

Emerging role of circANRIL expression in peripheral blood mononuclear cells as a promising liquid biopsy diagnostic biomarker for coronary artery disease

Gokce Akan (✉ akangokce@yahoo.com)

Muhimbili University of Health and Allied Sciences

Evarist Nyawawa

Jakaya Kikwete Cardiac Institute

Bashir Nyangasa

Jakaya Kikwete Cardiac Institute

Mehmet Kerem Turkcan

Columbia University

Erasto Mbugi

Muhimbili University of Health and Allied Sciences

Mohammed Janabi

Jakaya Kikwete Cardiac Institute

Fatmahan Atalar

Muhimbili University of Health and Allied Sciences

Research Article

Keywords: Coronary artery disease, 9p21.3, polymorphism, ANRIL, circANRIL, hsa_circ_0008574, Visceral adipose tissue, Peripheral mononuclear blood cells

Posted Date: February 1st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1300387/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Long-non-coding RNAs (lncRNAs) are increasingly recognized to play major roles in coronary artery disease (CAD). ANRIL, a well-known lncRNA is located on the 9p21.3 locus which is the first reported CAD risk locus. This study aimed to evaluate the association of ANRIL and its transcript variants expressions with susceptibility to CAD as well as to investigate the impact of 9p21.3 locus single nucleotide polymorphisms (SNPs) on their expressions in adipose tissues and peripheral blood mononuclear cells (PBMCs) of Tanzanian study cohort.

Methods: Expressions of ANRIL, and its variants; circANRIL (hsa_circ_0008574), NR003529, EU741058 and DQ485454 were detected by qRT-PCR in visceral adipose tissues (VAT) (epicardial adipose tissue (EAT) and mediastinal adipose tissue (MAT)), subcutaneous adipose tissue (SAT) and PBMCs of patients undergoing coronary artery bypass grafting due to CAD and non-CAD patients undergoing heart valve surgery. Five SNPs encompassing the 9p21.3 risk locus were genotyped in a cohort of 200 CAD patients and 220 non-CAD patients using qRT-PCR and their impact on gene expressions were further evaluated.

Results: ANRIL expression was significantly up-regulated, while the expression of circANRIL significantly down-regulated in CAD patients compared to non-CAD patients. Also, increased circANRIL expression was determined to be significantly inversely associated with the severity of CAD. Moreover, rs10757278 and rs10811656 polymorphisms were significantly associated with expressions of ANRIL, EU741058, NR003529 and circANRIL in VAT and PBMCs. Decreased circANRIL levels were correlated with aggressive clinical characteristics. Moreover, the random forest model was evaluated to feature importance of CAD prediction. Accordingly, the receiver operating characteristics (ROC) curve analysis was evaluated and the result suggested that circANRIL has high diagnostic accuracy. The area under the ROC curve (AUC) was 0.9808, the optimal cut-off value was 0.33, with a sensitivity of 1.0 and a specificity of 0.88.

Conclusion: We report the first data demonstrating the presence of, ANRIL and its transcript variants expressions in adipose tissues and PBMCs of CAD patients and the strong association between ANRIL expression and CAD. circANRIL having a synergetic effect with ANRIL plays a protective role in CAD pathogenesis. Therefore, altered circANRIL expression may become a potential diagnostic biomarker for early CAD diagnosis in Tanzania.

Introduction

Coronary artery disease (CAD) is a complex disease and it is the most common cardiovascular condition. CAD is characterized by plaque build-up of fat and cholesterol in the arteries [1]. These plaques, when they are unstable or become obstructive, lead to CAD, the clinical manifestations of atherosclerosis [2]. The risk of developing CAD is related to interplay of genetic, lifestyle and environmental factors, their effects are demonstrated to be mainly mediated through cardiovascular risk factors.

Obesity is known to be the most prevalent risk factor for CAD [3]. Regional body adipose tissue distribution, rather than total body adiposity is a well-known marker for cardiovascular risk [4]. Especially, visceral adipose tissue (VAT), including epicardial adipose tissues (EAT) and mediastinal (MAT), has been studied in the last decade for their relationship with CAD. EAT is the most studied cardiac adipose tissue and it has gained increasing attention due to its location. EAT is located between the visceral pericardium and the heart, also it is in contact with the coronary vessels. Recent studies showed that EAT has paracrine or endocrine activities and it can secrete inflammatory factors, such as TNF-alpha, IL-6, and adipocytokines and it has the ability to release and uptake free fatty acids, which plays a vital role in the development for CAD [5, 6]. Moreover, studies showed that the volume of EAT was demonstrated to be associated with the incidence of CAD [7-10]. Another cardiac adipose tissue is MAT, it is not close to the coronary vessels, but it is situated outside the pericardium that encloses the heart. Recently, the relationship between MAT and the development of CAD has been demonstrated and a study reported that the increased volume of MAT was associated with an enhanced CAD risk profile [12]. Also, the volume of MAT was positively correlated with plasma triacylglycerol and C-reactive protein which they are well-known risk factors and have a role in progression of the disease [13]. Furthermore, the fat depot's associated diseases such as CAD are related with excessive accumulation of subcutaneous adipose tissue (SAT), also an increase in the number of adipocytes in the SAT is accompanied by the presence of metabolic disorders [14].

As with most complex diseases, genetics play an important role for developing CAD. In recent genome-wide association studies (GWAS) have identified a new susceptibility region located on human chromosome 9p21.3 (Chr9p21.3) containing multiple single nucleotide polymorphisms (SNPs) associated with CAD. [15-18]. The nearest genes, approximately 100 kilobases (kb) away from the core CAD region, are two tumor suppressor genes (cyclin-dependent kinase inhibitors) *CDKN2A* and *CDKN2B* that play an important role in cell cycle regulation, apoptosis, senescence, aging, and inflammation which are processes strongly involved in atherogenesis. From a functional perspective *CDKN2A* and *CDKN2B* are potentially functional candidates genes to implicated in the pathogenesis of atherosclerosis [19-21].

Antisense non-coding RNA in the *INK4* locus (*ANRIL*, also known as *CDKN2BAS-1*) is another gene located at the same locus gene [22]. *ANRIL* (*ENSG00000240498*) spans a region almost 126.3 kb and *ANRIL* overlaps at its 5' end with *CDKN2B* and it has a role in the epigenetic regulation of the expression of adjacent protein-coding genes, including *MTAP*, *CDKN2A* and *CDKN2B* through multiple mechanisms, such as RNA interference, gene silencing, chromatin remodeling, or DNA methylation [23]. Recently, it was revealed that the risk locus could be responsible for regulating *ANRIL* expression to a certain extent, and also, 9p21.3 risk locus could have an impact *CDKN2A* and *CDKN2B* expression through *ANRIL* expression [24, 25]. Another study demonstrated that deletion of the targeted 9p21.3 risk locus of orthologous *ANRIL* interval in mice reduced expression of *CDKN2A* and *CDKN2B* in the heart and lead to excessive proliferation of vascular cells that contribute development of atherosclerosis [26].

A few investigations suggested that transcript variants of *ANRIL*, not only in linear transcript variants but also in circular transcript variant, could also be responsible for the development of CAD [27-30]. Most

studied linear transcript variants of *ANRIL* with CAD are as follows; the longest *ANRIL* transcript variant: *NR003529* (3857 bp, *ENST00000428597.6*), and 2 shorter variants of *ANRIL*; *DQ485454* (2659 bp, *ENST00000580576.6*), and *EU741058* (688 bp, exons 1, 5 to 7, and 13, *ENST00000455933.7*). Recently, *ANRIL* transcripts (*EU741058* and *NR003529*) were found to be significantly increased in peripheral blood mononuclear cells and atherosclerotic plaque tissue from patients of CAD compared to healthy individuals. Moreover, *ANRIL* expression was shown to be associated with CAD severity and with 9p21.3 risk genotypes [26]. Another study showed that subjects homozygous for the risk alleles of 9p21.3 exhibited increased expression levels of *ANRIL* short transcripts and decreased long variant expression levels, in comparison with healthy individuals carrying two copies of reference alleles [31].

Newly discovered circular transcript variant of *ANRIL*, circular antisense non-coding RNA in the *INK4* locus (*circANRIL* or *hsa_circ_0008574*), consisted of exons 5, 6 and 7, where exon 7 was non-canonically spliced to exon 5. Many researchers have focused considerable attention on circular RNAs (circRNAs) in recent years due to there are evidence circRNAs are emerging that they may have importance in the development of tissues or organs and also in disease pathogenesis [32].

The circRNAs have different structures from their linear counterparts, circRNAs have a unique structure of covalently closed continuous loop lacking 5' cap and 3' poly-adenylated tails [33]. These differences make circRNAs more stable from their linear counterparts and they gain resistance to degradation by exonuclease and RNases. Therefore, circRNAs are proposed as a new generation of predictive biomarkers and potential therapeutic targets for many diseases. The *circANRIL* was found in many different cell lines and in many primary cell types, including vascular smooth muscle cells, macrophages, heart, and vascular tissue [29, 30]. Burd *et al.* found that *circANRIL* expression is associated with atherosclerotic risk. In addition, they also reported that *circANRIL* is more stable than linear *ANRIL* [30]. Another study by Holdt *et al.* speculated that *circANRIL* is a prototype of a circRNA regulating ribosome biogenesis and conferring atheroprotection, and in the same study they suggested that *circANRIL* remains a potential therapeutic target for the treatment of atherosclerosis [29]. Moreover, there is evidence the *circANRIL* isoform (exons 5-7) has a synergistic effect and play a more effective and protective role of CAD by controlling rRNA maturation and modulating pathways of atherogenesis [24].

Taken together observations from all studies done so far pointed out the importance of *ANRIL* and/or transcript variants of *ANRIL* in the mechanism mediating the 9p21.3 association, and suggested that *ANRIL* could be a possible candidate gene of CAD at the 9p21.3 risk locus.

