# Genetic insights into resting heart rate and its role in cardiovascular disease

## Supplementary Information

Yordi J. van de Vegte#, Ruben N. Eppinga#, M. Yldau van der Ende, Yanick P. Hagemeijer, Yuvaraj Mahendran, Elias Salfati, Albert Smith, Vanessa Tan, Dan E. Arking, Ioanna Ntalla, Emil Appel, Claudia Schurmann, Ioanna Ntalla, Jennifer A. Brody, Rico Rueedi, Ozren Polasek, Delnaz Roshandel, Gardar Sveinbjornsson, Cecile Lecoeur, Claes Ladenvall, Jing Hua Zhao, Aaron Isaacs, Lihua Wang, Jian'an Luan, Shih-Jen Hwang, Nina Mononen, Kirsi Auro, Anne U. Jackson, Lawrence F. Bielak, Linyao Zeng, Nabi Shah, Maria Nethander, Archie Campbell, Tuomo Rankinen, Sonali Pechlivanis, Lu Qi, Wei Zhao, Federica Rizzi, Toshiko Tanaka, Antonietta Robino, Massimiliano Cocca, Leslie Lange, Martina Müller-Nurasyid, Carolina Roselli, Weihua Zhang, Marcus E. Kleber, Xiuqing Guo, Henry J. Lin, Francesca Pavani, Tessel E. Galesloot, Raymond Noordam, Yuri Milaneschi, Katharina E. Schraut, Marcel den Hoed, Frauke Degenhardt, Stella Trompet, Marten E. van den Berg, Giorgio Pistis, Yih-Chung Tham, Stefan Weiss, Xueling Sim, Hengtong L. Li, Peter J. van der Most, Ilja M. Nolte, Leo-Pekka Lyytikäinen, M. Abdullah Said, 23andMe Research Team, Daniel Witte, Carlos Iribarren, Lenore Launer, Susan Ring, Paul S. de Vries, Peter Sever, Allan Linneberg, Erwin P. Bottinger, Sandosh Padmanabhan, Bruce M. Psaty, Nona Sotoodehnia, Ivana Kolcic, The DCCT/EDIC Research Group, David O. Arnar, Daniel F. Gudbjartsson, Hilma Holm, Beverley Balkau, Claudia T. Silva, Christopher H. Newton-Cheh, Kjell Nikus, Perttu Salo, Karen Mohlke, Patricia A. Peyser, Heribert Schunkert, Mattias Lorentzon, Sandosh Padmanabhan, Jari Lahti, D.C. Rao, Marilyn Cornelis, Jessica D. Faul, Jennifer A. Smith, Katarzyna Stolarz-Skrzypek, Stefania Bandinelli, Maria Pina Concas, Gianfranco Sinagra, Thomas Meitinger, Melanie Waldenberger, Moritz F. Sinner, Konstantin Strauch, Graciela E. Delgado, Kent D. Taylor, Jie Yao, Luisa Foco, Olle Melander, Jacqueline de Graaf, Renée de Mutsert, Eco J. de Geus, Åsa Johansson, Peter K. Joshi, Lars Lind, Andre Franke, Peter W. Macfarlane, Kirill Tarasov, Nicholas Tan, Stephan B. Felix, E-Shyong Tai, Debra Q. Quek, Harold Snieder, Johan Ormel, Martin Ingelsson, Cecilia Lindgren, Andrew P. Morris, Olli T. Raitakari, Torben Hansen, Themistocles Assimes, Vilmundur Gudnason, Nicholas J. Timpson, Alanna C. Morrison, Patricia B. Munroe, David P. Strachan, Niels Grarup, Ruth J. Loos, Patricia B. Munroe, Susan R. Heckbert, Peter Vollenweider, Caroline Hayward, Andrew D. Paterson, Kari Stefansson, Philippe Froguel, Leif Groop, Nicholas J. Wareham, Cornelia M. van Duijn, Mary F. Feitosa, Nicholas J. Wareham, Christopher J. O'Donnell, Mika Kähönen, Markus Perola, Michael Boehnke, Sharon L. Kardia, Jeanette Erdmann, Colin N. Palmer, Claes Ohlsson, David J. Porteous, Johan G. Eriksson, Claude Bouchard, Susanne Moebus, Peter Kraft, David R. Weir, Daniele Cusi, Luigi Ferrucci, Sheila Ulivi, Giorgia Girotto, Adolfo Correa, Stefan Kääb, Annette Peters, John C. Chambers, Jaspal S. Kooner, Winfried März, Jerome I. Rotter, Andrew A. Hicks, J. Gustav Smith, Lambertus A. Kiemeney, Dennis O. Mook-Kanamori, Brenda W. Penninx, Ulf Gyllensten, James F. Wilson, Stephen Burgess, Johan Sundström, Wolfgang Lieb, J. Wouter Jukema, Mark Eijgelsheim, Edward L. Lakatta, Ching-Yu Cheng, Marcus Dörr, Tien-Yin Wong, Charumathi Sabanayagam, Albertine J. Oldehinkel, Harriette Riese, Terho Lehtimäki, Niek Verweij, Pim van der Harst.

