBV12-4	-4.378	0.998	0.000086
	KRTAP2-4	-4.338	1.052	0.00024
	CCL20	-4.141	0.549	2.1E-12
	IL1B	-4.122	0.556	5.0E-12
	OR14K1	-4.121	0.584	5.7E-11
	LBX1	-4.051	0.888	4.2E-5

†: Ifc: logarithmic fold change

Table 3 Top 10 differentially expressed lncRNA genes in gingival fibroblasts after CDKN2B-AS1 knockdown

	Gene	Log_2 Fold change	IfcSE†	P_{adj}(<0.005)
Top Up-regulated	*LINC00536*	7.192	1.129	4.40E-09
	RGS5-AS1	6.507	1.13	1.40E-07
	LINC02516	5.802	1.007	1.30E-07
	LOC112268408	5.793	1.106	0.000002
	LINC01091	5.597	0.634	1.20E-16
	ENSG00000224478	5.483	0.878	8.80E-09
	LINC02435	5.452	1.302	0.00019
	ENSG00000272094	5.385	1.336	0.00034
	NPSR1-AS1	5.274	1.347	0.00052
	ENSG00000257894	5.229	1.324	0.00046
Top Down-regulated	*NUP33-DT*	-6.618	1.043	5.00E-09
	ENSG00000255317	-5.662	0.544	7.10E-23
	KIF18B-DT	-5.528	0.7	1.60E-13
	ENSG00000272226	-5.498	1.166	0.000022
	ZBED9-AS1	-5.471	0.342	1.60E-53
	ENSG00000274460	-5.447	0.462	3.60E-29
	ENSG00000277511	-5.413	0.358	5.70E-48
	ENSG00000272366	-5.397	1.266	0.00014
	USP12-AS1	-5.141	1.3	0.00045
	LINC02112	-4.985	0.58	7.30E-16

†: Ifc: logarithmic fold change

CAPNS2 ($log_2FC = 5.65$ ($P = 9 \times 10^{-24}$; Table 2; Fig. 1A). This gene is predicted to be involved in proteolysis [provided by Alliance of Genome Resources, Apr 2022] and based on gene content similarity is part of the gene cluster 'Extracellular matrix organization' (PathCard). Western blotting with protein extract of CDKN2B-AS1 knocked-down gingival fibroblasts showed a significant upregulation in the protein expression of the CAPNS2 (fold change = 1.5, $P = 0.0063$) (Fig. 1C**&D**). We also noticed that many lncRNAs were differentially expressed following CDKN2B-AS1 knockdown and separately list the top 10 differentially expressed ncRNAs in Table 3.

We performed gene set enrichment analyses (GSEA) with the co-expression gene set tmod and the Molecular Signatures Database (MSigDB) gene sets reactome, hallmark, KEGG, and GO, by contrasting CDKN2B-AS1 knockdown cells to negative control transfected cells. The Reactome gene sets COLLAGEN_CHAIN_TRIMERIZATION (M27812) and COLLAGEN_BIOSYNTHESIS_AND_MODIFYING_ENZYMES (M26999) were most significantly upregulated, with $P_{adj} = 9.7 \times 10^{-5}$ (AUC = 0.66 and 0.67) and Hallmark gene set TNFA_SIGNALING_VIA_NFKB (M5890) was most significantly downregulated with $P_{adj} = 1.0 \times 10^{-5}$ (AUC = 0.60) (Fig. 1E-G). Additionally, Tmod gene sets "collagen, TGFB family (LI.M77) and extracellular matrix I (LI.M2.0) showed significant enrichment with $P_{adj} = 0.01$ (AUC = 0.67) and $P_{adj} = 0.002$ (AUC = 0.63), respectively (Fig. 1H**&J**).

CDKN2B-AS1 knockdown increased COL4A1 and COL6A1 protein levels in gingival fibroblasts

We validated the positive effect of CDKN2B-AS1 knockdown on the expression of the collagen genes *COL4A1* and *COL6A1* by Western blotting. *COL4A1* showed the strongest upregulation on the RNA level in our RNA-Seq data (FC = 12.13, $P_{adj} = 4.9 \times 10^{-25}$) and *COL6A1* (FC = 3.73, $P_{adj} = 1.1 \times 10^{-12}$) was found in a previous ChIRP-Seq experiments to be a direct regulatory target of CDKN2B-AS1. Western blotting with protein extract of CDKN2B-AS1 knocked-down gingival fibroblasts showed a significant upregulation in the protein expression of the collagen genes *COL4A1* (fold change = 1.74, $P = 0.0004$) and *COL6A1* (fold change = 1.57, $P = 0.0155$) (Fig. 1K). These results proved *COL4A1* and *COL6A1* repression by CDKN2B-AS1.

rs10757278-G allele reduced STAT1 binding in gingival fibroblasts

We next searched for putative causal variants within the CAD/periodontitis/MI risk haplotype block that were located

within chromatin stretches marked with biochemical modifications characteristic of regulatory DNA elements. The CAD GWAS lead SNP rs1333049 is in strong LD ($r^2 > 0.8$) with 55 common SNPs (minor allele frequency ≤ 0.05) in North-West European populations (1000 genomes population codes CEU and GBR). Of these, 24 SNPs located within chromatin elements with biochemical marks assigning them as putative regulatory elements (Fig. 2A, Supplementary Materials Table S3). We computationally analyzed whether the alternative alleles of these 24 SNPs changed predicted transcription factor binding sites (TFBSs). We found that Transfac and Jasper TF matrix files predicted TFBSs at two SNPs (rs10757278 and rs7859727) with the alternative alleles decreasing TF binding affinities (Table 1). rs10757278 locates within a STAT1 binding site with the common A allele being part of the STAT1 binding motif ($P = 5.8 \times 10^{-6}$) and a STAT1 matrix similarity of 94.8% (Fig. 2B). In transfac_2010.1 vertebrates matrix files, the alternative rs10757278-G allele 3,057-fold reduced binding affinity ($P = 0.0177$). rs10757278-G was described before as a putative causal variant for CAD affecting a STAT1 binding site (Harismendy et al. 2011). STAT1 is strongly expressed in gingival fibroblast (qRT-PCR threshold cycle [C_t] value for STAT1 = 20.45, $C_{t\,(GAPDH)} = 14.61$), indicating biological activity of STAT1 in this cell type (Supplementary Materials Fig. S4).

