Long Noncoding RNA \textit{ANRIL}: Lnc-ing Genetic Variation at the Chromosome 9p21 Locus to Molecular Mechanisms of Atherosclerosis

Lesca M. Holdt and Daniel Teupser*

Institute of Laboratory Medicine, University Hospital, LMU Munich, Munich, Germany

Ever since the first genome-wide association studies (GWAS) on coronary artery disease (CAD), the Chr9p21 risk locus has emerged as a top signal in GWAS of atherosclerotic cardiovascular disease, including stroke and peripheral artery disease. The CAD risk SNPs on Chr9p21 lie within a stretch of 58 kilobases of non-protein-coding DNA, containing the gene body of the long noncoding RNA (IncRNA) \textit{antisense non coding RNA in the INK4 locus (ANRIL)}. How risk is affected by the Chr9p21 locus in molecular detail is a matter of ongoing research. Here we will review recent advances in the understanding that \textit{ANRIL} serves as a key risk effector molecule of atherogenesis at the locus. One focus of this review is the shift in understanding that genetic variation at Chr9p21 not only affects the abundance of \textit{ANRIL}, and in some cases expression of the adjacent \textit{CDKN2A/B} tumor suppressors, but also impacts \textit{ANRIL} splicing, such that 3′-5′-linked circular noncoding \textit{ANRIL} RNA species are produced. We describe how the balance of linear and circular \textit{ANRIL} RNA, determined by the Chr9p21 genotype, regulates molecular pathways and cellular functions involved in atherogenesis. We end with an outlook on how manipulating circular \textit{ANRIL} abundance may be exploited for therapeutic purposes.

Keywords: lncRNA (long non-coding RNA), circRNA, GWAS (genome-wide association study), eQTL analysis, transcription, splicing, tumor suppressor proteins, cardiovascular diseases

INTRODUCTION

Since publication of the first genome-wide association studies (GWAS) of coronary artery disease (CAD) in 2007, Chr9p21 has emerged as the most significant risk locus associated with this frequent disease (1–4). The region contains a number of strongly interlinked SNPs within a stretch of 58 kilobases (kb) of non-protein-coding DNA. Later, the same haplotype block has been associated with other endpoints of atherosclerosis, such as stroke (5–11), peripheral artery disease (12–14), and also with different types of aneurysms (2, 8, 15, 16). Due to the availability of large study cohorts and the better resolution of genetic recombination in this region, it has now become clear that associations with other phenotypes at Chr9p21 fall in distinct haplotype blocks not overlapping with the CAD block (Figure 1A). Closely nearby, and proximal to the CAD locus, GWAS found associations with cancer, such as melanoma, glioma, basal cell carcinoma, and acute lymphoblastic leukemia [see (40) for review], and also with glaucoma, and diverse proliferative or
inflammatory diseases, such as endometriosis of the reproductive tract (41), periodontitis (42), and platelet reactivity (43). The region located distally to the CAD region contains a distinct haplotype block associated with type 2 diabetes (44, 45).

In the last 10 years, GWAS have been successfully used to increase the number of genetic loci implicated in CAD risk inheritance. The number of CAD risk loci in the genome rose from 56 by 2013 (24, 46–52) to 80 by 2015 (53–56), to 243 by 2017 (17). Concerning the Chr9p21 locus in these studies, the association rose steadily from $p = 5.40 \times 10^{-23}$ (rs4977575) (57), over $p = 4.68 \times 10^{-101}$ (rs4977574) (17) to $p = 8.8 \times 10^{-223}$ (rs4977574) (58). In populations of European descent, the allele frequency is very high (0.48), leading to the situation that approximately one-fourth of people are homozygous for the CAD risk alleles. CAD risk SNPs on Chr9p21 have recurrently been shown to have one of the top-ranking effect sizes [allelespecific odds ratio (OR) for CAD > 1.3] (3, 24). Despite the extent of effects, the Chr9p21 risk is independent of classically known CAD risk determinants, such as dyslipidemia, diabetes mellitus, age, and sex.

The Chr9p21 region contains at least 5 genes, which are, in part, tightly clustered and overlapping. These include the 3.8 kb long ANRIL non-coding RNA, and the tumor suppressors cyclin dependent kinase inhibitor CDKN2A/p16INK4A, CDKN2A/p14ARF, CDKN2B/p15INK4B, and methylthioadenosine phosphorylase (MTAP). ANRIL overlaps in antisense the full length of the p15 gene body, while sharing a bidirectional promoter with CDKN2A. Hence, it was also termed CDKN2B antisense RNA (CDKN2B-AS1). Only recently, the picture got even more complex: Advances in high-throughput sequencing and adaption in bioinformatics mapping of RNA reads to reference genomes have revealed that thousands of genes in our genome produce not only mature linear RNA but also 3’-5’ covalently linked circular RNAs (circRNAs) (59). So far, two studies have shown that a number of circular ANRIL (circANRIL) isoforms exist, comprised of different exons, whereby a downstream exon is fused to an upstream exon by the enzymatic activity of the spliceosome in a reaction termed “backsplicing” [see (60, 61) for review]. Circularizing exons in ANRIL stemmed mostly from middle parts of the lncRNA (Figure 1A), which are in part also shared by the linear ANRIL isoforms. CircANRIL was found not only in many different cell lines, but also in many primary cell types, including vascular smooth muscle cells (VSMCs) and macrophages, as well as in heart and vascular tissue (22, 36).

