

SCIENTIFIC REPORTS



OPEN

Gene-gene Interaction Analyses for Atrial Fibrillation

Honghuang Lin^{1,2}, Martina Mueller-Nurasyid^{3,4,5}, Albert V. Smith^{6,7}, Dan E. Arking⁸, John Barnard⁹, Traci M. Bartz¹⁰, Kathryn L. Lunetta^{1,11}, Kurt Lohman¹², Marcus E. Kleber¹³, Steven A. Lubitz^{14,15}, Bastiaan Geelhoed¹⁶, Stella Trompet^{17,18}, Maartje N. Niemeijer¹⁹, Tim Kacprowski^{20,21}, Daniel I. Chasman²², Derek Klarin^{23,24,25,26}, Moritz F. Sinner⁴, Melanie Waldenberger^{5,13,27}, Thomas Meitinger^{5,28,29}, Tamara B. Harris³⁰, Lenore J. Launer³⁰, Elsayed Z. Soliman³¹, Lin Y. Chen³², Jonathan D. Smith⁹, David R. Van Wagener⁹, Jerome I. Rotter³³, Bruce M. Psaty^{34,35}, Zhijun Xie², Audrey E. Hendricks^{1,36}, Jingzhong Ding³⁷, Graciela E. Delgado¹³, Niek Verweij¹⁶, Pim van der Harst¹⁶, Peter W. Macfarlane³⁸, Ian Ford³⁹, Albert Hofman¹⁹, André Uitterlinden⁴⁰, Jan Heeringa¹⁹, Oscar H. Franco¹⁹, Jan A. Kors⁴¹, Stefan Weiss^{20,21}, Henry Völzke^{21,42}, Lynda M. Rose²², Pradeep Natarajan^{15,23,24,26}, Sekar Kathiresan^{15,23,24,26}, Stefan Kääb^{4,5}, Vilundur Gudnason^{6,7}, Alvaro Alonso⁴³, Mina K. Chung⁹, Susan R. Heckbert^{35,44}, Emelia J. Benjamin^{1,45,46}, Yongmei Liu⁴⁷, Winfried März^{48,49,50}, Michiel Rienstra¹⁶, J. Wouter Jukema¹⁷, Bruno H. Stricker^{40,51}, Marcus Dörr^{21,52}, Christine M. Albert²² & Patrick T. Ellinor¹⁵

Received: 17 March 2016
Accepted: 28 September 2016
Published: 08 November 2016

Atrial fibrillation (AF) is a heritable disease that affects more than thirty million individuals worldwide. Extensive efforts have been devoted to the study of genetic determinants of AF. The objective of our study is to examine the effect of gene-gene interaction on AF susceptibility. We performed a large-scale association analysis of gene-gene interactions with AF in 8,173 AF cases, and 65,237 AF-free referents collected from 15 studies for discovery. We examined putative interactions between genome-wide SNPs and 17 known AF-related SNPs. The top interactions were then tested for association in an independent cohort for replication, which included more than 2,363 AF cases and 114,746 AF-free referents. One interaction, between rs7164883 at the *HCN4* locus and rs4980345 at the *SLC28A1* locus, was found to be significantly associated with AF in the discovery cohorts (interaction OR = 1.44, 95% CI: 1.27–1.65, $P = 4.3 \times 10^{-8}$). Eight additional gene-gene interactions were also marginally significant ($P < 5 \times 10^{-7}$). However, none of the top interactions were replicated. In summary, we did not find significant interactions that were associated with AF susceptibility. Future increases in sample size and denser genotyping might facilitate the identification of gene-gene interactions associated with AF.

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. ³Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany. ⁴Department of Medicine I, Ludwig-Maximilians-University Munich, Munich, Germany. ⁵DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, Germany. ⁶Icelandic Heart Association, Kopavogur, Iceland. ⁷Faculty of Medicine, University of Iceland, Reykjavik, Iceland. ⁸McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁹Cleveland Clinic, Cleveland, OH, USA. ¹⁰Department of Biostatistics, University of Washington, Seattle, WA, USA. ¹¹Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. ¹²Department of Biostatistical Sciences, Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA. ¹³Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Theodor-Kutzer-Ufer 1-3, 68167 Mannheim, Germany. ¹⁴Cardiac Arrhythmia Service, Massachusetts General Hospital, Boston, MA, USA. ¹⁵Harvard Medical School, Boston, MA, USA. ¹⁶Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands. ¹⁷Department of Cardiology, Leiden University Medical Center, the Netherlands. ¹⁸Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands. ¹⁹Department of Epidemiology, Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands. ²⁰Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt University Greifswald, Greifswald, Germany. ²¹DZHK (German Centre

