



DIGITAL ACCESS TO
SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU



HARVARD LIBRARY
Office for Scholarly Communication

Methylome-wide Association Study of Atrial Fibrillation in Framingham Heart Study

The Harvard community has made this
article openly available. [Please share](#) how
this access benefits you. Your story matters

Citation	Lin, H., X. Yin, Z. Xie, K. L. Lunetta, S. A. Lubitz, M. G. Larson, D. Ko, et al. 2017. "Methylome-wide Association Study of Atrial Fibrillation in Framingham Heart Study." Scientific Reports 7 (1): 40377. doi:10.1038/srep40377. http://dx.doi.org/10.1038/srep40377 .
Published Version	doi:10.1038/srep40377
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:30371050
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

SCIENTIFIC REPORTS



OPEN

Methylome-wide Association Study of Atrial Fibrillation in Framingham Heart Study

Received: 22 June 2016
Accepted: 05 December 2016
Published: 09 January 2017

Honghuang Lin^{1,2}, Xiaoyan Yin¹, Zhijun Xie^{2,3}, Kathryn L. Lunetta^{1,4}, Steven A. Lubitz^{5,6}, Martin G. Larson^{1,7}, Darae Ko⁹, Jared W. Magnani⁸, Michael M. Mendelson^{10,11}, Chunyu Liu^{1,4,10}, David D. McManus¹², Daniel Levy^{1,10}, Patrick T. Ellinor^{5,6} & Emelia J. Benjamin^{1,9,13,14}

Atrial fibrillation (AF) is the most common cardiac arrhythmia, but little is known about the molecular mechanisms associated with AF arrhythmogenesis. DNA methylation is an important epigenetic mechanism that regulates gene expression and downstream biological processes. We hypothesize that DNA methylation might play an important role in the susceptibility to develop AF. A total of 2,639 participants from the Offspring Cohort of Framingham Heart Study were enrolled in the current study. These participants included 183 participants with prevalent AF and 220 with incident AF during up to 9 years follow up. Genome-wide methylation was profiled using the Illumina Infinium HumanMethylation450 BeadChip on blood-derived DNA collected during the eighth examination cycle (2005–2008). Two CpG sites were significantly associated with prevalent AF, and five CpGs were associated with incident AF after correction for multiple testing (FDR < 0.05). Fourteen previously reported genome-wide significant AF-related SNP were each associated with at least one CpG site; the most significant association was rs6490029 at the *CUX2* locus and cg10833066 ($P = 9.5 \times 10^{-279}$). In summary, we performed genome-wide methylation profiling in a community-based cohort and identified seven methylation signatures associated with AF. Our study suggests that DNA methylation might play an important role in AF arrhythmogenesis.

More than 30 million people worldwide are currently affected by atrial fibrillation (AF)^{1,2}, the most common cardiac arrhythmia. AF is associated with a significantly increased risk for stroke³, dementia⁴, heart failure^{5–7}, myocardial infarction⁸, and death^{9–11}. In the past few years, genome-wide association studies (GWAS) have identified 14 genetic loci associated with AF^{12–15}. However, none of these loci is located in protein coding regions, and the molecular mechanisms underlying the associations remain largely unknown.

DNA methylation refers to the process of adding a methyl group to the cytosine-phosphate-guanine dinucleotides (CpG). DNA methylation stabilizes chromatin structure during transcription, which can regulate many downstream transcriptional processes. The methylation state can be transmitted through cell division, and can vary across multiple tissues¹⁶ or over an individual's lifetime. Recent studies found that alteration of DNA

¹National Heart Lung and Blood Institute's and Boston University's Framingham Heart Study, Framingham, MA, USA. ²Section of Computational Biomedicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA. ³TCM Clinical Basis Institute, Zhejiang Chinese Medicine University, Hangzhou, Zhejiang, China. ⁴Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. ⁵Cardiovascular Research Center, Massachusetts General Hospital, Charlestown, MA, USA. ⁶The Broad Institute of Harvard and MIT, Cambridge, MA, USA. ⁷Department of Mathematics and Statistics, Boston University, Boston, MA, USA. ⁸Division of Cardiology, Department of Medicine, UPMC Heart and Vascular Institute, University of Pittsburgh, Pittsburgh, PA, USA. ⁹Section of Cardiovascular Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA. ¹⁰Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, Bethesda, MD, USA. ¹¹Department of Cardiology, Boston Children's Hospital, Boston, MA, USA. ¹²Medicine and Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, USA. ¹³Section of Preventive Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA. ¹⁴Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA. Correspondence and requests for materials should be addressed to H.L. (email: hlin@bu.edu)

Characteristics	No AF (n = 2,236)	Prevalent AF (n = 183)	Incident AF (n = 220)
Women, n (%)	1285 (57%)	65 (36%)	89 (40%)
Age, year \pm SD	65 \pm 9	72 \pm 9	71 \pm 8
Height, inches \pm SD	66 \pm 4	67 \pm 4	66 \pm 4
Weight, pounds \pm SD	173 \pm 38	190 \pm 48	181 \pm 43
Current smoker, n (%)	206 (9%)	8 (4%)	15 (7%)
Systolic blood pressure, mm Hg	128 \pm 17	128 \pm 20	135 \pm 19
Diastolic blood pressure, mm Hg	74 \pm 10	70 \pm 10	72 \pm 10
Prevalent diabetes mellitus, n (%)	316 (14%)	60 (33%)	59 (27%)
Prevalent myocardial infarction, n (%)	18 (1%)	38 (21%)	8 (4%)
Prevalent heart failure, n (%)	68 (3%)	44 (24%)	14 (6%)
Antihypertensive treatment, n (%)	1016 (45%)	130 (71%)	142 (65%)

Table 1. Clinical characteristics of studied samples.

methylation is linked to many cardiovascular disease-related disorders, such as atherosclerosis^{17–20}, high blood pressure^{21,22}, and diabetes^{23–25}. Several CVD risk factors, including smoking²⁶, obesity²⁷, and alcohol consumption²⁸, are also associated with DNA methylation changes. However, there are no previous studies examining the relation of DNA methylation with prevalent or incident AF among community-based adults.

