MATERIALS AND METHODS

A myocardial infarction-associated SNP at 6p24 interferes with MEF2 binding and associates with *PHACTR1* expression levels in human coronary arteries

Mélissa Beaudoin[1,*], Rajat M. Gupta[2,3,*], Hong-Hee Won[4-7], Ken Sin Lo[1], Ron Do[4-7], Christopher A. Henderson[2], Claire Lavoie-St-Amour[1], Simon Langlois[1], Daniel Rivas[8], Stéphanie Lehoux[8], Sekar Kathiresan[4-7], Jean-Claude Tardif[1,9], Kiran Musunuru[2,3], Guillaume Lettre[1,9]

Affiliations
[1]Montreal Heart Institute, 5000 Bélanger Street, Montréal, Québec, H1T 1C8, Canada.
[2]Department of Stem Cell and Regenerative Biology, Harvard University, and Harvard Stem Cell Institute, Cambridge, Massachusetts 02138, USA.
[3]Division of Cardiovascular Medicine, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA.
[4]Center of Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA.
[5]Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA.
[6]Department of Medicine, Harvard Medical School, Boston, MA, USA.
[7]Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA.
[8]Lady Davis Institute for Medical Research, McGill University, 3755 Côte Sainte-Catherine, Montreal, Quebec, H3T 1E2, Canada.
[9]Université de Montréal, 2900 Boul. Édouard-Montpetit, Montréal, Québec, H3T 1J4, Canada.
*These authors contributed equally to this work.

Corresponding author
Guillaume Lettre, Tel: 514-376-3330, Fax: 514-593-2539, Email: guillaume.lettre@umontreal.ca.

Cohort description. Participants in this study were recruited from the Montreal Heart Institute (MHI) Biobank and all had four French-Canadian grandparents. Participant characteristics are summarized in **Supplemental Table I**. All cases had documented history of MI. For controls, we excluded patients with MI, percutaneous coronary intervention, coronary artery bypass graft surgery, transient ischemic attack or stroke, peripheral vascular disease, congestive heart failure and angina. Patients for DNA re-sequencing were described previously[1]. All participants gave written informed consent and the MHI ethics committee approved the project.

DNA genotyping and re-sequencing. We used several datasets to choose DNA markers to replicate and fine-map the association between MI and the *PHACTR1* locus in our patient population. First, we retrieved genotype data from the 1000 Genomes Project (60 individuals of Northern European ancestry (CEU))[2] in *PHACTR1* (±20 kb). We calculated linkage LD between rs12526453, the canonical *PHACTR1* SNP, and all markers in the 1000 Genomes Project-CEU dataset (accessed in June 2011). We identified 119 variants with a LD r^2 >0.2 with rs12526453, and using the software Tagger[3], selected 13 additional SNPs that could tag the 119 variants at a r^2 >0.85. Other variants were selected as described in **Supplemental Table III**. In total, 23 DNA sequence variants in *PHACTR1* were genotyped using the Sequenom iPLEX technology in 3,339 DNA samples. The genotype concordance rate estimated from DNA triplicates was >99.9%. The DNA re-sequencing project was described previously[1]. The mean coverage per sample for *PHACTR1* is 68X.

Statistical analysis. We performed quality-control filtering of the genotype data using PLINK[4]. Samples with genotyping success rate <90%, markers with genotyping success rate <95%, monomorphic markers, and markers outside the Hardy-Weinberg equilibrium (P<0.001) were excluded from downstream analyses. For genotype imputation, we removed markers with minor allele frequency (MAF) <1% and used the MaCH/minimac software[5] and the reference haplotypes from the European individuals re-sequenced in the 1000 Genomes Project (2012_March/ALL.chr6.phase1_release_v3.20101123, N=1092)[2, 6]. We kept for association testing only the markers with an imputation quality metric r2_hat >0.6. We performed association testing with mach2dat (http://www.sph.umich.edu/csg/abecasis/MACH/download/) using logistic regression under an additive genetic model; we corrected for sex, age (age at first MI for cases and at baseline for controls), age-squared, hypertension status, type 2 diabetes status and lipid-lowering drugs usage. We also used mach2dat/mach2qtl for conditional and secondary analyses to find independent association signals and to test association between *PHACTR1* SNPs and the main CAD/MI risk factors (hypertension, blood pressure, LDL- and HDL-cholesterol levels, type 2 diabetes, smoking). For quantitative phenotypes (LDL-C, HDL-C, systolic and diastolic blood pressures), phenotypes were corrected for sex, age, age-squared, MI status, anti-hypertensive and lipid-lowering drugs usage, and converted to Z-scores using inverse normal transformation.

