Molecular Psychiatry
https://doi.org/10.1038/s41380-021-01087-0

ARTICLE

Multi-ancestry genome-wide gene–sleep interactions identify novel loci for blood pressure

Heming Wang[1,2] et al.

Received: 11 June 2020 / Revised: 18 March 2021 / Accepted: 29 March 2021

Abstract

Long and short sleep duration are associated with elevated blood pressure (BP), possibly through effects on molecular pathways that influence neuroendocrine and vascular systems. To gain new insights into the genetic basis of sleep-related BP variation, we performed genome-wide gene by short or long sleep duration interaction analyses on four BP traits (systolic BP, diastolic BP, mean arterial pressure, and pulse pressure) across five ancestry groups in two stages using 2 degree of freedom (df) joint test followed by 1df test of interaction effects. Primary multi-ancestry analysis in 62,969 individuals in stage 1 identified three novel gene by sleep interactions that were replicated in an additional 59,296 individuals in stage 2 (stage 1 + 2 $P_{joint} < 5 \times 10^{-8}$), including rs7955964 (*FIGNL2/ANKRD33*) that increases BP among long sleepers, and rs73493041 (*SNORA26/C9orf170*) and rs10406644 (*KCTD15/LSM14A*) that increase BP among short sleepers ($P_{int} < 5 \times 10^{-8}$). Secondary ancestry-specific analysis identified another novel gene by long sleep interaction at rs111887471 (*TRPC3/KIAA1109*) in individuals of African ancestry ($P_{int} = 2 \times 10^{-6}$). Combined stage 1 and 2 analyses additionally identified significant gene by long sleep interactions at 10 loci including *MKLN1* and *RGL3/ELAVL3* previously associated with BP, and significant gene by short sleep interactions at 10 loci including C2orf43 previously associated with BP ($P_{int} < 10^{-3}$). 2df test also identified novel loci for BP after modeling sleep that has known functions in sleep–wake regulation, nervous and cardiometabolic systems. This study indicates that sleep and primary mechanisms regulating BP may interact to elevate BP level, suggesting novel insights into sleep-related BP regulation.

Introduction

Hypertension (HTN), including elevations in systolic blood pressure (SBP) and/or diastolic blood pressure (DBP), is a major risk factor for cardiovascular diseases, stroke, renal failure, and heart failure [1]. The heritability of HTN is estimated to be 30–60% in family studies [2, 3]. Recent well-powered large genome-wide association studies (GWAS) of blood pressure (BP) have identified over 1000 loci; however, in total these explain less than 3.5% of BP variation [4–16]. As complex traits are the likely result of an interplay between genes and the environment, gene–environment (G × E) interaction analyses have been proposed as a promising approach to explain additional heritability and identified novel loci for traits associated with cardiometabolic diseases [17, 18].

Long and short sleep durations are associated with elevated BP, possibly through effects on molecular pathways that influence neuroendocrine and vascular systems [19]. Recent multi-ancestry interaction analyses between genetic variants and sleep duration (gene–sleep for short) on blood lipid traits have identified novel loci and potentially distinct mechanisms for short- and long-sleep associated dyslipidemia, and suggest a modification effect of sleep–wake exposures on lipid biology [18]. We hypothesize that differences in sleep duration may also modify the effect of genetic factors on BP. Genome-

These authors contributed equally: Heming Wang, Raymond Noordam, Brian E. Cade, Karen Schwander, Thomas W. Winkler, Jiwon Lee, Yun Ju Sung, Amy R. Bentley

These authors jointly supervised this work: Dabeeru C. Rao, Susan Redline, Diana van Heemst

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41380-021-01087-0.

✉ Heming Wang
hwang@bwh.harvard.edu
✉ Diana van Heemst
d.van_heemst@lumc.nl

Extended author information available on the last page of the article

Published online: 15 April 2021

wide interaction study (GWIS) accounting for potential gene–sleep interactions may help identify novel BP loci and reveal new biological mechanisms that can be explored for treatment or prevention of HTN.