We hypothesized that DNA methylation might contribute to AF susceptibility. We performed a large-scale DNA methylation profiling study in participants from the Framingham Heart Study, and examined the association of DNA methylation with prevalent and incident AF. Given the importance of DNA methylation to gene expression, we investigated the association of DNA methylation signatures of AF with gene expression from the same group of participants. We also assessed the association of GWAS loci for AF with DNA methylation to examine potential implications of AF-related genetic variants on DNA methylation.

Results

Differential Methylation. Table 1 shows the descriptive characteristics of 2,639 eligible participants (mean age 65 \pm 9 years, 57% women), including 183 participants who developed AF before the eighth examination (prevalent AF), and an additional 220 participants who developed AF after the examination (incident AF) through follow-up in 2014.

Figure 1 shows the Manhattan plots of CpG sites associated with prevalent and incident AF. As shown in Table 2, two CpG were significantly associated with prevalent AF (FDR < 0.05). One significant CpG site was cg13639451 ($P = 1.1 \times 10^{-7}$, FDR = 0.03), located 1447 bp upstream of *WFIKKN2*, which encodes an endopeptidase inhibitor with a known role in muscle fiber development via inhibition of myostatin²⁹. The other one was cg07191189 ($P = 1.4 \times 10^{-7}$, FDR = 0.03), located 75 bp upstream of *STRN* (*Striatin*). The CpG site also was predicted to affect the binding of transcription factor TFII-I³⁰. Five CpG sites were associated with incident AF (FDR < 0.05). The most significant CpG site was cg26602477 ($P = 4.7 \times 10^{-9}$), located 207 bp downstream of *SSU72*. The results were similar when we additionally adjusted for different cell counts (Supplemental Table 1).

In our secondary analysis, we tested the associations of the seven significant prevalent or incident AF associated CpG sites adjusting for selected AF risk factors³¹, including smoking, height, weight, systolic blood pressure, diastolic blood pressure, prevalent diabetes mellitus, prevalent myocardial infarction, prevalent heart failure, and antihypertensive treatment. As shown in Supplemental Table 1, the results were similar to the primary analysis.

We also tested the association of methylation with all AF cases (combining prevalent and incident AF), but none of CpG site was significant after correction for multiple testing. The top CpG sites for all AF are listed in Supplemental Table 2.

We then examined if the inclusion of methylation profiles would improve the prediction of AF compared to traditional risk factors. Three models were tested: Model 1: Only included traditional risk factors associated with incident AF³¹; Model 2: Included traditional risk factors and 14 published AF-related genetic loci^{12–15}; Model 3: Included traditional risk factors, 14 known genetic loci associated with AF, and 5 CpG sites associated with incident AF. As shown in Fig. 2, the inclusion of genetic loci and methylation profiles modestly improved the prediction performance with area under curve (AUC) increasing from 0.729 (model 1) to 0.747 (model 2) and 0.764 (model 3).

DNA Methylation Associated with Gene Expression. DNA methylation is an important mechanism to regulate gene expression. We tested if any of the seven AF-related CpG sites was associated with gene expression or not. We stratified our analysis by *cis*-gene associations (defined as those within 1 Mb of the CpG site), and *trans*-genes (defined as those more than 1 Mb away from the CpG site or in different chromosomes). A total of 175 *cis*-associations and 124,936 *trans*-associations were found and tested. Therefore the significance cutoff was $0.05/175 = 2.9 \times 10^{-4}$ for *cis*-associations and $0.05/124,936 = 4.0 \times 10^{-7}$ for *trans*-associations. The most significant *cis*- and *trans*- genes for each CpG site are shown in Table 3.

None of the *cis*-associations reached significance after adjusting for multiple testing (all with $P > 2.9 \times 10^{-4}$). Two AF-related CpG sites (cg13639451 and cg15440392) were associated with the expression *trans*-genes ($P < 4.0 \times 10^{-7}$). cg13639451 was associated with the expression of *RPS18* ($P = 4.5 \times 10^{-20}$), and cg15440392 was associated with the expression of *GZMH* ($P = 7.0 \times 10^{-9}$).

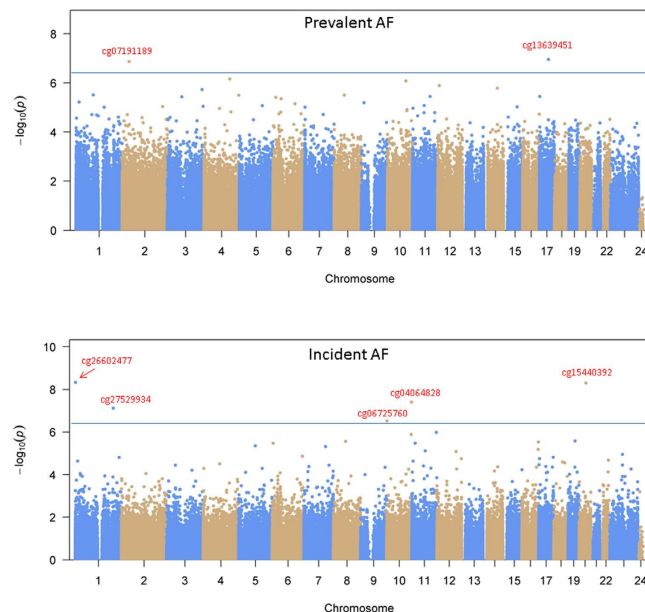


Figure 1. Manhattan plot of CpG sites associated with prevalent and incident AF. The x-axis represents the chromosome, and the y-axis represents the \log_{10} (p-value) of the associations with prevalent and incident AF. The horizontal line represents the significance cutoff (FDR = 0.05). We marked CpGs that were significantly associated with prevalent or incident AF (FDR < 0.05).