for Cardiovascular Research), partner site Greifswald, Greifswald, Germany. ²²Division of Preventive Medicine, Brigham and Women's Hospital, Boston MA, USA. ²³Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA. ²⁴Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA. ²⁵Department of Surgery, Massachusetts General Hospital, Boston, MA, USA. ²⁶Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA. ²⁷Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany. ²⁸Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany. ²⁹Institute of Human Genetics, Technische Universität München, Munich, Germany. ³⁰National Institute on Aging, National Institutes of Health, Bethesda, MD, USA. ³¹Epidemiological Cardiology Research Center, Wake Forest School of Medicine, Winston Salem, NC, USA. ³²Cardiovascular Division, Department of Medicine, University of Minnesota Medical School, Minneapolis, MN, USA. ³³Institute for Translational Genomics and Population Sciences (J.I.R.), Departments of Pediatrics and Medicine, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, USA. ³⁴Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology and Health Services, University of Washington, Seattle, WA, USA. ³⁵Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA. ³⁶Mathematical and Statistical Sciences, University of Colorado, Denver, Denver, CO, USA. ³⁷Department of Gerontology and Geriatric Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA. ³⁸Institute of Health and Wellbeing, College of Veterinary, Medical and Life Sciences, University of Glasgow, United Kingdom. ³⁹Robertson Center for Biostatistics, University of Glasgow, United Kingdom. ⁴⁰Department of Epidemiology & Internal Medicine, Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands. ⁴¹Department of Medical Informatics, Erasmus MC - University Medical Center Rotterdam, the Netherlands. ⁴²Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany. ⁴³Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA. ⁴⁴Department of Epidemiology, Cardiovascular Health Research Unit, University of Washington, Seattle, WA, USA. ⁴⁵Section of Cardiovascular Medicine and Preventive Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA. ⁴⁶Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA. ⁴⁷Department of Epidemiology & Prevention, Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA. ⁴⁸Synlab Academy, Synlab Services, GmbH P5,7, 68161 Mannheim, Germany. ⁴⁹Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria. ⁵⁰Medical Clinic V (Nephrology, Hypertensiology, Rheumatology, Endocrinology, Diabetology), Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany. ⁵¹Inspectorate of Health Care, Utrecht, the Netherlands. ⁵²Department of Internal Medicine B, University Medicine Greifswald, Greifswald, Germany. Correspondence and requests for materials should be addressed to H.L. (email: hlin@bu.edu)

Atrial fibrillation (AF) is the most common cardiac arrhythmia, estimated to affect about 33.5 million individuals globally¹. The heritability of AF, particularly lone AF, has long been established^{2–8}. Over the past few years, genome-wide association studies (GWAS) have successfully identified more than a dozen genetic loci associated with AF susceptibility^{9–14}. These loci include genes involved in cardiac signaling, cardiopulmonary development, and regulation of atrial action potential duration. However, all together, these loci still explain less than 5% of the heritability of AF¹⁵, whereas a large proportion of heritability remains unknown^{16,17}.

Epistasis refers to the interaction of multiple genes that might pose joint genetic effects. Epistasis plays a ubiquitous role in disease predisposition, conferring an increased risk in addition to the main effects for many complex diseases, such as breast cancer¹⁸ and coronary heart disease¹⁹. Gene-gene interactions play important roles in regulating various biological events and cellular behaviors^{20,21}. However, it remains unclear whether gene interactions contribute to the biological basis of AF.

The most straightforward approach to identifying interactions is to perform an exhaustive search of all the possible combinations of genetic variants and to test if any of them are significantly associated with AF. However, a major problem with such a comprehensive search is the huge computational burden. Assuming one million SNPs are genotyped in a typical GWAS, a complete search of a two-marker model would require testing 5×10^{11} pairs of SNPs. This number would further increase exponentially for multiple-SNP models. The cost of multiple testing corrections even in the 1 million marker scenario is extreme. For example, a Bonferroni correction requires $P < 1 \times 10^{-13}$ for significance in such a number of tests. As few SNP pairs will meet this threshold, false negatives are likely without massive sample sizes.

It has been suggested that at least one variant in significant gene-gene interactions tends to have a strong main effect²². We therefore sought to identify potential interactions between top AF susceptibility SNPs and other genome-wide variants in relation to AF by performing a meta-analysis of results from multiple studies.

Results

In total, our study included 8,173 AF cases and 65,237 AF-free referents of European ancestry from 15 studies. The clinical characteristics of the study participants are shown in Table 1.

Supplemental Figure 1 presents Q-Q plots for the interaction p-values of genome-wide SNPs with each of the AF-associated variants. The effect of population stratification was negligible, with genomic control λ ranging from 0.98 to 1.01.

Table 2 shows the most significant interactions ($P < 5 \times 10^{-7}$) that were associated with AF susceptibility. The top 10 interactions for each AF SNP are shown in Supplemental Table 1. None of interactions reached the significance after adjusting for multiple testing ($P < 5 \times 10^{-8}/17 = 2.8 \times 10^{-9}$). Only one interaction, SNP rs7164883 with rs4980345 exceeded the traditional genome-wide significance threshold ($P < 5 \times 10^{-8}$) for association with AF ($P = 4.3 \times 10^{-8}$). Both interacting SNPs are located in chromosome 15, 12Mb apart. The corresponding regional plot is shown in Fig. 1, and the forest plot of each contributing study is shown in Fig. 2. The SNP rs7164883 is located within the first intron of *HCN4*, and was also one of the top SNPs found to be significantly