Within the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Gene-Lifestyle Interactions Working Group [20], we investigate gene–sleep interactions on BP traits in 122,265 individuals from five ancestry groups. We perform GWIS using 2df joint test of main and interaction effects [21] followed by 1df test of interaction effect to identify novel gene–sleep interactions and gene–BP associations accounting for sleep duration.

Materials and methods

We performed genome-wide meta-analysis of gene–sleep interactions on four BP traits (SBP, DBP, mean arterial pressure [MAP], and pulse pressure [PP]) in 30 cohorts of five ancestry groups in two stages (Supplementary Notes). Stage 1 discovery analyses included 62,969 individuals of European (EUR), African (AFR), Asian (ASN), Hispanic (HIS), and Brazilian (BRZ) ancestries from 16 studies (Supplementary Tables 1–3). Stage 2 replication analyses included 59,296 individuals of EUR, AFR, ASN, and HIS ancestries from 14 additional studies (Supplementary Tables 4–6). We examined long total sleep time (LTST) and short total sleep time (STST) separately as lifestyle exposures. Given the heterogeneity of age, sleep duration, and BP levels across cohorts and ancestry groups, as well as differences in how sleep duration was assessed (Supplementary Tables 2 and 5), we followed procedures used in prior research [18] to categorize 20% of each sample as long sleepers and 20% as short sleepers based on responses to questionnaires, accounting for age and sex variability within each cohort (Supplementary Methods).

The overall study design is described in Fig. 1. To screen for both gene–sleep interactions and genetic main effect on BP accounting for sleep duration, we performed GWIS using 2df joint test of main and interaction effects adjusting for age, sex, population structure, and other cohort-specific covariates in each ancestry of each cohort using various software such as ProbABEL [22], MMAP, and R package sandwich [23] (Supplementary Table 3). Since BMI is associated with both sleep and BP [24, 25], we performed another GWIS additionally adjusted for BMI to identify genetic loci through biological pathways independent of obesity. We then conducted 2 df joint fixed-effects meta-analysis of the combined main and interaction effects (P_{joint}) using Manning et al.'s method implemented in the METAL software [21] across multi-ancestry in stage 1 and stage 2 separately. Secondary ancestry-specific meta-analyses were performed restricted to EUR and AFR groups. We performed extensive study-level and meta-level quality controls (QCs) using the R package EasyQC [26] as described in Supplementary Methods.

Genetic variants with $P_{joint} < 10^{-6}$ in stage 1 were followed up in stage 2 replication analyses and subsequently meta-analyzed with stage 1 summary statistics. The replication significance threshold was defined as stage 2 $P_{joint} < 0.05$ and stage $1 + 2$ $P_{joint} < 5 \times 10^{-8}$, with consistent directions of association effects. To maximize the statistical power, we also performed genome-wide combined stage 1 and 2 meta-analyses in multi-ancestry and EUR groups using a stricter significant threshold ($P_{joint} < 3.125 \times 10^{-9}$), after Bonferroni correction for two independent BP traits, two exposures, with and without BMI adjustment, in two groups.

We then investigate the interaction effect with sleep for the significant novel ($r^2 < 0.1$ and >1 Mb from any previously identified BP locus) and known BP loci (≤1 Mb) using 1df test (P_{int}). Novel gene–sleep interactions were identified with stage $1 + 2$ $P_{int} < 10^{-3}$ accounting for the number of independent loci. We compared the risk effects on BP of loci significantly interact with sleep in individuals with LTST, STST, and normal sleep duration (60% of the sample; Supplementary Methods). The variance of four BP traits explained by the SNP main and interaction effects was estimated using summary statistics in combined analyses using the R package VarExp [27].

Significant novel loci were followed up for bioinformatics analyses. We annotated functional effects for the novel loci using HaploReg [28], Regulome [29], and GTex (v8) [30] database. Genes under the association regions were mapped using PLINK 2.0 [31] and SNPsea [32] software and were interrogated for associated phenotypes, Mendelian diseases, and druggable targets using PheGeni [33], OMIM [34], and DGIdb [35] database. Tissue and pathway enrichment analyses were performed using online software FUMA [36].