Type	CpG site	Chr	Position ⁺	Closest gene	Distance	Effect size	SE*	P value	FDR*
Prevalent AF	cg13639451	17	48911157	<i>WFIKKN2</i>	1447 bp	-0.010	0.002	1.1×10^{-7}	0.030
	cg07191189	2	37193690	<i>STRN</i>	75 bp	0.011	0.002	1.4×10^{-7}	0.030
Incident AF	cg26602477	1	1476845	<i>SSU72</i>	207 bp	-8.289	1.415	4.7×10^{-9}	0.001
	cg15440392	20	36156634	<i>BLCAP</i>	301 bp	4.808	0.823	5.1×10^{-9}	0.001
	cg04064828	10	134002751	<i>DPYSL4</i>	0	-28.669	5.218	3.9×10^{-8}	0.006
	cg27529934	1	205054684	<i>RBBP5</i>	585 bp	-8.291	1.542	7.5×10^{-8}	0.008
	cg06725760	10	1102461	<i>WDR37</i>	314 bp	6.655	1.300	3.1×10^{-7}	0.027

Table 2. Most significant CpG sites associated with AF (FDR < 0.05). ⁺NCBI Genome Build 37. *SE: Standard error; FDR: false discovery rate⁵³.

DNA Methylation Associated with AF-related Genetic Loci. DNA methylation may be regulated by both genetic and environmental factors. Fourteen AF-related SNPs have been previously reported^{12–15}, however, their association with nearby DNA methylation has not been studied yet. For each AF-related SNP, we studied its association with the methylation of CpG sites within 1 Mb. A total of 6042 SNP-CpG pairs were identified and tested; therefore the significance cutoff was defined as $P < 0.05/6042 = 8.3 \times 10^{-6}$. As shown in Table 4, all the AF-related SNPs were significantly associated with methylation of at least one CpG site. The most significant association was between rs6490029 at the *CUX2* locus and cg10833066 ($P = 9.5 \times 10^{-279}$). As shown in Fig. 3, the methylation level of cg10833066 increased with increasing copies of the “A” allele of rs6490029, which also was associated with higher AF risk according to previous GWAS¹³.

Discussion

DNA methylation is an important process regulating gene expression without changing DNA sequence. We performed genome-wide methylation profiling of more than 2600 participants from the Framingham Heart Study Offspring cohort. Two CpG sites were found to be significantly associated with prevalent AF, and five other CpG sites were associated with incident AF. We also examined the association between methylation and GWAS loci for AF, and found that all AF loci were associated with the methylation of at least one CpG site, suggesting DNA methylation might be a possible mechanism through which AF-specific genetic variations affect gene regulation.

Interestingly, the most significant CpG site for prevalent AF (cg13639451) was also associated with the expression of the ribosomal protein *RPS18* ($P = 4.5 \times 10^{-20}$). The gene has been shown to bind to *CAMK2D*³², a gene encoding a calcium/calmodulin-dependent protein kinase II that plays an important role in excitation-contraction coupling in the heart³³. The other significant CpG site for prevalent AF, cg07191189, was located 75 bp upstream of *STRN*, which encodes striatin, a calmodulin-dependent scaffolding protein. Striatin has been shown to bind directly to caveolin-1³⁴, encoded by *CAVI*, which is involved in cardiac development³⁵. Further, the *CAVI* locus

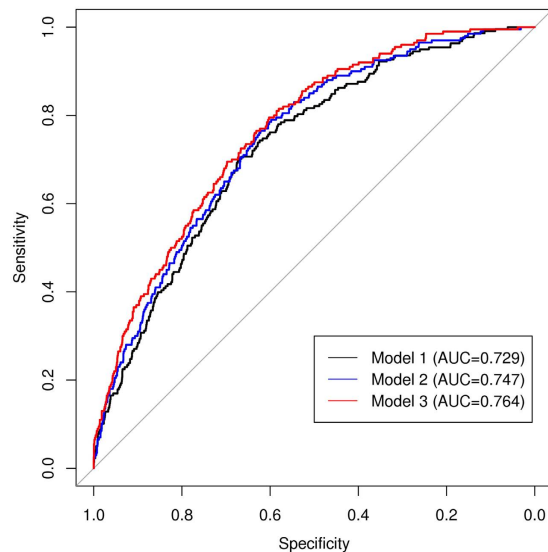


Figure 2. Receiver operating characteristic (ROC) curves of three models to predict incident AF. Model 1: Only included traditional risk factors; Model 2: Included traditional risk factors and 14 AF-related genetic loci; Model 3: Included traditional risk factors, 14 AF-related genetic loci, and 5 AF-related CpG sites. The inclusion of genetic loci and methylation profiles modestly improved the prediction performance with area under curve (AUC) increasing from 0.729 (model 1) to 0.747 (model 2) and 0.764 (model 3).