	Study	Group	n	Age, years	Men, %	HTN, %	BMI, kg/m ²	Diabetes, %	MI, %	CHF, %
Discovery	AFNET/KORA	Cases	448	51.0 ± 7.6	68	41	28.1 ± 5.2	8	1	2
		Referents	438	55.8 ± 7.3	50	45	27.7 ± 4.3	4	0	1
	AGES	Cases	399	78.6 ± 5.6	58	90	27.2 ± 4.4	13	6	9
		Referents	2,560	76.1 ± 5.4	36	78	27.0 ± 4.5	11	5	1
	ARIC	Cases	1,420	56.8 ± 5.4	57	41	28.2 ± 5.3	13	9	8
		Referents	7,633	53.8 ± 5.6	45	24	26.8 ± 4.7	8	3	3
	CCF	Cases	496	58.8 ± 10.7	76	58	30.2 ± 6.2	6	0	8
		Referents	2,971	28.4 ± 22.2	38	–	–	–	–	–
	CHS	Cases	1,011	72.7 ± 5.4	44	62	26.4 ± 4.5	14	0	0
		Referents	2,190	72.0 ± 5.3	36	52	26.2 ± 4.4	11	0	0
	FHS	Cases	956	71.9 ± 12.3	57	68	28.3 ± 5.5	13	9	3
		Referents	7,470	51.8 ± 15.7	44	31	27.2 ± 5.3	5	2	1
	HealthABC	Cases	129	74.4 ± 2.9	63	87	26.3 ± 3.9	–	–	–
		Referents	1,532	73.7 ± 2.8	52	62	26.6 ± 4.1	–	–	–
	LURIC	Cases	361	66.4 ± 27.7	72	75	27.7 ± 4.2	45	33	30
		Referents	2,598	62.2 ± 10.7	70	73	27.4 ± 4.0	40	44	18
	MGH/MIGEN	Cases	366	53.4 ± 10.5	81	23	27.8 ± 5.0	3	1	3
		Referents	911	47.9 ± 8.8	53	–	–	–	–	–
	PREVEND	Cases	272	61.3 ± 9.4	67	68	27.7 ± 4.4	7	12	3
		Referents	3,277	48.4 ± 12.1	50	34	26.0 ± 4.2	4	2	0.1
PROSPER	Cases	505	76.0 ± 3.5	58	64	27.1 ± 4.3	11	19	0	
	Referents	4,739	75.3 ± 3.3	47	62	26.8 ± 4.2	10	13	0	
RS-I	Cases	954	72.6 ± 8.5	46	65	26.7 ± 3.7	17	11	14	
	Referents	4,691	68.3 ± 8.8	40	53	26.2 ± 3.6	9	6	7	
RS-II	Cases	146	71.5 ± 9.7	54	78	27.2 ± 4.3	22	12	6	
	Referents	2,011	64.3 ± 7.6	45	59	27.2 ± 4.2	12	4	1	
SHIP	Cases	106	62.0 ± 10.3	63	55	29.6 ± 5.1	23	13	27	
	Referents	1,815	49.0 ± 14.4	47	24	27.2 ± 4.5	11	3	9	
WGHS	Cases	959	58.2 ± 7.6	–	40	27.2 ± 5.3	5	–	–	
	Referents	19,897	53.9 ± 4.9	–	23	25.8 ± 4.9	2	–	–	
Replication	UK Biobank	Cases	2,363	62.3 ± 5.8	70	52	29.1 ± 5.5	13	13	12
		Referents	114,746	56.7 ± 7.9	47	21	27.5 ± 4.8	5	2	0

Table 1. Clinical characteristics of the participating studies. HTN – hypertension, defined as systolic blood pressure ≥ 140 mmHg, or diastolic blood pressure ≥ 90 mmHg, or antihypertensive treatment. BMI – body mass index; Diabetes – diabetes mellitus; MI – myocardial infarction; CHF – heart failure. “–” Signifies data not available.

associated with AF in our previous study¹⁰. The SNP rs4980345 was located within the tenth intron of *SLC28A1*. SNP rs4980345 was not associated with AF ($P = 0.78$) in marginal analyses from the prior meta-analysis¹⁰.

As shown in Table 2, eight additional interactions also showed suggestive association with AF ($P < 5 \times 10^{-7}$). These interactions included two each with rs12415501 (*NEURL*), rs3807989 (*CAVI*), and rs1448818 (*PITX2*). There was one marginally significant interaction each with rs10821415 (*C9orf3*) and rs2106261 (*ZFH3*).

We also tested the association of rs2106261 at the *ZFH3* locus and rs2200733 at the *PITX2* locus with AF, which was recently reported to be associated with AF in a meta-analysis of three Chinese samples (OR = 5.36, $P = 8.0 \times 10^{-24}$)²³. The interaction, however, was not significant in any of the 16 studies included in the present paper, or in our meta-analysis (all with $P > 0.05$).

We then tried to replicate our findings in an independent cohort, UK Biobank, which included more than 2000 AF cases and 11,000 AF-free referents. As shown in Table 2, none of significant interactions from discovery phase were replicated (all with $P > 0.05/9 = 0.0056$).

Discussion

In the past decade, increasing evidence has suggested that the genetic predisposition is an important factor that contributes to AF as well as many other cardiovascular diseases^{24,25}. Due to the enormous number of association tests, few studies have been performed to investigate the associations of gene interactions with AF susceptibility. By restricting our analyses to interactions with known AF loci, we limited the multiple testing burden in our analysis and sought to examine the potential mechanisms by which variants at top loci contribute to AF susceptibility. One genome-wide significant gene interaction with AF, rs7164883 at the *HCN4* locus and rs4980345 at the *SLC28A1* locus, was found. Eight additional interactions were also marginally significant ($P < 5 \times 10^{-7}$), but did not withstand multiple testing correction. However, none of the top interactions were significant in the replication phase. It is noteworthy that the ORs of suggestive interactions from the replication cohort were very