# Shared first author

## Supplementary Information

#### **Supplementary Figures**

**Supplementary Figure 1:** quantile–quantile (QQ) plot for the GWAS of RHR in A) the UK Biobank and B) the IC-RHR.

Supplementary Figure 2: Network plot of DEPICT gene set enrichment analyses way.

Supplementary Figure 3: ECG-wide heatmap and single cell gene expression dotplot of RHR SNPs.

Supplementary Figure 4: ECG-wide Mendelian randomization analyses of RHR SNPs.

**Supplementary Figure 5:** Forestplot of the results of the association between the genetic risk score of RHR and all-cause mortality across different sets of SNPs, effect sizes, *P* value thresholds, populations and follow-up lengths.

**Supplementary Figure 6:** Scatterplots of the Mendelian randomization analyses between genetically predicted RHR and mortality and longevity within the UK Biobank.

**Supplementary Figure 7:** Scatterplots of the Mendelian randomization analyses between genetically predicted RHR and cardiovascular diseases within the UK Biobank.

**Supplementary Figure 8:** Dose-response curve of the non-linear Mendelian randomization analyses between genetically predicted RHR and cardiovascular diseases within the UK Biobank.

**Supplementary Figure 9:** Scatterplots of the Mendelian randomization analyses between genetically predicted RHR and cardiovascular diseases within the CARDIoGRAMplusC4D, AFGen or MEGASTROKE cohorts.

**Supplementary Figure 10:** Scatterplots of the Mendelian randomization analyses between genetically predicted RHR and blood pressure phenotypes within the ICBP consortium

#### **Supplementary Tables**

**Supplementary Table 1:** Sensitivity analysis for the two-sample Mendelian randomization analysis between RHR and dilated cardiomyopathy.

#### **Supplementary Data**

Supplementary Data 1: Study characteristics of the cohorts.

Supplementary Data 2: 493 Genome-wide significant RHR SNPs.

**Supplementary Data 3:** Comparison of previously RHR associated loci and/or genetic variants identified in the studies from Guo *et al.*, Eppinga *et al.*, and Den hoed *et al.*, with the genetic variants associated with RHR in the current study.

**Supplementary Data 4:** Chow-test for all genome-wide significant RHR SNPs to assess differences of effect estimates between participants taking RHR-altering medication or with a history of any cardiovascular disease versus those who did not.

Supplementary Data 5: Genetic correlation between RHR and previously performed GWAS's.

Supplementary Data 6: List of RHR variants associated with previously discovered variants.

Supplementary Data 7: List of coding variants.

Supplementary Data 8: List of functional eQTL genes.

Supplementary Data 9: List of DEPICT genes.

Supplementary Data 10: List of gene annotations for all identified genes.