rs7859727-T locates within a predicted GATA binding site ($P = 0.0006$). In transfac_2010.1 vertebrates matrix files, the alternative rs10757278-C allele 836-fold reduced binding affinity ($P = 0.549$). However, RNA-Seq data of healthy gingival biopsies (Richter et al. 2019) showed GATA expression in gingival tissues below detection limit, indicating that this tissue does not express GATA.

Allele specific STAT1 binding at rs10757278 was described by Chromatin Immunoprecipitation (ChIP) in lymphoblastoid cells (LCL) before (Harismendy et al. 2011). We validated rs10757278 allele specific STAT1 binding with protein extract isolated from gingival fibroblasts and performed a STAT1 antibody EMSA with DNA probes that contained either rs10757278 ref-A allele or alt-G allele. Using protein extract of gingival fibroblasts, STAT1 binding at DNA probes with rs10757278 alt-G allele was 14.4% reduced compared to the ref-A allele (Fig. 2C**&D**). The observation of increased STAT1 binding at the G allele corresponded with the previous observation in LCL cells.

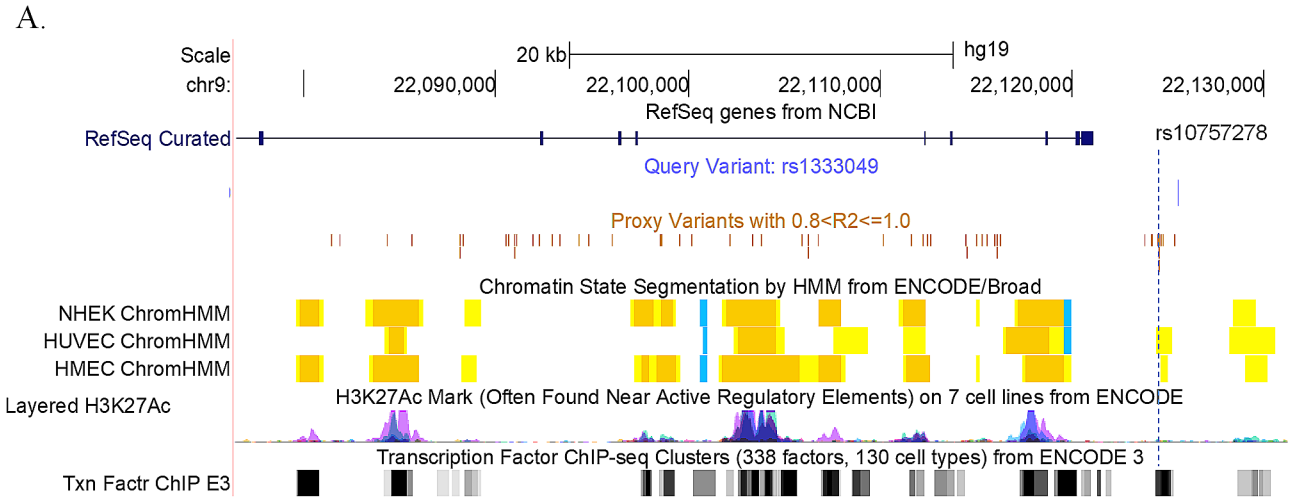
STAT1 binding at rs10757278 repressed gene activity

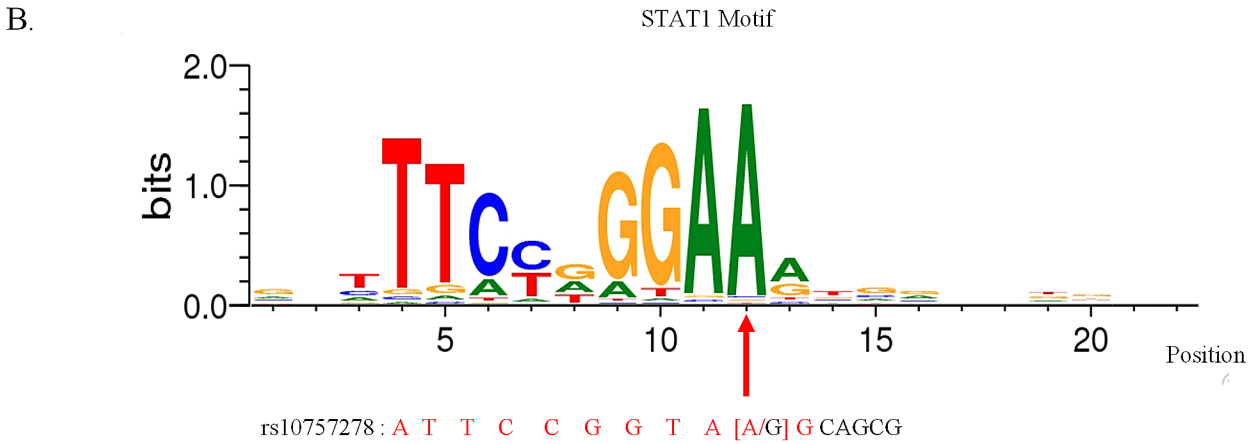
We tested the regulatory effect direction of the DNA element at rs10757278 in gingival fibroblasts using luciferase reporter gene assays. The DNA sequence containing the STAT1 binding allele rs10757278-A decreased luciferase activity in gingival fibroblasts 41.2%, when compared with the common G-allele ($P = 0.0056$; Fig. 3A). We validated this finding in HeLa cells, because STAT1 is also expressed in this cell type, which can be efficiently transfected in vitro. Luciferase reporter gene transfection into HeLa cells confirmed decreased luciferase activity in the background of the rs10757278-A allele compared with the alternative G-allele ($P = 0.0063$; Fig. 3B). These results corresponded with GTEx data that also showed reduced CDKN2B-AS1 expression in homozygous carriers of the A allele compared to the G- allele ($P = 3.9 \times 10^{-6}$; Fig. 3C). Taken together, these results demonstrated that the reference rs10757278-A allele is part of a biological functional STAT1 binding site that acts as a transcriptional repressor.