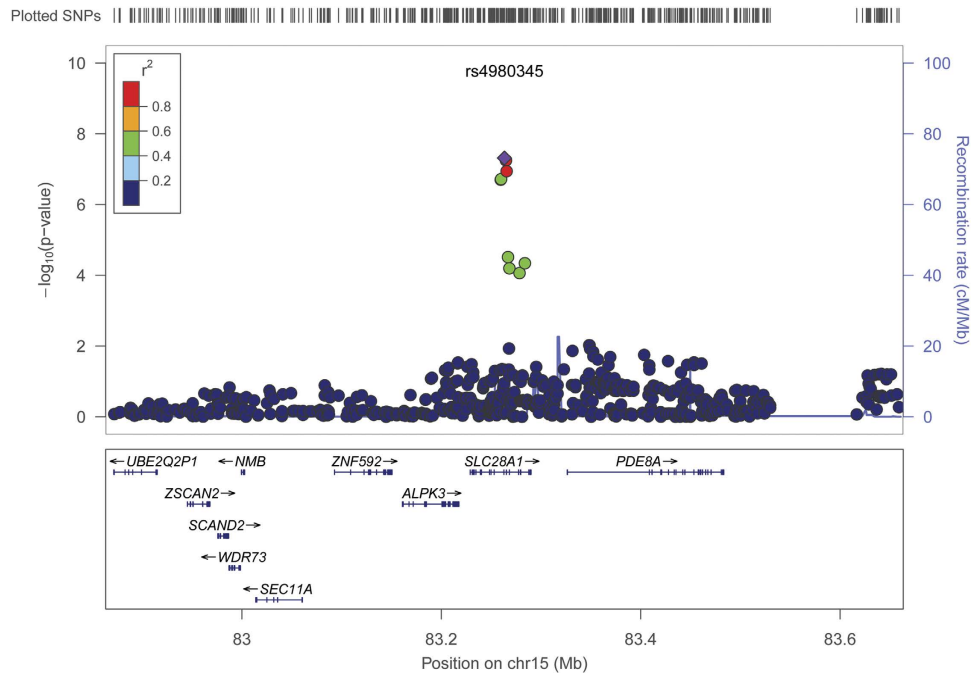


Figure 1. Regional plot of the interaction between rs7164883 and SNPs close to rs4980345. Each dot represents one SNP. The x-axis represents the chromosomal position, whereas the y-axis represents the $-\log_{10}(P)$ of the association of the interaction between rs7164883 and each SNP with AF.

moderate. The most significant interaction from the discovery cohorts, rs7164883 with rs4980345, was even in the reverse direction in the replication cohort. Given that the replication cohort has similar genetic background to the discovery cohorts, the discrepancy indicates that these suggestive interactions are unlikely to be true AF-related interactions.

Our analyses were restricted to interactions with loci previously found to have a main effect association with AF. The underlying assumption of our approach is that interactions with significant effects tend to have observable main effects in at least one of the interacting SNPs²². However, it is possible that two variants without main effects might have large interaction effects. Our analysis will not identify such interactions. A variety of other methods have been developed to account for the enormous number of interactions between variants in genetic association studies^{26,27}. One approach is to employ prior biological knowledge to limit the search space²⁸. Gene interactions have been discovered through experimental assays. These might be used to guide the search of potential variant interactions. Additionally, it has been recognized that many known genetic interactions were enriched with well-studied pathways, and could only happen under certain conditions²⁹, which might introduce additional bias to the analysis. In fact, none of the top interactions identified in the present study was reported in known interaction databases³⁰, suggesting that the interaction between some variants may arise through some other intermediate pathway.

We did not detect a recently reported interaction with AF by Huang and colleagues²³. This interaction involved rs2106261 at the *ZFH3* locus and rs2200733 at the *PITX2* locus. SNP rs2106261 was the most significant SNP at the *ZFH3* locus associated with AF in our earlier meta-analysis⁹. SNP rs2200733 was one of the top SNPs at the *PITX2* locus, and is in complete linkage disequilibrium ($r^2 = 1.0$) with the most significant SNP rs6817105, the SNP we tested in this study. One possible explanation for the discrepancy between the findings of the two studies is the difference in allele frequency between the Asian population studied by Huang²³ vs. the European ancestry population we studied (18% vs 28% for rs2106261, and 45% vs 16% for rs2200733, respectively). The effect of allelic difference and linkage disequilibrium could be amplified when the interaction was tested, suggesting that population stratification should be considered when comparing the results from studies based on different ethnicities.

We acknowledge several limitations of our study. All study participants in our study are of European ancestry, thus it is unclear whether our findings are relevant for populations of other ancestries. Furthermore, our analysis was restricted to two-variant interactions. However, it is possible that some interactions might involve more than two variants. Although our current study included more than 8,000 AF cases and 65,000 referents, it is possible that we did not have sufficient power to identify meaningful interactions for AF. We are currently expanding our AFGen Consortium to include additional cohorts, not only participants of European ancestry, but also participants of African ancestry and Asian ancestry. With the increasing sample size, we might be able to identify significant interactions in the future. In addition, we are currently imputing genotypes from individual studies to emerging reference panels such as the Haplotype Reference Consortium³¹, which is expected to provide better resolution to identify interacting variants. Given that our current study only tested interactions with known AF

