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

Honghuang Lin  
Framingham Heart Study

Let us know how access to this document benefits you.
Follow this and additional works at: https://escholarship.umassmed.edu/oapubs

Part of the Cardiology Commons, Cardiovascular Diseases Commons, Molecular Biology Commons, and the Molecular Genetics Commons

Repository Citation

Creative Commons License
This work is licensed under a Creative Commons Attribution 4.0 License.
This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in Open Access Articles by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
Methylome-wide Association Study of Atrial Fibrillation in Framingham Heart Study

Honghuang Lin\textsuperscript{1,2}, Xiaoyan Yin\textsuperscript{1}, Zhijun Xie\textsuperscript{2,3}, Kathryn L. Lunetta\textsuperscript{4,5}, Steven A. Lubitz\textsuperscript{5,6}, Martin G. Larson\textsuperscript{1,7}, Darae Ko\textsuperscript{9}, Jared W. Magnani\textsuperscript{8}, Michael M. Mendelson\textsuperscript{10,11}, Chunyu Liu\textsuperscript{1,4,10}, David D. McManus\textsuperscript{12}, Daniel Levy\textsuperscript{1,10}, Patrick T. Ellinor\textsuperscript{5,6} & Emelia J. Benjamin\textsuperscript{1,5,13,14}

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 \textit{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)\textsuperscript{1,2}, the most common cardiac arrhythmia. AF is associated with a significantly increased risk for stroke\textsuperscript{4}, dementia\textsuperscript{4}, heart failure\textsuperscript{5–7}, myocardial infarction\textsuperscript{5}, and death\textsuperscript{5,10}. In the past few years, genome-wide association studies (GWAS) have identified 14 genetic loci associated with AF\textsuperscript{10–15}. 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 of 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 tissues\textsuperscript{16} or over an individual’s lifetime. Recent studies found that alteration of DNA

\textsuperscript{1}National Heart Lung and Blood Institute’s and Boston University’s Framingham Heart Study, Framingham, MA, USA. \textsuperscript{2}Section of Computational Biomedicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA. \textsuperscript{3}TCM Clinical Basis Institute, Zhejiang Chinese Medicine University, Hangzhou, Zhejiang, China. \textsuperscript{4}Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. \textsuperscript{5}Cardiovascular Research Center, Massachusetts General Hospital, Charlestown, MA, USA. \textsuperscript{6}The Broad Institute of Harvard and MIT, Cambridge, MA, USA. \textsuperscript{7}Department of Mathematics and Statistics, Boston University, Boston, MA, USA. \textsuperscript{8}Division of Cardiology, Department of Medicine, UPMC Heart and Vascular Institute, University of Pittsburgh, Pittsburgh, PA, USA. \textsuperscript{9}Section of Cardiovascular Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA. \textsuperscript{10}Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, Bethesda, MD, USA. \textsuperscript{11}Department of Cardiology, Boston Children’s Hospital, Boston, MA, USA. \textsuperscript{12}Medicine and Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, USA. \textsuperscript{13}Section of Preventive Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA. \textsuperscript{14}Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA. Correspondence and requests for materials should be addressed to H.L. (email: hhlin@bu.edu)
Two AF-related CpG sites (cg13639451 and cg15440392) were associated with the expression of trans-genes. The prediction performance with area under curve (AUC) increasing from 0.729 (model 1) to 0.747 (model 2) and incident AF. As shown in Fig. 2, the inclusion of genetic loci and methylation profiles modestly improved the prediction performance of prevalent and incident AF associations for each CpG site were found and tested. Therefore the significance cutoff was \( P < 10^{-7} \) for both prevalent and incident AF association. 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 ± 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 myostatin. The other one was cg07191189 (\( P = 1.4 \times 10^{-7}, \) FDR = 0.03), located 75 bp upstream of STRN (Striatin). The CpG site was also predicted to affect the binding of transcription factor TFIIH [19]. 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 association of the seven significant prevalent or incident AF-associated CpG sites with gene expression. 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 the CpG sites was significant after correction for multiple testing. The top CpG sites for prevalent and incident AF are listed in Table 1.

**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 × 10^{-4} for cis-associations and 0.05/124,936 = 4.0 × 10^{-7} for trans-associations. The most significant cis- and trans-genes for each CpG site are shown in Table 2.

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 of RPS18 (\( P = 4.5 \times 10^{-3} \)), and cg15440392 was associated with the expression of GZMH (\( P = 7.0 \times 10^{-4} \)).