A major focus in exploring how risk is effected by Chr9p21 has been on whether genetic variation affected expression of genes at the locus in cis (Figure 1A) or whether it elicited gene expression changes in trans. Top CAD-associated SNPs lie within the distal parts of long linear ANRIL isoforms (Figure 1A) and several studies have shown that they co-localize with sequences marked by chromatin modifications, RNA polymerase II transcription patterns and DNA motifs characteristic of bona-fide transcriptional enhancers (19, 35, 62–65). Using expression quantitative trait locus (eQTL) analyses in patient samples, several groups have by now investigated if the risk alleles at the locus were associated with the expression of specific target genes in cis (cis-eQTLs). Whereas studies investigating ANRIL expression have mostly used quantitative PCRs (qPCRs) targeting different exons from the lncRNA, expression of p14, p15, p16, or MTAP has either been investigated using genome-wide expression arrays or isoform-specific qPCRs. Here, we focus on studies investigating eQTLs in atherosclerosis cohorts but do not cover studies related to other phenotypes, such as cancer, which are reviewed elsewhere (66).

CIS-eQTLs AT Chr9p21

ANRIL expression at Chr9p21 is complex and at least 20 linear isoforms as well as multiple circular isoforms have been reported [www.ensembl.org, (22, 36, 39)]. In principal, linear and circular isoforms can be distinguished by the fact that the latter derive from a backsplice event, where splicing of a downstream exon (e.g., exon 7) to an upstream exon (e.g., to exon 5) can be detected. Backsplicing of ex7-5 was the most common event observed in our own study in peripheral blood monocytes (36). Concordantly, Burd and colleagues have reported dominant backsplicing isoforms spanning ex14-4 in peripheral blood T lymphocytes (22). In both studies, exon 1 and exons 17-20 were not contained in circularized ANRIL (Table 1). Thus, for classification reasons, results from studies targeting these exons will be referred to as proximal linear isoforms (containing the first ANRIL exons) and long linear isoforms (containing the distal exons 17-20) (Table 1). Since both linear and circular ANRIL may contain exons from the middle portion of the lncRNA (e.g., exons 4-16), a clear distinction as to whether linear or circular isoforms were investigated cannot be made in cases where these exons were targeted by qPCRs which were non-specific for backsplice junctions (Table 1).

As one of the first studies on Chr9p21, Jarinova et al. have shown that ANRIL expression was induced by the CAD risk SNP rs1333049 in peripheral blood monocytes (PBMCs). No significant effects on CDKN2A or on CDKN2B were recorded in that study (19). Over the years, comparable quantifications of these genes followed in whole blood, peripheral blood T lymphocytes, lymphoblastoid cells lines, aortic smooth muscle cells (SMCs) and in different tissue samples that are known to have a role in atherosclerosis. For example, vascular tissues such as carotid atherosclerotic plaque samples, samples from aorta, mammary artery, and from the heart ventricles have been analyzed, but also tissues like subcutaneous or omental fat have been used (Table 1). Of the 23 cis-eQTL studies conducted in the Chr9p21 CAD region to date, 16 investigated different isoforms of ANRIL, out of which 10 used assays targeting proximal ANRIL exons, 8 used assays targeting the middle region, 6 used assays targeting downstream linear ANRIL exons, and two investigated backsplices contained in circANRIL (Table 1). Complicating a clear-cut interpretation, in the different studies, different risk genotypes were used to indicate risk haplotypes. The expression of CDKN2A and of CDKN2B was investigated in 18 studies and MTAP in 10 studies (Table 1).

Overall, 80% of the studies investigating ANRIL expression found an association with the Chr9p21 genotype. Here, a trend
FIGURE 1 | ANRIL and the CAD risk locus at Chr9p21. (A) Genomic mapping of SNPs identified in a region ± 300 kb around the top-ranking sentinel SNP rs4977574 based on data of the most recent large CAD GWAS (17). Chromosome ideogram and zoom-in onto RefSeq transcripts for ANRIL, CDKN2A, CDKN2B, and MTAP.
higher expression of the proximal and distal exons contained in linear ANRIL in patients carrying the CAD-risk allele was observed (7 of 10 and 3 of 6 studies). In contrast, circular ANRIL was downregulated in the two published studies in patients carrying the Chr9p21 risk haplotype. No clear tendency was observed when assays targeting the middle region of ANRIL were used (Table 1). This is likely explained by the fact that these assays target both linear and circular, ANRIL isoforms, which seem to be inversely regulated. With respect to the tumor suppressor genes contained at the Chr9p21 locus, 78 and 67% of the studies failed to find an association of CDKN2A and CDKN2B with Chr9p21, respectively. When reporting an association, specifically CDKN2B was down-regulated in the majority of studies (94%), yet its expression was not always anticorrelating with ANRIL expression (19, 21, 22, 29). MTAP expression was not associated with the Chr9p21 genotype in any of the published studies. Overall, the picture emerges that circular ANRIL and CDKN2B tend to be down-regulated in patients carrying the risk allele, whereas linear ANRIL isoforms tend to be inversely regulated (Figure 1B). It is currently unclear, why expression of p16 or of p14 and p16 were in many cases positively correlated with ANRIL (19, 21, 22, 27, 29, 32, 34, 65). Also, MTAP, which was not associated with Chr9p21 (Table 1), was in some conditions anticorrelating to ANRIL, but not in all cases or contexts (20, 34, 67). SNPs in ANRIL can hypothetically affect enhancers in both directions, either by disrupting transcription factor binding sites in open chromatin (68) or by increasing enhancer activity through yet unknown primary effects (24, 65).