AF SNP		Interacting SNP							Interaction effects			Replication		
SNP	Closest gene	SNP	Locus	Closest gene	Location	Coding allele	CAF ^s	Meta P value	OR ⁺	95% CI [†]	P value	OR ⁺	95% CI [†]	P value
rs7164883	<i>HCN4</i>	rs4980345	15q25.3	<i>SLC28A1</i>	Intron	T	0.06	0.78	1.44	1.27–1.65	4.3×10^{-8}	0.94	0.74–1.20	0.64
rs10821415	<i>C9orf3</i>	rs1492056	3p14.1	<i>MITF</i>	Intergenic	A	0.43	0.37	1.15	1.09–1.21	1.4×10^{-7}	0.91	0.84–0.99	0.04
rs12415501	<i>NEURL</i>	rs699801	1p31.1	<i>CRYZ</i>	Intergenic	T	0.45	0.83	1.19	1.12–1.27	1.9×10^{-7}	1.01	0.90–1.13	0.91
rs2106261	<i>ZFHX3</i>	rs12652090	5q34	<i>TENM2</i>	Intergenic	A	0.11	0.43	1.31	1.18–1.45	2.1×10^{-7}	1.02	0.87–1.20	0.80
rs1448818	<i>PITX2 (2)</i>	rs693832	8p21.1	<i>MIR3622B</i>	Intergenic	C	0.11	0.36	1.27	1.16–1.39	3.1×10^{-7}	1.03	0.89–1.19	0.72
rs3807989	<i>CAV1</i>	rs3802477	9q22.33	<i>GABBR2</i>	Intron	C	0.05	0.30	1.35	1.20–1.52	3.6×10^{-7}	1.05	0.86–1.27	0.64
rs3807989	<i>CAV1</i>	rs2327995	6p22.3	<i>ATXN1</i>	Intron	G	0.27	0.74	1.15	1.09–1.22	4.3×10^{-7}	1.05	0.95–1.15	0.32
rs1448818	<i>PITX2 (2)</i>	rs2328452	20p11.23	<i>RIN2</i>	Intergenic	G	0.88	0.92	1.25	1.15–1.37	4.7×10^{-7}	0.94	0.82–1.08	0.37
rs12415501	<i>NEURL</i>	rs7946907	11p15.2	<i>SPON1</i>	Intron	A	0.54	0.05	1.18	1.11–1.26	4.7×10^{-7}	1.07	0.96–1.20	0.23

Table 2. Most significant interactions associated with AF ($P < 5 \times 10^{-7}$). ^sCAF: coding allele frequency; ⁺OR: odds ratio; [†]CI: confidence interval.

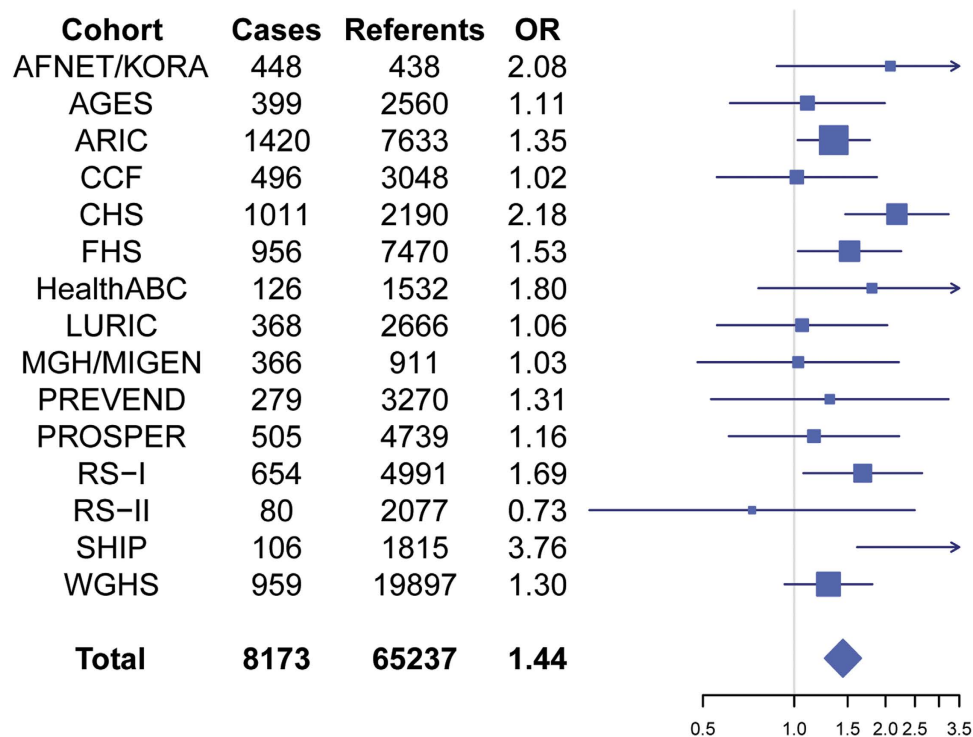


Figure 2. Forest plot of the association of interaction between rs7164883 and rs4980345 with AF in each study. Each line represents the 95% confidence interval, and the size is proportional to the number of cases. OR: odds ratio.

loci, we are also planning to expand our analyses to all interactions with the increasing sample size and more advanced computational methods.

In summary, we identified one genome-wide significant gene-gene interaction that was associated with AF susceptibility, suggesting that gene interactions might be involved in the development of AF. However, the finding was not replicated. Future work in functional genomics and efficient algorithms for epistasis analysis will likely facilitate the discovery of additional novel and high-order interactions that contribute to AF.