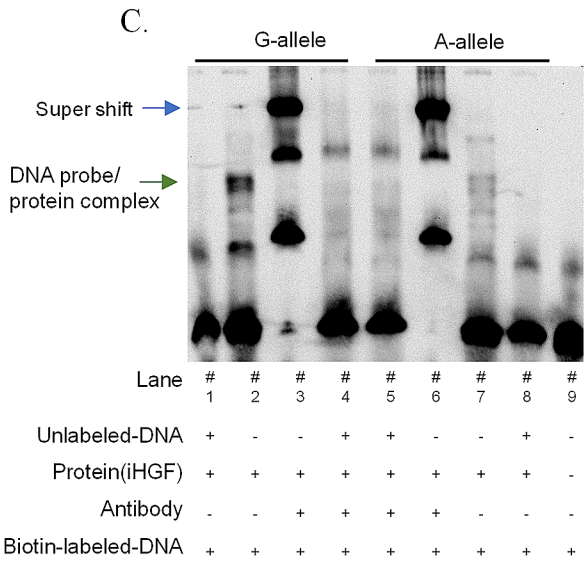
Discussion

In the current work, we provide evidence that *CDKN2B-AS1* represses expression of collagen genes in gingival fibroblasts and that *CDKN2B-AS1* is under negative control of the inflammatory transcription factor STAT1.

A previous study, which used the identical set of GapmeRs for CDKN2B-AS1 knockdown in the kidney cell line HEK293 and subsequently combined chromatin immune RNA precipitation followed by sequencing (ChIRP-Seq) with genomewide expression profiling, found that CDKN2B-AS1 directly contacted and regulated collagen gene expression (Alfeghaly et al. 2021). This study showed upregulation of the genes *COL6A1* and *COL12A1* and concluded that these collagen genes were direct targets of CDKN2B-AS1. However, this study did not find any enriched pathways in the list of direct trans-regulated genes. The immortalized kidney cell line HEK293 has a different expression pattern and differentiation state compared with primary gingival fibroblasts. Accordingly, it may not fully represent the biological functions, which CDKN2B-AS1 has in gingival fibroblasts, where this gene is naturally expressed at comparatively high levels. Therefore, to validate the findings of this previous study, we used the identical set of gapmers to reduce CDKN2B-AS2 transcript levels in gingival fibroblasts. Here, we gave evidence that collagen synthesis pathways were the most enriched gene sets in response to suppression of CDKN2B-AS1 transcript levels. We also proved on the protein level that in gingival fibroblasts CDKN2B-AS1 negatively controls the collagen genes *COL6A1*, but also *COL4A1*, which was the strongest upregulated collagen gene in our study. Moreover, as a novel finding, we detected that *CAPNS2*, which encodes the Calpain Small Subunit 2, was the most upregulated gene after CDKN2B-AS1 knockdown. Calpains are calcium-activated


A.
Scale
20 kb
hg19
chr9:
22,090,000
22,100,000
22,110,000
22,120,000
22,130,000
RefSeq genes from NCBI
RefSeq Curated
rs10757278
Query Variant: rs1333049
Proxy Variants with 0.8<R2<=1.0
Chromatin State Segmentation by HMM from ENCODE/Broad
NHEK ChromHMM
HUVEC ChromHMM
HMEC ChromHMM
Layered H3K27Ac
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE
Transcription Factor ChIP-seq Clusters (338 factors, 130 cell types) from ENCODE 3
Txn Factr ChIP E3


B.
STAT1 Motif
2.0
1.0
0.0
bits
5
10
15
20
Position
rs10757278 : A T T C C G G T A [A/G] G CAGCG


C.
G-allele
A-allele
Super shift
DNA probe/
protein complex
Lane
1
2
3
4
5
6
7
8
9
Unlabeled-DNA + - - + + - - + -
Protein(iHGF) + + + + + + + + -
Antibody - - + + + + - - -
Biotin-labeled-DNA + + + + + + + + +