This work was approved by the Institutional Review Board of Washington University in St. Louis and complies with all relevant ethical regulations. For each of the participating cohorts, the appropriate ethics review board approved the data collection and all participants provided informed consent. All summary results are available in dbGaP (phs000930.v1.p1).

Results

GWIS

The Miami and QQ plots of stage 1 2df GWIS in multi-ancestry, EUR and AFR groups are provided in

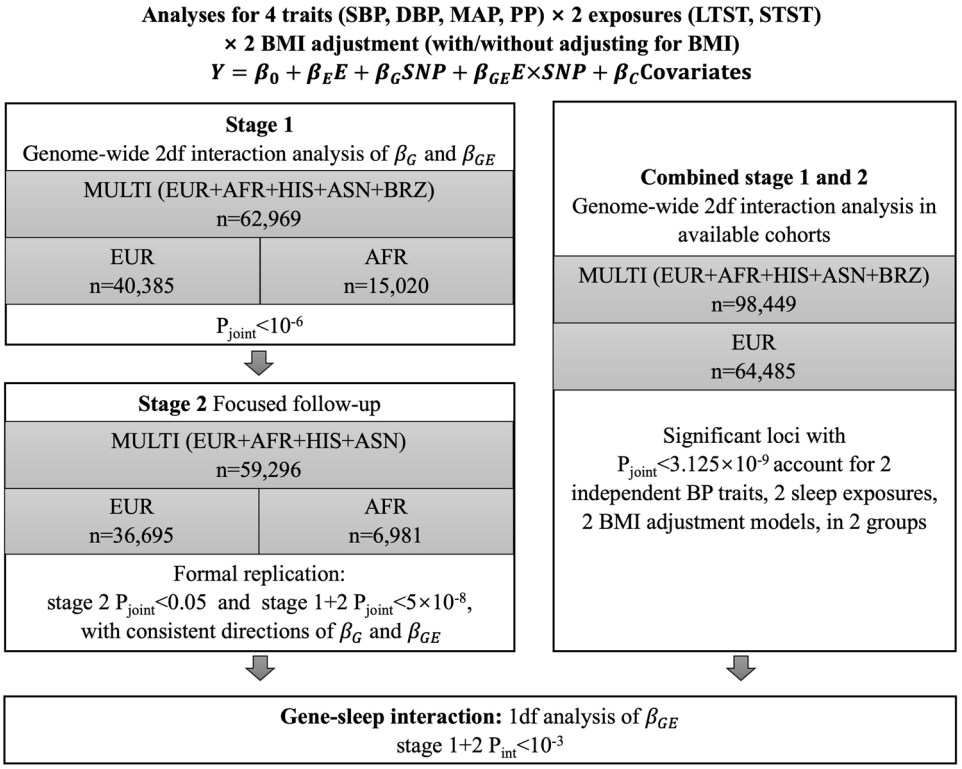


Fig. 1 Study overview. Two-stage genome-wide interaction analyses were performed using 2df joint test to screen for both gene-sleep interactions and genetic main effect on BP accounting for sleep. Formally replicated loci in two-stage analyses and Bonferroni corrected significant loci in combined stage 1 and 2 analyses were followed by 1 df test of interaction effect.

Supplementary Figs 1–6. 1976 genetic variants with $P_{joint} < 10^{-6}$ were followed up for replication analyses. Of these, 1081 variants were available in stage 2 cohorts and passed quality control, of which 268 (24.8%) variants showed $P_{joint} < 0.05$.

Our primary two-stage analyses in the multi-ancestry group formally replicated one novel locus (*FIGNL2/ANKRD33*; Table 1) and eight known loci (*ULK4, CHIC2, PRDM8/FGF5, IGFBP1/IGFBP3, PIK3CG, PDP1/CDH17, GPR20,* and *ADAMTS8*; Supplementary Table 7) in 2df gene-LTST interaction analyses, and two novel loci (*SNORA26/C9orf170* and *KCTD15/LSM14A*; Table 1) and eight known loci (*ULK4, CHIC2, PRDM8/FGF5, IGFBP1/IGFBP3, PIK3CG, PDP1/CDH17, ADAMTS8,* and *SH2B3*, Supplementary Table 7) in 2df gene–STST interaction analyses (stage 2 $P_{joint} < 0.05$ and stage $1 + 2$ $P_{joint} < 5 \times 10^{-8}$). The regional association plots are shown in Supplementary Fig. 7.