In summary, many studies document cis-eQTLs for ANRIL or, separately, for CDKN2B (35). Throughout, from the existing data, it can be concluded that these effects are cell-type specific and combinatorial. Of note, many studies have investigated only very small cohorts and those, simultaneously testing both ANRIL and CDKN2B in larger cohorts (>1000 samples) identified much stronger effects of Chr9p21 on ANRIL than on CDKN2B (13, 33, 36). This observation might be explained by the haplotype block structure of the region, where effects of CAD lead SNPs are located within ANRIL, but bleed through due to linkage disequilibrium, resulting in more subtle concomitant effects on CDKN2B expression. Another possibility is that the Chr9p21 genotype impacts transcription enhancers at the locus which contact and activate gene promoters affecting CAD. The consequences of such contacts would not be expected to be captured through traditional non-allelic RNA expression analysis. In fact, when allelic expression control through 3D-enhancer looping was specifically measured in a separate study in human coronary aortic SMCs (64), physical contacts of CAD variant-containing enhancers in the locus and the promoters of CDKN2A, CDKN2B, and ANRIL were corroborated.

Taken together, these data suggest that genetic variation within the core 9p21 CAD region relates to differential expression not only of ANRIL, but in specific cells or conditions, also of the CDKN2A/B tumor suppressors encoded in the locus. While either of these factors could potentially increase cell proliferation, or lead to unscheduled senescence, or elicit out of context inflammatory signaling, as far as based on work with cells in vitro, no study in humans or in mouse models has been able to decisively implicate a downstream effector pathway in vivo.

**TRANS-eQTLs AT Chr9p21 AND MOLECULAR FUNCTIONS OF ANRIL IN TRANSCRIPTIONAL REGULATION**

As opposed to cis effects, two eQTL studies have so far detected modest and tissue-selective differential expression of dozens of genes associated with Chr9p21 genotype with genome-wide significance (19, 27). Affected genes were from a broad range of classes (AVPR2, PEAK1, FBLN1, KALRN, DAZL, STA2U, HLA-DQA1, BTNL8, PLEKHA6, TDGF1) in whole blood (19) and different, non-overlapping gene sets linked to tissue wounding, cell migration and inflammatory response, when analyzing heart tissue, plaques, aortas, and arteries (27).

Other, and in part, larger studies in vascular tissue (20), peripheral blood mononuclear cells (PBMC, n = 2280) (33) and in blood monocytes (n = 1490) (23) reported no significant expression association.

Though not directly comparable, another study showed that in macrophages cultured in vitro under stress-bearing IFNy and LPS stimulation, the CAD risk genotype led to differential up- and downregulation of target genes outside the Chr9p21 locus and yet distinct from the previously mentioned studies (IL1B, IL12B, CASP5, CCL8, MT1A, MT1E, MUCL1, TNIP3, VCAN, ENPP2, NDP, CD163) (30). Also ANRIL knockdown in cultured cell lines (69–72) and overexpression of linear ANRIL affected the expression of non-overlapping gene sets in the genome in trans (33, 36).

How ANRIL exerts trans-regulation is not known, and despite a study that showed a physical interaction of ANRIL with promoters of target genes (33), this role is likely not a classical function as enhancer RNA [eRNA (73)], because it involved both up- and down-regulated genes, and was suggested to involve sequence homology (33). In the case of ANRIL, trans-regulation of target genes was ascribed to an ALU motif.
### TABLE 1 | Chr9p21 cis-eQTLs in patients.

| References | CAD SNPs | Cells or tissues | Sample size | Assay | Proximal exons 1-5* | Middle exons 4-16 | Distal exons 17-20 | CDKN2B | CDKN2A | MTAP |
|------------|----------|-----------------|-------------|-------|-------------------|------------------|------------------|--------|--------|-------|
| (18)       | rs10757278 | Peripheral blood T cells (healthy) | 170         | qPCR  | ex1-5             | ex13             | n.d.             | n.d.   | n.d.   | n.d.  |
| (19)       | rs10757272, rs4977574, rs2891168, rs1330049, rs1333049, rs10757278, rs10111277 | Whole blood (healthy) | 120         | qPCR  | ex1-5             | ex13             | n.d.             | n.d.   | n.d.   | n.d.  |
|            |          | Whole blood (CAD) | 42          | Microarray | n.d.             | ex13             | ex15-16          | n.d.   | n.d.   | n.d.  |
| (20)       | rs2891168, rs2383207, rs2383206, rs1330049, rs1333049, rs10757278, rs10111277 | Lymphoblastoid cell lines from HapMap | 233         | Microarray | ex1-3             | ex6,7,14         | n.d.             | ex17-19 | n.d.   | n.d.  |
|            |          | Vascular tissue | 294         | Microarray | ex1-3             | ex6,7,14         | n.d.             | ex17-19 | n.d.   | n.d.  |
| (13)       | SNPs rs10757274, rs2383206, rs10757278, rs10736005 | PBMCs (healthy, CAD) | 1098        | qPCR  | ex1-5             | ex4-5            | n.d.             | ex18-19 | n.d.   | n.d.  |
|            |          | Whole blood (healthy, CAD) | 769         | qPCR  | ex1-5             | ex4-5            | n.d.             | ex18-19 | n.d.   | n.d.  |
|            |          | Vascular tissue | 41          | qPCR  | ex1-5             | ex4-5            | n.d.             | ex18-19 | n.d.   | n.d.  |
| (21)       | rs10757274, rs10757278, rs1330049, rs2383206, rs3217992, rs704859, rs4966892, rs564398 and others | Whole blood (healthy) | 487         | qPCR  | ex1-2             | n.d.             | n.d.             | n.d.   | n.d.   | n.d.  |
|            |          | Peripheral blood T cells (healthy) | 106         | qPCR  | ex1-2             | n.d.             | ex14-5 circular  | ex18-19 | n.d.   | n.d.  |
| (22)       | rs1075727 | PBMCs (healthy/subclinical CAD) | 1490        | Microarray | n.d.             | n.d.             | n.d.             | n.d.   | n.d.   | n.d.  |
| (23)       | rs4977574 | Omental adipose tissue (healthy, CAD) | 2430        | Microarray | n.d.             | n.d.             | n.d.             | n.d.   | n.d.   | n.d.  |
| (24)       | rs10757274, rs2383206, rs2383207 | Vascular tissue (CAD) | 57          | qPCR  | n.d.             | n.d.             | n.d.             | n.d.   | n.d.   | n.d.  |
| (25)       | rs4977574 | aortic SMCs (healthy) | 79          | qPCR  | n.d.             | n.d.             | n.d.             | n.d.   | n.d.   | n.d.  |
| (26)       | rs1330049 | Heart (healthy/subclinical) | 108         | Microarray | ex1-2             | n.d.             | n.d.             | n.d.   | n.d.   | n.d.  |
|            |          | Combined analysis of myocardium, vascular tissues (healthy, CAD) | 406         | Microarray | ex1-2             | n.d.             | n.d.             | n.d.   | n.d.   | n.d.  |