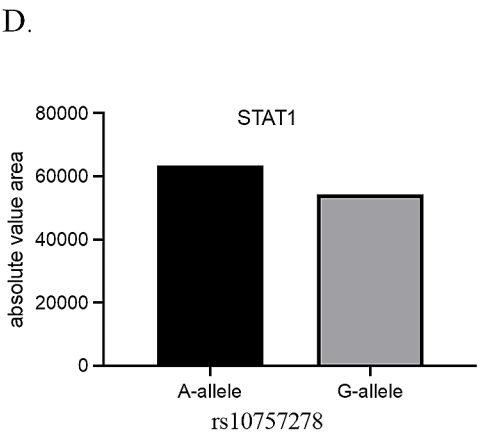

D.
STAT1
80000
60000
40000
20000
0
absolute value area
A-allele
G-allele
rs10757278

Fig. 2 rs10757278 is a functional SNP within the chr9p21.3 risk haploblock and localizes to a STAT1 binding site. **(A)** 55 common SNPs (MAF ≤ 0.05) in CEU and GBR populations in strong LD with the GWAS lead SNP rs1333049 and 24 SNPs located within chromatin elements that correlate with regulatory functions of gene expression, which is indicated by chromatin state segmentation for 3 cell types (data from ENCODE; orange = predicted strong enhancer, yellow = weak enhancer, blue = insulator). Some proxy SNPs locate in H3K4me1 and H3K27ac methylation marks, which are often associated with regulation of gene transcription and within TFBS that were determined from ENCODE ChIP-Seq data. The position of rs10757278 is marked with a dashed line. **(B)** The DNA sequence at rs10757278-A allele shares a matrix similarity of 95% with the STAT1 transcription factor (TF) binding motif. **(C)** EMSA was performed with rs10757278 allele-specific oligonucleotide probes and nuclear protein extract from gingival fibroblast cells. Binding of STAT1 antibody to allele-specific probes is shown in lane 2 and 7. The supershift caused by STAT1 antibody binding to the DNA probe-protein complex is seen in lane 3 and 6. Unlabeled DNA was added in lanes 1 and 8 to verify that the band shift was antibody-specific. **(D)** Absolute value area of the antibody-specific bands. In the background of rs10757278-G allele, STAT1 binding to the allele-specific oligonucleotide probe was reduced 14.4% compared to the A-allele

cysteine proteases that act as part of numerous intracellular signaling pathways. Of particular interest, calpain activity is required for differentiation/activation of fibroblasts, which lay down extracellular collagen matrix (ECM) proteins (e.g. collagen). In response to tissue injury, calpain is activated and promotes, in addition to the expression and release of proinflammatory cytokines (Ji et al. 2016), the activation and differentiation of fibroblasts, thereby promoting the production of collagens (Scaffidi et al. 2002). In contrast it was shown that inhibition of calpains interrupt the early steps of fibroblast activation and differentiation, thereby attenuating the production of collagen (Kim et al. 2019; Letavernier et al. 2012). Therefore, collagen production and accumulation as seen in repeated injury, chronic inflammation and wound healing requires precise regulation to avoid fibrosis and to maintain barrier tissue function. Our data indicate CDKN2B-AS1 regulates CAPNS2 activity.

Additionally, our gene pathway enrichment analysis also revealed, in response to CDKN2B-AS1 repression, a significant downregulation of the pathway 'TNFA Signaling via NFKB' including significant repression of the genes *IL1A* and *−1B*. Consistent with this observation, it was previously shown that knockdown of CDKN2B-AS1 transcripts in endothelial cells inhibited TNFA induced IL6 and IL8 expression (Zhou et al. 2016). It has long been known that endogenous TNFA down-regulates collagen synthesis during normal wound healing (Regan et al. 1993) and that TNFA inhibits collagen-alpha gene expression in cultured fibroblasts (Buck et al. 1996). Considering the association of CDKN2B-AS1 with severe, progressive periodontitis, our data imply that CDKN2B-AS1 is a molecular regulator that aligns TNFA signaling and collagen synthesis in gingival fibroblasts.

Our data also confirmed the previously described role of rs10757278, a risk SNP for CAD (Tcheandjieu et al. 2022) and MI (Helgadottir et al. 2007a), as being a putative causal variant of this risk haplotype block, which impairs binding of the IFNG responsive signal transducer STAT1 (Harismendy et al. 2011). We confirmed STAT1 binding at this SNP sequence and showed that the rs10757278-G allele reduced STAT1 binding and increased reporter gene expression in gingival fibroblasts. GTEx data also reported increased CDKN2B-AS1 expression in homozygous carriers of the G-allele compared to the A-allele, confirming our result.

Validation of the results of previous work through an independent approach in a cell type in which CDKN2B-AS1 is naturally and comparatively highly expressed represents a significant value of our work by emphasizing these results and placing them in a new functional context. This context lies in the regulation of collagen synthesis in a barrier tissue, possibly in response to an inflammatory phase.

Periodontitis is characterized by recurrent and prolonged inflammation and gingival bleeding. Periodontal healing after active inflammation requires reconstruction of the gingival barrier tissues. These wound healing processes include tissue formation and tissue remodeling, which follow but partially overlap with inflammation (Yen et al. 2018). To fully restore the tissue barrier, these two processes of healing from gingival bleeding involves collagen deposition and collagen remodeling, respectively.

The CDKN2B-AS1 rs10757278-G allele is also associated with increased risk for MI (Helgadottir et al. 2007a). Of note, *COL4A1*, which is regulated by CDKN2B-AS1 shown in the current study for gingival fibroblasts and also previously for HEK293 cells (Alfeghaly et al. 2021), is also a risk gene for MI (Nikpay et al. 2015). Collagen is a critical component of atherosclerotic lesions and constitutes up to 60% of total plaque protein (Rekhter et al. 1996; Smith 1965). High collagen contributes to plaque structural integrity and mechanical "strength". Therefore, a deficit of collagen reinforcement leads to plaque weakness and vulnerability (Burleigh et al. 1992; Lee and Libby 1997), making atherosclerotic plaque prone to rupture (Rekhter 1999), increasing the risk for MI. The in vitro results of the current study provide a functional link between the two MI risk genes *CDKN2B-AS1* and *COL4A1*. Our study implies that the rs10757278-G allele leads to reduced CDKN2B-AS1 repression by reducing STAT1 binding. As a result, higher CDKN2B-AS1 would in turn lead to increased collagen repression. This could destabilize atherosclerotic plaque and weaken the gingival tissue barrier, independently increasing the risk for MI and periodontitis.