In secondary ancestry-specific two-stage analyses, we formally replicated one known BP locus (*INSR*) in 2df gene-STST interaction analyses restricted to EUR individuals (stage 2 $P_{joint} < 0.05$ and stage $1 + 2$ $P_{joint} < 5 \times 10^{-8}$; Supplementary Table 7). We additionally identified three novel loci (*TRPC3/KIAA1109*, *ANK*, and *RP11-322L20.1/RP11-736P16.1*) in 2df gene–LTST interaction analyses restricted to AFR individuals (stage 1 $P_{joint} < 5 \times 10^{-8}$ and stage 2 $P_{joint} < 0.05$, with consistent directions of main effects; Supplementary Table 8). The regional association plots are shown in Supplementary Fig. 8. However, these three variants did not survive our formal replication criteria of stage $1 + 2$ $P_{joint} < 5 \times 10^{-8}$, possibly reflecting heterogeneity between discovery and replication cohorts.

Genome-wide combined stage 1 and stage 2 meta-analyses (Miami and QQ plots in Supplementary Figs 9–12) additionally identified nine novel and four known BP loci in 2df gene–LTST interaction analyses; and 11 novel and three known BP loci in 2df gene–STST interaction analysis ($P_{joint} < 3.125 \times 10^{-9}$; Supplementary Tables 9 and 10). The regional association plots of the 20 novel loci are shown in Supplementary Fig. 13. Replication in independent datasets is needed to validate these unreported loci. Additional loci that were genome-wide significant ($3.125 \times 10^{-9} < P_{joint} < 5 \times 10^{-8}$) are also summarized in Supplementary Tables 11 and 12.

Interactions with sleep

We then investigated the 1df gene–sleep interaction effects of the 26 novel and 18 known loci identified in the two-stage or combined analyses. Among the formally replicated loci in multi-ancestry two-stage analyses, one novel locus rs7955964 (*FIGNL2/ANKRD33*) showed a genome-wide significant 1df SNP × LTST interaction (stage $1 + 2$ $P_{int} < 5 \times 10^{-8}$; Table 1) with risk effect on BP only present in long sleepers (Fig. 2A). Two novel loci, rs73493041 (*SNORA26/C9orf170*) and rs10406644 (*KCTD15/LSM14A*), showed genome-wide significant 1df SNP × STST interactions (stage $1 + 2$ $P_{int} < 5 \times$

Table 1 Replicated novel BP loci significantly associated with sleep duration.