Supplementary Data II: Results of gene set enrichment analyses by DEPICT.

Supplementary Data 12: Results of tissue enrichment analysis by DEPICT.

Supplementary Data 13: Effect of RHR SNPs on the ECG.

**Supplementary Data 14:** Mean scaled expression per gene and tissue from the Single-nucleus RNA sequencing data obtained from the healthy human heart.

**Supplementary Data 15:** Definitions of mortality and cardiovascular disease phenotypes in the UK Biobank.

**Supplementary Data 16:** Definitions of cardiovascular disease phenotypes in the CARDIoGRAMplusC4D, AFGen and MEGASTROKE consortia.

**Supplementary Data 17:** Results of the two-sample Mendelian randomization analyses of RHR on mortality within the UK Biobank.

**Supplementary Data 18:** Additional sensitivity analyses of the two-sample Mendelian randomization analyses of RHR on mortality within the UK Biobank.

**Supplementary Data 19:** Single SNP exposure, outcome and exposure-outcome associations between RHR and mortality.

**Supplementary Data 20:** Association between genetic risk scores of RHR and all-cause mortality across different sets of SNPs, effect sizes, *P* value thresholds, populations and follow-up lengths.

**Supplementary Data 21:** Results of the non-linear Mendelian randomization estimates between genetically predicted RHR and all-cause mortality and cardiovascular diseases in the UK Biobank.

**Supplementary Data 22:** Localized average causal effects on all-cause mortality and cardiovascular diseases in the UK Biobank for 30 quantiles of RHR.

**Supplementary Data 23:** Results of the Mendelian randomization between RHR and cardiovascular diseases.

**Supplementary Data 24:** Additional sensitivity analyses of the two-sample Mendelian randomization analyses of RHR and cardiovascular diseases.

**Supplementary Data 25:** Single SNP exposure, outcome and exposure-outcome associations between RHR (effect sizes IC-RHR) and cardiovascular disease (UK Biobank).

**Supplementary Data 26:** Single SNP exposure, outcome and exposure-outcome associations between RHR (effect sizes UK Biobank) and cardiovascular diseases (CARDIoGRAMplusC4D, AFGen and MEGASTROKE consortia).

**Supplementary Data 27:** Results of the two-sample multivariable Mendelian randomization analyses between resting heart rate, atrial fibrillation, blood pressure traits and stroke.

**Supplementary Data 29:** Sensitivity analyses in the two-sample multivariable MR between resting heart rate, atrial fibrillation, blood pressure traits and stroke.

**Supplementary Data 29:** Results of the Mendelian randomization between RHR and blood pressure phenotypes within the ICBP consortium.

**Supplementary Data 30:** Additional sensitivity analyses of the Two-sample Mendelian randomization analyses of RHR and blood pressure phenotypes within the ICBP consortium.

**Supplementary Data 31:** List of Wald estimates with significant ( $P < 1.01 \times 10^{-4}$ ) associations with the cardiovascular outcomes.

Supplementary Figure 1: quantile-quantile (QQ) plot for the GWAS of RHR in A) the UK Biobank and B) the IC-RHR



Quantile-quantile (QQ) plot for the GWAS of RHR within A) the UK Biobank B) the 99 cohorts of the IC-RHR. The genomic intercept indicated a possibility of population stratification for the UK Biobank GWAS. However, the attenuation ratio statistic indicated polygenicity to be the main cause of the observed inflation of test statistics for the UK Biobank GWAS of RHR. The X-axis shows the expected distribution in  $-\log_{10}(P-value)$ . The red line follows expected P-values from a theoretical  $\chi_2$ -distribution, whereas the black line follows the observed P-values in the current GWAS.



Supplementary Figure 2: Network plot of DEPICT gene set enrichment analyses way.