Thus, the aim of the present study was first to assess whether the adipose tissue and peripheral blood mononuclear cells (PBMCs) expressions of *ANRIL* and *ANRIL* transcript variants were associated with CAD susceptibility, and to investigate the impact of 9p21.3 locus variations on the expression of *ANRIL* and *ANRIL* transcript variants in PBMCs, in adipose tissues which play a role in the development of CAD (VAT (EAT and MAT) and SAT) in Tanzanian CAD patients.

Materials And Methods

Subjects: This study was performed at Jakaya Kikwete Cardiac Institute, Dar es Salaam, Tanzania. Participants enrolled in the study were selected among patients admitted to the cardiology department outpatient clinic for symptoms of angina, dyspnea and chest discomfort at the time of diagnosis, and patients admitted to the cardiology department for operation. A total of 420 subjects including 200 CAD patients (175 patients with CAD from outpatient clinic and 25 patients undergoing coronary artery bypass grafting (CABG) due to CAD) and 220 non-CAD patients (195 patients without CAD from outpatient clinic and 25 patients undergoing heart valve operation) were enrolled to investigate the association between the polymorphisms located on the 9p21.3 CAD risk locus and CAD susceptibility. Then 25 CAD patients undergoing CABG and 25 patients undergoing heart valve operation were included to detect ANRIL and its transcript variants expressions in adipose tissues and PBMCs. CAD was defined as $\geq 50\%$ luminal narrowing in at least one coronary artery from outpatient clinic and it was defined as $\geq 50\%$ luminal narrowing in at least two coronary artery from cardiology department surgery clinic. CAD patients from cardiology department surgery clinic underwent bypass surgery due to at least two vessel disease according to the number of significantly affected vessels using the Coronary Artery Surgery Study classification [34]. Only the patients with normal coronary angiogram were included in the non-CAD patients. All subjects enrolled in this study were of Tanzanian origin.

Body mass index (BMI) was calculated as a ratio of mass in kilogram divided by the square of height in meters. Diagnosis of hypertension was defined as resting blood pressure $\geq 140/90$ mm Hg or current use of antihypertensive therapy. Diagnosis of diabetes was defined either by 1999 World Health Organization (WHO) criteria 11 or self-report of being previously diagnosed as diabetic. Diagnosis of dyslipidemia was defined as low-density lipoprotein (LDL) cholesterol ≥ 3 mmol/l or the use of hypolipidemic agents.

A detailed information of subjects on demographics, medical history, current medication and CAD risk factors such as diabetes, hypertension, smoking *etc.* were completed through personal interviews. Oral antidiabetic and lipid lowering drugs which may interfere with gene expression were stopped three days before the operation and sample collections. PBMCs and adipose tissue samples were obtained after overnight (12 hours) fasting. Fasting venous blood samples were centrifuged at 1500xg at 25°C for 10 minutes promptly and separated serum and plasma were stored at -80°C for further analyses.

Biochemical measurements: Serum total cholesterol (TC) and high-density lipoprotein cholesterol (HDL) and triglycerides (TG) were measured by routine enzymatic endpoint methods (Analyzer A15 Biosystems, Philippines). Fasting glucose was determined using the enzymatic reference method with glucose oxidase. Very low density lipoprotein cholesterol (VLDL) and LDL were calculated in keeping with Friedewald's Formula [35].

Tissue biopsies: Approximately 30-50 mg of EAT, MAT and SAT were collected during operation in the form of biopsy. EAT samples were collected from the fat surrounding to the heart, within the pericardium and MAT samples were collected from the fat within the mediastinum, outside the pericardial sac. SAT

samples were collected from retrosternal region. All adipose tissues and the tissue samples were then immediately frozen in liquid nitrogen and stored at -80°C prior to total RNA preparation.

DNA Extraction: Genomic DNA was obtained from peripheral blood leukocytes of all participants by the use of MagnaPure DNA Isolation robot (Roche, Germany) and DNA quantity was determined by NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, Wilmington, Delaware USA). The extracted DNA was stored at -80°C and used later for the downstream analysis.

Genotyping: The SNPs rs10757274, rs2383207, rs2383206, rs10811656 and rs10757278 in all participants were detected with the use of LightSNiP typing assays (TIB MolBiol, Berlin, Germany), employing quantitative real-time polymerase chain reaction (QRT-PCR) amplifications with melting curve analysis. The reactions were performed on a LightCycler®480 II Real-Time PCR system (Roche-Germany), following the recommendations of the manufacturer.

Total RNA extraction and quantitation: 30 mg of EAT, MAT and SAT samples were obtained from the subgroup including 25 CAD patients undergoing CABG operation and 25 non-CAD patients undergoing heart valve operation. Total RNA from EAT, MAT and SAT were extracted with the Trizol™ Reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. RNA quantity was determined by NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, Wilmington, Delaware USA). After 1 μg of RNA was reverse transcribed in 20 μl total volume using OligodT as primers and Revert Aid First Strand cDNA Synthesis Kit (Fermentas, Canada). Following cDNA synthesis, the quality of cDNA was assessed by polymerase chain reaction (PCR) amplification of the housekeeping gene β -Globin. All cDNA samples were stored at -20°C prior to gene expression analysis. Moreover, PBMCs were isolated from 5 ml anticoagulated peripheral blood, and stored at -80°C until processed. Total RNA in PBMCs was extracted with Trizol™ Reagent (Invitrogen, Carlsbad, CA, USA) and the concentration of RNA was quantified using NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, Wilmington, Delaware USA) and was reverse-transcribed into cDNA as described above.

Gene expression analysis by quantitative real-time PCR (QRT-PCR): The expression levels of *ANRIL*, *circANRIL*, *NR003529*, *EU741058* and *DQ485454* were detected by qRT-PCR using a LightCycler®480 II Real-Time PCR system (Roche, Germany). Ten-fold dilutions of cDNA synthesized from total RNA were used. All samples were amplified in duplicate and the mean values were obtained for further calculations. Primers and probes designed using the soft-ware Universal Probe Library and primers synthesized by IDT (Skokie-USA) and probes synthesized by Roche Applied Science (Roche-Germany).

Data analysis: Statistical analysis was performed using SPSS software (Statistical Package for the Social Sciences 25.0, SPSS Inc, Chicago, IL, USA). Quantitative variables were expressed as mean \pm standard deviation (S.D.), and qualitative variables were expressed as percentages. The allelic frequencies of these polymorphisms between the non-CAD and CAD patients and the demographic characteristics between the groups were estimated for categorical variables were compared using Chi square (χ^2). Hardy-Weinberg equilibrium (HWE) was assessed by Fischer's exact test. The gene

expression data were obtained as Cycle Threshold (CT) values (CT = cycle number at which logarithmic PCR plots cross a calculated threshold line). The expression of each gene was compared between depots using the $2^{\Delta\Delta CT}$ method ($\Delta\Delta CT = CT$ of the target gene - CT of the housekeeping gene), respectively. The differences of normal and non-normal distributed continuous variables were compared using Student's t test and Mann-Whitney U test respectively. To evaluate differences between groups, the data were undergo log transformation to satisfy ANOVA criteria and then subjected to one-way ANOVA with Tukey's post-hoc analysis. The correlations between the gene expression and risk factors of CAD were evaluated with the Spearman correlation test. The statistical software package MedCalc Statistical Software (version 16.2, Ostend, Belgium) was used for multiple logistic regression analysis and the receiver operating characteristic (ROC) analysis to examine expressions of target genes in the PBMCs can be used as a biomarker for CAD. A logistic regression classifier implemented in the scikit-learn library [36] was used to perform leave-one-out cross-validation and to calculate the p-value using the method by DeLong *et al.* [37]. Normalized coefficient magnitudes was used as a means for calculating feature importances. Figures were generated in Python, using the matplotlib library [38]. Statistical significance was taken as $p < 0.05$.

Results

Clinical and anthropometric characteristics of the study population: The clinical and anthropometric characteristics of the study group were presented in Table 1. There was not significant difference between the groups in terms of age. When compared the lipid profiles, serum TG, TC, LDL were higher in the all CAD patients than in the non-CAD patients ($p < 0.05$, respectively), whereas serum HDL levels were significantly higher among non-CAD patients ($p < 0.05$). The average fasting plasma glucose of the CAD patients were significantly higher than that of non-CAD patients ($p < 0.05$). CAD patients had higher levels of systolic blood pressure and diastolic blood pressure. There was also a higher prevalence of obesity, hypertension, diabetes, hyperlipidemia and smokers in CAD patients compared to non-CAD patients ($p < 0.05$). In addition, male gender was predominant in CAD patients (76%). Hypertension was the more commonly associated clinical condition in CAD patients (92%), while the least common one was determined to be the hyperlipidemia (56%). The further data analysis demonstrated that male gender, obesity, hypertension, diabetes mellitus, smoking and hyperlipidemia were the predominating risk factors for CAD in Tanzanian population.