Exposure	rsID	Gene(s)	Chr: position (Build 37)	Alleles (E/A)	EAF	Trait	Stage	BMI adjustment	N	β_{SNP}	SE_{SNP}	β_{Int}	SE_{Int}	P_{Joint}	P_{Int}
LTST	rs7955964	*FIGNL2, ANKRD33*	12:52281279	A/T	0.896	MAP	1	Without BMI	18583	−0.400	0.290	2.711	0.582	1.75×10^{-6}	1.34×10^{-6}
								with BMI	18583	−0.462	0.279	2.768	0.546	2.48×10^{-7}	2.10×10^{-7}
							2	Without BMI	12335	−0.621	0.289	2.163	0.632	2.52×10^{-3}	5.49×10^{-4}
								with BMI	12327	−0.519	0.281	2.210	0.613	1.72×10^{-3}	2.66×10^{-4}
							1 + 2	Without BMI	29985	−0.517	0.208	2.505	0.433	1.11×10^{-7}	4.40×10^{-9}
								with BMI	29957	−0.500	0.201	2.577	0.413	6.74×10^{-9}	2.94×10^{-10}
STST	rs73493041	*SNORA26, C9orf170*	9:89849304	T/C	0.959	DBP	1	Without BMI	36858	−0.725	0.229	2.336	0.471	4.65×10^{-7}	3.6×10^{-7}
								with BMI	36858	−0.723	0.219	2.235	0.456	5.16×10^{-7}	5.18×10^{-7}
							2	Without BMI	24413	−0.763	0.321	1.888	0.705	5.44×10^{-3}	9.43×10^{-3}
								with BMI	24385	−0.704	0.335	1.875	0.704	1.09×10^{-2}	1.27×10^{-2}
							1 + 2	Without BMI	61271	−0.724	0.185	2.213	0.387	3.62×10^{-8}	1.30×10^{-8}
								with BMI	61243	−0.709	0.182	2.132	0.381	7.15×10^{-8}	2.58×10^{-8}
	rs10406644	*KCTD15, LSM14A*	19:34595645	A/G	0.095	PP	1	Without BMI	15021	0.542	0.275	−3.194	0.605	1.26×10^{-7}	1.35×10^{-7}
								with BMI	12921	0.565	0.306	−3.382	0.677	5.23×10^{-7}	4.81×10^{-7}
							2	Without BMI	11401	1.142	0.587	−2.702	1.163	4.59×10^{-2}	2.02×10^{-2}
								with BMI	11373	1.102	0.582	−2.533	1.155	6.08×10^{-2}	2.83×10^{-2}
							1 + 2	Without BMI	26422	0.648	0.249	−3.067	0.536	1.39×10^{-8}	7.59×10^{-9}
								with BMI	24294	0.678	0.271	−3.135	0.584	8.56×10^{-8}	4.35×10^{-8}

10^{-8}; Table 1) with risk effects on BP only present in short sleepers (Fig. 2B and C). Those effects were largely consistent across cohorts. In the EUR population, the aggregate main effects of these three loci explained up to 0.016% of the variation of four BP traits, while the gene–LTST and –STST interaction effects additionally explained 0.002–0.01% and 0.005–0.027% of the variation (Supplementary Table 13). In the AFR population, the aggregate main effect of these three loci explained 0.116–0.188% of the variation of four BP traits, while the gene–LTST and –STST interaction effects additionally explained 0.375–0.784% and 0.162–0.254% of the variation (Supplementary Table 13). Given the limited sample sizes in the AFR group, the estimation of BP variation in AFR is likely inflated.

In the two-stage analyses restricted to AFR individuals, one novel loci rs111887471 (*TRPC3/KIAA1109*) showed significant 1df SNP × LTST interaction with risk effect on BP only present in long sleepers (stage 1 + 2 $P_{int} = 2 \times 10^{-6}$; Supplementary Table 8 and Supplementary Fig 14A).

Among the loci identified in combined stage 1 and stage 2 analyses, eight novel loci (*LINC01720/AL138927.1, RYR2, SEMA4F/HK2, DPP10/DDX18, PDZRN3/CNTN3, LEKR1/LINC00880, FSTL5, AC008558.1/HTR1A*, and *ZFPM2*; Supplementary Table 9) and two previously reported BP loci (*MKLN1* and *RGL3/ELAVL3*; Supplementary Table 10) showed significant 1df interactions with LTST (Pint $< 1 \times 10^{-3}$). The risk effects on BP in long sleepers differed from the effects in normal or short sleepers (Supplementary Fig 14A). Nine novel loci (*GJA4, PSRC1/MYBPHL, AL033381.3/FOXQ1, PTPRN2, ERICH1, AL162384.1/IL33, FRMD4A, RP11-408B11.2*, and *TTC6*; Supplementary Table 9) and one previously reported BP locus (*C2orf43*; Supplementary Table 10) showed significant 1df interactions with STST ($P_{int} < 10^{-3}$; Supplementary Table 9-10). The risk effects on BP in short sleepers differed from the effects in normal or long sleepers (Supplementary Fig 14B).

We also looked up the previously validated 362 BP loci [4–15] and 113 sleep duration loci [37] in the combined analyses, but none of these showed significant 1df interactions after accounting for multiple comparisons ($P_{int} > 10^{-4}$; Supplementary Tables 14–17).