Materials and Methods

Study participants. Our discovery phase included individuals of European ancestry from 15 studies. These studies included the German Competence Network for Atrial Fibrillation/Cooperative Research in the Region of Augsburg (AFNET/KORA), Age, Gene/Environment Susceptibility Study (AGES) Reykjavik, Atherosclerosis Risk in Communities study (ARIC), Cleveland Clinic Lone AF GeneBank Study (CCAF), Cardiovascular Health Study (CHS), Framingham Heart Study (FHS), Health, Aging and Body Composition Study (HealthABC), Ludwigshafen Risk and Cardiovascular Health Study (LURIC), Massachusetts General Hospital Atrial Fibrillation Study (MGH), Prevention of Renal and Vascular Endstage Disease Study (PREVEND), the PROspective Study of Pravastatin in the Elderly at Risk study (PROSPER), Rotterdam Study (RS-I, RS-II), Study of Health in Pomerania

AF SNP	Closest Gene	Locus	Coding allele	Non-coding allele	Coding allele frequency	Source
rs6666258	KCNN3	1q21.3	C	G	0.3	13
rs3903239	PRRX1	1q24.2	G	A	0.45	9
rs4642101	CAND2	3p25.1	T	G	0.28	10
rs1448818	PITX2 (2)	4q25	C	A	0.17	14
rs6817105	PITX2 (1)	4q25	C	T	0.12	9,11
rs4400058	PITX2 (3)	4q25	A	G	0.12	14
rs6838973	PITX2 (4)	4q25	T	C	0.47	14
rs13216675	GJA1	6q22.31	C	T	0.27	10
rs3807989	CAV1	7q31.2	A	G	0.44	9
rs10821415	C9orf3	9q22.32	A	C	0.40	9
rs12415501	NEURL	10p14	T	C	0.11	10
rs10824026	SYNPO2L	10q22.2	G	A	0.14	9
rs6490029	CUX2	12q24.11	A	G	0.23	10
rs10507248	TBX5	12q24.21	G	T	0.23	10
rs1152591	SYNE2	14q23.2	A	G	0.45	9
rs7164883	HCN4	15q24.1	G	A	0.09	9
rs2106261	ZFH3	16q22.3	T	C	0.16	9,12

Table 3. The list of 17 AF top SNPs. It includes four independent SNPs at the *PITX2* locus (with a number in parenthesis), and one top SNP at remaining 13 AF loci.

(SHIP), and Women’s Genome Health Study (WGHS). The replication phase was performed on UK Biobank. The study protocol was approved by the internal review boards of Ludwig Maximilian University of Munich, University of Iceland, University of Minnesota, Cleveland Clinic, University of Washington, Boston University Medical Campus, Wake Forest School of Medicine, Heidelberg University, Massachusetts General Hospital, University Medical Center Groningen, Leiden University Medical Center, Erasmus MC - University Medical Center Rotterdam, University Medicine Greifswald, and Brigham and Women’s Hospital. The study was performed in accordance with the approved guidelines. All participants provided written informed consent to participate in genetic research.

AF ascertainment. Details about AF ascertainment were described in previous publications^{9,10,14}. Briefly, at each study, we combined evidence from a variety of sources to determine AF status, including electrocardiograms, Holter recordings, rhythm cards, medical records, and/or hospital discharge diagnostic codes. To achieve higher statistical power, we did not distinguish prevalent and incident AF cases, but combined them as individuals with a history of AF.

Genotyping. Genotyping was performed independently in each study, using either Affymetrix SNP arrays or Illumina SNP arrays⁹, and then imputed to ~2.5 million SNPs in the HapMap II release 22 CEU panel to obtain a comprehensive set of SNPs across the genome. Detailed information regarding genotyping platforms, quality control metrics, and imputation methods for each study has been described previously^{9,10,12–14}.

Known AF-associated variants. The known AF-associated variants were selected from recent GWAS meta-analysis results^{9,10,14}. Ellinor *et al.* reported nine AF loci in the meta-analysis of AF⁹. Sinner *et al.* reported five additional loci that were marginally significant in the earlier analysis but became genome-wide significant when combining with additional studies¹⁰. Lubitz *et al.* performed conditional analysis on the known *PITX2* locus¹⁴, and identified three additional independent SNPs within the locus. This resulted in a total of 17 AF-associated SNPs with genome-wide significance. The SNPs included one top SNP at each AF locus, and three additional independent SNPs at the *PITX2* locus. The full list of 17 SNPs is shown in Table 3.

Statistical Analyses. A multivariable logistic regression model was used to test the associations of interacting SNPs with AF. Each interaction was comprised of one of the 17 AF SNPs, and one SNP from the ~2.5 million imputed HapMap Phase II SNPs. We assumed a multiplicative interaction effect as follows:

$$\log(\text{odds AF}) = \beta_0 + \beta_{\text{age}} \text{age} + \beta_{\text{sex}} \text{sex} + \beta_{\text{PCs}} \text{PCs} + \beta_1 \text{SNP}_{\text{TopAF}} + \beta_2 \text{SNP}_2 + \beta_{\text{int}} \text{SNP}_{\text{TopAF}} * \text{SNP}_2$$

in which β_1 and β_2 are the main effects for the known AF SNP and the SNP to be tested, respectively. β_{int} represents the effect of the interaction between the AF SNP and the SNP to be tested. PCs represent principal components as necessary in each study to account for population structure. The model was also adjusted for *age* at DNA draw and *sex*, two factors that contribute significantly to AF risk. Studies with multiple study centers also adjusted for site. In order to account for the family correlation in FHS, we used generalized estimating equations (GEE) as implemented in the “geepack” R package. The association of each interaction with AF was adjusted for the independence working correlation structure in FHS, where each pedigree was a cluster in the robust variance estimate for the effect of interest.

