COPY NUMBER VARIANTS: DISTRIBUTION IN PATIENTS WITH CORONARY ATHEROSCLEROSIS

K. Gancheva1, A. Postadjian2, D. Brazma1, C. Grace3, A. Chanalaris3, E. Nacheva1, M.D. Apostolova1
Institute of Molecular Biology “Acad. Rumen Tsanev”, Laboratory for Medical and Biological Research, Bulgarian Academy of Sciences, Sofia, Bulgaria
University Hospital “St. Anna”, Clinic of Cardiology, Sofia, Bulgaria2
Royal Free and UCL Medical School London, Department of Haematology, Cyto genetic Laboratory, London, UK3
Correspondence to: Margarita Apostolova
E-mail: margo@obzor.bio21.bas.bg

ABSTRACT
The aim of this study was to investigate the distribution of CNVs in patients with coronary atherosclerosis and to assess the association between them. A total number of 31 subjects (13 Females and 18 Males) were involved in the study. They were divided into two groups according to the clinical diagnosis. The first group consisted of 21 patients with non-ST segment elevation ACS (unstable angina and non ST elevation myocardial infarction) and the second – from 10 healthy subjects. The number of CNVs observed using aCGH kit was 334. One hundred and twenty six (37.73%) are newly observed, 153 out of all 334 were from gene coding regions. The genes, which contain newly described CNVs, and their products were found to have role in cellular metabolism, regulation of transcription, transport and signal transduction. The present study suggests that there is a relation between CNVs described by us and the possible processes involved in the development of CAD. These observations need to be verified on a larger group of patients to clarify the role of these possible links.

Keywords: atherosclerosis, array comparative genomic hybridization (aCGH), copy number variations (CNVs)
Abbreviations: aCGH: array comparative genomic hybridization; ACS: acute coronary syndromes; CAD: coronary artery disease; CNVs: copy number variations; CTNND1: cadherin-associated protein- delta1; ECG: electrocardiogram; FR: fluorescence ratios; GO: gene ontology; NCBI36: National Center for Biotechnology Information (Human database; Build 36.1); OMIM: Online Mendelian Inheritance in Man; qPCR: quantitative polymerase-chain reaction; SD: standard deviation; SNPs: single nucleotide polymorphisms; TCAG: The Centre for Applied Genomics; TIGR: The Institute for Genomic Research; TNF-α: Tumor necrosis factor–alpha; UCSC hg18: University of California, Santa Cruz, Genome Bioinformatics, version hg18);

Introduction
Atherosclerosis is a chronic, multifocal immunoinflammatory, fibroproliferative disease of medium–sized and large arteries mainly driven by lipid accumulation. Coronary artery disease (CAD) involves two distinct processes: a fixed and barely reversible process that causes gradual luminal narrowing slowly over decades (atherosclerosis) and a dynamic, and potentially reversible process that punctuates the slow progression in a sudden, and unpredictable way, causing rapid complete or partial coronary occlusion (thrombosis, vasospasm or both). Generally, atherosclerosis predominates in lesions responsible for clinically silent or chronic stable angina, whereas thrombosis constitutes the critical component of culprit lesions responsible for the acute coronary (unstable angina, acute myocardial infarction) or cerebral (transitory ischemic attack, stroke) syndromes (18). Although there was a significant progress in the treatment of atherothrombosis and its life-threatening manifestations, still they are the leading cause of death in the European community (1.5 million cases of death per year). According to the official data of the National Center for Health Information for 2006, the highest percent of the patients hospitalized in Bulgaria were those with diseases of circulatory system (International Classification of Diseases, 10th version) – 14.8% or 244 979 patients. The international comparisons of the mortality-rate using standardized coefficients to eliminate the influence of the different aged structures of the population in different countries showed that Bulgaria is on the first place for mortality-rate caused by the diseases of cardiovascular system among the 27 countries-members of European community1.

High-throughput technologies facilitate identification of genetic, genomic, proteomic, and metabolomic markers of CAD risk that may find a place in clinically useful prediction algorithms. The operation of intricate networks of genes, environmental factors, and gene-by-environmental interactions further complicates our understanding of the genetic components of atherosclerosis (15).

Until recently, single nucleotide polymorphisms (SNPs) were thought to be the predominant form of genomic variation and to account for phenotypic variation in patients with CAD and in normal subjects as well. However with the advent and

1 statistical data base, WHO, June 2007 http://data.euro.who.int/hfadb
application of array-based comparative genomic hybridization (aCGH) assays which allow analysis of the genome at a significantly higher resolution than previously possible, scientists, have demonstrated that humans are much more genetically variable than previously accounted. In two different publication in 2004 (13, 23) researches found hundreds of genomic regions that varied significantly with respect to the number of copies (CNVs). Since then, the observations of Iafrate (13) and Sebat (23) have been replicated and expanded (4, 5, 12, 14, 20, 21, 24, 28, 29). CNVs vary greatly in size, with variants ranging from insertions or deletions of under 1 kb (commonly described as indels) to several Mb in length (7). Like other types of genetic variation, some gene CNVs have been associated with susceptibility or resistance to disease (1, 3, 6, 11, 19, 22, 25).