We used the sequence kernel association test (SKAT) program[7] to test association between rare (MAF <1%) and exonic *PHACTR1* sequence variants (**Supplementary Table III**, ss836901033, ss836901061 and ss836901090 were polymorphic) and MI risk, correcting for sex, age, age-squared, hypertension status,

diabetes status, and lipid-lowering drug usage. Replication analysis of the original SKAT results was performed in an independent subset of the MHI Biobank that included 870 MI cases and 1494 controls ascertained using the same criteria[1]. We carried out all other statistical analyses with the R statistical package (*www.r-project.org/).*

Cell culture. We purchased pooled HUVEC from Lonza and cultured them in EBM-2 media supplemented with EGM-2 SingleQuot kit. Only cells between passage 2 and 6 were used in our experiments. For VEGF treatment, we plated $1.5x10^5$ HUVEC in a 6-well plate format until they reach confluence. Then, cells were starved overnight in EBM-2 media supplemented with EGM-2 SingleQuots (except VEGF). Following starvation, different amounts of VEGF (2-100ng/mL) and different induction times (5-60 minutes) were tested. For TNFα treatment, we plated $1.5x10^5$ HUVEC in a 6-well plate format for three days and then added 10ng/mL of TNFα (from Peprotech) for 16 hours. For the shear stress experiments, HUVEC were plated in 60 mm dishes with a 0.2% gelatin coating and allowed to reach confluence. Low (5 dynes/cm^2) and high (15 dynes/cm^2) shear stress was applied during 6 or 24 hours using a cone plate apparatus or parallel plate chamber[8]. For gene knockdowns, we transfected HUVEC with siRNA against *MEF2A* (sc-35894) and *MEF2C* (sc-38062) according to the manufacturer's protocol. Briefly, we plated cells one day before transfecting them. We mixed 2μL of 10μM siRNA with 6μL of transfection reagent for 40 minutes at room temperature. After washing the cells with the transfection medium, we incubated them with the transfection mix for 4 hours. We then replaced the transfection medium by EGM2 medium.

RNA extraction and gene expression. For the tissue panel expression, we purchased total RNA from Clontech and Stratagene. We extracted RNA from HUVEC using the RNeasy Plus mini kit from Qiagen. We obtained human coronary arteries from the "Réseau d'Échanges de Tissus et d'Échantillons Biologiques" (RÉTEB) biorepository at the Montreal Heart Institute. All donors provided written informed consent. For the right coronary artery tissues, we extracted RNA from 100 mg of tissue with the Ribopure Kit from Ambion. RNA quality and concentration were measured by Agilent RNA 6000 Nano II assays (Agilent Technologies) on an Agilent 2100 Bioanalyzer. We reverse transcribed 1 μg of total RNA using random primers and the MultiScribe Reverse Transcriptase from Applied Biosystems. We performed qPCR analysis using Platinum SYBR Green qPCR SuperMix-UDG (Life Technologies) on a Mx3000P system from Stratagene or on the Eco Illumina qPCR system (Montreal Biotech) with the following thermal profile: 10 minutes at 95°C, and 40 cycles of: 30 seconds at 95°C, 30 seconds at 55°C and 45 seconds at 72°C following by a melt curve. Expression levels were measured and normalized in relation to the expression levels of the house-keeping gene TATA box-binding protein (*TBP*) or hypoxanthine-guanine phosphoribosyltransferase (*HPRT*) using the ΔΔCt method[9]. The primer sequences are: *PHACTR1* forward (5'-CTGTTGTGATTGAATGTGATGAC-3') and reverse (5'-CCTGTTGCTGAGTTTGATGG-3'); *KLF2* forward (5'-CAAGACCTACACCAAGAGTTCG-3') and reverse (5'-CATGTGCCGTTTCATGTGC-3'); *NFKB1* forward (5'-GACAACTATGAGGTCTCTGG-3') and reverse (5'-ATCACTTCAATTGCTTCGG-3'); *TBP* forward (5'-CGAATATAATCCCAAGCGGTTT-3') and reverse (5'-GTGGTTCGTGGCTCTCTTATCC-3'); *HPRT* forward (5'-TGGCGTCGTGATTAGTGATG-3') and reverse (5'-CAGAGGGCTACAATGTGATGG-

3'); *MEF2A* forward (5'-ATGGTTGTGAGAGCCCTGATG-3') and reverse (5'-AAAAGTTCTGAGGTGGCAGAC-3'); *MEF2C* forward (5'-TTCCAGTATGCCAGCACCG-3') and reverse (5'-GGCCCTTCTTTCTCAACGTCTC-3').