A limitation of the current study is that we have not shown that CDKN2B-AS1 directly binds to the collagen genes in

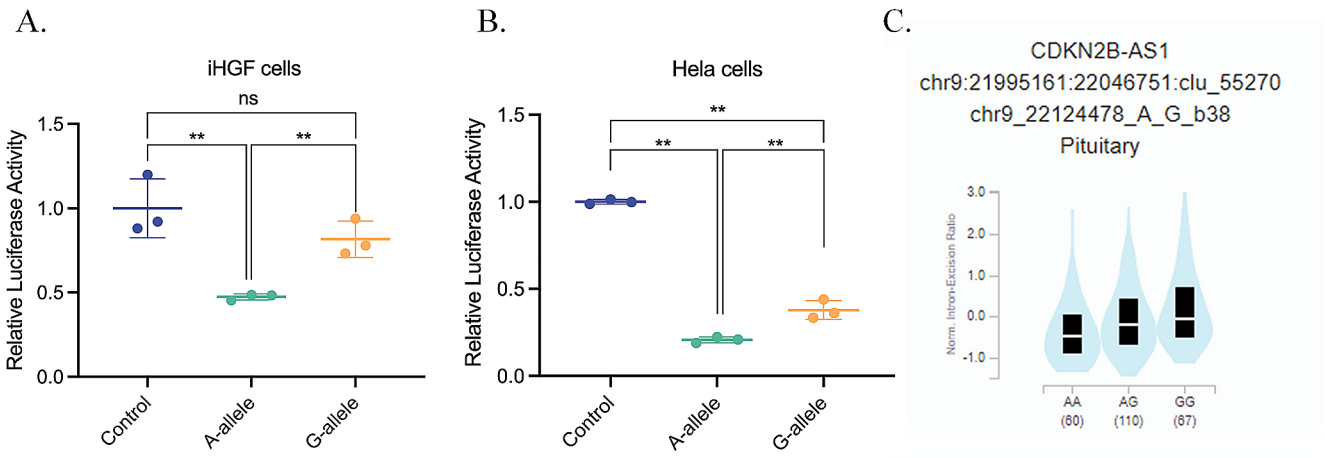


Fig. 3 rs10757278 STAT1 binding A-allele reduces reporter gene activity. **(A, B)** rs10757278-A significantly reduced luciferase activity in gingival fibroblasts (fold change = 1.72-fold, $P = 0.0056$) and in HeLa cells (fold change = 1.83-fold, $P = 0.0063$). (Control = empty pGL4.24 plasmid, $*p < 0.05$; $**p < 0.01$). **(C)** GTEx data indicate a cis-eQTL effect for rs10757278 on *CDKN2B*-AS1 expression (tissue: pituitary gland) with rs10757278-A significantly reducing *CDKN2B-AS1* expression compared to the G allele ($p = 3.9 \times 10^{-6}$)

gingival fibroblasts, e.g. by RNA immune precipitation. However, we consider our finding that CDKN2B-AS1 knockdown in gingival cells upregulated collagen genes and the results from our protein blots as a validation of the previous ChIRP-Seq results obtained in HEK293 cells (Alfeghaly et al. 2021). Another limitation of our study was that we silenced the expression of all major CDKN2B-AS1 isoforms in parallel by using a mixture of four LNA-GapmeRs that hybridized to individual regions of the major CDKN2B-AS1 isoforms exon1 (all isoforms), exon17-18 (NR isoform), exon12-13 (DQ isoform), and exon7-13 (EU isoform). It is possible that only one particular isoform regulates the expression of collagen genes. Future work is needed to determine whether a specific isoform or all major isoforms contribute to this process.

In conclusion, the results provide an explanation for the molecular mechanisms underlying the shared genetic risk for periodontitis and MI at CDKN2B-AS1. This mechanism implies that the effect on each disease is independent of the presence of the other diseases, an effect known as horizontal pleiotropy. Validation of this explanation of the epidemiological connection between periodontitis and MI, e.g. by correlating the carriage of the risk allele rs10757278-G in MI patients with the collagen content of their atherosclerotic plaques, could give final proof of the causality of the proposed genetic effect. We conclude that this genetic effect is caused by reduced STAT1 binding affinities at the risk G allele, leading to increased CDKN2B-AS1 expression with corresponding increased TNFA signaling, as suggested by our data. This would cause a prolongation of the inflammatory reaction and thereby contribute to disease progression.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00439-024-02674-1.

Author contributions Study conception and design: Arne S Schaefer. Material preparation: Weiwei Shi, Henrik Dommisch. Data collection and analysis were performed by [Weiwei Shi], [January 3rd Weiner], [Avneesh Chopra] and [Dieter Beule]. The first draft of the manuscript was written by [Weiwei Shi] and [Arne S Schaefer] and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by a research grant of the German Research Foundation (Deutsche Forschungsgemeinschaft SCHA 1582/14 – 1), to AS, and by the Chinese Scholarship Council, CSC202108440129. All authors gave their final approval. They agree to be accountable for all aspects of the work.
Open Access funding enabled and organized by Projekt DEAL.

Data availability RNA-Seq datasets are accessible in the Gene Expression Omnibus GEO under the accession number GSE266478 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE266478.

Declarations

Conflict of interest The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

Ethics approval Ethics Review Board of the Charité – Universitätsmedizin Berlin approved this study (#EA4/207/17). Informed consent was obtained from all individual participants included in the study.