Table 1: Baseline characteristics of the study groups

Variables	CAD (n=200)	non-CAD (n=220)	P Value*
Age (year)	60.04±7.67	58.64±10.27	0.580
Sex/ male (%)	19 (76%)	10(40%)	0.010
Weight (kg)	84.56±11.58	67.52±14.56	0.001
Height (m)	1.60±0.088	1.65±0.073	0.291
BMI (kg/m²)	32.36±4.70	23.84±3.98	0.001
Systolic BP (mmHg)	133.88±28.89	122.64±24.56	0.037
Diastolic BP (mmHg)	91.40±15.54	77.08±16.33	0.003
Glucose (mmol/L)	8.01±2.14	4.13±0.78	0.001
Cholesterol (mmol/L)	5.88±0.95	3.95±0.72	0.001
HDL (mmol/L)	0.84±0.28	1.36±0.30	0.001
LDL (mol/L)	3.97±0.76	2.56±0.66	0.001
VLDL (mmol/L)	1.19±0.44	0.52±0.28	0.001
TG (mmol/L)	2.63±1.31	0.98±0.41	0.001
Obesity (%)	18 (72%)	0	0.001

	23 (92%)	11 (44%)	0.001
Hypertension (%)			
	17 (68%)	1 (4%)	0.001
Diabetes (%)			
Hyperlipidemia (%)	14 (56%)	0	0.001
Smoking (%)	17 (68%)	7 (28%)	0.005

Data: mean \pm SD *Comparisons of differences between mean values of two groups unpaired Student t-test was used

Allele and genotypic association of different SNPs: The allele and genotype frequencies of five SNPs (rs10757274, rs2383207, rs2383206, rs10811656, and rs10757278) were analyzed for all participants including 200 CAD patients and 220 non-CAD patients and the results are presented in Supplementary Table 1. All SNPs were at Hardy-Weinberg equilibrium in both groups (all p values >0.05). Significant differences were observed in genotype and allele frequencies of rs10757274, rs2383206, rs10811656, and rs10757278 variants between CAD and non-CAD patients ($p < 0.005$).

The genotype frequencies of rs10757274, rs2383206, rs10811656, and rs10757278 SNPs remained significant when analyzed in the subgroup including 25 CAD patients undergoing CABG operation and 25 non-CAD patients undergoing heart valve operation (Table 2). Furthermore, the risk alleles rs10757274 G allele, rs2383206 G allele, rs10811656 T allele, and rs10757278 G allele were found statistically significant (OR=5.79, 95% CI= 2.32-14.43, OR=3.80, 95% CI= 1.65-8.74, OR=6.16, 95% CI=2.41-15.75 and OR=5.67 95% CI= 2.14-14.99, respectively) in CAD patients compared to non-CAD patients of the subgroup (Table 2). The analyses of allelic and genotypic showed that the rs2383207 SNP is not associated CAD risk in both groups.

ANRIL and ANRIL splice variants expression levels in adipose tissues and PBMCs from CAD patients and non-CAD patients: For expression analysis, we included 25 CAD patients undergoing CABG and 25 non-CAD patients undergoing valve replacement. The expression levels of *ANRIL*, and *ANRIL transcripts*; *circANRIL*, *NR003529*, *EU741058* and *DQ485454* were studied in EAT, MAT, SAT and PBMCs (Figure 1). *ANRIL* expression levels were significantly up-regulated in PBMCs of the CAD patients compared to non-CAD patients (fold change=1.6, $p < 0.001$). Although *ANRIL* expression levels in EAT, MAT and SAT were found increased in CAD patients compared to non-CAD patients, the differences were not significant (Figure 1A).

circANRIL was significantly down-regulated in PBMCs of CAD patients compared to non-CAD patients (fold change=5.3, $p < 0.001$) (Figure 1E) while no significant differences were determined in EAT, MAT and SAT. The associations of *ANRIL* and *circANRIL* expressions with the CAD severity (double stenotic vessels disease (n=13) versus triple stenotic vessels disease (n=12)) were also evaluated. A statistically

significant difference in the expression levels of *circANRIL* was determined between the two groups ($p < 0.05$) (Figure 2) but no statistically significant difference was observed for *ANRIL* expression ($p > 0.05$).

Moreover, the expression levels of *NR003529*, *EU741058* and *DQ485454* in EAT, MAT, SAT and in PBMCs showed no statistical differences between CAD patients and non-CAD patients ($p > 0.05$, respectively). The expression level of *EU741058* in PBMCs was down-regulated in CAD patients compared to non-CAD (fold change ≈ 0.8) but the difference was not significant.

Table 2: The genotypic and allelic frequency distributions of SNPs on chromosome 9p21.3 in the subgroup

SNP	Genotypic Frequencies n (%)		P-Value	Allelic Frequencies			X ²	OR/CI(95%)	P-Value
	CAD(n=25)	non-CAD (n=25)		Allele	CAD(n=25)	non-CAD (n=25)			
rs10757274									
AA	3(12)	17(68)							
AG	16(64)	7(28)	0.001	A/G	0.44/0.56	0.82/0.18	15.49	5.79/2.32-14.43	0.001
GG	6(24)	1(4)							
rs2383207									
AA	3(12)	12(48)							
AG	13(52)	2(8)	0.563	A/G	0.38/0.62	0.52/0.48	1.98	1.76/0.79-3.91	0.159
GG	9(36)	11(44)							
rs2383206									
AA	3(12)	14(56)							
AG	13(52)	7(28)	0.001	A/G	0.38/0.62	0.70/0.30	10.31	3.80/1.65-8.74	0.001
GG	9(36)	4(16)							
rs10811656									
CC	6(24)	19(76)							
CT	11(44)	4(16)	0.002	C/T	0.46/0.54	0.84/0.16	15.87	6.16/2.41-15.75	0.001
TT	8(32)	2(8)							
rs10757278									
AA	7(28)	19(76)							
AG	12(48)	5(16)	0.001	A/G	0.52/0.48	0.86/0.14	13.51	5.67/2.14-14.99	0.001
GG	6(24)	1(8)							

The subgroup includes 25 CAD patients undergoing CABG and 25 non-CAD patients undergoing valve placement. OR: Odd Ratio, CI: Confidence Interval *The genotypic and allelic frequency distributions of polymorphisms between the groups were compared using χ^2 and HWE test. In all cases differences were considered significant at $p < 0.05$.

Associations between 9p21.3 risk locus genotypes and ANRIL and ANRIL transcripts expression levels: To better understand the relationship between *ANRIL* and CAD, we next evaluated the potential effects of 9p21.3 risk locus SNPs on the expression levels of *ANRIL* and *ANRIL* transcript variants in PBMCs and adipose tissues of the subgroup including 25 CAD patients undergoing CABG and 25 non-CAD patients undergoing valve placement. The expression levels of *ANRIL* in PBMCs were significantly higher in the risk genotype carriers of rs10757278 and rs10811656 (GA and GG for rs10757278 and CT and TT for rs10811656) in CAD patients compared to wild type carriers ($p=0.004$ and $p=0.013$, respectively). Also, the expression levels of *ANRIL* transcript variants; *NR003529* and *EU741058* in EAT, MAT and PBMCs were significantly higher in CAD patients carrying the risk genotype of rs10757278 and rs10811656 compared to wild type carriers, while there was no difference in *DQ485454* ($p=0.001$, $p=0.007$, $p=0.028$, $p=0.006$, $p=0.002$, $p=0.019$ and $p>0.05$, respectively). However, the expression levels of *circANRIL* in PBMCs were significantly down-regulated in rs10757278 and rs10811656 risk genotype carriers compared to wild type carriers ($p=0.001$ and $p=0.01$, respectively).

Impacts of CAD risk factors on the expression levels of ANRIL transcripts: The correlation analyses between expression levels of candidate genes and risk factors of CAD were analyzed by the Spearman correlation test. *ANRIL* expression in PBMCs was positively correlated with BMI, glucose level, total cholesterol, TG and LDL ($r=0.362$, $p=0.01$, $r=0.325$ $p=0.021$, $r=0.323$ $p=0.02$, $r=0.444$ $p=0.001$ and $r=0.460$ $p=0.001$, respectively) but negatively associated with HDL ($r=0.304$ $p=0.032$).

The *circANRIL* expression levels in PBMCs were negatively correlated with BMI, glucose level, total cholesterol, TG, LDL, Systolic BP and Diastolic BP ($r=0.531$, $p=0.001$, $r=0.547$ $p=0.001$, $r=0.599$ $p=0.001$, $r=0.558$ $p=0.001$, $r=0.535$ $p=0.001$, $r=0.363$ $p=0.009$ and $r=0.469$ $p=0.001$, respectively) but positively associated with HDL ($r=0.583$ $p=0.001$).

ANRIL and *circANRIL* were co-regulated with most of the risk factors of CAD such as lipid levels, blood pressure and glucose levels. These positive and negative correlation results with the risk factors, which mentioned above, suggested that *ANRIL* and *circANRIL* expressions regulate the risk factors leading to the CAD development and *ANRIL* and *circANRIL* may serve as indicator genes in CAD patients.

Future importance of the variables: Machine learning (ML) is a highly effective method for disease prediction using machine learning techniques. They are able to capture the complex interactions between predictors and outcomes during the data process, and they would provide a new and novel discernment towards the disease [39]. Recent studies showed that random forest (RF) is the most efficient algorithm for the prediction of CAD than the other algorithms and it gives consistently better accuracy of the prediction system [40-42]. We measured the importance of different features for the risk factors of CAD together with the expression levels of *ANRIL* and its transcript variants by the mean decrease impurity (Gini importance) of all decision trees in a tuned RF model. The importance of included variables obtained from the tuned RF model is presented in Figure 3. As expected, age, systolic BP, BMI and smoking were among the top risk factors. In addition we observed that expression levels of *ANRIL* and *circANRIL* in PBMCs were also among the top risk factors.

Improvement in the diagnostic value: Multivariate logistic regression analyses revealed that weight (OR=0.91, 95%CI=0.85-0.96), height (OR=3.54, 95%CI=1.60-7.81), BMI (OR=0.61, 95%CI=0.46-0.80), systolic BP (OR=0.98, 95%CI=0.96-1.0), smoking (OR=5.46, 95%CI=1.62-18.35), diastolic BP (OR=0.95, 95%CI=0.90-0.99), sex (OR=0.22, 95%CI=0.06-0.71) and the expression levels of *ANRIL* (OR=2.05, 95%CI=1.74-2.35) and *circANRIL* (OR=13.63, 95%CI=3.74-49.60) in PBMCs were potential biomarkers for CAD.