(Continued)
| References | CAD SNPs | Cells or tissues | Sample size | Assay | ANRIL 17 linear and dozens of circular isoforms | CDKN2B | CDKN2A | MTAP |
|------------|----------|-----------------|-------------|-------|-----------------------------------------------|--------|--------|------|
| (28)       | rs10757274 | whole blood (healthy, CAD) | 205 | qPCR | ex1-5 | n.d. | n.d. | ex17-18 | n.d. |
|            |           |                 |             |       | ex4-5 | n.d. | n.d. | ex17-18 | n.d. |
| (29)       | rs496892, rs564398, rs6475606, rs1063192, rs10811650, rs10738607 (prox. ANRIL); rs10965235, rs3731217 (dist. ANRIL); rs20835028 (CDKN2B); rs564398, rs10965224 (short ANRIL); rs2811712, rs7855162 (dist. ANRIL); rs10965235 (CDKN2A); rs1333049 | Whole blood (healthy, CAD) | 57 | qPCR | ex1-2 | n.d. | n.d. | ex17-18 | n.d. |
| (30)       | rs7044859, rs496892, rs7865618, rs1333049 | PBMCs differentiated to macrophages (healthy, CAD) | 68 | microarray, qPCR | n.d. | n.d. | n.d. | n.d. | n.d. |
| (31)       | rs1333049 | Primary VSMCs (arterial from umbilical cords, healthy) | 69 | qPCR, Western, histochemistry | n.d. | ex14-15 | n.d. | n.d. | n.d. |
| (32)       | rs944799 (ex12-13 ANRIL); rs94479, rs598664, rs1002878 (ANRIL); rs598664 (CDKN2B); rs10965228 (ANRIL) | PBMCs (healthy, CAD) | 281 | qPCR | n.d. | ex12-13 | n.d. | n.d. | n.d. |
| (33)       | rs10757274, rs2383206, rs10757278 | PBMCs (healthy, CAD); Whole blood (healthy, CAD); Vascular tissues | 1933 | microarray, qPCR | ex 1-5 | ex7-13 | n.d. | ex18-19 | n.d. |
| (34)       | rs10757278, rs10757274, rs2383206, rs1333049, rs4977574 | Whole blood; plasma/healthy, CAD | 200 | qPCR (blood), ELISA (plasma) | ex1-5 | n.d. | n.d. | n.d. | n.d. |

Continued...
| References | CAD SNPs | Cells or tissues | Sample size | Assay | ANRIL 17 linear and dozens of circular isoforms | CDKN2B | CDKN2A | MTAP |
|------------|----------|-----------------|-------------|-------|-----------------------------------------------|-------|-------|------|
|            |          |                 |             |       | Proximal exons 1-5* | Middle exons 4-16 | Distal exons 17-20 | p15 | p16 | p14 |
| (36)       | rs10757274, rs2383206, rs2383207, rs10757278 | PBMCs (healthy, CAD) | 1933 | Microarray, qPCR | n.d. | n.d. | ex7-5 circular | n.d. | #   | n.d. |
|            |          | Whole blood (healthy, CAD) | 1933 | n.d. | n.d. | ex7-5 circular | n.d. | n.d. | n.d. |
|            |          | Vascular tissues | 218 | n.d. | n.d. | ex7-5 circular | n.d. | n.d. | n.d. |
| (37)       | rs1537371, rs1333040 | Vascular tissues, fat, muscle, blood (CAD) | 600 | RNAseq | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. |
| (15)       | rs10965215, rs10738605 | PBMCs (healthy, CAD) | 66 | qPCR | n.d. | ex7 | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. |
| (17)       | rs1537371 | Aorta (CAD) | 600 | RNAseq | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. |
| Total number of relevant studies | 10 | 8 | 2 | 6 | 19 | 18 | 10 |
| % of studies showing upregulation | 70 | 37 | 0 | 50 | 11 | 6 | 0 |
| % of studies showing downregulation | 30 | 50 | 100 | 17 | 47 | 17 | 0 |
| % of studies showing unchanged expression | 40 | 63 | 0 | 67 | 63 | 78 | 100 |

CAD-relevant SNPs encoded in Chr9p21 and defined as cis-eQTLs for regulation of ANRIL, p15, p16, p14, and/or MTAP through gene expression profiling with microarrays or targeted qPCR analysis. Upregulation of target gene expression (green), downregulation (red), no effect on RNA abundance (gray), effect not determined (n.d.). Effects on different isoforms of ANRIL (linear or circular) is specified when exon-specific qPCR analysis was performed targeting proximal, middle or distal ANRIL exons (ex). In total 23 studies are reviewed here. References: (13, 17–38). Note that the percentages (%) of studies showing up- and down-regulation of ANRIL isoforms do not necessarily add up to 100%, because different ANRIL exons were quantified in different studies, not all classes of ANRIL transcripts were analyzed in each study, and because one study can report on both up- and down-regulation of different isoforms belonging to the same class of ANRIL transcripts (proximal/middle/distal).