Type	CpG site	Most significant <i>cis</i> -gene for each AF-related CpG site				Most significant <i>trans</i> -gene for each AF-related CpG site			
		Gene	Effect size	SE	P value	Gene ⁺	Effect size	SE	P value
Prevalent AF	cg13639451	<i>HILS1</i>	0.36	0.14	9.1×10^{-3}	<i>RPS18</i>	1.85	0.20	4.5×10^{-20}
	cg07191189	<i>STRN</i>	0.35	0.12	4.8×10^{-3}	<i>LGSN</i>	0.77	0.18	2.3×10^{-5}
Incident AF	cg26602477	<i>C1orf159</i>	-0.35	0.16	3.1×10^{-2}	<i>PTPN2</i>	-1.37	0.31	1.2×10^{-5}
	cg15440392	<i>RPN2</i>	0.21	0.12	7.8×10^{-2}	<i>GZMH</i>	-2.18	0.37	7.0×10^{-9}
	cg04064828	<i>DPYSL4</i>	0.56	0.23	1.7×10^{-2}	<i>ABCA4</i>	-0.98	0.20	7.9×10^{-7}
	cg27529934	<i>PLEKHA6</i>	0.26	0.09	5.6×10^{-3}	<i>ADAMTS2</i>	0.40	0.09	1.3×10^{-5}
	cg06725760	<i>ZMYND11</i>	-0.22	0.21	2.8×10^{-1}	<i>GZMH</i>	-3.01	0.72	3.2×10^{-5}

Table 3. Association of AF-specific CpG sites with the most significant *cis*-gene expression and *trans*-gene expression. *SE: Standard error. ⁺All the most significant genes are located in different chromosomes from the corresponding CpG sites. *SE: Standard error.

was associated with AF in our previous GWAS¹². Given that all known GWAS loci for AF are located in intergenic or intronic regions, our results suggest that methylation might act as an important bridge to link genetic variation and disease susceptibility.

The most significant CpG site for incident AF was cg26602477, located 207 bp away from *SSU7*, a gene encoding a protein phosphatase that regulates the dephosphorylation process of RNA polymerase³⁶. Another significant CpG site cg15440392 is located within *BLCAP*, which encodes a bladder cancer-associated protein. The gene is highly conserved through evolution and is expressed in a variety of human tissues including cardiac³⁷. However, the implications of these genes for AF risk are still unknown.

The AF-related SNP rs6490029 was significantly associated with CpG site cg10833066, which is located ~400 kb upstream of *ALDH2*, a gene encoding aldehyde dehydrogenase. Aldehyde dehydrogenase plays an important role in the metabolism of alcohol through catalyzing the oxidation of aldehydes into carboxylic acids. Previous studies have proven that excessive alcohol usage is associated with increased risk of AF^{38–40}, suggesting that DNA methylation might be an important factor in AF susceptibility.

It is worth noting that there is no overlap between CpG sites that were associated with prevalent and incident AF. Several reasons might be responsible for the lack of overlap. Unlike genetic variations, the relations between DNA methylation and AF are reciprocal⁴¹, which means that not only DNA methylation could affect the susceptibility of AF, the disease could also influence one's methylation profile. Given that DNA methylation was profiled from whole blood collected during a routine examination, the methylation profile in samples with prevalent AF could have been influenced by AF for an extended period of time. In contrast, for samples with incident AF, the methylation profile could be one of factors that influenced their susceptibility to develop AF. In addition, there

AF SNP	Cloest gene to AF SNP	CpG site	Chr	Position ⁺	Cloest gene to CpG site	Distance to AF SNP	Effect size	SE*	P value	Methylation to AF risk [§]
rs6490029	<i>CUX2</i>	cg10833066	12	111,807,467	<i>FAM109A</i>	109,010 bp	0.1380	0.0034	9.5×10^{-279}	↑
rs10824026	<i>SYNPO2L</i>	cg02286717	10	75,415,704	<i>SYNPO2L</i>	5,504 bp	-0.0578	0.0018	1.3×10^{-190}	↑
rs1152591	<i>SYNE2</i>	cg23250157	14	64,679,961	<i>SYNE2</i>	887 bp	-0.0109	0.0005	2.8×10^{-97}	↓
rs7164883	<i>HCN4</i>	cg06757333	15	73,655,217	<i>HCN4</i>	3,043 bp	-0.0357	0.0019	1.3×10^{-71}	↓
rs3807989	<i>CAV1</i>	cg12739419	7	116,140,593	<i>CAV2</i>	45,648 bp	0.0233	0.0013	1.2×10^{-65}	↓
rs4642101	<i>CAND2</i>	cg24848339	3	12,840,334	<i>CAND2</i>	1,889 bp	-0.0112	0.0006	1.8×10^{-63}	↑
rs12415501	<i>NEURL</i>	cg12662887	10	105,343,920	<i>NEURL1</i>	19,146 bp	0.0562	0.0034	7.5×10^{-59}	↑
rs6666258	<i>KCNN3</i>	cg06221963	1	154,839,813	<i>KCNN3</i>	25,545 bp	0.0714	0.0052	6.6×10^{-42}	↑
rs10507248	<i>TBX5</i>	cg10233830	12	114,701,413	<i>TBX5</i>	95,680 bp	-0.0311	0.0031	4.9×10^{-23}	↑
rs10821415	<i>C9orf3</i>	cg13792694	9	97,865,232	<i>FANCC</i>	151,773 bp	-0.0033	0.0004	7.1×10^{-20}	↓
rs3903239	<i>PRRX1</i>	cg09010107	1	170,638,807	<i>PRRX1</i>	69,490 bp	0.0116	0.0016	1.1×10^{-13}	↑
rs6817105	<i>PITX2</i>	cg03587884	4	111,642,146	<i>PITX2</i>	63,622 bp	-0.0153	0.0022	1.0×10^{-11}	↓
rs2106261	<i>ZFH3</i>	cg06618356	16	73,097,364	<i>ZFH3</i>	45,744 bp	-0.0096	0.0016	1.1×10^{-9}	↓
rs13216675	<i>GJA1</i>	cg05720511	6	123,043,994	<i>PKIB</i>	591,665 bp	0.0024	0.0005	1.6×10^{-6}	↓

Table 4. Most significant CpG site associated with each AF GWAS locus. All loci were significantly associated methylation status ($P < 8.3 \times 10^{-6}$). ⁺NCBI Genome Build 37. ^{*}SE: Standard error. [§]Indicate if AF risk allele was the same allele to increase methylation. “↑” Represents the AF risk allele would increase the methylation level, whereas “↓” represents the AF risk allele would decrease the methylation level.