Associations with other relevant traits

2df two-stage and combined analyses total identified 26 novel loci for BP with or without significant 1df interactions (three formally replicated in multi-ancestry two-stage analyses, three in AFR two-stage analyses, and 20 in combined analyses). We looked up the associations between those loci with cardiovascular diseases, stroke, chronic kidney disease, and self-reported and objective (derived from 7-day accelerometry) sleep traits using publicly available genome-

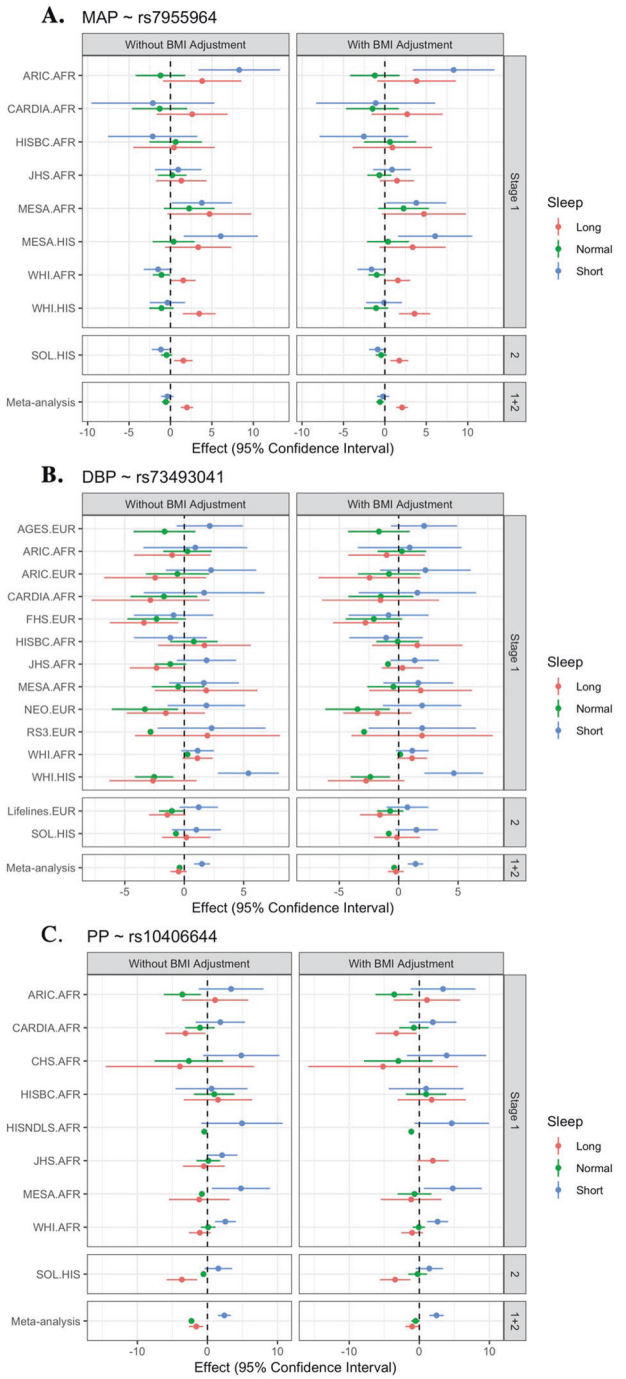


Fig. 2 Forest plots of effects on BP in long, normal, and short sleepers at three replicated novel loci in the multi-ancestry population. A Effects of rs7955964 on MAP in long, normal, and short sleepers across cohorts. **B** Effects of rs73493041 on DBP in long, normal, and short sleepers across cohorts. **C** Effects of rs10406644 on PP in long, normal, and short sleepers across cohorts.

wide summary statistics from large GWAS (Supplementary Tables 18–23). One of the replicated loci rs73493041 (*SNORA26/C9orf170*) was associated with self-reported chronotype (morningness vs eveningness) ($P = 9.1 \times 10^{-6}$; Supplementary Table 22). Among the other novel loci, rs17036094 (*PSRC1/MYBPHL*) was associated with coronary artery disease and myocardial infarction ($P \le 0.005$; Supplementary Table 19), and rs140526840 (*FSTL5*) was associated with chronic kidney disease ($P = 0.006$; Supplementary Table 21),

Bioinformatics analyses

All of the 26 novel variants were mapped to intergenic or intronic regions using HaploReg [28], including four in promoter histone marks, 11 in enhancer histone marks, 10 in DNAse, three altered the binding sites of regulatory proteins, and two conserved elements (Supplementary Table 24).