The null hypothesis was that the interaction term, $\beta_{\text{int}} = 0$. Each study estimated and provided β_{int} and a robust estimate of standard error $\text{SE}(\beta_{\text{int}})$ for each SNP interacting with each of the 17 AF-associated SNPs. Thus, we performed 17 interaction GWAS. The study-specific interaction regression parameter estimates r were then meta-analyzed using METAL³², applying a fixed effects approach weighted for the inverse of the variance. The effect of interaction was presented as an interaction odds ratio (OR), i.e., $\exp(\beta_{\text{int}})$. Given that we performed the genome-wide test for 17 SNPs, we defined significant interactions as those with a P -value less than 2.8×10^{-9} ($=5 \times 10^{-8}/17$ SNPs tested).

In the replication phase, we tested the association of significant or suggestive interactions ($P < 5 \times 10^{-7}$) in an independent cohort, UK Biobank. An interaction was replicated if it had the same direction of effect as the discovery, and the association $P < 0.05/N$, where N was the number of tests.

References

- 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).
- Wolff, L. Familial auricular fibrillation. *New Eng J Med* **229**, 396–398 (1943).
- Fox, C. S. *et al.* Parental atrial fibrillation as a risk factor for atrial fibrillation in offspring. *JAMA* **291**, 2851–2855, doi: 10.1001/jama.291.23.2851291/23/2851 [pii] (2004).
- Arnar, D. O. *et al.* Familial aggregation of atrial fibrillation in Iceland. *Eur Heart J* **27**, 708–712 (2006).
- Ellinor, P. T., Yoerger, D. M., Ruskin, J. N. & MacRae, C. A. Familial aggregation in lone atrial fibrillation. *Hum Genet* **118**, 179–184 (2005).
- Darbar, D. *et al.* Familial atrial fibrillation is a genetically heterogeneous disorder. *Journal of the American College of Cardiology* **41**, 2185–2192 (2003).
- Kozłowski, D. *et al.* Lone atrial fibrillation: what do we know? *Heart* **96**, 498–503, doi: 10.1136/hrt.2009.176321 (2010).
- Kiliszczek, M. *et al.* The 4q25, 1q21, and 16q22 polymorphisms and recurrence of atrial fibrillation after pulmonary vein isolation. *Archives of medical science. AMS* **12**, 38–44, doi: 10.5114/aoms.2015.48284 (2016).
- 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: nature06007 [pii]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: ng.416 [pii]10.1038/ng.416 (2009).
- Ellinor, P. T. *et al.* Common variants in KCNN3 are associated with lone atrial fibrillation. *Nat Genet* **42**, 240–244, doi: 10.1038/ng.537 (2010).
- Lubitz, S. A. *et al.* Novel genetic markers associate with atrial fibrillation risk in Europeans and Japanese. *J Am Coll Cardiol* **63**, 1200–1210, doi: 10.1016/j.jacc.2013.12.015 (2014).
- Lubitz, S. A. *et al.* Association between familial atrial fibrillation and risk of new-onset atrial fibrillation. *JAMA* **304**, 2263–2269, doi: 10.1001/jama.2010.1690 (2010).
- Maher, B. Personal genomes: The case of the missing heritability. *Nature* **456**, 18–21, doi: 456018a [pii]10.1038/456018a (2008).
- Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753, doi: nature08494 [pii]10.1038/nature08494 (2009).
- Onay, V. U. *et al.* SNP-SNP interactions in breast cancer susceptibility. *BMC Cancer* **6**, 114, doi: 10.1186/1471-2407-6-114 (2006).
- Nelson, M. R., Kardia, S. L., Ferrell, R. E. & Sing, C. F. A combinatorial partitioning method to identify multilocus genotypic partitions that predict quantitative trait variation. *Genome research* **11**, 458–470, doi: 10.1101/gr.172901 (2001).
- Legrain, P., Wojcik, J. & Gauthier, J. M. Protein–protein interaction maps: a lead towards cellular functions. *Trends Genet* **17**, 346–352 (2001).
- Walhout, A. J. & Vidal, M. Protein interaction maps for model organisms. *Nat Rev Mol Cell Biol* **2**, 55–62 (2001).
- Kooperberg, C. & Leblanc, M. Increasing the power of identifying gene x gene interactions in genome-wide association studies. *Genet Epidemiol* **32**, 255–263, doi: 10.1002/gepi.20300 (2008).
- Huang, Y. *et al.* Molecular Basis of Gene–Gene Interaction: Cyclic Cross-Regulation of Gene Expression and Post-GWAS Gene–Gene Interaction Involved in Atrial Fibrillation. *PLoS Genet* **11**, e1005393, doi: 10.1371/journal.pgen.1005393 (2015).
- Bielecka-Dabrowa, A. *et al.* Differences in biochemical and genetic biomarkers in patients with heart failure of various etiologies. *Int J Cardiol* **221**, 1073–1080, doi: 10.1016/j.ijcard.2016.07.150 (2016).
- Gluba, A., Banach, M., Mikhailidis, D. P. & Rysz, J. Genetic determinants of cardiovascular disease: the renin-angiotensin-aldosterone system, paraoxonases, endothelin-1, nitric oxide synthase and adrenergic receptors. *In vivo* **23**, 797–812 (2009).
- Ritchie, M. D. *et al.* Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *American journal of human genetics* **69**, 138–147, doi: S0002-9297(07)61453-0 [pii]10.1086/321276 (2001).
- Chung, Y., Lee, S. Y., Elston, R. C. & Park, T. Odds ratio based multifactor-dimensionality reduction method for detecting gene-gene interactions. *Bioinformatics* **23**, 71–76, doi: bt1557 [pii]10.1093/bioinformatics/bt1557 (2007).
- Carlson, C. S., Eberle, M. A., Kruglyak, L. & Nickerson, D. A. Mapping complex disease loci in whole-genome association studies. *Nature* **429**, 446–452, doi: 10.1038/nature02623 (2004).
- Gavin, A. C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
- Cowley, M. J. *et al.* PINA v2.0: mining interactome modules. *Nucleic Acids Res* **40**, D862–D865, doi: 10.1093/nar/gkr967 (2012).
- Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74, doi: 10.1038/nature15393 (2015).
- Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191, doi: 10.1093/bioinformatics/btq340 (2010).