To measure the expression of genes located nearby *PHACTR1* rs9349379 (1-Mb on either side) in human coronary arteries, we used the following PCR primers:

Gene name	Forward primer	Reverse primer
EDN1	ACTTCTGCCACCTGGACATCA	CTCCAAGGCTCTCTTGGACCTA
TBC1D7	AAGTGGGGTTTCGTGGAGTT	CAATGCACGGTACATGGACG
GFOD1	CCAGACAGATCGCTGTCAAA	GTAGCCCTCCTCGATCAGC
SIRT5	TGGTTTGGAGAAAACCTGGA	CCCTGGAAATGAAACCTGAA
NOL7	AACGCTCCTGAAGGAGAAGA	TGGAGTCATTTCCTTTTTCACA
RANBP9	ACAGCCCAGTTGGAAATCAG	GTCTTCCACTGTGGCAAATG
MCUR1 / CDC90A	AATTTTCAGCCCTCAGAGCA	TTCCAGCAGCTTCTTTTCGT
HIVEP1	TTTATGGAGCTGCCTTTTGG	TTTGAAACTTCTGCCCCATT

High-resolution melting (HRM) genotyping. We extracted genomic DNA from blood samples that correspond to the human coronary artery samples used in the gene expression analysis, and performed high-resolution melting to genotype SNP rs9349379 (A/G). We purchased HPLC-purified primers from IDT DNA and used them at a final concentration of 500nM: forward (5'-GGTAATAAATATGTCTATGCCCTTGA-3') and reverse (5'-AAAACTCAGCTCGTGGAAAA-3'). We amplified 10ng of genomic DNA to obtain a product size of 87 bp using the MBI Evolution 5x Evagreen qPCR mix with the following thermal profile: 2 minutes at 50°C, 15 minutes at 95°C, and 40 cycles of: 10 seconds at 95°C, 30 seconds at 60°C and 15 seconds at 72°C. The high-resolution melt curve was generated and analyzed using the Eco Illumina qPCR system. This genotyping assay was validated on 12 genomic DNA samples in triplicates from the MHI Biobank where rs9349379 had already been genotyped: we obtained 100% concordance between HRM-based and Sequenom iPLEX-based genotypes.

Electromobility shift assay (EMSA). EMSA were performed with HUVEC nuclear extract obtained using NE-PER nuclear and cytoplasmic extraction reagents from Pierce following the manufacturer's instructions with $1x10^7$ cells. We purchased oligonucleotides from IDT DNA and biotinylated them following the instructions of biotin 3' end DNA labelling kit from Pierce. The sequences are: Probe C forward (5'-TTGAGATCATCTAAAAA*TAGCTTAAAATCATTG-3'), reverse (5'-CAATGATTTTAAGCTAT*TTTTAGATGATCTCAA-3'); Probe A forward (5'-TTGAGATCATATAAAAA*TAGCTTAAAATCATTG-3'), reverse (5'-CAATGATTTTAAGCTAT*TTTTATATGATCTCAA-3') and Probe G forward (5'-TTGAGATCATATAAAAG*TAGCTTAAAATCATTG-3'), reverse (5'-CAATGATTTTAAGCTAC*TTTTATATGATCTCAA-3'). The stars highlight the rs9349379 position and we underlined the nucleotide substitution used to create the known MEF2 consensus binding site in Probe C. The efficiency of labelling for each oligonucleotide was tested by a dot-test that confirmed that all oligonucleotides were labeled similarly. Complementary biotin-labeled oligonucleotides were mixed together in equimolar

concentration, denature for 5 minutes at 95°C and then allowed to anneal by slow cooling at room temperature. Each EMSA reaction was performed with the LightShift Chemiluminescent EMSA kit from Pierce according to the manufacturer's instructions. For the competition assay, 100-fold excess of unlabeled annealed probe was added and for the supershift assay, 1.8µg of antibody (anti-MEF2 from Santa Cruz Technologies) was used. These reactions were incubated at room temperature for 20 minutes before the addition of each biotin-labeled probe at a concentration of 200fmol. Following a second incubation of 30 minutes, 5µL of loading buffer was added to each reaction. During incubation, a 4% non-denaturing polyacrylamide (19:1) gel was pre-run for 1hr at 100V. Then, each reaction was loaded and migrated for 45min at 100V. After electrophoretic transfer to a Biodyne B nylon membrane from Pierce (1hr at 360mA) and cross-linking, the biotin-labeled oligonucleotides were detected by chemiluminescence according to the protocol.