Among the three replicated novel loci, rs73493041 (*SNORA26/C9orf170*) was an eQTL for *GAS1* in suprapubic skin using GTEx (v8) [30] (Supplementary Table 25). Using PLINK pruning and SNPsea [32], rs7955964 (*SNORA26/C9orf170*) was mapped to a region of 10 genes (Supplementary Table 26), including *ANKRD33* and *NR4A1*, implicated in sleep–wake control regulation and the neurovascular system [38, 39]. Rs10406644 (*KCTD15/LSM14A*) was mapped to a region overlapping with nine genes (Supplementary Table 27), including *KCTD15* and *CHST8*, previously associated with adiposity traits and involved in neurodevelopmental and neuropsychiatric diseases [40–42] (see Discussion).

Among the other 23 novel loci, four variants showed strong eQTL evidence across various tissues such as blood and adipose tissue (Supplementary Table 25). 14 loci were mapped to genes with known functions in cardiac and nervous systems (e.g., *TRPC3* [43], *RYR2* [44], *ANK2* [45], *GJA4* [46], and *SORT1* [47]) and associated with other cardiometabolic (e.g., *HTR1A* [48], *PSRC1* [49], *PSKH1* [50]), inflammatory (e.g., *IL33* [51]), cognition (e.g., *FRMD4A* [52]) and psychiatric traits (e.g., *NFATC3* [53]) (Supplementary Tables 26 and 27).

In total, 11 novel loci harbored genes implicated in Mendelian syndromes such as ventricular tachycardia and cryptogenic cirrhosis. 13 loci harbored one or more genes with potential drug targets (Supplementary Tables 26 and 27).

We performed tissue and pathway enrichment analyses using annotated genes under novel association regions using FUMA [36] (Supplementary Tables 28 and 29). Genes under the association regions in gene–LTST interaction analyses were enriched in multiple artery and cardiac muscle related pathways (Supplementary Table 30).

Discussion

We performed genome-wide gene–sleep interaction analyses on BP using 122,265 individuals from five ancestry groups

in 30 studies in two stages, using a 2df joint test of main and interaction effects followed 1df test investigation of interaction effects. Primary 2df GWIS in multi-ancestry group identified three novel loci that were replicated in additional samples (stage 1 + 2 $P_{joint} < 5 \times 10^{-8}$). Secondary ancestry-specific 2df GWIS additionally identified three novel loci with weak replication evidence in AFR. Combined stage 1 and 2 analyses identified another 20 novel loci after accounting for multiple comparisons ($P_{joint} < 3.125 \times 10^{-9}$), which require external replication. The associations were largely unchanged after additionally adjusting for BMI.

The emergence of novel loci after considering gene–sleep interactions suggests an important modifying role of sleep on BP regulation, which involves both central and peripheral regulation (including the brain, adrenal glands, kidneys, and vasculature). Insufficient or short sleep can increase BP through effects on elevating sympathetic nervous system activity and altering hypothalamic-pituitary-adrenal (HPA) axis activities, leading to hormonal changes, endothelial dysfunction, insulin resistance, and systemic inflammation [19, 54]. The mechanisms underlying the association between long sleep duration and BP are less well understood, and may reflect circadian misalignment in a 24-h period, including disrupted sleep–wake cycle, a misalignment of internal biological clocks with the external environment, and desynchronized central and peripheral clocks in tissues relevant for BP control [55]. The importance of circadian control of BP is evident by the normal nocturnal decline ("dipping") in BP. Non-dipping of BP, associated with increased mortality, is observed with both sleep disturbances and abnormalities of sodium transport in the kidney [56, 57]. Our data suggest that sleep and renal and neuroendocrine control of BP may interact to influence susceptibility to HTN. The novel loci found by gene–LTST and gene–STST interaction analyses were distinct, supporting the different mechanisms of short and long sleep modifying BP. Similarly, in prior gene–sleep interaction analyses for blood lipids, LTST and STST each also modified gene effects in a non-overlapping pattern [18].