References

Alfeghaly C, Sanchez A, Rouget R, Thuillier Q, Igel-Bourguignon V, Marchand V, Branlant C, Motorin Y, Behm-Ansmant I, Maenner S (2021) Implication of repeat insertion domains in the trans-activity of the long non-coding RNA ANRIL. Nucleic Acids Res 49:4954–4970. https://doi.org/10.1093/nar/gkab245

Ambrosini G, Groux R, Bucher P (2018) PWMScan: a fast tool for scanning entire genomes with a position-specific weight matrix. Bioinformatics 34:2483–2484. https://doi.org/10.1093/bioinformatics/bty127

Bochenek G, Hasler R, El Mokhtari NE, Konig IR, Loos BG, Jepsen S, Rosenstiel P, Schreiber S, Schaefer AS (2013) The large non-coding RNA ANRIL, which is associated with atherosclerosis, periodontitis and several forms of cancer, regulates ADIPOR1, VAMP3 and C11ORF10. Hum Mol Genet 22:4516–4527. https://doi.org/10.1093/hmg/ddt299

Buck M, Houglum K, Chojkier M (1996) Tumor necrosis factor-alpha inhibits collagen alpha1(I) gene expression and wound healing in a murine model of cachexia. Am J Pathol 149:195–204

Burleigh MC, Briggs AD, Lendon CL, Davies MJ, Born GV, Richardson PD (1992) Collagen types I and III, collagen content, GAGs and mechanical strength of human atherosclerotic plaque caps: span-wise variations. Atherosclerosis 96:71–81. https://doi.org/10.1016/0021-9150(92)90039-j

Chopra A, Mueller R, Weiner J 3rd, Rosowski J, Dommisch H, Grohmann E, Schaefer AS (2021) BACH1 binding links the genetic risk for severe periodontitis with ST8SIA1. J Dent Res 220345211017510. https://doi.org/10.1177/00220345211017510

Consortium CAD, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, Goldstein BA, Stirrups K, Konig IR, Cazier JB, Johansson A, Hall AS, Lee JY, Willer CJ, Chambers JC, Esko T, Folkersen L, Goel A, Grundberg E, Havulinna AS, Ho WK, Hopewell JC, Eriksson N, Kleber ME, Kristiansson K, Lundmark P, Lyytikainen LP, Rafelt S, Shungin D, Strawbridge RJ, Thorleifsson G, Tikkanen E, Van Zuydam N, Voight BF, Waite LL, Zhang W, Ziegler A, Absher D, Altshuler D, Balmforth AJ, Barroso I, Braund PS, Burgdorf C, Claudi-Boehm S, Cox D, Dimitriou M, Do R, Consortium D, Consortium C, Doney AS, El Mokhtari N, Eriksson P, Fischer K, Fontanillas P, Franco-Cereceda A, Gigante B, Groop L, Gustafsson S, Hager J, Hallmans G, Han BG, Hunt SE, Kang HM, Illig T, Kessler T, Knowles JW, Kolovou G, Kuusisto J, Langenberg C, Langford C, Leander K, Lokki ML, Lundmark A, McCarthy MI, Meisinger C, Melander O, Mihailov E, Maouche S, Morris AD, Muller-Nurasyid M, Mu TC, Nikus K, Peden JF, Rayner NW, Rasheed A, Rosinger S, Rubin D, Rumpf MP, Schafer A, Sivananthan M, Song C, Stewart AF, Tan ST, Thorgeirsson G, van der Schoot CE, Wagner PJ et al (2013) Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 45:25–33. https://doi.org/10.1038/ng.2480

DeLuca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire MD, Williams C, Reich M, Winckler W, Getz G (2012) RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics 28:1530–1532. https://doi.org/10.1093/bioinformatics/bts196

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21. https://doi.org/10.1093/bioinformatics/bts635

Ewels P, Magnusson M, Lundin S, Kaller M (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Folkersen L, Kyriakou T, Goel A, Peden J, Malarstig A, Paulsson-Berne G, Hamsten A, Hugh W, Franco-Cereceda A, Gabrielsen A, Eriksson P (2009) consortia P Relationship between CAD risk genotype in the chromosome 9p21 locus and gene expression. Identification of eight new ANRIL splice variants. PLoS One 4:e7677. https://doi.org/10.1371/journal.pone.0007677

Freitag-Wolf S, Munz M, Wiehe R, Junge O, Graetz C, Jockel-Schneider Y, Staufenbiel I, Bruckmann C, Lieb W, Franke A, Loos BG, Jepsen S, Dommisch H, Schaefer AS (2019) Smoking modifies the genetic risk for early-onset Periodontitis. J Dent Res 98:1332–1339. https://doi.org/10.1177/0022034519875443

Garcia-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Gotz S, Tarazona S, Dopazo J, Meyer TF, Conesa A (2012) Qualimap: evaluating next-generation sequencing alignment data. Bioinformatics 28:2678–2679. https://doi.org/10.1093/bioinformatics/bts503

Harismendy O, Notani D, Song X, Rahim NG, Tanasa B, Heintzman N, Ren B, Fu XD, Topol EJ, Rosenfeld MG, Frazer KA (2011) 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. Nature 470:264–268. https://doi.org/10.1038/nature09753

Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, Jonasdottir A, Sigurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levey AI, Backman VM, Matthiasdottir S, Jonsdottir T, Palsson S, Einarsdottir H, Gunnarsdottir S, Gylfason A, Vaccarino V, Hooper WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgeirsson G, Thorsteinsdottir U, Kong A, Stefansson K (2007a) A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 316:1491–1493. https://doi.org/10.1126/science.1142842

Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, Sigurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levey AI, Backman VM, Matthiasdottir S, Jonsdottir T, Palsson S, Einarsdottir H, Gunnarsdottir S, Gylfason A, Vaccarino V, Hooper WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgeirsson G, Thorsteinsdottir U, Kong A, Stefansson K (2007b) A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 316:1491–1493

Holdt LM, Hoffmann S, Sass K, Langenberger D, Scholz M, Krohn K, Finstermeier K, Stahringer A, Wilfert W, Beutner F, Gielen S, Schuler G, Gabel G, Bergert H, Bechmann I, Stadler PF, Thiery J, Teupser D (2013) Alu elements in ANRIL non-coding RNA at chromosome 9p21 modulate atherogenic cell functions through trans-regulation of gene networks. PLoS Genet 9:e1003588. https://doi.org/10.1371/journal.pgen.1003588