**Table 1. Clinical characteristics of studied samples.**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No AF (n = 2,236)</th>
<th>Prevalent AF (n = 183)</th>
<th>Incident AF (n = 220)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women, n (%)</td>
<td>1285 (57%)</td>
<td>65 (36%)</td>
<td>89 (40%)</td>
</tr>
<tr>
<td>Age, year ± SD</td>
<td>65 ± 9</td>
<td>72 ± 9</td>
<td>71 ± 8</td>
</tr>
<tr>
<td>Height, inches ± SD</td>
<td>66 ± 4</td>
<td>67 ± 4</td>
<td>66 ± 4</td>
</tr>
<tr>
<td>Weight, pounds ± SD</td>
<td>173 ± 38</td>
<td>190 ± 48</td>
<td>181 ± 43</td>
</tr>
<tr>
<td>Current smoker, n (%)</td>
<td>206 (9%)</td>
<td>8 (4%)</td>
<td>15 (7%)</td>
</tr>
<tr>
<td>Systolic blood pressure, mm Hg</td>
<td>128 ± 17</td>
<td>128 ± 20</td>
<td>135 ± 19</td>
</tr>
<tr>
<td>Diastolic blood pressure, mm Hg</td>
<td>74 ± 10</td>
<td>70 ± 10</td>
<td>72 ± 10</td>
</tr>
<tr>
<td>Prevalent diabetes mellitus, n (%)</td>
<td>316 (14%)</td>
<td>60 (33%)</td>
<td>59 (27%)</td>
</tr>
<tr>
<td>Prevalent myocardial infarction, n (%)</td>
<td>18 (1%)</td>
<td>38 (21%)</td>
<td>8 (4%)</td>
</tr>
<tr>
<td>Prevalent heart failure, n (%)</td>
<td>68 (3%)</td>
<td>44 (24%)</td>
<td>14 (6%)</td>
</tr>
<tr>
<td>Antihypertensive treatment, n (%)</td>
<td>1016 (45%)</td>
<td>130 (71%)</td>
<td>142 (65%)</td>
</tr>
</tbody>
</table>
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 reported12–15, 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 $P < \frac{0.05}{6042} = 8.3 \times 10^{-6}$. 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 CAMK2D12, 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 CAV1, which is involved in cardiac development35. Further, the CAV1 locus

---

**Table 2. Most significant CpG sites associated with AF (FDR < 0.05).**

<table>
<thead>
<tr>
<th>Type</th>
<th>CpG site</th>
<th>Chr</th>
<th>Position</th>
<th>Closest gene</th>
<th>Distance</th>
<th>Effect size</th>
<th>SE*</th>
<th>P* value</th>
<th>FDR*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalent AF</td>
<td>cg13639451</td>
<td>17</td>
<td>48911157</td>
<td>WFKKNN2</td>
<td>1447 bp</td>
<td>-0.010</td>
<td>0.002</td>
<td>$1.1 \times 10^{-7}$</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>cg07191189</td>
<td>2</td>
<td>3719390</td>
<td>STRN</td>
<td>75 bp</td>
<td>0.011</td>
<td>0.002</td>
<td>$1.4 \times 10^{-7}$</td>
<td>0.030</td>
</tr>
<tr>
<td>Incident AF</td>
<td>cg26602477</td>
<td>1</td>
<td>1476845</td>
<td>SSL2</td>
<td>207 bp</td>
<td>-8.289</td>
<td>1.415</td>
<td>$4.7 \times 10^{-9}$</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>cg15440392</td>
<td>2</td>
<td>36156634</td>
<td>BLCAP</td>
<td>301 bp</td>
<td>-8.280</td>
<td>1.415</td>
<td>$4.7 \times 10^{-9}$</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>cg04064828</td>
<td>10</td>
<td>13400275</td>
<td>DPYSLS4</td>
<td>0</td>
<td>-8.280</td>
<td>1.415</td>
<td>$4.7 \times 10^{-9}$</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>cg27529934</td>
<td>1</td>
<td>205054684</td>
<td>RB5BP5</td>
<td>585 bp</td>
<td>-8.291</td>
<td>1.542</td>
<td>$7.5 \times 10^{-8}$</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>cg06725760</td>
<td>10</td>
<td>1102461</td>
<td>WDR37</td>
<td>314 bp</td>
<td>1.300</td>
<td>3.100</td>
<td>$1.3 \times 10^{-7}$</td>
<td>0.027</td>
</tr>
</tbody>
</table>

*NCBI Genome Build 37. SE: Standard error; FDR: false discovery rate53.
was associated with AF in our previous GWAS\textsubscript{12}. 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 polymerase\textsuperscript{36}. 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 cardiac\textsuperscript{37}. 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\textsuperscript{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 reciprocal\textsuperscript{41}, 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