Acknowledgements

The study was supported by the German Federal Ministry of Education and Research (BMBF) in the context of the German National Genome Research Network (NGFN), the German National Competence network on atrial fibrillation (AFNET), the Leducq Foundation (07-CVD 03), the D.W. Reynolds Foundation Clinical Cardiovascular Research Center at Johns Hopkins University, and the Bioinformatics for the Functional Analysis of Mammalian Genomes program (BFAM) by grants to Stefan Käab (01GS0499, 01GI0204, 01GS0838). The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research

(BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. Dr. Sinner is supported by the German Heart Foundation. The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C,), R01HL087641, R01HL59367, R01HL086694, and R01HL111314; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. Dr. Alonso is supported by grant 16EIA26410001 from the American Heart Association. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. This CHS research was supported by NHLBI contracts N01-HC-85239, N01-HC-85079, N01-HC-85080, N01-HC-85081, N01-HC-85082, N01-HC-85083, N01-HC-85084, N01-HC-85085, N01-HC-85086; N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, HHSN268200800007C, HHSN268201200036C and NHLBI grants HL080295, HL087652, HL103612, HL105756, HL120393 with additional contribution from NINDS. Additional support was provided through AG-023629, AG-15928, AG-20098, and AG-027058 from the NIA. See also <http://www.chs-nhlbi.org/pi.htm>. The Cleveland Clinic AF study was supported by NIH grants to Drs. Chung, Barnard, Smith, and Van Wagoner (R01HL090620, R01HL111314), an NIH/NCRR CTSA grant (1UL1RR024989), Heart and Vascular Institute Philanthropic grant, Department of Cardiovascular Medicine, Cleveland Clinic (Chung), and a Leducq Fondation grant 07-CVD-03 (Van Wagoner). The Framingham Heart Study is conducted and supported by the NHLBI in collaboration with Boston University (Contract No. N01-HC-25195; HHSN268201500001I). This project also was supported by NIH grants to Drs. Ellinor, Benjamin, and Lunetta (1R01HL092577) and Benjamin and Ellinor (1R01HL128914). LURIC was supported by the 7th Framework Program (AtheroRemo, grant agreement number 201668 and RiskyCAD, grant agreement number 305739) of the EU. The PREVENT study is supported by the Dutch Kidney Foundation (grant E0.13), the National Institutes of Health (grant 2R01LM010098), The Netherlands organization for health research and development (NWO-Groot grant 175.010.2007.006, ZonMw grant 90.700.441), and the Dutch Inter University Cardiology Institute Netherlands (ICIN), and the Netherlands Heart Foundation (grant NHS2010B280). Dr. M. Rienstra is supported by a grant from the Netherlands Organization for Scientific Research (Veni grant 016.136.055). There are no relations with industry. The PROSPER study was supported by an investigator initiated grant obtained from Bristol-Myers Squibb. Prof. Dr. J. W. Jukema is an Established Clinical Investigator of the Netherlands Heart Foundation (grant 2001 D 032). Support for genotyping was provided by the seventh framework program of the European commission (grant 223004) and by the Netherlands Genomics Initiative (Netherlands Consortium for Healthy Aging grant 050-060-810). SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED)' funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg- West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG and the Caché Campus program of the InterSystems GmbH. The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. The Rotterdam Study is supported by the Erasmus MC and Erasmus University Rotterdam; the Netherlands Organisation for Scientific Research (NWO); the Netherlands Organisation for Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly (RIDE) the Netherlands Genomics Initiative (NGI); the Ministry of Education, Culture and Science; the Ministry of Health Welfare and Sport; the European Commission (DG XII); and the Municipality of Rotterdam. Dr. Ellinor is supported by grants from the National Institutes of Health (HL104156, HL105780, HL065962). Dr. Ellinor is also supported by 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, and a Doris Duke Charitable Foundation Clinical Scientist Development Award 2014105.

Author Contributions

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

Additional Information

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

Competing financial interests: Dr. Ellinor serves as the principal investigator on a grant from Bayer HealthCare to the Broad Institute related to atrial fibrillation. Dr. Psaty serves on the DSMB of a clinical trial funded by the device manufacturer (Zoll LifeCor) and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson.

How to cite this article: Lin, H. *et al.* Gene-gene Interaction Analyses for Atrial Fibrillation. *Sci. Rep.* **6**, 35371; doi: 10.1038/srep35371 (2016).

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) 2016