To test comparisons of the diagnostic value of expression levels of *ANRIL* and *circANRIL* to top risk factors which are clinical features and observed in the RF model in the development of CAD, receiver operating characteristic (ROC) curve analysis was performed and the area under curve (AUC) was calculated.

Three models for CAD prediction based on clinical features and expressions of *ANRIL* and *circANRIL* were built. The first model (clinical model) consisted of CAD risk factors: weight, height, BMI, systolic BP, diastolic BP, smoking and age. The second model (clinical+*ANRIL* expression model) consisted of the clinical model and *ANRIL* expression, and the last model (clinical+*circANRIL* expression model) *circANRIL* expression was included in the clinical model.

The AUC value was 0.844 for clinical model (95% CI: 0.724-0.963, Optimal Cut-off: 0.55, Specificity 0.88, Sensitivity: 0.80), further the introduction of *ANRIL* expression increased AUC from 0.844 to 0.912 (95% CI 0.821-1.0, Optimal Cut-off: 0.61, Specificity 0.96, Sensitivity: 0.84, $p=0.02$) (Figure 4A), while the *circANRIL* expression into the clinical model including the CAD risk factors (age, weight, height, BMI, systolic BP and smoking) the area under curve (AUC) significantly increased from 0.844 to 0.980 (95% CI 0.953-1.0, Optimal Cut-off: 0.33, Specificity 0.88, Sensitivity: 1.0, $p=0.009$) (Figure 4B).

Finally, ROC analyses suggested that the detection of *ANRIL* expression and *circANRIL* expression together with risk factors of CAD exhibited a higher diagnostic performance compared with the detection of risk factors only. This result implied that the combination of *ANRIL* and *circANRIL* expressions in PBMCs has a great potential to be sensitive, and reliable biomarker possibly having a higher diagnostic value for CAD.

Discussion

While a big part of the human genome has already been transcribed, only $\approx 2\%$ of the genome appears coding for protein, the rest being deemed “non-coding.” Recent studies demonstrated that the human genome is prevalently transcribed and produced many thousands of non-coding RNAs (ncRNAs) such as microRNAs, small interfering RNAs, and various classes of long ncRNAs. A lot of ncRNAs play a vital role in cellular processes and they work with other biomolecules to modulate in gene regulatory networks and identification of the role of ncRNAs in the disease process could increase our understanding of the pathogenesis of complex diseases [43, 44]. Moreover, ncRNAs have been considered within the molecular mechanisms affecting different diseases, particularly age-related complex disorders. Accumulating studies indicate that several long non-coding RNAs (lncRNAs) which are a member of the ncRNAs play

important roles as key regulators of gene regulation in the development of cardiovascular diseases [45, 46].

ANRIL is one of the most important antisense transcripts lncRNA, receiving extensive attention as potentially novel biological regulator. Many roles have been attributed to *ANRIL* including DNA damage response, DNA repair, epigenetic regulation of *CDKN2A/B*, controlling the cell cycle checkpoints, apoptosis and RNA splicing *etc* [47]. A growing number of recent studies confirm that *ANRIL* plays pivotal roles in diverse physiological and pathological processes in CAD [22, 24-31]. Moreover, the polymorphisms on 9p21.3 region are responsible for regulating transcript variants of *ANRIL* and expression of *ANRIL* in CAD [47]. The SNPs rs10757278 and rs10811656 are located in one of *ANRIL* gene enhancers binding site for *STAT1*. The risk alleles of rs10757278/G and rs10811656/T disrupt a binding site for *STAT1* and increases the expression of *ANRIL* [47].

We herein aimed to ascertain whether the *ANRIL* and *ANRIL* transcript variants were involved in the manifestation of CAD susceptibility and *ANRIL* transcripts are associated with the 9p21.3 CAD risk in the Tanzanian patients. The present study was reported for the first time the expression of *ANRIL* and different transcript variants of *ANRIL* in PBMCs and in various adipose tissues.

Here we reported that differential expressions of *ANRIL* and its transcript variants in the CAD patients compared to non-CAD patients. *ANRIL* expression levels were significantly increased in PBMCs of CAD patients, but no statistically significant expression was observed in adipose tissues. There was also no statistical difference in terms of the expression levels of NR003529, EU741058 and DQ485454 in EAT, MAT, SAT and also in PBMCs between CAD and non-CAD patients. On the other hand, we discovered as the main findings of our study; circ*ANRIL* expression was increased in PBMCs of CAD patients compared with non-CAD patients, and it well-discriminated CAD patients from non-CAD patients, and circ*ANRIL* level was associated with increased CAD severity.

Our study demonstrated that up-regulated *ANRIL* expression could possibly be involved in CAD pathogenesis. This is in concordance with other studies which investigated the association of *ANRIL* expression and CAD in different populations [48, 49]. Moreover, it was demonstrated that increased expression of *ANRIL* lead to increased inflammation, CAD risk, severity and poor prognosis of CAD [50-52]. *ANRIL* expression was also showed to be associated with inflammation cytokines; monocyte chemoattractant protein-1 and interleukin-10 [52] and the knockdown of *ANRIL* significantly promoted cell proliferation and tubule formation and inhibited inflammatory activation and apoptosis of human umbilical vein endothelial cell [52]. In contrast, a study from Yang *et al.* reported no significant association between CAD and *ANRIL* expression in the Chinese Han population [53].

Furthermore, we observed a significant down-regulation of circ*ANRIL* expression in CAD patients compared to non-CAD patients and we suggested that circ*ANRIL* may play an atheroprotective role in CAD pathogenesis, which would indicate a synergetic effect between circ*ANRIL* and *ANRIL*. Holdt *et al.* also suggested that, circ*ANRIL* has a protective role against atherosclerosis [29]. Moreover, recent studies showed that increased circ*ANRIL* expression in human primary smooth muscle cells and vascular

endothelial cells in a rat model of CAD could induce apoptosis and inhibit the proliferation and thereby increased *circANRIL* expression could decreased to susceptibility of CAD [54, 55].

But we failed to detect the same significance of *ANRIL* and *circANRIL* expression levels in adipose tissues. One potential explanation of this result would be that *ANRIL* and *circANRIL* could be differentially expressed in adipose tissues and PBMCs.

We also found decreased expression levels of *EU741058* in PBMCs of CAD patients compared to non-CAD patients but the differences was a borderline statistical significance. Previously, a few studies reported that *EU741058* variant down-regulated in CAD patients and these results in agreement with our finding [53, 56].

Further analysis revealed that rs10757278 and rs10811656 were altered the expression levels of *ANRIL* and *ANRIL* transcript variants not only in PBMCs but also in adipose tissues. This result emphasized that rs10757278 and rs10811656 could controlling the expression levels of *ANRIL* and *ANRIL* transcript variants.

The *circANRIL* expression levels in PBMCs were significantly down-regulated in the CAD risk genotype carriers of rs10757278 (GA and GG) and rs10811656 (CT and TT) compared to wild type carriers. Also, the risk genotypes carriers of rs10757278 (GA and GG) and rs10811656 (CT and TT) had significantly higher *ANRIL* expression in EAT, MAT and in PBMCs of CAD patients. A recent study by Liu *et al.* showed that the expression of *ANRIL* significantly decreased in rs10757278 risk allele carriers in T cells of 170 healthy individuals [57]. The potential explanations of the diverse results between the research by Liu *et al.*, and our study might be the samples used were T cells from healthy individuals and in our research the samples were EAT, MAT and PBMCs from patients suffering from CAD [57].

Also, the expression levels of *NR003529* and *EU741058* in EAT, MAT and in PBMCs were statistically significantly higher in the CAD patients carrying the risk genotype of rs10757278 and rs10811656 compared to wild type carriers, but the expression of *DQ485454* remained unaffected. This finding indicated that *NR003529* and *EU741058* were abundantly transcribed in CAD risk genotype carriers leading to the development of CAD.

It is well known that EAT is an important component of VAT due to its contiguity to the coronary arteries and the effects of endocrine and paracrine activity secreting pro-inflammatory and anti-inflammatory cytokines and chemokines, therefore it has been suggested to influence coronary atherosclerosis development [58, 59]. In recent years MAT also has gained increasing attention on account of the contributions of development of atherosclerosis [60, 61]. Our results also showed that the clinical importance of EAT and also MAT in CAD.

Another finding in our study was the increased expression levels of *circANRIL* in PBMCs which was also associated with less severity of CAD.

The *circANRIL* expression was statistically significantly decreased in triple stenotic vessels disease compared to double stenotic vessels disease (2 fold). A previous study reported that *EU741058* and *NR00359* expressions were pre-dominantly associated with severity of CAD [27]. Although, we could not find the association between *EU741058* and *NR00359* expression and CAD severity but we successfully showed that the *circANRIL* expression was directly associated with severity of CAD.

In clinical practice, the increased levels of glucose, TC, TG, LDL and BMI, decreased HDL levels and also hypertension were always being attributed as major risk factors for CAD. In our study, we observed that these risk factors of CAD correlated with *ANRIL* and *circANRIL* expression levels in the blood. This co-regulation suggested that *ANRIL* and *circANRIL* regulate the risk factors leading to the CAD development and they may well serve as early disease indicators in CAD patients.

Although the CAD mortality rate is high, the chance of survival is higher if the diagnosis is made early enough in an accurate and efficient way. Therefore, the experts attempt to find a way for the accurate and timely prediction of CAD by using new statistical techniques, such as data mining which would help to identify the risk models of the disease, to recognize the disease patterns and the influencing factors. The feature importance selection methods give us an opportunity to increase the accuracy of the prediction of disease compared to traditional methods of classification. Especially, the combination of the feature importance selection algorithm such as RF, which has better accuracy for CAD prediction than other algorithms, is necessary for the efficient prediction of CAD [42, 62]. We examined the importance of different features for the risk factors of CAD together with the expression levels of *ANRIL* and its transcript variants and we observed that age, systolic BP, BMI, and smoking were among the top risk factors as we expected. In addition, the expression levels of *ANRIL* and *circANRIL* in PBMCs were also among the top risk factors.