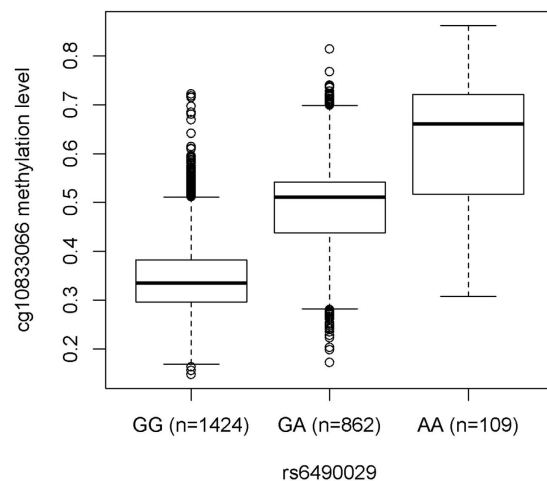


Figure 3. Association of SNP rs6490029 (*CUX2*) with the methylation of cg10833066. The boxplot indicates the minimum, 25%, 50%, 75% and the maximum methylation level for each genotype. Outliers were marked as points. Samples with one or two “A” alleles at rs6490029 tended to have a higher methylation level. The number of samples with each genotype was marked as well.

was a slight difference between the numbers of incident AF cases compared to prevalent AF cases, which might also have contributed to the variation in the methylation profile identified.

We acknowledge several limitations in our study. DNA methylation was collected at a single examination, so we were unable to investigate longitudinal changes in methylation profiles. Moreover, the DNA methylation profile was measured from whole blood, which could vary from levels in atria or specific white blood cell types. However, invasive specimen collection is not feasible in a community-based study. The participants in our study are largely middle-age to older and are of European decent, so it is unclear whether our findings are generalizable to other ages or races/ethnicities. As an observational study we cannot exclude residual confounding, or determine causal relationships. Given that many other cohorts have very limited numbers of AF cases with methylation profiling, and they are usually derived from different types of cells, we were unable to replicate our findings in the current study. Future studies with multiple examinations and population diversity might provide better understanding of methylation signatures for AF and are required to verify our findings.

In conclusion, we examined the association of DNA methylation with AF in a moderately large community-based cohort, and identified multiple methylation signatures associated with AF. Our results suggest that DNA methylation might represent an important bridge to link genetic variations with AF susceptibility. Future validation might uncover AF-specific methylation regulation mechanisms, and potentially lead to the identification of novel therapeutic targets for better treatment of AF.

Materials and Methods

Study Samples. The Framingham Heart Study is a three-generation community-based cohort initiated in 1948. The Framingham Offspring cohort was recruited in 1971, and consisted of 5124 participants who are the offspring and the spouses of offspring of the Original cohort⁴². The present study was focused on Offspring cohort participants who attended the eighth examination (2005–2008). All participants gave written informed consent and the study was approved by the Institutional Review Boards of National Human Genome Research Institute and Boston University Medical Center, and all experiments were performed in accordance with relevant guidelines and regulations.

AF Ascertainment. AF was ascertained from a combination of multiple sources. Each participant was asked about his/her cardiovascular history and a 12-lead electrocardiogram was obtained during clinic visits scheduled every 4–8 years. Additional information also was solicited during surveillance interviews biennially^{43,44}, and from cardiovascular disease-related hospitalizations and clinician visits. At least two Framingham Heart Study cardiologists reviewed all electrocardiograms available from study visits or in- and outpatient records to adjudicate incident AF.

DNA Methylation Quantification. The fasting peripheral whole blood was collected using Gentra Puregene Blood Kits (QIAGEN, Venlo, Netherlands) in a single examination (the eighth examination). The genomic DNA was then bisulfite-treated, amplified and hybridized to the Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA) according to the manufacturer's standard protocols⁴⁵. The methylation assay was performed at two centers and normalized separately, and pooled together after adjusting for batch effects between two centers.

The methylation status was represented by the β value as a continuous variable between 0 and 1 representing the proportion of methylation at each CpG site. The raw data were normalized and corrected for the background noise by "DASEN" R package⁴⁶. We excluded probes with detection P -values less than 0.01. The whole probe was removed if $\geq 1\%$ of samples had missing values. A sample was removed if $\leq 95\%$ probes had P -value < 0.01 . We removed probes overlapping with known genetic polymorphisms from the 1000 Genomes Project Phase 1⁴⁷. Our analysis focused on autosomal chromosomes. A total of 443,252 CpG sites were included in the current study.

Gene Expression Profiling. Framingham gene expression profiling has been described in detail^{48–50}. Briefly, total RNA was isolated from fasting peripheral whole blood collected during clinic visits. RNA was then amplified and reverse transcribed into cDNA, which was hybridized to the Human Exon 1.0st Array (Affymetrix, Santa Clara, CA) according to standardized protocols. We used Robust Multi-array Average⁵¹ method to normalize and summarize the raw data. The gene annotations were obtained from Affymetrix NetAffx Analysis Center (version 31). Only the most reliable probe sets derived from RefSeq and GenBank records were used in this study, corresponding to 17,873 distinct transcripts^{48–50}.