Hubberten M, Bochenek G, Chen H, Hasler R, Wiehe R, Rosenstiel P, Jepsen S, Dommisch H, Schaefer AS (2019) Linear isoforms of the long noncoding RNA CDKN2B-AS1 regulate the c-myc-enhancer binding factor RBMS1. Eur J Hum Genet 27:80–89. https://doi.org/10.1038/s41431-018-0210-7

Ji J, Su L, Liu Z (2016) Critical role of calpain in inflammation. Biomed Rep 5:647–652. https://doi.org/10.3892/br.2016.785

Kim DH, Beckett JD, Nagpal V, Seman-Senderos MA, Gould RA, Creamer TJ, MacFarlane EG, Chen Y, Bedja D, Butcher JT, Mitzner W, Rouf R, Hata S, Warren DS, Dietz HC (2019) Calpain 9 as a therapeutic target in TGFbeta-induced mesenchymal transition and fibrosis. Sci Transl Med 11. https://doi.org/10.1126/scitranslmed.aau2814

Lee RT, Libby P (1997) The unstable atheroma. Arterioscler Thromb Vasc Biol 17:1859–1867. https://doi.org/10.1161/01.atv.17.10.1859

Letavernier E, Zafrani L, Perez J, Letavernier B, Haymann JP, Baud L (2012) The role of calpains in myocardial remodelling and heart failure. Cardiovasc Res 96:38–45. https://doi.org/10.1093/cvr/cvs099

Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P (2015) The Molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst 1:417–425. https://doi.org/10.1016/j.cels.2015.12.004

Love MI, Huber W, Anders S (2014) Moderated estimation of Fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8

Machiela MJ, Chanock SJ (2015) LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinform 31:3555–3557. https://doi.org/10.1093/bioinformatics/btv402

Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. Nat Rev Genet 10:155–159. https://doi.org/10.1038/nrg2521

Munz M, Richter GM, Loos BG, Jepsen S, Divaris K, Offenbacher S, Teumer A, Holtfreter B, Kocher T, Bruckmann C, Jockel-Schneider Y, Graetz C, Munoz L, Bhandari A, Tennstedt S, Staufenbiel I, van der Velde N, Uitterlinden AG, de Groot L, Wellmann J, Berger K, Krone B, Hoffmann P, Laudes M, Lieb W, Franke A, Dommisch H, Erdmann J, Schaefer AS (2018) Genome-wide association meta-analysis of coronary artery disease and periodontitis reveals a novel shared risk locus. Sci Rep 8:13678. https://doi.org/10.1038/s41598-018-31980-8

Munz M, Wohlers I, Simon E, Reinberger T, Busch H, Schaefer AS, Erdmann J (2020) Qtlizer: comprehensive QTL annotation of GWAS results. Sci Rep 10:20417. https://doi.org/10.1038/s41598-020-75770-7

Myocardial Infarction Genetics C, Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, Anand S, Engert JC, Samani NJ, Schunkert H, Erdmann J, Reilly MP, Rader DJ, Morgan T, Spertus JA, Stoll M, Girelli D, McKeown PP, Patterson CC, Siscovick DS, O'Donnell CJ, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Melander O, Altshuler D, Ardissino D, Merlini PA, Berzuini C, Bernardinelli L, Peyvandi F, Tubaro M, Celli P, Ferrario M, Fetiveau R, Marziliano N, Casari G, Galli M, Ribichini F, Rossi M, Bernardi F, Zonzin P, Piazza A, Mannucci PM, Schwartz SM, Siscovick DS, Yee J, Friedlander Y, Elosua R, Marrugat J, Lucas G, Subirana I, Sala J, Ramos R, Kathiresan S, Meigs JB, Williams G, Nathan DM, MacRae CA, O'Donnell CJ, Salomaa V, Havulinna AS, Peltonen L, Melander O, Berglund G, Voight BF, Kathiresan S, Hirschhorn JN, Asselta R, Duga S, Spreafico M, Musunuru K, Daly MJ, Purcell S, Voight BF, Purcell S, Nemesh J, Korn JM, McCarroll SA, Schwartz SM, Yee J, Kathiresan S, Lucas G, Subirana I, Elosua R, Surti A, Guiducci C, Gianniny L, Mirel D, Parkin M, Burtt N, Gabriel SB, Samani NJ, Thompson JR, Braund PS, Wright BJ, Balmforth AJ, Ball SG et al (2009) Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 41:334–341. https://doi.org/10.1038/ng.327

Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikainen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, de Andrade M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Huang J, Jalilzadeh S, Kessler T, Konig IR, Lannfelt L, Lieb W, Lind L, Lindgren CM, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FU, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF et al (2015) A comprehensive 1,000 genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 47:1121–1130. https://doi.org/10.1038/ng.3396

Rankin CR, Lokhandwala ZA, Huang R, Pekow J, Pothoulakis C, Padua D (2019) Linear and circular CDKN2B-AS1 expression is associated with inflammatory bowel Disease and participates in intestinal barrier formation. Life Sci 231:116571. https://doi.org/10.1016/j.lfs.2019.116571

Regan MC, Kirk SJ, Hurson M, Sodeyama M, Wasserkrug HL, Barbul A (1993) Tumor necrosis factor-alpha inhibits in vivo collagen synthesis. Surgery 113:173–177

Rekhter MD (1999) Collagen synthesis in atherosclerosis: too much and not enough. Cardiovasc Res 41:376–384. https://doi.org/10.1016/s0008-6363(98)00321-6