The ROC analysis was done accordingly and the analyses showed that a good predictive value of *circANRIL* (exon 5-7) expression together with risk factors of CAD exhibited a higher diagnostic performance compared to CAD risk factors only. The *circANRIL* expression levels in PBMCs may be a new and non-invasive diagnostic tool for the diagnosis of CAD. As a screening tool in clinical practice, *circANRIL* has potential diagnostic value and is worthy of clinical promotion [63].

ANRIL became one of the most popular lncRNA since the discovery of 9p21.3 region as a susceptibility region for CAD. The SNPs in 9p21.3 region are located 4-bp apart in the STAT1 binding site of *ANRIL* gene [64]. The risk SNPs might well disrupt the binding site of STAT1 and in turn this would affect the expression of *ANRIL* leading to increased enhancer activity through the following possible mechanism (Figure 5). It is well known that, *ANRIL* regulates the expression of protein-coding genes via a physically interaction with the specific site, including the Alu element to CBX7 and SUZ12 components of the Polycomb complex (PRC1 and PRC2). *ANRIL* acts as a pivotal regulator and affects the expression of *CDKN2A* and *CDKN2B* tumor suppressor genes in 9p21.3 locus through same mechanism [65, 66] (Figure 5). On the other hand, in animal model it was demonstrated that *ANRIL* upregulates vascular endothelial growth factor (VEGF) and activated Nuclear Factor kappa B (NF- κ B) signaling pathway [67]

Moreover, the *circANRIL*, a circular transcript variant of *ANRIL*, binds and negatively regulates Pescadillo Ribosomal Biogenesis Factor 1 (*PES1*) protein in the *PeBoW* complex known as an essential complex to ribosome biogenesis in macrophages and vascular smooth muscle cells. And this negatively regulation induces *p53* activation and nucleolar stress, resulting in the inhibition of proliferation and induction of *p53* in proliferating cells [29] (Figure 5).

Taken together, our results emphasized that, the remarkable changes of *ANRIL* and *circANRIL* expression levels could be influenced by the progression of atherosclerosis in CAD by directly or indirectly ways. *ANRIL* and *circANRIL* could be a feature as potential biomarkers as well as therapeutic targets and also they might be a useful disease marker.

Conclusions

In conclusion, we genetically analyzed the expression levels of *ANRIL* and its transcript variants in adipose tissues and in PBMCs, and we further analyzed the associations between expression levels of *ANRIL* and its transcript variants and the variations of SNPs located on 9p21.3 region. To the best of our knowledge, this is the first study showing the expression levels of *ANRIL* and its transcript variants in VAT in CAD patients, and suggesting presence of an association between the expression levels of *ANRIL* and its transcript variants in VAT and CAD susceptibility.

Our results revealed that *ANRIL* expression levels in PBMCs were significantly increased while *circANRIL* expression levels were significantly decreased in CAD patients when compared to non-CAD patients. Additionally, *circANRIL* expression was found to be related with the severity of CAD. The rs10757278 and rs10811656 SNPs were also determined to be associated with not only *ANRIL* but also transcript variants of *ANRIL* (*EU741058*, *NR003529* and *circANRIL*). Not last, the RF model and ROC analyses showed that measuring the expression levels of *circANRIL* (exons 5 to 7) might offer a well reliable and sensitive prognostic value and allow for early detection and better monitor treatment response or disease recurrence for CAD. Taken together, our findings can provide a new liquid biomarker to be used together with the clinical decision-making and improve the early diagnosis and treatment of cardiovascular diseases.

Abbreviations

CAD: Coronary Artery Disease, VAT:Visceral Adipose Tissue, EAT:Epicardial adipose Tissue, MAT: Mediastinal adipose tissue, SAT: Subcutaneous Adipose Tissue, PBMCs: Peripheral Blood Mononuclear Cells, WHO: World Health Organization, GWAS: Genome-Wide Association Studies, SNP: Single Nucleotide Polymorphism, LD: Linkage Disequilibrium, BMI: Body Mass Index, BP: Blood Pressure, TC: Total Cholesterol, HDL-C:High-Density Lipoprotein Cholesterol, TG: Triglycerides, LDL-C:Low-Density Lipoprotein Cholesterol, VLDL-C: Very Low-Density Lipoprotein Cholesterol, QRT-PCR: Quantitative Real-Time Polymerase Chain Reaction, HWE: Hardy-Weinberg Equilibrium, CT: Cycle Threshold, ML:Machine learning, RF:Random Forest, AUC:Area Under the Curve, ROC:Receiver Operating Characteristic,

SD:Standard Deviations, OR: Odd Ratio, CI:Confidence Interval, ncRNA: Noncoding RNA, IncRNA: Long Noncoding RNA, PCR:Polycomb Complex,VEGF:Vascular Endothelial Growth Factor

Declarations

Ethics approval and consent to participate: This study was approved by the Research Ethics Committee of Muhimbili University of Health and Allied Sciences (Ref. Number 2017-05-25/AEC/Vol.XII/66 Date 25.05.2017).

Consent for publication: Not applicable

Availability of data and materials: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contribution: G.A. and F.A. conceived the study. F.A., M.J. and G.A. designed the study methodology. M.J., E.N., and B.N. performed the clinical arm of the study. G.A. performed the experiments. G.A., F.A. and M.K.T. contributed to the statistical analysis and interpretation of the results. G.A. wrote the original draft of the manuscript. F.A. edited, and reviewed the manuscript. F.A., E.M., and M.J. supervised the study. All authors have read and approved this final manuscript.

Acknowledgements: The authors gratefully acknowledge all the support from CAD patients and Jakaya Kikwete Cardiac Institute staff. In addition, we wish to thank individuals who gave us valuable help with the patient consent forms, sample collection and continuous support for the project: Miss Sarah Isaya Haule, Mr Ismael Adolf Chatita and Mr. Ally Luhaga from Muhimbili of Health and Allied Sciences Genetics Laboratory.

References

1. Mannarino E, Pirro M. Molecular biology of atherosclerosis. *Clin Cases Miner Bone Metab.* 2008; 5:57-62
2. Cullen P, Rauterberg J, Lorkowski S. The pathogenesis of atherosclerosis Review. *Handb Exp Pharmacol.* 2005;170: 3-70.
3. Yusuf S, Hawken S, Ôunpuu S, Bautista L, Franzosi MG, Commerford P, Lang CC, Rumboldt Z, Onen CL, Lisheng L, Tanomsup S, Wangai P, Razak F, Sharma AM, Anand SS: Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: a case–control study. *Lancet* 2005; 366:1640–1649.

4. Janssen I, Katzmarzyk PT, Ross R. Body mass index, waist circumference, and health risk: evidence in support of current national institutes of health guidelines. *Arch Intern Med* 2002; 162:2074–2079.
5. Gruzdeva O, Uchasova E, Dyleva Y, Borodkina D, Akbasheva O, Belik E, et al. Relationships between epicardial adipose tissue thickness and adipo-fibrokinase indicator profiles post-myocardial infarction. *Cardiovasc Diabetol*. 2018; 17:40.
6. Li Y, Liu B, Li Y, et al. Epicardial fat tissue in patients with diabetes mellitus: a systematic review and meta-analysis. *Cardiovasc Diabetol*. 2019; doi:10.1186/s12933-019-0807-3
7. Mahabadi AA, Massaro JM, Rosito GA, et al. Association of pericardial fat, intrathoracic fat, and visceral abdominal fat with cardiovascular disease burden: the Framingham Heart Study. *Eur Heart J*. 2009; doi:10.1093/eurheartj/ehn573
8. Ding J, Hsu FC, Harris TB, et al. The association of pericardial fat with incident coronary heart disease: the Multi-Ethnic Study of Atherosclerosis (MESA). *Am J Clin Nutr*. 2009; doi:10.3945/ajcn.2008.27358
9. Iacobellis G, Ribaudo MC, Assael F, Vecci E, Tiberti C, Zappaterreno A, Di Mario U, Leonetti F. Echocardiographic epicardial adipose tissue is related to anthropometric and clinical parameters of metabolic syndrome: a new indicator of cardiovascular risk. *J Clin Endocrinol Metab* 2003; 88:5163–8.
10. Taguchi R, Takasu J, Itani Y, Yamamoto R, Yokoyama K, Watanabe S, Masuda Y. Pericardial fat accumulation in men as a risk factor for coronary artery disease. *Atherosclerosis* 2001; 157:203–9.
11. Yerramasu A, Dey D, Venuraju S, Anand DV, Atwal S, Corder R, Berman DS, Lahiri A. Increased volume of epicardial fat is an independent risk factor for accelerated progression of sub-clinical coronary atherosclerosis. *Atherosclerosis* 2012; 220:223–30.
12. Sicari R, Sironi AM, Petz R, Frassi F, Chubuchny V, De Marchi D, Positano V, Lombardi M, Picano E, Gastaldelli A: Pericardial rather than epicardial fat is a cardiometabolic risk marker: an MRI vs echo study. *J Am Soc Echocardiogr* 2011; 24:1156–1162.
13. Cheung L, Gertow J, Werngren O, et al. Human mediastinal adipose tissue displays certain characteristics of brown fat. *Nutr Diabetes*. 2013; doi:10.1038/nutd.2013.6
14. Gruzdeva O, Borodkina D, Uchasova E, Dyleva Y, Barbarash O. Localization of fat depots and cardiovascular risk. *Lipids Health Dis*. 2018; doi: 10.1186/s12944-018-0856-8.
15. Helgadóttir A, Thorleifsson G, Manolescu A, Gretarsdóttir S, Blondal T, Jonasdóttir A, Jonasdóttir A, Sigurdsson A, Baker A, Palsson A. *et al*. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 2007;316: 1491-1493.
16. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR. *et al*. A common allele on chromosome 9 associated with coronary heart disease. *Science* 2007; 316:1488-1491
17. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Richard D, Meitinger T, Baraud P, Wichmann E. *et al*. WTCCC and the Cardiogenics Consortium. Genomewide association analysis of coronary artery disease. *N Engl J Med*. 2007;357: 443-453.