<table>
<thead>
<tr>
<th>Type</th>
<th>CpG site</th>
<th>Most significant cis-gene for each AF-related CpG site</th>
<th>Most significant trans-gene for each AF-related CpG site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalent AF</td>
<td>cg13639451</td>
<td>HILS1 0.36 0.14 9.1 × 10^{-3}</td>
<td>RPS18 1.85 0.20 4.5 × 10^{-20}</td>
</tr>
<tr>
<td></td>
<td>cg07191189</td>
<td>STRN 0.35 0.12 4.8 × 10^{-3}</td>
<td>LGSN 0.77 0.18 2.3 × 10^{-5}</td>
</tr>
<tr>
<td>Incident AF</td>
<td>cg26602477</td>
<td>Clof159 −0.35 0.16 3.1 × 10^{-2}</td>
<td>PTPN2 −1.37 0.31 1.2 × 10^{-5}</td>
</tr>
<tr>
<td></td>
<td>cg15440392</td>
<td>RPN2 0.21 0.12 7.8 × 10^{-3}</td>
<td>GZMH −2.18 0.37 7.0 × 10^{-9}</td>
</tr>
<tr>
<td></td>
<td>cg04064828</td>
<td>DYS14 0.56 0.23 1.7 × 10^{-3}</td>
<td>ABCA4 −0.98 0.20 7.9 × 10^{-5}</td>
</tr>
<tr>
<td></td>
<td>cg27529934</td>
<td>PLEKHA6 0.26 0.09 5.6 × 10^{-3}</td>
<td>ADAMTS2 0.40 0.09 1.3 × 10^{-5}</td>
</tr>
<tr>
<td></td>
<td>cg06725760</td>
<td>ZMYND11 −0.22 0.21 2.8 × 10^{-3}</td>
<td>GZMH −3.01 0.72 3.2 × 10^{-5}</td>
</tr>
</tbody>
</table>

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.
Table 4. Most significant CpG site associated with each AF GWAS locus. All loci were significantly associated methylation status (P < 8.3 × 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.