Genetic Profiling. Genetic variation was profiled by Affymetrix 550k Array (Affymetrix, Santa Clara, CA) as previously reported¹². We excluded variants with call rates less than 97%, Hardy–Weinberg Equilibrium P -values less than 1.0×10^{-6} , or minor allele frequencies less than 0.01. The remaining variants were imputed to HapMap r22 CEU panel by Mach (v 1.0.15)⁵².

Statistical Analyses. Our primary analysis tested the association between methylation and AF status. For prevalent AF analysis, we hypothesized that AF status could affect methylation level. We used linear mixed effects regression models to test the association between the methylation level of each CpG with prevalent AF. We specified the methylation level as the dependent measure and the AF status at the eighth examination as the exposure. For this analysis, both incident AF cases and participants with no history of AF served as the reference group. For the incident AF analysis, we hypothesized that DNA methylation could be associated with future risk of AF. We thus used Cox proportional hazards models to relate the methylation level at each CpG to incident AF (censored at the last follow-up time or death). Participants with prevalent AF at baseline were excluded from this analysis.

Both analyses were adjusted for age, sex, assay site, and pedigree structure in Framingham. The false discovery rate (FDR)⁵³ was used to correct for multiple testing, and statistical significance was claimed if the FDR was less than 0.05, corresponding to $P < 1.4 \times 10^{-7}$. For significant associations, we additionally adjusted for the proportion of six cell types imputed via a reference panel using the Houseman method⁵⁴, including CD8⁺ T cells, CD4⁺ T cell, natural killer cells, B cells, monocytes and granulocytes, to account for cell count heterogeneity between study groups. In our secondary analysis, we tested the association of top CpG sites with AF adjusting for additional AF risk factors³¹, including smoking, height, weight, systolic blood pressure, diastolic blood pressure, prevalent diabetes mellitus, prevalent myocardial infarction, prevalent heart failure, and antihypertensive treatment. In an additional exploratory analysis, we combined prevalent and incident AF cases to test the association of each CpG with AF.

We also developed a combined methylation score from CpG sites significantly associated with incident AF. The score for sample i is defined as $S_i = \sum_{j=1}^n \beta_j M_{ij}$, where n is the number of CpG sites significantly associated with incident AF, β_j is the estimate of effect size for CpG site j , and M_{ij} is the methylation level at CpG site j for sample i . Similarly, we developed a combined genetic score from 14 top AF loci that were published previously^{12–15}. We then combined the genetic and methylation scores together with traditional AF risk factors, and investigated their association with incident AF using Cox proportional hazards models.

Association with Gene Expression. For each of the CpG sites associated with prevalent or incident AF, we examined the association of methylation with gene expression, stratified by *cis*-genes (defined as those within 1 Mb of the CpG site), and *trans*-genes (defined as those more than 1 Mb away from the CpG site or in different

chromosomes). The association was tested by linear mixed effects regression models, whereas gene expression was treated as the dependent measure, and the methylation status was treated as the exposure, adjusted for age, sex, and family structure. We used Bonferroni correction to adjust for multiple testing, and the significance was defined as $P < 0.05/N$, where N is the number of tests.

Association with AF-related Genetic Loci. We examined the association between genetic variations associated with AF and DNA methylation. Our analysis was limited to the 14 SNPs that were previously reported to associate with AF susceptibility by GWAS^{12,13}. The loci included rs10821415 (*C9orf3*), rs10824026 (*SYNPO2L*), rs1152591 (*SYNE2*), rs2106261 (*ZFHX3*), rs3807989 (*CAV1*), rs3903239 (*PRRX1*), rs6666258 (*KCNN3*), rs6817105 (*PITX2*), rs7164883 (*HCN4*), rs4642101 (*CAND2*), rs13216675 (*GJA1*), rs12415501 (*NEURL*), rs10507248 (*TBX5*), and rs6490029 (*CUX2*). Linear mixed effects regression models were used to test the association between each genetic variant and DNA methylation levels of CpGs within 1 Mb of AF SNPs. Methylation levels were treated as the dependent measures and genetic variants were treated as the exposures. The analysis was adjusted for age, sex, and family structure. We used Bonferroni correction to account for multiple testing, and the significance was claimed if the P value of the association was less than $0.05/N$, where N was determined as the number of tests.