18. Wu C, Gong Y, Yuan J, Gong H, Zou Y, Ge J. Identification of shared genetic susceptibility locus for coronary artery disease, type 2 diabetes and obesity: a meta-analysis of genome-wide studies. *Cardiovasc Diabetol*. 2012; doi:10.1186/1475-2840-11-68
19. Holdt LM, Sass K, Gäbel G, Bergert H, Thiery J, Teupser D. Expression of Chr9p21 genes CDKN2B (p15(INK4b)), CDKN2A (p16(INK4a), p14(ARF)) and MTAP in human atherosclerotic plaque. *Atherosclerosis*. 2011; doi: 10.1016/j.atherosclerosis.2010.06.029.
20. Baker DJ, Jin F, van Deursen JM. The yin and yang of the Cdkn2a locus in senescence and aging. *Cell Cycle*. 2008; doi:10.4161/cc.7.18.6687
21. Rathi S, Danford I, Gudiseva HV, et al. Molecular Genetics and Functional Analysis Implicate CDKN2BAS1-CDKN2B Involvement in POAG Pathogenesis. *Cells*. 2020; doi:10.3390/cells9091934
22. Pilbrow AP, Folkersen L, Pearson JF, et al. The chromosome 9p21.3 coronary heart disease risk allele is associated with altered gene expression in normal heart and vascular tissues. *PLoS One*. 2012;7(6): e39574. doi:10.1371/journal.pone.0039574
23. Amaral PP, Dinger ME, Mercer TR, Mattick JS. The eukaryotic genome as an RNA machine. *Science*. 2008; 319:1787-1789.
24. Cheng J, Cai MY, Chen YN, et al. Variants in ANRIL gene correlated with its expression contribute to myocardial infarction risk. *Oncotarget*. 2017; doi:10.18632/oncotarget.14721
25. Congrains A, Kamide K, Oguro R, Yasuda O, Miyata K, Yamamoto E, Kawai T, Kusunoki H, Yamamoto H, Takeya Y, Yamamoto K, Onishi M, Sugimoto K, *et al*. Genetic variants at the 9p21 locus contribute to atherosclerosis through modulation of ANRIL and CDKN2A/B. *Atherosclerosis*. 2012; 220:449-455.
26. Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, Blow MJ, Cohen JC, Rubin EM, Pennacchio LA. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature*. 2010; 464:409-412.
27. Holdt LM, Beutner F, Scholz M, Gielen S, Gabel G, Bergert H, Schuler G, Thiery J, Teupser D. ANRIL expression is associated with atherosclerosis risk at chromosome 9p21. *Arterioscler Thromb Vasc Biol*. 2010; 30:620-627.
28. Folkersen L, Kyriakou T, Goel A, Peden J, Malarstig A, Paulsson-Berne G, Hamsten A, Hugh W, Franco-Cereceda A, Gabrielsen A, Eriksson P, consortia P. Relationship between CAD risk genotype in the chromosome 9p21 locus and gene expression. Identification of eight new ANRIL splice variants. *PLoS One*. 2009; 4:7677.
29. Holdt LM, Stahringer A, Sass K, et al. Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nat Commun*. 2016; doi:10.1038/ncomms12429
30. Burd CE, Jeck WR, Liu Y, Sanoff HK, Wang Z, Sharpless NE. Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS genetics*. 2010; 6:1001233.
31. Jarinova O, Stewart AF, Roberts R, Wells G, Lau P, Naing T, Buerki C, McLean BW, Cook RC, Parker JS, McPherson R. Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. *Arterioscler Thromb Vasc Biol*. 2009; 29:1671-1677.

32. Kong Y, Hsieh CH, Alonso LC. ANRIL: A lncRNA at the CDKN2A/B Locus With Roles in Cancer and Metabolic Disease. *Front Endocrinol (Lausanne)*. 2018; doi:10.3389/fendo.2018.00405
33. Zaiou M. The Emerging Role and Promise of Circular RNAs in Obesity and Related Metabolic Disorders. *Cells*. 2020; doi: 10.3390/cells9061473.
34. Weiner DA, Ryan TJ, McCabe CH, Chaitman BR, Sheffield LT, Fisher LD, Tristani F. Value of exercise testing in determining the risk classification and the response to coronary artery bypass grafting in three-vessel coronary artery disease: a report from the Coronary Artery Surgery Study (CASS) registry. *Am J Cardiol*. 1987;60: 262–266.
35. Knopfholz J, Disserol CC, Pierin AJ, et al. Validation of the friedewald formula in patients with metabolic syndrome. *Cholesterol*. 2014; doi:10.1155/2014/261878
36. Pedregosa, Fabian, *et al.* "Scikit-learn: Machine learning in Python." *Journal of machine learning research*.2011; 2825-2830.
37. DeLong, Elizabeth R, David M, DeLong, Daniel L. Clarke-Pearson. "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach." *Biometrics* 44.3 1988; 837-845.
38. Hunter, John D. "Matplotlib: A 2D graphics environment." *Computing in science & engineering* 9.3 2007; 90.
39. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med Inform Decis Mak*. 2019; doi: 10.1186/s12911-019-1004-8
40. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *Eur Heart J*. 2017; doi:10.1093/eurheartj/ehw302,
41. Ambale-Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, Gomes AS, Folsom AR, Shea S, Guallar E, Bluemke DA, Lima JAC. Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis. *Circ Res*. 2017; doi:10.1161/ CIRCRESAHA.117.311312,
42. Jiang Y, Zhang X, Ma R, Wang X, Liu J, Keerman M, Yan Y, Ma J, Song Y, Zhang J, He J, Guo S, Guo H. Cardiovascular Disease Prediction by Machine Learning Algorithms Based on Cytokines in Kazakhs of China. *Clin Epidemiol*. 2021; doi: 10.2147/CLEP.S313343
43. Zhang P, Wu W, Chen Q, Chen M. Non-Coding RNAs and their Integrated Networks. *J Integr Bioinform*. 2019; doi: 10.1515/jib-2019-0027.
44. Awan HM, Shah A, Rashid F, Shan G. Primate-specific Long Non-coding RNAs and MicroRNAs. *Genomics Proteomics Bioinformatics*. 2017; doi: 10.1016/j.gpb.2017.04.002.
45. Holdt LM, Kohlmaier A, Teupser D. Molecular functions and specific roles of circRNAs in the cardiovascular system. *Noncoding RNA Res*. 2018; doi: 10.1016/j.ncrna.2018.05.002
46. Gandhi S, Ruehle F, Stoll M. Evolutionary Patterns of Non-Coding RNA in Cardiovascular Biology. *Noncoding RNA*. 2019; doi:10.3390/ncrna5010015.

47. Chi JS, Li JZ, Jia JJ, Zhang T, Liu XM, Yi L. Long non-coding RNA ANRIL in gene regulation and its duality in atherosclerosis. *J Huazhong Univ Sci Technolog Med Sci.* 2017; doi: 10.1007/s11596-017-1812-y.
48. Rahimi E, Ahmadi A, Boroumand MA, Mohammad Soltani B, Behmanesh M. Association of ANRIL Expression with Coronary Artery Disease in Type 2 Diabetic Patients. *Cell J.* 2018; doi: 10.22074/cellj.2018.4821.
49. Arslan S, Berkan Ö, Lalem T, Özbilüm N, Göksel S, Korkmaz Ö, Çetin N, DevauxY; Cardioline™ network. Long non-coding RNAs in the atherosclerotic plaque. *Atherosclerosis.* 2017; doi:10.1016/j.atherosclerosis.2017.10.012.
50. Yao Hu, Jing Hu. Diagnostic value of circulating lncRNA ANRIL and its correlation with coronary artery disease parameters. *Braz J Med Biol Res.* 2019; doi: 10.1590/1414-431X20198309
51. Zheng Qin, Danan Liu. Circulating lncRNA ANRIL level positively correlates with disease risk, severity, inflammation level and poor prognosis of coronary artery disease. *Int J Clin Exp Med* 2019; 12:8964-8970
52. Liu X, Li S, Yang Y, Sun Y, Yang Q, Gu N, Li J, Huang T, Liu Y, Dong H, Sun S, Fu G, Wu J, Yu B. The lncRNA ANRIL regulates endothelial dysfunction by targeting the let-7b/TGF- β 1 signalling pathway. *J Cell Physiol.* 2021; doi: 10.1002/jcp.29993.
53. Yang J, Gu L, Guo X, Huang J, Chen Z, Huang G, Kang Y, Zhang X, Long J, Su L. LncRNA ANRIL Expression and ANRIL Gene Polymorphisms Contribute to the Risk of Ischemic Stroke in the Chinese Han Population. *Cell Mol Neurobiol.* 2018; doi: 10.1007/s10571-018-0593-6.
54. Song CL, Wang JP, Xue X, Liu N, Zhang XH, Zhao Z, Liu JG, Zhang CP, Piao ZH, Liu Y, Yang YB. Effect of Circular ANRIL on the Inflammatory Response of Vascular Endothelial Cells in a Rat Model of Coronary Atherosclerosis. *Cell Physiol Biochem.* 2017; doi: 10.1159/000478918.
55. Liu Y, Yang Y, Wang Z, Fu X, Chu XM, Li Y, Wang Q, He X, Li M, Wang K, Wang JX, Li PF, Yu T. Insights into the regulatory role of circRNA in angiogenesis and clinical implications. *Atherosclerosis.* 2020; doi: 10.1016/j.atherosclerosis.2020.02.017.
56. Yari M, Bitarafan S, Broumand MA, Fazeli Z, Rahimi M, Ghaderian SMH, Mirfakhraie R, Omrani MD. Association between Long Noncoding RNA ANRIL Expression Variants and Susceptibility to Coronary Artery Disease. *Int J Mol Cell Med.* 2018; doi: 10.22088/IJMCM.BUMS.7.1.1.
57. Liu Y, Sanoff HK, Cho H, Burd CE, Torrice C, Mohlke KL, Ibrahim JG, Thomas NE, Sharpless NE. INK4/ARF transcript expression is associated with chromosome 9p21 variants linked to atherosclerosis. *PLoS One.* 2009; doi:10.1371/journal.pone.0005027.
58. Marchington JM, Pond CM. Site-specific properties of pericardial and epicardial adipose tissue: the effects of insulin and high-fat feeding on lipogenesis and the incorporation of fatty acids in vitro. *Int J Obes* 1990; 14: 1013-1022
59. Wronska A, Kmiec Z. Structural and biochemical characteristics of various white adipose tissue depots. *Acta Physiol (Oxf)* 2012; 205: 194-208