<table>
<thead>
<tr>
<th>AF SNP</th>
<th>Closest gene to AF SNP</th>
<th>CpG site</th>
<th>Chr</th>
<th>Position*</th>
<th>Closest gene to CpG site</th>
<th>Distance to AF SNP</th>
<th>Effect size</th>
<th>SE</th>
<th>P value</th>
<th>Methylation to AF risk$</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs6490029</td>
<td>CUX2</td>
<td>cg10833066</td>
<td>12</td>
<td>111,807,467</td>
<td>FAM109A</td>
<td>109,010 bp</td>
<td>0.1380</td>
<td>0.0034</td>
<td>9.5 × 10^{-27}</td>
<td>↑</td>
</tr>
<tr>
<td>rs10824026</td>
<td>SYNPO2L</td>
<td>cg02286717</td>
<td>10</td>
<td>75,415,704</td>
<td>SYNPO2L</td>
<td>5,504 bp</td>
<td>-0.0578</td>
<td>0.0018</td>
<td>1.3 × 10^{-16}</td>
<td>↓</td>
</tr>
<tr>
<td>rs1152591</td>
<td>SYNE2</td>
<td>cg23250157</td>
<td>14</td>
<td>64,679,961</td>
<td>SYNE2</td>
<td>887 bp</td>
<td>-0.0109</td>
<td>0.0005</td>
<td>2.8 × 10^{-10}</td>
<td>↓</td>
</tr>
<tr>
<td>rs7164883</td>
<td>HCN4</td>
<td>cg06757333</td>
<td>15</td>
<td>73,655,217</td>
<td>HCN4</td>
<td>3,043 bp</td>
<td>-0.0357</td>
<td>0.0019</td>
<td>1.3 × 10^{-14}</td>
<td>↓</td>
</tr>
<tr>
<td>rs3807989</td>
<td>CAV1</td>
<td>cg12739419</td>
<td>7</td>
<td>116,140,593</td>
<td>CAV2</td>
<td>45,648 bp</td>
<td>0.0233</td>
<td>0.0013</td>
<td>1.2 × 10^{-14}</td>
<td>↓</td>
</tr>
<tr>
<td>rs4642101</td>
<td>CAND2</td>
<td>cg24848339</td>
<td>3</td>
<td>12,840,334</td>
<td>CAND2</td>
<td>1,889 bp</td>
<td>-0.0112</td>
<td>0.0006</td>
<td>1.8 × 10^{-10}</td>
<td>↑</td>
</tr>
<tr>
<td>rs12419501</td>
<td>NEURL</td>
<td>cg12662887</td>
<td>10</td>
<td>105,343,920</td>
<td>NEURL1</td>
<td>19,146 bp</td>
<td>0.0562</td>
<td>0.0034</td>
<td>7.5 × 10^{-10}</td>
<td>↑</td>
</tr>
<tr>
<td>rs6666258</td>
<td>KCN3</td>
<td>cg06221963</td>
<td>1</td>
<td>154,839,813</td>
<td>KCN3</td>
<td>25,545 bp</td>
<td>0.0714</td>
<td>0.0052</td>
<td>6.6 × 10^{-12}</td>
<td>↑</td>
</tr>
<tr>
<td>rs10507248</td>
<td>TBX5</td>
<td>cg10233830</td>
<td>12</td>
<td>114,701,413</td>
<td>TBX5</td>
<td>95,680 bp</td>
<td>-0.0311</td>
<td>0.0031</td>
<td>4.9 × 10^{-13}</td>
<td>↑</td>
</tr>
<tr>
<td>rs10821415</td>
<td>C9orf5</td>
<td>cg13792694</td>
<td>9</td>
<td>97,865,232</td>
<td>FANCC</td>
<td>151,773 bp</td>
<td>-0.0033</td>
<td>0.0004</td>
<td>7.1 × 10^{-16}</td>
<td>↓</td>
</tr>
<tr>
<td>rs3902329</td>
<td>PRRX1</td>
<td>cg09010107</td>
<td>12</td>
<td>170,638,807</td>
<td>PRRX1</td>
<td>69,490 bp</td>
<td>0.0116</td>
<td>0.0016</td>
<td>1.1 × 10^{-15}</td>
<td>↑</td>
</tr>
<tr>
<td>rs6817105</td>
<td>PITX2</td>
<td>cg03587884</td>
<td>4</td>
<td>111,640,593</td>
<td>PITX2</td>
<td>63,622 bp</td>
<td>-0.0153</td>
<td>0.0022</td>
<td>1.0 × 10^{-13}</td>
<td>↓</td>
</tr>
<tr>
<td>rs2106261</td>
<td>ZFHX3</td>
<td>cg06618356</td>
<td>7</td>
<td>137,926,323</td>
<td>ZFHX3</td>
<td>19,146 bp</td>
<td>-0.0096</td>
<td>0.0016</td>
<td>1.1 × 10^{-11}</td>
<td>↓</td>
</tr>
<tr>
<td>rs13216675</td>
<td>GJA1</td>
<td>cg05720511</td>
<td>10</td>
<td>123,043,994</td>
<td>PKIB</td>
<td>59,665 bp</td>
<td>0.0024</td>
<td>0.0005</td>
<td>1.6 × 10^{-13}</td>
<td>↓</td>
</tr>
</tbody>
</table>

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 cohort. 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 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 HumanMethylation-450 BeadChip (Illumina, San Diego, CA) according to the manufacturer’s standard protocols. 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 package. We excluded probes with detection P-values less than 0.01. The whole probe was removed if ≥1% of samples had missing values. A sample was removed if ≤95% probes had P-value < 0.01. We removed probes overlapping with known genetic polymorphisms from the 1000 Genomes Project Phase 1. 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. 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 Average 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.

Genetic Profiling. Genetic variation was profiled by Affymetrix 550 k Array (Affymetrix, Santa Clara, CA) as previously reported. 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).

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) 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 × 10−7. For significant associations, we additionally adjusted for the proportion of six cell types imputed via a reference panel using the Houseman method, 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 factors, 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 \( S_i = \sum_{j=1}^{n} \beta_j^+ M_j \), where n is the number of CpG sites significantly associated with incident AF, \( \beta_j^+ \) is the estimate of effect size for CpG site j, and \( M_j \) 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. 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. The loci included rs10821415 (C9orf3), rs110824026 (SYNPO2L), rs1152591 (SYNE2), rs2106261 (ZFHX3), rs3807989 (CAVI), rs3903239 (PRRX1), rs6666258 (KCN3N), rs6817105 (PITX2), rs7164883 (HCN4), rs4642101 (CAND2), rs13216675 (OIA1), 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**

Methylome-wide Association Study of Atrial Fibrillation in Framingham: Lin, H.

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


Acknowledgements

The Framingham Heart Study is funded by National Institutes of Health contract N01-HC-25195; HHSN268 201500011L. 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 (2RO1HL092577) and Benjamin and Ellinor (1RO1HL128914). 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 (14CDV01). 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


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