*(ANRIL linearity determined by PCR forward primer residing in exons 1, 2, or 3). *(ANRIL circularity determined by PCR primers detecting backsplicing between exons 10-2, 5-intron3, 6-intron3, 6-4, 7-4, 14-4, 10-4, 12-4, 13-4, 14-4, 16-4, 6-5, 7-5, 8-5, 10-5, 14-5 16-5, 19-5, 7-6, 10-6, 14-6, 16-13, 16-19 (22, 36, 39).

# (cis-eQTL locating in enhancer element, but with unspecified direction of effect on expression).
in both ANRIL and the target gene promoters (33). Similarly, an independent study found that ANRIL did not only silence its targets, but unexpectedly also upregulated target genes: For example, proinflammatory interleukins IL6/8 were found to be co-stimulated by ANRIL and YY1, a transcription-regulating factor that bound to the ANRIL RNA, especially in the context of TNFa/NFkB signaling (70). Therefore, opposite to what could have been expected from the reported physical interaction of ANRIL with proteins from the repressive Polycomb group complexes (74), ANRIL might be an activator, at least for some trans-regulated genes (33, 70) (see chapter 4 for details). Whether circANRIL, beyond regulating rRNA maturation, is involved in primary transcriptional control, alone or via impacting linear ANRIL’s function, is not known (36). Nevertheless, it is interesting to note that circANRIL isoforms linked to CAD are produced from exons located in the middle of the ANRIL gene (22, 36), and as such do not include the ALU motif, which is important for gene trans-regulation by linear ANRIL and is located more distally in the gene (33). Thus, variation in ANRIL RNA at the molecular level (linear vs. circular) might impose a fundamental alteration in ANRIL effector function, while not offering any explanation per se on how linear ANRIL regulates genes, as scaffold for promoter-activating complexes, or as decoy/inhibitor of repressive chromatin-modifying complexes. Conservatively speaking, it seems possible that Chr9p21 CAD risk genotypes affects genomic expression both in cis and in trans, and linear ANRIL RNA may be one, but not the sole, important effector molecule for how the Chr9p21 locus transduces such effects (Figure 1B).

**CORRELATION OF CHR9P21 GENES WITH ATHEROSCLEROSIS SEVERITY IN HUMANS AND MOUSE MODELS**

Another piece of evidence for a functional role of ANRIL in determining CAD risk stems from correlation analysis with disease features in patient cohorts. Aside of the genetic association, ANRIL levels were often increased in CAD patients, and not only in atherosclerotic plaque tissue, but also in circulating PBMCs or whole blood. Here, linear ANRIL levels were positively correlated with the severity of atherosclerosis (13, 29, 75) whereas circANRIL was anticorrelated (36) (Figure 1B). Thus, while the genotype of Chr9p21 determines the production of atherogenic (linear) over antiatherogenic ANRIL RNA species (circular), CAD and peripheral artery disease-dependent changes may additionally feed into ANRIL regulation. For CDKN2B, two studies reported a correlation of the expression with atherosclerosis severity (34, 76), where the direction of the correlation (downregulation in plaques) was consistent from what could be expected from the association results. But another study reported increased p16INK4A levels to positively correlate with inflammation markers in plaques instead of anticorrelation (25). Together, results from association as well as correlation analyses have established ANRIL IncRNA as prime candidate at the Chr9p21 locus.

**MOLECULAR FUNCTION OF ANRIL AND CDKN2A/B IN ATHEROGENESIS**

ANRIL belongs to the group of long non-coding RNAs and as such has been suggested to act as a molecular scaffold of chromatin-modifying complexes that control gene expression through modifying histone tails. Specifically, ANRIL was found to physically interact with the CBX7 protein inside the PRC1 Polycomb complex, one of the major gene repression complexes in cells (74). Knockdown of members of this Polycomb group complex led to increased expression of the CDKN2A and CDKN2B tumor suppressors in the Chr9p21 locus. Also, ongoing RNA polymerase II transcription was important for the association of the Polycomb proteins with the locus, indicative of the importance of RNA for recruitment. It was concluded that ANRIL’s function may be, at least in part, to repress the CDKN2A and CDKN2B tumor suppressors. As a consequence increased ANRIL levels are thought to promote overproliferation and to be incompatible with senescence onset, a major function of CDKN2A/B. As described in chapter 2, other work has shown that recruitment of the Polycomb complexes may account also for how ANRIL regulates genes in trans on a genome-wide level: Overexpression of linear ANRIL isoforms in cultured cells was found to promote pro-atherogenic cell functions, such as proliferation and reduced apoptosis, and to trigger the differential expression of hundreds of genes, in this case without affecting CDKN2A/B suppressors. Results from that study therefore questioned whether ANRIL regulated these tumor suppressor genes in cis at all (36, 77).