60. Atalar F, Gormez S, Caynak B, Akan G, Tanriverdi G, Bilgic-Gazioglu S, GunayD, Duran C, Akpinar B, Ozbek U, Buyukdevrim AS, Yazıcı Z. Mediastinal adiposetissue expresses a pathogenic profile of 11 β -hydroxysteroid dehydrogenase Type1, glucocorticoid receptor, and CD68 in patients with coronary artery disease. *Cardiovasc Pathol*. 2013; doi: 10.1016/j.carpath.2012.07.006.
61. Atalar F, Gormez S, Caynak B, Akan G, Tanriverdi G, Bilgic-Gazioglu S, Gunay D, Duran C, Akpinar B, Ozbek U, Buyukdevrim AS, Yazici Z. The role of mediastinal adipose tissue 11 β -hydroxysteroid d ehydrogenase type 1 and glucocorticoid expression in the development of coronary atherosclerosis in obese patients with ischemic heart disease. *Cardiovasc Diabetol*. 2012; doi: 10.1186/1475-2840-11-115.
62. Nancy P, Swaminathan B, Navina K, Nandhine B and Lokesh P. Tuned Random Forest Algorithm for Improved Prediction of Cardiovascular Disease. *International Journal of Recent Technology and Engineering (IJRTE)*. 2020; doi:10.35940/ijrte.A1599.059120
63. Fang J, Pan Z, Wang D, Jialan Lv, Dong Y, Xu R, Jin Y, Sheng J, Yin X, Xie X, Wang X, Guo X. Multiple Non-coding ANRIL Transcripts Are Associated with Risk of Coronary Artery Disease: a Promising Circulating Biomarker. *Journal of Cardiovascular Translational Research*. 2020; doi.org/10.1007/s12265-020-10053-0.
64. Harismendy O, Notani D, Song X, Rahim NG, Tanasa B, Heintzman N, Ren B, Fu XD, Topol EJ, Rosenfeld MG. 9p21 DNA variants associated with coronary artery disease impair interferon- γ signalling response. *Nature* 2011; doi: 10.1038/nature09753
65. Yap KL, Li S, Mun~ oz-Cabello AM, Raquz S, Zeng L, Mujitaba S, *et al*. Molecular interplay of the non-coding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* 2010;38:662–74.
66. Chi JS, Li JZ, Jia JJ, Zhang T, Liu XM, Yi L. Long non-coding RNA ANRIL in gene regulation and its duality in atherosclerosis. *J Huazhong Univ Sci Technolog Med Sci*. 2017; 37:816-822.
67. Zhang B, Wang D, Ji TF, Shi L, Yu JL. Overexpression oflncRNA ANRIL up-regulates VEGF expression and promotesangiogenesis of diabetes mellitus combined with cerebral infarction by activating NF-kappaB signaling pathway in a. rat model. *Oncotarget*. 2017; 8:17347–17359.

Figures

Figure 1

Relative expression levels of ANRIL, NR003529, EU741058, DQ485454 andcicrANRIL in EAT, MAT, SAT and in PBMCs among subgroup includes 25 CAD patients undergoing CABG and 25 non-CAD patients undergoing valve placement. A: Relative expression levels of ANRIL in CAD and non-CAD patients; B: Relative expression levels of NR003529 in CAD and non-CAD patients; C: Relative expression levels of EU741058 in CAD and non-CAD patients; D: Relative expression levels of DQ485454 in CAD and non-CAD

patients and E: Relative expression levels of circANRIL in CAD and non-CAD patients. AU: Arbitrary Unit. PDMC: Peripheral Blood Mononuclear Cells. *p<0.05

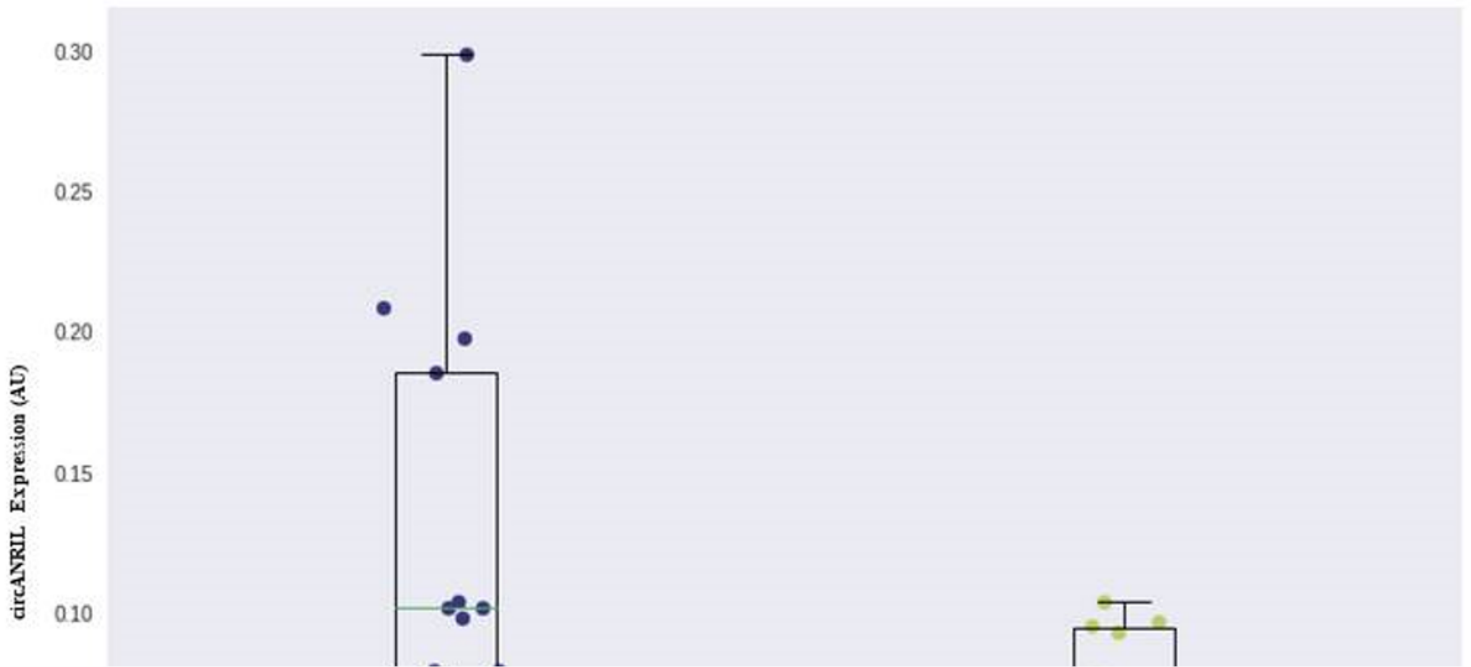


Figure 2

Association of expression levelsof circANRIL with severity of CAD.