Pathway analysis identified 1.471 significantly enriched gene-sets relevant for resting heart rate. Enriched genesets were further clustered on the basis of the correlation between scores for all genes using an Affinity Propagation method as provided by DEPICT. The 155 meta-gene set clusters are shown. Each cluster was named according to the name of the most central gene set as identified using the Affinity Propagation method. Identified meta-clusters were compared to the clusters found in the study of Eppinga et al. and were determined to be new if not a single cluster within the meta-cluster had been identified before. A red border around the node indicates a newly discovered meta-geneset, a black border a previously discovered geneset. P-value is provided in a single hue blue scale, the strength of the correlation is provided by the color and width of the edges. **Supplementary Figure 3:** ECG-wide heatmap and single cell gene expression dotplot of RHR associated SNPs.



The ECGenetics browser was used to gain insights in the electrophysiological effect of the RHR SNPs and were tested for their association with non-normalized (left panel) and normalized ECG association patterns (middle panel). All SNPs associated with at least one point on the ECG at a Bonferonni-corrected P-value of  $0.05/493/1000 = 1 \times 10-7$  are shown. Effects were aligned to the most positively associated allele across all time points, in which red indicates a positive effect, blue a negative effect and yellow indicating no effect. Single nucleus RNA data from the study of Tucker et al. was queried for all identified genes to gain insights in transcriptional and cellular diversity of RHR gene expression. The right panel shows a dotplot detailing information of single cell gene expression for the most likely candidate gene, with the dot size detailing the percentage of cells which showed expression for the gene and the blue hue the mean scaled expression. Supplementary Figure 4: ECG-wide Mendelian randomization analyses of RHR associated SNPs



.



The ECG genetics browser was used to gain insights in the total effect of the 493 RHR variants on ECG morphology. An ECG-wide MR approach (inverse variance weighted fixed-effects model) was used and this figure shows the results on the non-normalized (panel A) and normalized association pattern (panel B). The X-axis shows the time in ms, the Y-axis the signed -log<sub>10</sub>(P-values). The dotted black lines are the average ECG amplitude of the full cohort as analyzed in the study by Verweij et al. The red lines the P-value for association with each time point of the ECG (n=500 timepoints) on a log<sub>10</sub> scale, signed to show direction of association. Supporting data is provided in Supplementary Data 12.

**Supplementary Figure 5:** Forestplot of the results of the association between the genetic risk score of RHR and all-cause mortality across different sets of SNPs, effect sizes, *P* value thresholds, populations and follow-up lengths.



Genetic risk scores of resting heart rate were constructed using newly discovered variants within the full meta-analyses with the independent effect sizes of the IC-RHR (blue), the effect sizes of the UK Biobank (red) and using the previously discovered variants (green) at five P value thresholds ( $1 \times 10-8, 5 \times 10-8, 1 \times 10-7, 1 \times 10-6, 1 \times 10-5$ ). The darkness of the color represents the strictness of the inclusion P-value threshold, with lighter color meaning a more liberal threshold.

The association with all-cause mortality was tested in different subsets of the UK Biobank (all individuals, ncases = 396,183, ncontrols = 16,289; individuals which were in the UK Biobank interim release from May 2015 and included in the GWAS by Eppinga et al., ncases = 113,102, ncontrols = 4.953; those that were not in the not UK Biobank interim release and therefore not included in the GWAS by Eppinga et al. (ncases = 28,3081, ncontrols = 11,336). The results are shown in panel A).

The analyses were re-performed using mortality data up until the date available in the study of Eppinga et al. (All individuals, ncases = 405,373, ncontrols = 7,099; individuals which were in the UK Biobank interim release from May 2015, ncases = 115,956, ncontrols = 2,099); those that were not in the not UK Biobank interim release and therefore not included in the GWAS by Eppinga et al. (ncases = 289,417, ncontrols = 5,000;. The results are shown in panel B).