In humans, CNVs encompass more DNA than SNPs and may be responsible for a substantial amount of human phenotypic variability and disease susceptibility (8, 19).

There is only one review describing the possible significance of CNVs in the CAD’s development (17). The authors described few genes that were found in other wide gene observations (19, 31) and had a role in different cardiovascular diseases (17). Currently, the role of CNVs found in CAD is poorly understood and it would be of great interest to discover genetic associations between cardiovascular disease and CNVs. We sought to assess the association between CNVs and CAD in patients with acute coronary syndromes (ACS) from Bulgarian population.

Materials and Methods

The information and genetic samples provided by all individuals in this study were obtained with written informed consent in accordance with the approval of the institutional Ethics Committees of the “St. Anna Hospital”.

Between December 2006 and July 2007, a total number of 31 subjects (13 Females and 18 Males) were involved in the study. Patients were divided into two groups according to the clinical diagnosis. The first group consisted of 21 patients with non-ST segment elevation ACS (unstable angina and non ST elevation myocardial infarction). Unstable angina patients had ischemic chest pain at rest within the preceding 48 hours that had developed in the absence of an extracardiac precipitating cause with either ST segment depression of >0.1 mV or T-wave inversion in two or more contiguous leads on the presenting 12 lead ECG. Patients with non ST elevation myocardial infarction had similar diagnostic criteria along with elevation of serum troponin T, without the evolution of pathological q-waves. The control group comprised 10 healthy volunteers with the same age distribution, without angina symptoms and with normal physical examination and stress test. Subjects with acute or chronic inflammatory diseases, malignancies, renal insufficiency, severe liver disease, on immunosuppressive and antibiotic treatment, were excluded from the study. Patients with acute ST elevation myocardial infarction, diabetes mellitus, and history of myocardial infarction, surgical intervention or major trauma within the preceding month were also excluded. In patients with ACS blood was collected after admittance to the intensive care unit at “St. Anna Hospital” before starting the anti-ischemic and anti-coagulation therapy.

DNA preparations

Genomic DNA was extracted from whole peripheral blood cells using sodium extraction protocol (16). The purity of the isolated DNA was confirmed with an A260/280 ratio of 1.80-2.00 for all samples tested. The DNA quality was checked on agarose gel, followed by ethidium bromide staining (30).

Array CGH

Commercially available Human CGH Microarray Kit 105A was received from Agilent Technologies (USA). The features of the kit are:

1) 99,000 + coding and non-coding human sequences represented;
2) Probes annotated against build NCBI36 (UCSC hg18);
3) 21.7kb overall median probe spacing (18.9kb);

The array slides have been hybridized for each sample, swapping Cy3 and Cy5 fluorochromes for test and sex mis-matched reference DNA. Full details of the array construction and layout are available on the Agilent Technologies web page at www.chem.agilent.com. DNA preparation, labeling, hybridization, and washing procedures followed the manufacturer’s protocols (Agilent Technologies, USA).

Sex mis-matched reference DNA

Both male and female DNA was purchased from Promega UK (Promega, Southampton, UK). The DNA represented a mixture derived from six unrelated disease-free individuals of the same sex, but of unknown ethnic origin, medical history and age (2).

Imaging, data analysis and presentation

All procedures were performed as described previously by Brazma et al., (2007) (2). In brief, following hybridization slides were scanned at 10 µm resolution in a two-channel laser scanner - GenePix 4100B (MDS Analytical Technologies, USA), using GenePix v.6 software generating image (.tiff) and spot analysis (.gpr) files. Median average spot and background intensities were extracted from the .gpr file and further processed with TIGR Midas software (The Institute for Genomic Research, www.tigr.org) to correct background, average spot intensities, dye swap, rogue results, and for normalization. Fluorescence ratios (FR) for each of the spots on the slide for each sample were displayed graphically using Formatter2 software developed by Cytogenetic Laboratory, Department of Haematology, Royal Free and UCL Medical School London) either for single or multiple samples with thresholds set at 3SD from the average. Limits were calculated for any spot beyond which a deviation of the FR from unity could be considered amplified or deleted. This technique allowed us to compare genome profiles with varying quality in multiple patients’ samples (2).

BIOTECHNOL. & BIOTECHNOL. EQ. 23/2009/1
Results and Discussion

In order to understand the potential significance of CNVs in the atherosclerosis we have studied 21 Bulgarian patients with ACS (unstable angina pectoris and/or myocardial infarction) and healthy individuals. Using commercial aCGH kit we observed 334 different CNVs, which frequencies are more then 25% in the patients with atherothrombosis. There are 27 different CNVs in the control subject with frequency more than 25%. Eleven of those 27 CNVs discovered in the control group were the same as in the group of atherosclerotic patients, and were excluded from further analysis.