Feature Importances for CAD Prediction

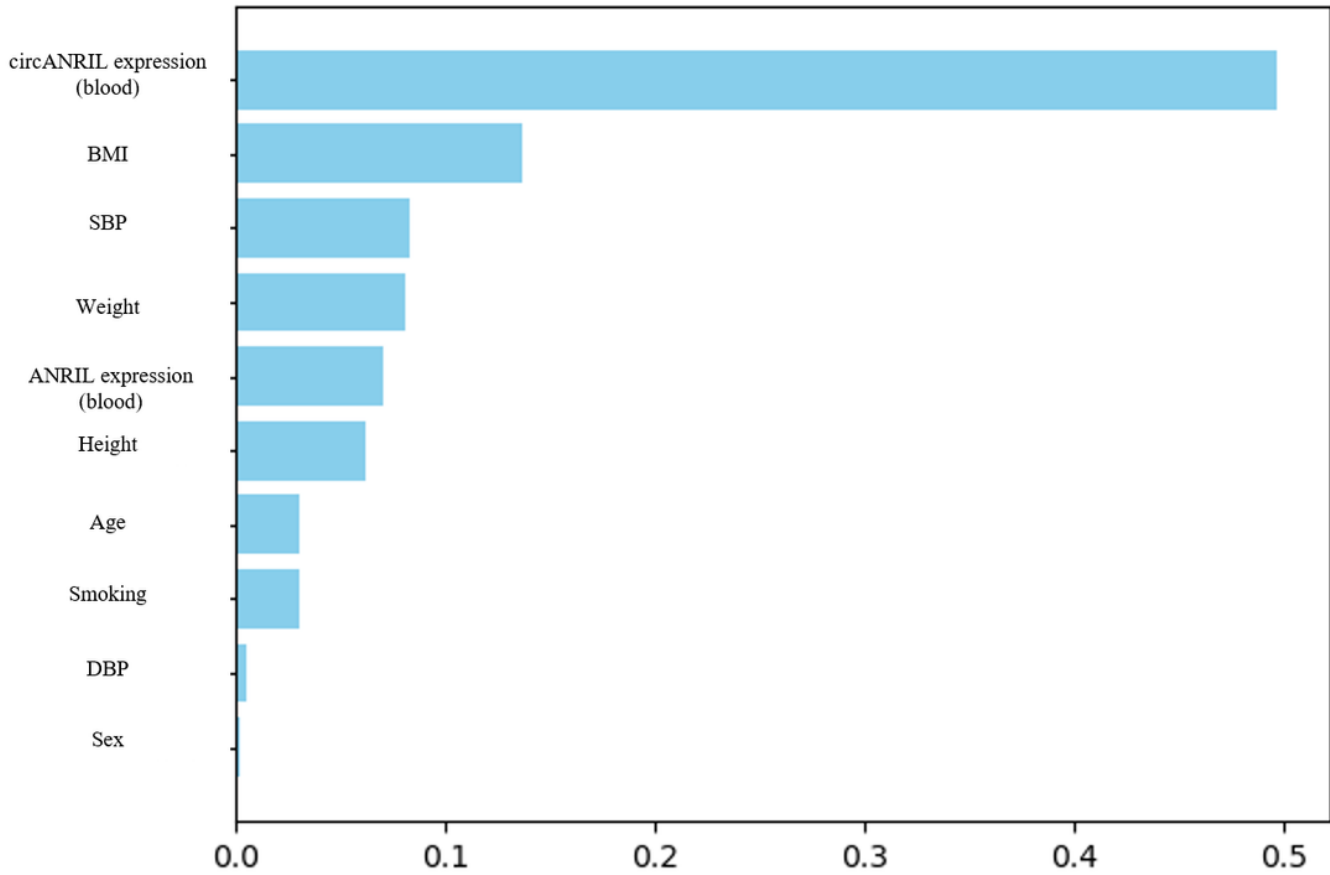


Figure 3

Feature importance analysis of included variables obtained from a tuned random forest model in Tanzanian CAD patients. BMI: body mass index, SBP: systolic blood pressure, DBP: diastolic blood pressure.

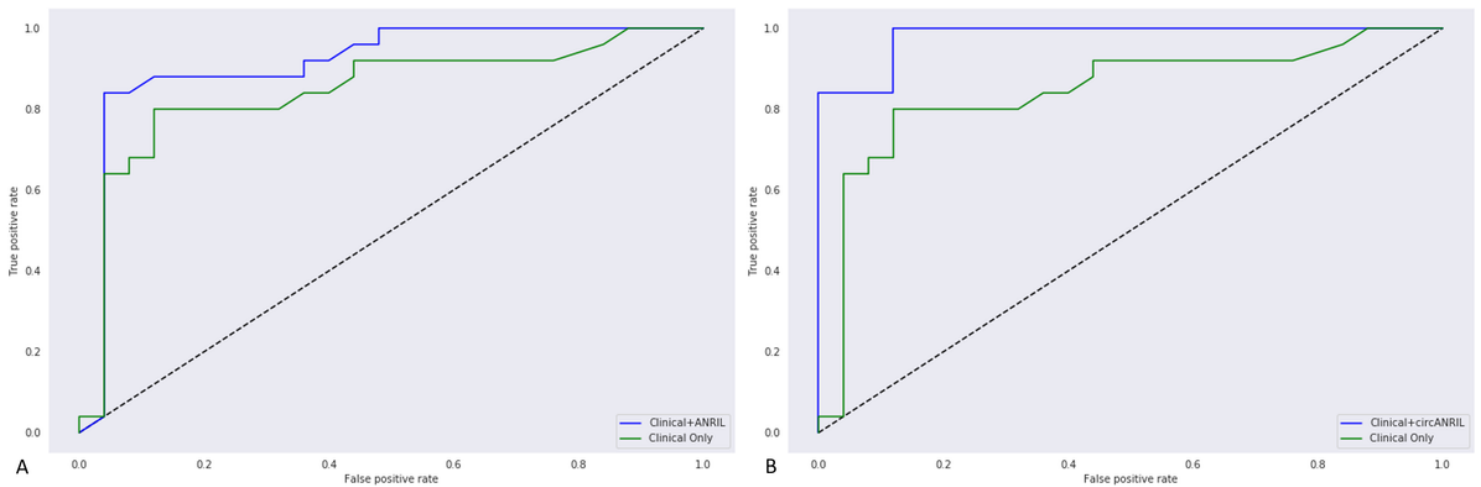


Figure 4

ROC Curve Comparisons between models. A: clinical model and clinical + ANRIL expression model, B: clinical model and clinical + circANRIL expression model. AUC: area under curve.

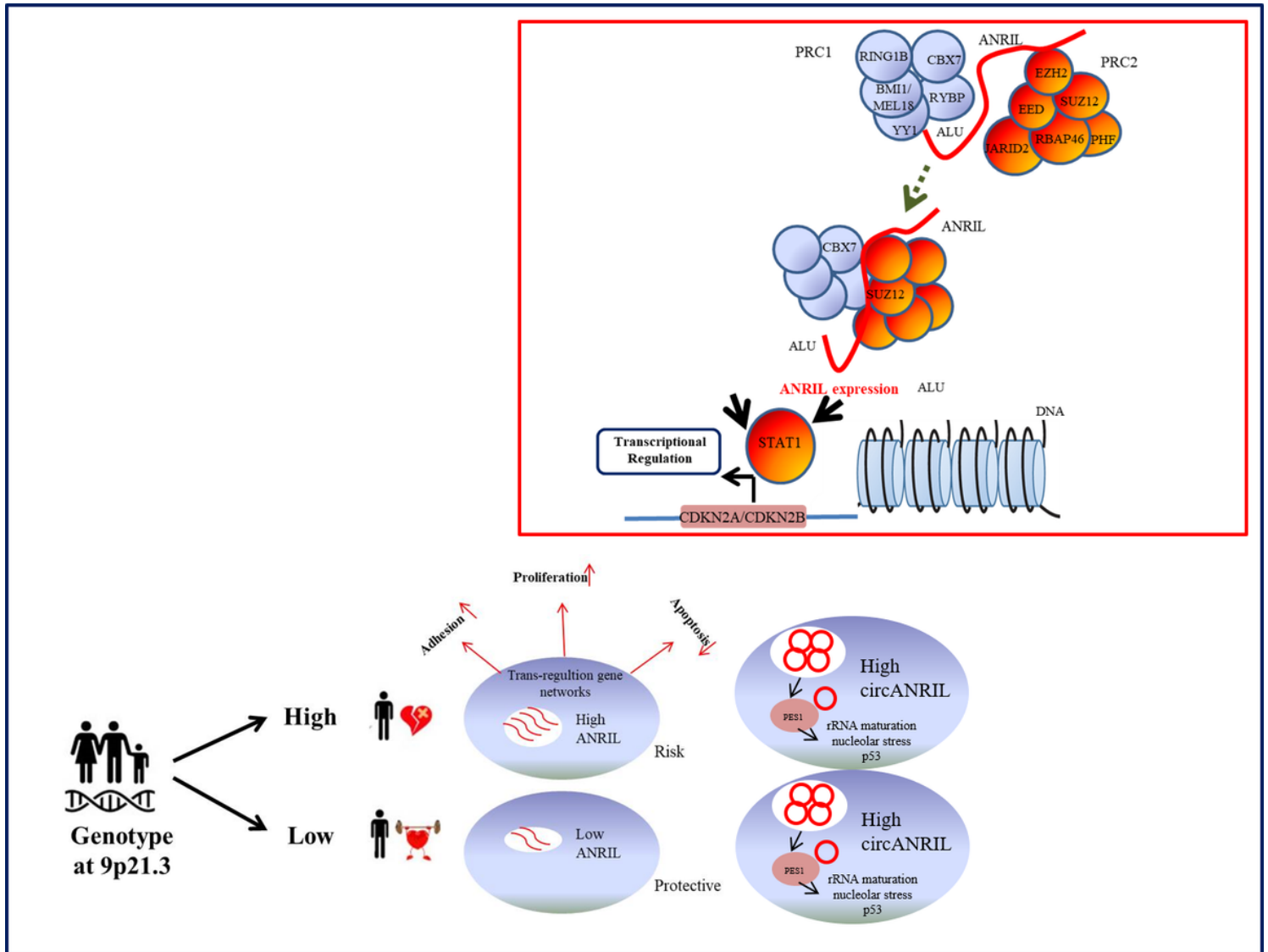


Figure 5

Model of linear ANRIL and circANRIL functions according to the genotype of 9p21.3. The risk alleles lead to up-regulation of ANRIL and down-regulation of circANRIL. Increased ANRIL expression regulated of the expression of adjacent protein-coding genes, including CDKN2A and CDKN2B leading to pro-atherogenic cell properties (increased cell adhesion, increased proliferation, decreased apoptosis) through CBX7 and SUZ12 function in the Polycomb complex. circANRIL binds PES1 protein and impairs ribosome biogenesis, leading to activation of p53 and a subsequent increase in apoptosis and a decrease in proliferative rate. Abbreviations: PES1: Pescadillo Ribosomal Biogenesis Factor 1, CBX7: Chromobox homolog 7, SUZ12: Polycomb Repressive Complex 2 Subunit, p53: tumor protein 53.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable1.docx](#)