CRISPR/Cas9 genome editing. The rs9349379 targeted clones were generated as previously described[10]. We summarize the methods below. An efficient and specific protospacer sequence in proximity to the rs9349379 SNP in human DNA was chosen using the Zhang Lab CRISPR Design website (http://crispr.mit.edu). A plasmid based on gRNA_Cloning Vector (https://www.addgene.org/41824/) was used to express the guide RNA with the chosen protospacer sequence from the U6 promoter. pCas9_GFP (https://www.addgene.org/44719/) was used to co-express a human codon-optimized Cas9 gene with a C-terminal nuclear localization signal and GFP from the CAG promoter.

The HUES 9 human embryonic stem cell (hESC) line was grown in feeder-free adherent culture in chemically defined mTeSR1 (STEMCELL Technologies) supplemented with penicillin/streptomycin on plates pre-coated with Geltrex matrix (Invitrogen). Cells were disassociated into single cells with Accutase (Invitrogen), and 10 million cells were electroporated with 25 µg of each of the two CRISPR/Cas9 plasmids in a single cuvette and replated. The cells were collected from the culture plates 48 hours post-electroporation by Accutase treatement and resuspended in PBS. Cells expressing GFP were collected by FACS (FACSAria II; BD Biosciences) and replated on 10-cm tissue culture plates at 30,000 cells/plate to allow for recovery in growth media. Post-FACS, the cells were allowed to recover for 7-10 days, after which single colonies were manually picked and dispersed and replated individually to wells of 96-well plates. Colonies were screened by PCR and Sanger sequencing to identify insertion-deletion (indels) at the rs9349379 target site. One clone had a heterozygous 34 bp deletion flanking the target site (**Figure 4A**) and was chosen for expansion and differentiation along with a wild-type clone from the same 96-well plate.

In three separate experiments, the del34 and wild-type clonal hESC lines were differentiated into endothelial cells with a previously described 10-day protocol[11, 12]. Pure population of endothelial cells were isolated at the end of the 10 days with CD144-positive selection using magnetic beads (Miltenyi Biotech). RNA was harvested from both wild-type and del34 endothelial cells, transcribed into cDNA, and analyzed for *PHACTR1* expression using a TaqMan qPCR probe (Applied Biosystems).

Material and Methods References

1. Beaudoin M, Lo KS, N'Diaye A, Rivas MA, Dube MP, Laplante N, Phillips MS, Rioux JD, Tardif JC, Lettre G. Pooled DNA resequencing of 68 myocardial infarction candidate genes in french canadians. *Circulation. Cardiovascular genetics*. 2012;5:547-554
2. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature*. 2010;467:1061-1073
3. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. *Nature genetics*. 2005;37:1217-1223
4. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. Plink: A tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*. 2007;81:559-575
5. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. Mach: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic epidemiology*. 2010;34:816-834
6. 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491:56-65
7. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *American journal of human genetics*. 2011;89:82-93
8. Ramkhelawon B, Rivas D, Lehoux S. Shear stress activates extracellular signal-regulated kinase 1/2 via the angiotensin ii type 1 receptor. *FASEB J.* 2013;27:3008-3016
9. Schmittgen TD, Livak KJ. Analyzing real-time pcr data by the comparative c(t) method. *Nature protocols*. 2008;3:1101-1108
10. Ding Q, Regan SN, Xia Y, Oostrom LA, Cowan CA, Musunuru K. Enhanced efficiency of human pluripotent stem cell genome editing through replacing talens with crisprs. *Cell stem cell*. 2013;12:393-394
11. White MP, Rufaihah AJ, Liu L, Ghebremariam YT, Ivey KN, Cooke JP, Srivastava D. Limited gene expression variation in human embryonic stem cell and induced pluripotent stem cell-derived endothelial cells. *Stem cells*. 2013;31:92-103
12. Rufaihah AJ, Huang NF, Jame S, Lee JC, Nguyen HN, Byers B, De A, Okogbaa J, Rollins M, Reijo-Pera R, Gambhir SS, Cooke JP. Endothelial cells derived from human ipscs increase capillary density and improve perfusion in a mouse model of peripheral arterial disease. *Arteriosclerosis, thrombosis, and vascular biology*. 2011;31:e72-79