This figure shows that the discrepancy in the results between the current and our previous study is likely due to the MR-approach used (Twosample vs. One-sample, respectively). Using genetic variants associated with RHR in the current study and effect sizes from the IC-RHR (shown in blue), we did not find that a) liberating the P value threshold for inclusion of RHR variants to  $1 \times 10^{-5}$  (HR 1.017, 95% CI 0.993-1.041, P=0.16), b) assessing the association in only individuals included in the UK Biobank interim release of May 2015 (HR 1.027, 95% CI 0.976-1.082, P=0.30) or c) the combination of the previous two options (HR 1.025, 95% CI 0.982-1.069, P=0.25) contribute to the discrepancy of the results as we still did not find an association between genetically predicted RHR and all-cause mortality. Using genetic variants associated with RHR in our previous study, we find evidence for a significant association with all-cause mortality when loosening the P-value threshold for inclusion to  $P < 1 \times 10^{-5}$  while assessing the association within the individuals which were included in the UK Biobank interim release and hence in the discovery GWAS (HR 1.171, 95% CI 0.64-1.121, P=6.91 × 10<sup>-6</sup>), while this was not true when testing the association in individuals not included in the UK Biobank interim release (IR 0.994, 95% CI 0.962-1.027, P=0.71). This makes it likely that the MR approach (One- vs Two-sample) rather than genetic variant selection is the reason for the discrepancy between the current and previous results describing the association between genetically predicted RHR and all-cause mortality. Scaling back the followup length did not alter the results part from broadening the confidence intervals.

The Y-axes show hazard ratios and 95% confidence intervals. GV = genetic variant; IC-RHR = International cohorts for resting heart rate; UKB = UK Biobank; HR = hazard ratio; CI = confidence interval; IQR = inter-quartile range.

**Supplementary Figure 6:** Scatterplots of the Mendelian randomization analyses between genetically predicted RHR and mortality and longevity within the UK Biobank.















Scatter plots of the Mendelian randomization analyses between genetically predicted RHR and mortality and longevity within the UK Biobank. The variants' effect size and standard error on RHR (obtained from the IC-RHR meta-analysis) are displayed on the X-axis, the variants' effect size and standard error on (major causes of) of mortality or longevity on the Y-axis. The blue line is the regression line of the inverse variance weighted multiplicative random effects meta-analysis, the green line of the MR contamination mixture model, the red line of the MR-Egger analysis, the orange line of the MR Lasso method and the purple line of the weighted median method. SNP denotes single nucleotide polymorphism, MR denotes Mendelian randomization.

**Supplementary Figure 7:** Scatterplots of the Mendelian randomization analyses between genetically predicted RHR and cardiovascular diseases within the UK Biobank.





Scatter plots of the Mendelian randomization analyses between genetically predicted RHR and cardiovascular diseases within the UK Biobank. The variants' effect size and standard error on RHR (obtained from the IC-RHR meta-analysis) are displayed on the X-axis, the variants' effect size and standard error on (major causes of) of mortality or longevity on the Y-axis. The blue line is the regression line of the inverse variance weighted multiplicative random effects meta-analysis, the green line of the MR contamination mixture model, the red line of the MR-Egger analysis, the orange line of the MR Lasso method and the purple line of the weighted median method. SNP denotes single nucleotide polymorphism, MR denotes Mendelian randomization.



**Supplementary Figure 8:** Dose-response curve of the non-linear Mendelian randomization analyses between genetically predicted RHR and cardiovascular diseases within the UK Biobank.



The graphs represent the non-linear Mendelian randomization estimates. Shown are the dose-response curve between genetically predicted RHR and all-cause mortality and cardiovascular diseases in the UK Biobank study. The comparisons are conducted within strata and therefore the graph provides information on the expected average change in the outcome if a person with a RHR of (say) 70 bpm instead had a RHR value of 90 bpm. Consequences of the expected change in the outcome can only be made if the individuals with a RHR of 70 bpm are otherwise similar to those with a RHR of 90 bpm. The gradient at each point of the curve is the localized average causal effect. Shaded areas represent 95% confidence intervals.