Rekhter MD, O'Brien E, Shah N, Schwartz SM, Simpson JB, Gordon D (1996) The importance of thrombus organization and stellate cell phenotype in collagen I gene expression in human, coronary atherosclerotic and restenotic lesions. Cardiovasc Res 32:496–502

Richter GM, Kruppa J, Munz M, Wiehe R, Hasler R, Franke A, Martins O, Jockel-Schneider Y, Bruckmann C, Dommisch H, Schaefer AS (2019) A combined epigenome- and transcriptome-wide association study of the oral masticatory mucosa assigns CYP1B1 a central role for epithelial health in smokers. Clin Epigenetics 11:105. https://doi.org/10.1186/s13148-019-0697-y

Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, Eliceiri KW (2017) ImageJ2: ImageJ for the next generation of scientific image data. BMC Bioinformatics 18:529. https://doi.org/10.1186/s12859-017-1934-z

Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B (2004) JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res 32:D91–D94. https://doi.org/10.1093/nar/gkh012

Sayols S, Scherzinger D, Klein H (2016) dupRadar: a Bioconductor package for the assessment of PCR artifacts in RNA-Seq data. BMC Bioinformatics 17:428. https://doi.org/10.1186/s12859-016-1276-2

Scaffidi AK, Mutsaers SE, Moodley YP, McAnulty RJ, Laurent GJ, Thompson PJ, Knight DA (2002) Oncostatin M stimulates proliferation, induces collagen production and inhibits apoptosis of human lung fibroblasts. Br J Pharmacol 136:793–801. https://doi.org/10.1038/sj.bjp.0704769

Schaefer AS, Richter GM, Groessner-Schreiber B, Noack B, Nothnagel M, El Mokhtari NE, Loos BG, Jepsen S, Schreiber S (2009) Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. PLoS Genet 5:e1000378

Smith EB (1965) The influence of age and atherosclerosis on the Chemistry of Aortic Intima.2. Collagen and Mucopolysaccharides. J Atheroscler Res 5:241–248. https://doi.org/10.1016/s0368-1319(65)80065-5

Statello L, Guo CJ, Chen LL, Huarte M (2021) Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol 22:96–118. https://doi.org/10.1038/s41580-020-00315-9

Tcheandjieu C, Zhu X, Hilliard AT, Clarke SL, Napolioni V, Ma S, Lee KM, Fang H, Chen F, Lu Y, Tsao NL, Raghavan S, Koyama S, Gorman BR, Vujkovic M, Klarin D, Levin MG, Sinnott-Armstrong N, Wojcik GL, Plomondon ME, Maddox TM, Waldo

SW, Bick AG, Pyarajan S, Huang J, Song R, Ho YL, Buyske S, Kooperberg C, Haessler J, Loos RJF, Do R, Verbanck M, Chaudhary K, North KE, Avery CL, Graff M, Haiman CA, Le Marchand L, Wilkens LR, Bis JC, Leonard H, Shen B, Lange LA, Giri A, Dikilitas O, Kullo IJ, Stanaway IB, Jarvik GP, Gordon AS, Hebbring S, Namjou B, Kaufman KM, Ito K, Ishigaki K, Kamatani Y, Verma SS, Ritchie MD, Kember RL, Baras A, Lotta LA, Regeneron Genetics C, Consortium CAD, Biobank J, Million Veteran P, Kathiresan S, Hauser ER, Miller DR, Lee JS, Saleheen D, Reaven PD, Cho K, Gaziano JM, Natarajan P, Huffman JE, Voight BF, Rader DJ, Chang KM, Lynch JA, Damrauer SM, Wilson PWF, Tang H, Sun YV, Tsao PS, O'Donnell CJ, Assimes TL (2022) Large-scale genome-wide association study of coronary artery disease in genetically diverse populations. Nat Med 28:1679–1692. https://doi.org/10.1038/s41591-022-01891-3

Thomas-Chollier M, Hufton A, Heinig M, O'Keeffe S, Masri NE, Roider HG, Manke T, Vingron M (2011) Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. Nat Protoc 6:1860–1869. https://doi.org/10.1038/nprot.2011.409

WTCCC (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661–678

Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou MM (2010) Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Mol Cell 38: 662–74. doi: S1097-2765(10)00335-7 [pii] https://doi.org/10.1016/j.molcel.2010.03.021

Yen YH, Pu CM, Liu CW, Chen YC, Chen YC, Liang CJ, Hsieh JH, Huang HF, Chen YL (2018) Curcumin accelerates cutaneous wound healing via multiple biological actions: the involvement of TNF-alpha, MMP-9, alpha-SMA, and collagen. Int Wound J 15:605–617. https://doi.org/10.1111/iwj.12904

Young MD, Wakefield MJ, Smyth GK, Oshlack A (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol 11:R14. https://doi.org/10.1186/gb-2010-11-2-r14

Zhao Y, Li H, Fang S, Kang Y, Wu W, Hao Y, Li Z, Bu D, Sun N, Zhang MQ, Chen R (2016) NONCODE 2016: an informative and valuable data source of long non-coding RNAs. Nucleic Acids Res 44:D203–D208. https://doi.org/10.1093/nar/gkv1252

Zhou X, Han X, Wittfeldt A, Sun J, Liu C, Wang X, Gan LM, Cao H, Liang Z (2016) Long non-coding RNA ANRIL regulates inflammatory responses as a novel component of NF-kappaB pathway. RNA Biol 13:98–108. https://doi.org/10.1080/15476286.2015.1122164

Zyla J, Marczyk M, Domaszewska T, Kaufmann SHE, Polanska J, Weiner J (2019) Gene set enrichment for reproducible science: comparison of CERNO and eight other algorithms. Bioinformatics 35:5146–5154. https://doi.org/10.1093/bioinformatics/btz447