To assess what proportion of the variant regions discovered in this study was novel, we compared our data with those of the Database of Genomic Variants (TCAG, http://projects.tcag.ca/variation/). Two hundred and eight (62.3%) CNVs overlapped loci reported in the TCAG. Another 126 regions were not represented in the TCAG database. These results validate the capability of our array to detect known copy-number alterations, and indicate that 37.73% (126 out of all-334) genomic regions we identified are novel. From all 334 CNVs observed 153 (45.8%) contained genes listed in OMIM. Fifty-four (42.9%) from the newly described CNVs are from coding regions. The distribution of these 153 CNVs among the chromosomes, according to the location, is presented on Fig. 1. Imbalances with features of CNVs were detected mostly on chromosome 19.

Fig. 1. Chromosome distribution chart
The localization of the genes with observed copy number variations (frequencies more then 25%, n=153) among the chromosomes in patients with acute coronary syndromes. Each chromosome is represented by a bar and each gene is shown by a cross symbol.

In order to investigate the potential biological consequences of the CNVs observed, we further conducted a gene ontology (GO) analysis. The distribution by their tissue expression profile showed that the main part is predominantly expressed in the brain (111 out of 153), and 78 in the heart. Some of the novel CNVs are in the gene regions that may have an important role in development of CAD.

A search for common structural features of the genomic loci shown to contain CNVs unique to this cohort of CAD patients showed several traits. Genomic profiling of the genes that overlap the novel CNVs, using the GoTree Machine (32), revealed an enrichment for genes involved in the cell adhesion and function, ion-binding and control of enzyme activity. For better understanding the importance of these genes and their role in the processes of atherosclerosis, we examined and described them according to: i) molecular function; and ii) role in various biological processes.

In dependence of their molecular function genes can be divided in a few groups that have different activities: ion binding, nucleotide binding, protein binding, transcriptional factors, etc. (Fig. 2 and Fig. 3). According to the ion binding activity there are CNVs in the gene coding proteins binding calcium, potassium, magnesium, zinc and copper. It is known that the alteration in the expression and function of ion channels occurs in myogenic and non-myogenic vascular cells (smooth muscle and endothelial cells) during the development of atherosclerosis (9). The vascular ion channels play a key role in regulation of vascular cell contractility and homeostasis. By neuronal stimulation calcium channels contribute to cross-membrane signaling and maintain the contractility of the vascular smooth muscle cells. Potassium channels are sensitive to voltage changes and mechanical stress in the endothelial cells. Disorder of these ion channels and downstream events, may cause a dysfunction of the vascular wall and muscle tones, maintenance, strength and duration of contractions, thus affecting the cardiac function.

Fig. 2. Number of gens with observed novel CNVs, coding proteins grouped by its molecular function, observed in patients with acute coronary syndrome.

Fig. 4 presents the number of gens with observed CNVs coding proteins, grouped by their participation in different biological processes. Most of the proteins are involved in cellular metabolism and its regulation, regulation of transcription, transport and signal transduction. More important of the genes with enriched copy number, within the
patients with atherosclerosis, was the gene encoding catenin (cadherin-associated protein- delta1, CTNND1). Catenins are proteins found in complexes with cadherin cell adhesion molecules of human cells. Cadherins are a super-family of calcium dependent cell-cell adhesion proteins that form the bases of adherence junction (10). These junctions play a critical role in small and large arteries, including coronary and carotid arteries. Cell-cell contacts and extracellular matrix protein network play an important role in the regulation of the stability of atherosclerotic plaques. Enrichment of the CTNND1 can lead to formation of additional catenin-cadherin complexes, which play a pivotal function in adhesion processes and transduction of different cell signals. Another gene with enriched copy number is that for sphingosine kinase, an enzyme that catalyzes a key step in the pathway by which TNF-α stimulates the expression of endothelial cell adhesion molecules (26). These signals can activate endothelial, smooth muscle, and peripheral blood mononuclear cells, and trigger the generation of proinflammatory cytokines and adhesion molecules (27), exacerbating atherosclerosis.

![Fig. 3. Tree Chart of GO categories for significantly changed gene numbers in the novel CNVs compared to the Human reference gene set. Categories in “cursive” represent significantly changed genes compared to the reference gene set (Hypergeometric test, p<0.001).](image)

![Fig. 4. Number of gens with observed novel CNVs, coding proteins grouped by its participation in different biological process, observed in patients with acute coronary syndrome.](image)

There are continuously arising data of novel CNVs and we could not be sure if the CNVs observed in our study are directly connected with coronary artery diseases. Our observations need to be further verified by qPCR on a larger group of patients to clarify the possible link between the observed CNVs and the pathogenesis of the disease. However, with the lack of complete overlap between the CNVs detected in our study and those identified elsewhere, and the hypothesis that thousands of CNVs exist in the genome, and they are yet to be described, this comprehensive study is an early step toward
better understanding of CNVs within the human population, and more studies are needed to examine the functions of CNVs and their role in different diseases.

Conclusions
In this study we examined the diversity of CNVs in 21 patients with ACS and 10 healthy subjects, and demonstrated the presence of 126 newly observed genomic variance out of 334 in genes, which products are involved in developing of the CAD. Identifying the genes and factors that render patients more susceptible to ACS may help in the treatment and management of CAD.

Acknowledgments
Part of this work was supported by the National Science Fund of Bulgaria (Grant N G2/04).

REFERENCES