**Supplementary Figure 9:** Scatterplots of the Mendelian randomization analyses between genetically predicted RHR and cardiovascular diseases within the CARDIoGRAMplusC4D, AFGen or MEGASTROKE cohorts.







Scatter plots of the Mendelian randomization analyses between genetically predicted RHR and cardiovascular diseases within the CARDIOGRAMplusC4D, AFGen or MEGASTROKE cohorts. The variants' effect size and standard error on RHR (obtained from the UK Biobank GWAS) and are displayed on the X-axis, the variants' effect size and standard error on (major causes of) of mortality or longevity on the Y-axis. The blue line is the regression line of the inverse variance weighted multiplicative random effects meta-analysis, the green line of the MR contamination mixture model, the red line of the MR-Egger analysis, the orange line of the MR Lasso method and the purple line of the weighted median method. SNP denotes single nucleotide polymorphism, MR denotes Mendelian randomization.

**Supplementary Figure 10:** Scatterplots of the Mendelian randomization analyses between genetically predicted RHR and blood pressure phenotypes within the ICBP consortium.



Scatter plots of the Mendelian randomization analyses between genetically predicted RHR and blood pressure phenotypes within the ICBP consortium, after subtraction of the effect sizes from individuals from the UK Biobank from the total ICBP effect sizes. The variants' effect size and standard error on RHR (obtained from the UK Biobank GWAS) and are displayed on the X-axis, the variants' effect size and standard error on blood pressure phenotypes on the Y-axis. The plots on the left display the results before MR-Steiger filtering, the plots on the right after MR steiger filtering. The discrepancy between the results before and after MR-Steiger filtering for systolic blood pressure and pulse pressure indicate that the association between the RHR associated genetic variants and pulse pressure is unlikely mediated through RHR entirely. The blue line is the regression line of the inverse variance weighted multiplicative random effects meta-analysis, the green line of the MR contamination mixture model, the red line of the MR-Egger analysis, the orange line of the MR Lasso method and the purple line of the weighted median method. SNP denotes single nucleotide polymorphism, MR denotes Mendelian randomization.

Method	Nsnp	Beta	Se	P-value	OR	95% CI	95% CI	Noutcome	Ncontrol
						min	plus		
Inverse variance weighted (fixed									
effects)	377	0.378	0.077	9.82 × 10 <sup>-07</sup>	1.459	1.254	1.697	824	411648
Inverse variance weighted									
(multiplicative random effects)	377	0.378	0.081	2.89 × 10 <sup>-06</sup>	1.459	1.245	1.709	824	411648
MR Egger	377	0.274	0.156	0.08	1.315	0.968	1.786	824	411648
Weighted median	377	0.188	0.130	0.15	1.207	0.935	1.558	824	411648
MR Lasso	377	0.389	0.079	7.55 × 10 <sup>-07</sup>	1.475	1.265	1.721	824	411648
MR contamination mixture model	377	0.549	0.158	4.98 × 10 <sup>-04</sup>	1.731	1.271	2.357	824	411648

Supplementary Table 1: Sensitivity analysis for the Two-sample Mendelian randomization analysis between RHR and dilated cardiomyopathy

We corrected for potential reversed causation through exclusion of 96 SNPs that showed a minimal association (P<0.05) with Q-R upslope at -18ms of the R peak, which has been proven to be a biomarker for DCM. This corresponds to the "pval\_241" column in the Non-RR interval corrected ECG associations (column IR) in Online Table 11. Similar to the main analysis, we found evidence for some balanced horizontal pleiotropy (P index = 8.66% (CI= 0.0-20.4); Cochran's Q = 411.65, df = 367; P = 0.09), but not for unbalanced horizontal pleiotropy (Q-Q' = 0.66, df = 1, P = 0.42; MR-Egger intercept 0.006  $\pm$  0.006, P = 0.44). There was no evidence for weak instrument bias in the MR-Egger regression ( $F_{GX} = 0.96$ ).