How does circular ANRIL, whose abundances is reduced in CAD patients, fit into this model? Both in human peripheral blood T-lymphocytes, as well as in PBMCs, whole blood and endarterectomy plaque tissue, circANRIL isoforms were found to be downregulated in samples from CAD patients carrying the Chr9p21 risk allele (22, 36). In an initial model, it was suggested that the production of circANRIL from central ANRIL exons would shortend the linear ANRIL IncRNA and, thereby, impaired linear ANRIL’s function in epigenetic control of target genes (22). In a second study, a more primary role was found for circANRIL that was, furthermore, independent of linear ANRIL (36). Here, circANRIL was found to be 10-fold more abundant than linear ANRIL. Mass-spectrometric analysis of proteins interacting with circANRIL showed that it bound to PES1 protein, a member of the evolutionarily conserved PeBoW complex. This complex is essential for proper rRNA-processing, that is the excision of RNA spacer elements from pre-ribosomal rRNA precursors. CircANRIL inhibited the activity of the PeBoW complex, as deduced from the accumulation of insufficiently processed (and non-functional) 26S and 32S pre-rRNA intermediates when circANRIL was overexpressed (36). A deficit in rRNA maturation caused nucleolar stress and p53 activation, culminating in inhibition of cell proliferation and in an increase in apoptosis. Notably, the observed functions of circANRIL were inverse to that of linear ANRIL and, as shown by genomic knockout of linear ANRIL exons, independent from the presence of these IncRNA isoforms. Thus, experimental evidence from expression analysis in vivo and from genetic experiments both indicated
that circANRIL was anti-atherogenic. Together, linear ANRIL
congrues overproliferation, and circular ANRIL protects from
overproliferation, suggesting that the genotype of Chr9p21 is
important to determine the balance of linear and circular ANRIL
levels in SMCs and macrophages, and that a dominance of linear
ANRIL in this ratio, even when small, over decades skews for
CAD (36) (Figure 1B).

Whether suppressing linear ANRIL or boosting circularization
is sufficient to protect from atherosclerotic cues in vivo is
matter of ongoing research. The fact that ANRIL RNA is
not conserved beyond primates complicates the functional in vivo
analysis of the Chr9p21 locus. So far, insight on how
CAD is controlled by Chr9p21 through genetic modeling
in mouse mutants is fragmented. The genetic elements of
Chr9p21 and their relative positioning are overall syntenically
conserved in mouse chromosome 4. So far, only one study
has investigated, if deletion of a 70 kb long portion of mouse
Chr4 corresponding to the CAD haplotype block in humans
had an effect on atherosclerosis in vivo (78). This region
contains a multi-exon lncRNA, AK148321, which is, however,
likely not corresponding to human ANRIL. Mutant mice (78)
developed tumors, reminiscent of tumorigenesis associated with
mutation in the Chr9p21 region. But despite some metabolic
changes in the mutant mice and enhanced platelet activation,
no significant change in atherosclerotic fatty lesion formation
was observed (78), putting in question the validity of this mouse
model for studying ANRIL-driven aterogenesis. On the other
hand, the mutants did develop more vascular aneurysms (79),
supporting that some aspects of CAD were indeed contained in
the noncoding mouse sequence.

Overall, the picture is not yet fully clear. While the genetic
data from mice support the importance of individual noncoding
gene regulatory elements and of some of the protein-coding tumor
suppressors for regulation of atherosclerosis and other CAD
entities, whether the lncRNA encoded in the locus regulates CAD
mechanistically via epigenetic regulating the neighboring
tumor suppressors in cis has not been determined. Nevertheless,
mouse genetics remains an interesting research avenue to explore
some aspects of Chr9p21 biology, at least relating to aneumus,
cancer, and glaucoma formation.

SUMMARY

Starting from a GWAS signal for CAD in a “gene desert”
on Chr9p21 in 2007, research in the last decade has firmly
established this region as strongest genetic factor of human
atherosclerosis and has contributed to a better understanding of
the underlying pathophysiology. The picture has emerged that
one of the major routes how this locus controls atherosclerosis
risk is through regulating the expression of the lncRNA ANRIL
in cis, where the risk allele leads to high levels of linear ANRIL
but decreases circular ANRIL expression. Linear ANRIL has
been established as molecular scaffold guiding epigenetic protein
complexes and promoting pro-atherogenic cells functions. On
the contrary, circularization shifts ANRIL function toward
controlling ribosomal RNA processing and controlling protein
translation thereby promoting athero-protection (Figure 1B).
The molecular mechanisms of how the ratio of linear and
circular ANRIL is controlled by the genotype at the locus are
currently not resolved and it will be important to determine
which gene regulatory elements within the ANRIL gene are
caused by stular CAD risk SNPs. Experimentally exploring
details of the molecular effector mechanisms for linear ANRIL
and for circular ANRIL will be paramount, but this task will
not be trivial because linear and circular ANRIL isoforms
always co-exist and in part share the same sequence. Not last,
more nuanced relations between Chr9p21 genotype and gene
expression output can be expected to be found in the future if,
for example, analyses were to take into account cell type-specific
and context (stress, inflammation, senescence)-specific effects,
aspect that whole tissue expression profiling is currently missing.
Additionally, although it is early days, measuring the levels of
circANRIL/linear ANRIL, might offer a prognostic value and
help improve CAD risk stratification or allow to better monitor
treatment response or disease recurrence. Yet, since circANRIL
levels are reduced in plaque tissue, and since circANRIL has been
found to be anti-atherogenic with or without co-existing linear
ANRIL, increasing circANRIL abundance in patients could also
be of therapeutic relevance. Expressing circANRIL levels in the
cells of the vasculature in CAD disease models might, therefore,
be a promising next step to exploit the accumulated knowledge
on the Chr9p21 CAD risk locus.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual
contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

We thank Bernd Northoff for bioinformatics analyses and his aid
in preparing the Figure. This work was in part funded by the
German Research Foundation (DFG) as part of the Collaborative
Research Center CRC1123 Atherosclerosis–Mechanisms and
Networks of Novel Therapeutic Targets (project B1) and by the
Leducq-foundation CADgenomics.