References

- Go, A. S. *et al.* Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* **285**, 2370–2375 (2001).
- Chugh, S. S. *et al.* Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation* **129**, 837–847, doi: 10.1161/CIRCULATIONAHA.113.005119 (2014).
- Kannel, W. B., Wolf, P. A., Benjamin, E. J. & Levy, D. Prevalence, incidence, prognosis, and predisposing conditions for atrial fibrillation: population-based estimates. *Am J Cardiol* **82**, 2N–9N (1998).
- Ott, A. *et al.* Atrial fibrillation and dementia in a population-based study. The Rotterdam Study. *Stroke* **28**, 316–321 (1997).
- Wang, T. J. *et al.* Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: the Framingham Heart Study. *Circulation* **107**, 2920–2925 (2003).
- Krahn, A. D., Manfreda, J., Tate, R. B., Mathewson, F. A. & Cuddy, T. E. The natural history of atrial fibrillation: incidence, risk factors, and prognosis in the Manitoba Follow-Up Study. *Am J Med* **98**, 476–484 (1995).
- Stewart, S., Hart, C. L., Hole, D. J. & McMurray, J. J. A population-based study of the long-term risks associated with atrial fibrillation: 20-year follow-up of the Renfrew/Paisley study. *Am J Med* **113**, 359–364 (2002).
- Soliman, E. Z. *et al.* Atrial Fibrillation and the Risk of Myocardial Infarction. *JAMA internal medicine*, doi: 10.1001/jamainternmed.2013.11912 (2013).
- Chen, L. Y. *et al.* Atrial fibrillation and the risk of sudden cardiac death: the atherosclerosis risk in communities study and cardiovascular health study. *JAMA internal medicine* **173**, 29–35, doi: 10.1001/2013.jamainternmed.744 (2013).
- Benjamin, E. J. *et al.* Impact of atrial fibrillation on the risk of death: the Framingham Heart Study. *Circulation* **98**, 946–952 (1998).
- Schnabel, R. B. *et al.* 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. *Lancet* **386**, 154–162, doi: 10.1016/S0140-6736(14)61774-8 (2015).
- Ellinor, P. T. *et al.* Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nat Genet* **44**, 670–675, doi: 10.1038/ng.2261 (2012).
- Sinner, M. F. *et al.* Integrating genetic, transcriptional, and functional analyses to identify 5 novel genes for atrial fibrillation. *Circulation* **130**, 1225–1235, doi: 10.1161/CIRCULATIONAHA.114.009892 (2014).
- Gudbjartsson, D. F. *et al.* Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature* **448**, 353–357, doi: 10.1038/nature06007 (2007).
- Benjamin, E. J. *et al.* Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry. *Nat Genet* **41**, 879–881, doi: 10.1038/ng.416 (2009).
- Ma, B. *et al.* Predicting DNA methylation level across human tissues. *Nucleic Acids Res* **42**, 3515–3528, doi: 10.1093/nar/gkt1380 (2014).
- Hiltunen, M. O. & Yla-Herttuala, S. DNA methylation, smooth muscle cells, and atherogenesis. *Arterioscler Thromb Vasc Biol* **23**, 1750–1753 (2003).
- Ying, A. K. *et al.* Methylation of the estrogen receptor-alpha gene promoter is selectively increased in proliferating human aortic smooth muscle cells. *Cardiovasc Res* **46**, 172–179 (2000).
- Fitzpatrick, D. R. & Wilson, C. B. Methylation and demethylation in the regulation of genes, cells, and responses in the immune system. *Clin Immunol* **109**, 37–45 (2003).
- Lund, G. & Zaina, S. Atherosclerosis risk factors can impose aberrant DNA methylation patterns: a tale of traffic and homocysteine. *Curr Opin Lipidol* **20**, 448–449, doi: 10.1097/MOL.0b013e3283309928 (2009).
- Alikhani-Koopaei, R., Fouladkou, F., Frey, F. J. & Frey, B. M. Epigenetic regulation of 11 beta-hydroxysteroid dehydrogenase type 2 expression. *J Clin Invest* **114**, 1146–1157 (2004).
- Friso, S. *et al.* Epigenetic control of 11 beta-hydroxysteroid dehydrogenase 2 gene promoter is related to human hypertension. *Atherosclerosis* **199**, 323–327, doi: 10.1016/j.atherosclerosis.2007.11.029 (2008).
- Gallou-Kabani, C. & Junien, C. Nutritional epigenomics of metabolic syndrome: new perspective against the epidemic. *Diabetes* **54**, 1899–1906 (2005).
- Maier, S. & Olek, A. Diabetes: a candidate disease for efficient DNA methylation profiling. *J Nutr* **132**, 2440S–2443S (2002).
- Ling, C. & Groop, L. Epigenetics: a molecular link between environmental factors and type 2 diabetes. *Diabetes* **58**, 2718–2725, doi: 10.2337/db09-1003 (2009).
- Toledo-Rodriguez, M. *et al.* Maternal smoking during pregnancy is associated with epigenetic modifications of the brain-derived neurotrophic factor-6 exon in adolescent offspring. *Am J Med Genet B Neuropsychiatr Genet*, doi: 10.1002/ajmg.b.31109 (2010).
- Dick, K. J. *et al.* DNA methylation and body-mass index: a genome-wide analysis. *Lancet* **383**, 1990–1998, doi: 10.1016/S0140-6736(13)62674-4 (2014).
- Choi, J. Y. *et al.* Association between global DNA hypomethylation in leukocytes and risk of breast cancer. *Carcinogenesis* **30**, 1889–1897, doi: 10.1093/carcin/bgp143 (2009).
- Hill, J. J., Qiu, Y., Hewick, R. M. & Wolfman, N. M. Regulation of myostatin *in vivo* by growth and differentiation factor-associated serum protein-1: a novel protein with protease inhibitor and follistatin domains. *Molecular endocrinology* **17**, 1144–1154, doi: 10.1210/me.2002-0366 (2003).
- Pique-Regi, R. *et al.* Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. *Genome Res* **21**, 447–455, doi: 10.1101/gr.112623.110 (2011).

31. Alonso, A. *et al.* Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. *Journal of the American Heart Association* **2**, e000102, doi: 10.1161/JAHA.112.000102 (2013).
32. Mishra-Gorur, K., Singer, H. A. & Castellot, J. J., Jr. The S18 ribosomal protein is a putative substrate for Ca²⁺/calmodulin-activated protein kinase II. *J Biol Chem* **277**, 33537–33540, doi: 10.1074/jbc.C200342200 (2002).
33. Maier, L. S. & Bers, D. M. Role of Ca²⁺/calmodulin-dependent protein kinase (CaMK) in excitation-contraction coupling in the heart. *Cardiovasc Res* **73**, 631–640, doi: 10.1016/j.cardiores.2006.11.005 (2007).
34. Gaillard, S., Bartoli, M., Castets, F. & Monneron, A. Striatin, a calmodulin-dependent scaffolding protein, directly binds caveolin-1. *FEBS Lett* **508**, 49–52 (2001).
35. Mangoni, M. E. *et al.* Functional role of L-type Cav1.3 Ca²⁺ channels in cardiac pacemaker activity. *Proc Natl Acad Sci USA* **100**, 5543–5548, doi: 10.1073/pnas.0935295100 (2003).
36. St-Pierre, B. *et al.* Conserved and specific functions of mammalian ssu72. *Nucleic Acids Res* **33**, 464–477, doi: 10.1093/nar/gki171 (2005).
37. Galeano, F. *et al.* Human BLCAP transcript: new editing events in normal and cancerous tissues. *International journal of cancer* **127**, 127–137, doi: 10.1002/ijc.25022 (2010).
38. Djousse, L. *et al.* Long-term alcohol consumption and the risk of atrial fibrillation in the Framingham Study. *Am J Cardiol* **93**, 710–713, doi: 10.1016/j.amjcard.2003.12.004 (2004).
39. Mukamal, K. J., Tolstrup, J. S., Friberg, J., Jensen, G. & Gronbaek, M. Alcohol consumption and risk of atrial fibrillation in men and women: the Copenhagen City Heart Study. *Circulation* **112**, 1736–1742, doi: 10.1161/CIRCULATIONAHA.105.547844 (2005).
40. Frost, L. & Vestergaard, P. Alcohol and risk of atrial fibrillation or flutter: a cohort study. *Archives of internal medicine* **164**, 1993–1998, doi: 10.1001/archinte.164.18.1993 (2004).
41. Tao, H., Shi, K. H., Yang, J. J. & Li, J. Epigenetic mechanisms in atrial fibrillation: New insights and future directions. *Trends Cardiovasc Med* **26**, 306–318, doi: 10.1016/j.tcm.2015.08.006 (2016).
42. Kannel, W. B., Feinleib, M., McNamara, P. M., Garrison, R. J. & Castelli, W. P. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* **110**, 281–290 (1979).
43. Wolf, P. A., Abbott, R. D. & Kannel, W. B. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* **22**, 983–988 (1991).
44. Piccini, J. P. *et al.* Incidence and prevalence of atrial fibrillation and associated mortality among Medicare beneficiaries, 1993–2007. *Circ Cardiovasc Qual Outcomes* **5**, 85–93, doi: 10.1161/CIRCOUTCOMES.111.962688 (2012).
45. Bibikova, M. *et al.* High density DNA methylation array with single CpG site resolution. *Genomics* **98**, 288–295, doi: 10.1016/j.ygeno.2011.07.007 (2011).
46. Pidsley, R. *et al.* A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC genomics* **14**, 293, doi: 10.1186/1471-2164-14-293 (2013).
47. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073, doi: 10.1038/nature09534 (2010).
48. Joehanes, R. *et al.* Gene expression signatures of coronary heart disease. *Arterioscler Thromb Vasc Biol* **33**, 1418–1426, doi: 10.1161/ATVBAHA.112.301169 (2013).
49. Lin, H. *et al.* Whole blood gene expression and interleukin-6 levels. *Genomics* **104**, 490–495, doi: 10.1016/j.ygeno.2014.10.003 (2014).
50. Lin, H. *et al.* Whole blood gene expression and atrial fibrillation: the framingham heart study. *PLoS ONE* **9**, e96794, doi: 10.1371/journal.pone.0096794 (2014).
51. Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics (Oxford, England)* **4**, 249–264 (2003).
52. Abecasis, G. R. *MACH 1.0*, <http://www.sph.umich.edu/csg/abecasis/MaCH/index.html> (2008).
53. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)* **57**, 289–300 (1995).
54. Houseman, E. A. *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* **13**, 86, doi: 10.1186/1471-2105-13-86 (2012).

Acknowledgements

The Framingham Heart Study is funded by National Institutes of Health contract N01-HC-25195; HHSN268 201500001I. The laboratory work for this investigation was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, and by a Director's Challenge Award, National Institutes of Health (DL, PI). This project also was supported by NIH grants to Drs. Ellinor, Benjamin, and Lunetta (2R01HL092577) and Benjamin and Ellinor (1R01HL128914). Dr. Ellinor is supported by grants from the National Institutes of Health (K24HL105780), an Established Investigator Award from the American Heart Association (13EIA14220013) and by support from the Fondation Leducq (14CVD01). Dr. Lubitz is supported by NIH grants K23HL114724. This work was supported by Grant 2014105 and Grant 2015084 from the Doris Duke Charitable Foundation. Dr. Mendelson is partly supported by a family foundation award from the Tommy Kaplan Fund, Department of Cardiology, Boston Children's Hospital. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the U.S. Department of Health and Human Services.

Author Contributions

H.L., S.A.L., P.T.E. and E.J.B. drafted the manuscript. H.L. and X.Y. performed the main analysis. Z.X., K.L.L., S.A.L., M.G.L., D.K., J.W.M., M.M.M., C.L. and D.D.M. participated in the analysis and interpreted the results. D.L., P.T.E. and E.J.B. supervised the study.

Additional Information

Competing financial interests: Dr. Ellinor is the PI on a grant from Bayer HealthCare to the Broad Institute focused on the genetics and therapeutics of atrial fibrillation.

Supplementary information accompanies this paper at <http://www.nature.com/srep>

How to cite this article: Lin, H. *et al.* Methyloome-wide Association Study of Atrial Fibrillation in Framingham Heart Study. *Sci. Rep.* **7**, 40377; doi: 10.1038/srep40377 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017