REFERENCES

1. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, et al.
A common allele on chromosome 9 associated with coronary heart disease.
Science (2007) 316:1488–91. doi: 10.1126/science.1142447
2. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, et al. A common variant on chromosome 9p21
affects the risk of myocardial infarction. Science (2007) 316:1491–3.
doi: 10.1126/science.1142842
3. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al.
Genomewide association analysis of coronary artery disease. N Engl J Med.
(2007) 357:443–53. doi: 10.1056/NEJMoa072366
4. Wellcome Trust Case Control C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000
shared controls. *Nature* (2007) 447:661–78. doi: 10.1038/nature05911

5. Larson MG, Atwood LD, Benjamin EJ, Cupples LA, D’Agostino RBSr, Fox CS, et al. Framingham Heart Study 100K project: genome-wide associations for cardiovascular disease outcomes. *BMJ Med. Genet.* (2007) 8 Suppl 1:55. doi: 10.1136/bmj.h145

6. Matarin M, Brown WM, Singleton A, Hardy JA, Narkiewicz K, Almgren P, et al. The myocardial infarction associated CDKN2A/CDKN2B locus on chromosome 9p21 is associated with stroke independently of coronary events in patients with hypertension. *J. Hypertens.* (2009) 27:769–73. doi: 10.1097/1HJH.0b013e32832f67eb

7. Wahlstrand B, Orho-Melander M, Delling L, Kjeldsen S, Narkiewicz K, Almgren P, et al. The myocardial infarction associated CDKN2A/CDKN2B locus on chromosome 9p21 is associated with stroke independently of coronary events in patients with hypertension. *J. Hypertens.* (2009) 27:769–73. doi: 10.1097/1HJH.0b013e32832f67eb

8. Helgadottir A, Thorleifsson G, Magnussson KP, Gretarsdottir S, et al. Functional analysis of the chromosome 9p21.3 coronary artery infarction risk. *Nat Genet.* (2008) 40:217–24. doi: 10.1038/ng.72

9. Gschwendtner A, Bevan S, Cole JW, Plourde A, Matarin M, Ross-Adams H, et al. Sequence variants on chromosome 9p21 confer risk for atherosclerotic stroke. *Ann. Neurol.* (2009) 65:531–9. doi: 10.1002/ana.21590

10. International Stroke Genetics C, Wellcome Trust Case Control C, Bellenguez M, Bevan S, Gschwendtner A, Spencer CC, et al. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat. Genet.* (2012) 44:328–33. doi: 10.1038/ng.1081

11. Dichgans M, Malik R, Konig IR, Rosand J, Clarke R, Gretarsdottir S, et al. Association between chromosome 9p21 variants and the ANRIL gene correlated with its expression contribute to myocardial disease risk locus. *PLoS ONE* (2010) 5:e100899. doi: 10.1371/journal.pone.0100899

12. Cluett C, McDermott MM, Guraká I, Ferrucci L, Bandinelli S, Miljovic I, et al. The same sequence variant on chromosome 9p21.3 coronary artery disease risk locus in Asian Indians. *Thromb Haemost.* (2014) 111:960–9. doi: 10.1160/CIRCULATIONAHA.111.064881

13. Zollbricht C, Gras S, Henner B, Lebeck K, et al. Expression pattern in human macrophages dependent on 9p21.3 coronary artery disease risk locus. *Atherosclerosis* (2013) 227:244–9. doi: 10.1016/j.atherosclerosis.2012.12.030

14. Motterle A, Pu X, Wood H, Xiao Q, Gor S, Ng FL, et al. Functional analyses of coronary artery disease associated variation on chromosome 9p21 in vascular smooth muscle cells. *Hum Mol Genet.* (2012) 21:4021–9. doi: 10.1093/hmg/dds224

15. Varshney A, Barlow GM, Gao X, Qiu H, et al. Resequencing and clinical associations of the 9p21.3 region: a comprehensive investigation in the Framingham heart study. *Circulation* (2013) 127:799–810. doi: 10.1161/CIRCULATIONAHA.112.111559

16. Zou Y, Wang L, Li H, Wang W, Wei Y, Li W, et al. Methylation of p15INK4b and expression of ANRIL on chromosome 9p21 are associated with coronary artery disease. *PLoS ONE* (2017) 12:e0179193. doi: 10.1371/journal.pone.0179193

17. Stengel BM, Shih J, Wang J, et al. Cardiometabolic risk loci share downstream cis- and trans- regulation of gene networks. *PLoS Genet.* (2013) 9:e1003588. doi: 10.1371/journal.pgen.1003588

18. Shanker J, Arvind P, Jambunathan S, Nair J, Kakkar V. Genetic analysis of the 9p21.3 CAD risk locus in Asian Indians. *Thromb Haemost.* (2014) 111:960–9. doi: 10.1080/08957061.2013.766825

19. Miller CL, Pijan M, Wang T, Nguyen T, Cohnain A, Lee JD, et al. Integrative functional genomics identifies regulatory mechanisms at coronary artery disease locus. *Nat Commun.* (2016) 7:12079. doi: 10.1038/ncomms12079

20. Kostek C, Kao CL, et al. Circular non-coding RNA ANRIL modulates ribosomal RNA transcript expression is associated with chromosome 9p21 variants linked to atherosclerosis. *PLoS ONE* (2014) 9:e105207. doi: 10.1371/journal.pone.0105207
Sarkar D, Oghabian A, Bodiyabalu PK, Joseph WR, Leung EY, Finlay GJ, et al. Multiple isoforms of ANRIL in melanoma cells: structural complexity suggests variations in processing. *Int J Mol Sci.* 2018:1578. doi: 10.3390/ijms18071378

40. Jeck WR, Siebold AP, Sharpless NE. Review: a meta-analysis of GWAS and age-associated diseases. *Aging Cell* (2012) 11:727–31. doi: 10.1111/j.1474-9726.2012.00871.x

41. Nakaoka H, Gurumurthy A, Hayano T, Ahmadloo S, Omer WH, Yoshihara K, et al. Allelic imbalance in regulation of ANRIL through chromatin interaction at 9p21 endometrosis risk locus. *PLoS Genet.* (2016) 12:e1005893. doi: 10.1371/journal.pgen.1005893

42. Schaefer AS, Richter GM, Groessner-Schreiber B, Noack B, Nothnagel M, El Mokhtari NE, et al. Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. *PLoS Genet.* (2009) 5:e1000378. doi: 10.1371/journal.pgen.1000378

43. Musunuru K, Post WS, Herzog W, Shen H, O’Connell JR, McArule PF, et al. Association of single nucleotide polymorphisms on chromosome 9p21.3 with platelet reactivity: a potential mechanism for increased vascular disease. *Circ Cardiovasc Genet.* (2010) 3:445–53. doi: 10.1161/CIRCGENETICS.109.923508

44. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* (2007) 316:1341–5. doi: 10.1126/science.1142382

45. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Mit Lu, Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, et al. Genotype-phenotype association identifies loci for type 2 diabetes and triglyceride levels. *Science* (2007) 316:1331–6. doi: 10.1126/science.1142358

46. Erdmann J, Grosshennig A, Braund PS, Konig IR, Hengstenberg C, Hall AS, et al. New susceptibility loci for coronary artery disease on chromosome 3q22.3. *Nat. Genet.* (2009) 41:280–2. doi: 10.1038/ng.307

47. Myocardial Infarction Genetics Consortium, Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, et al. Genomic association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat. Genet.* (2009) 41:334–41. doi: 10.1038/ng.327

48. Tregouet DA, Konig IR, Erdmann J, Munteanu A, Braund PS, Hall AS, et al. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nat. Genet.* (2009) 41:283–5. doi: 10.1038/ng.314

49. Consortium I KC. Large-scale gene-centric analysis identifies novel variants for coronary artery disease. *PLoS Genet.* (2011) 7:e1002260. doi: 10.1371/journal.pgen.1002260

50. Wang Z, Jacobs KR, Yeager M, Hutchinson A, Sampson J, Chatterjee N, et al. Improved imputation of common and uncommon SNPs with a new reference set. *Nat. Genet.* (2011) 44:6–7. doi: 10.1038/ng.1044

51. Consortium CAD, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat. Genet.* (2013) 45:25–33. doi: 10.1038/ng.2480

52. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.* (2015) 47:1121–30. doi: 10.1038/ng.3396

53. Myocardial Infarction Genetics Consortium, CAEIC, Sztitziel NO, Stirrups KE, Masca NG, Erdmann J, et al. Coding variation in ANGPTL4, LPL, and SVEP1 and the risk of coronary disease. *Circulation* (2010) 122:433–43. doi: 10.1161/CIRCRESAHA.111.960443

54. Almontashiri NA, Antoine D, Zhou X, Vilmundarson RO, Zhang SX, Hao KN, et al. 9p21.3 coronary artery disease risk variants disrupt TEAD transcription factor-dependent transforming growth factor beta regulation of p16 expression in human aortic smooth muscle cells. *Circulation* (2015) 132:1969–78. doi: 10.1161/CIRCULATIONAHA.114.015023

55. Bochenek G, Hauser R, El Mokhtari NE, Konig IR, Loos BG, Jepsen S, et al. 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.* (2013) 22:4516–27. doi: 10.1093/hmg/ddt299

56. Baal, Y, Nie S, Jiang G, Zhou Y, Zhu M, Zhao Y, et al. Regulation of CARD8 expression by ANRIL and association of CARD8 single nucleotide polymorphism rs2043211 (p.C100) with ischemic stroke. *Stroke* (2014) 45:383–8. doi: 10.1161/STROKEAHA.113.003393

57. Yamashita K, Muto T, Ibukiyama S, Osumi Y, Nishio T, Yoshimi S, et al. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 37 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* (2010) 38:662–74. doi: 10.1016/j.molcel.2010.03.021

58. Aslan A, Berkman O, Lalem T, Ozbilum N, Goksel S, Korkmaz O, et al. Long non-coding RNAs in the atherosclerotic plaque. *Atherosclerosis* (2017) 266:176–81. doi: 10.1016/j.atherosclerosis.2017.10.012
76. Kojima Y, Downing K, Kundu R, Miller C, Dewey F, Lancero H, et al. Cyclin-dependent kinase inhibitor 2B regulates efferocytosis and atherosclerosis. *J Clin Invest.* (2014) 124:1083–97. doi: 10.1172/JCI70391

77. Sato K, Nakagawa H, Tajima A, Yoshida K, Inoue I. ANRIL is implicated in the regulation of nucleus and potential transcriptional target of E2F1. *Oncol Rep.* (2010) 24:701–7. doi: 10.3892/or_00000910

78. Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, et al. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature* (2010) 464:409–12. doi: 10.1038/nature08801

79. Loinard C, Basatemur G, Masters L, Baker L, Harrison J, Figg N, et al. Deletion of chromosome 9p21 noncoding cardiovascular risk interval in mice alters Smad2 signaling and promotes vascular aneurysm. *Circ Cardiovasc Genet.* (2014) 7:799–805. doi: 10.1161/CIRCGENETICS.114.000696

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Holdt and Teupser. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.