Using the 1df test, we identified three novel gene–sleep interactions that were formally replicated in primary multi-ancestry analyses (stage 1 + 2 $P_{int} < 5 \times 10^{-8}$). Among those, rs7955964 (*FIGNL2/ANKRD33*) only increased MAP in long sleepers (Fig. 2A). In the association region under this locus, *ANKRD33* is expressed in retinal photoreceptors and the pineal gland and acts as a transcriptional repressor for CRX-activated photoreceptor gene regulation [38]. Given the importance of light in the central regulation of circadian rhythms, long sleep—a circadian disruptor—may interact with this gene to influence BP [56]. In addition, *NR4A1* (that also maps to this locus) is a member of the nuclear hormone receptor family, which regulates neurohormonal systems including dopamine and norepinephrine and cardiac stress responses [39, 58]. Its expression is influenced by an array of stimuli, including those influence nutrient sensing. Our findings suggest that perturbed sleep and circadian rhythms may also alter the effects of this gene, increasing BP.

Rs10406644 (*KCTD15/LSM14A*) only increased PP in short sleepers. *KCTD15* is implicated in both renal (nephron) development and adiposity, possibly through effects on Wnt signaling and neural crest development. Short sleep can lead to hypothalamic-adrenal-cortisol dysfunction, and potentially may amplify the effects of this gene on metabolism and kidney function to increase BP [59, 60]. This locus also maps to *CHST8* which is associated with adiposity traits [40, 41] as well as to *GPI* that functions in glucose metabolism and immune system pathways [61, 62].

Rs73493041 (*SNORA26/C9orf170*) only increased DBP in short sleepers. Rs73493041 was an eQTL for *GAS1*, a pleiotropic regulator of cellular homeostasis and widely expressed in the central nervous system [63, 64]. The risk allele was also significantly associated with self-reported eveningness chronotype ($P = 9.1 \times 10^{-6}$; Supplementary Table 22), a circadian phenotype associated with increased cardiometabolic and neuropsychiatric disorders [65]. Short sleep may magnify cardiometabolic dysfunction associated with delayed sleep timing.

Given the high prevalence of HTN in African Americans, there is a critical need to identify modifiable risk factors. Notably, African Americans have poorly controlled HTN as well as circadian abnormalities in BP regulation [66]. They also have a higher prevalence of short and long sleep duration compared to individuals of European ancestry [67, 68], likely due to combinations of social–environmental exposures and genetic and epigenetic susceptibility [69]. In AFR specific gene–LTST analyses, we identified a novel SNP–LTST interaction at rs111887471 (*TRPC3/KIAA1109*) with risk effect on SBP only present in long sleepers ($P_{int} = 2 \times 10^{-6}$; Supplementary Fig. 14). *TRPC3* has been shown to play an important role in cardiac ion (Na^+ and Ca^{2+}) homeostasis [43]. The association observed in AFR may reflect differences in BP control with individuals of African ancestry having greater sodium sensitivity [70], with BP effects amplified by disrupted circadian rhythm regulation due to long sleep [57].

Combined stage 1 and 2 analyses additionally identified significant gene–LTST interactions at MKLN1, RGL3/ELAVL3, LINC01720/AL138927.1, RYR2, SEMA4F/HK2, DPP10/DDX18, PDZRN3/CNTN3, LEKR1/LINC00880, FSTL5, AC008558.1/HTR1A, ZFPM2, and significant gene–STST interactions at C2orf43, GJA4, PSRC1/MYBPHL, AL033381.3/FOXQ1, PTPRN2, ERICH1, AL162384.1/IL33, FRMD4A, RP11-408B11.2, and TTC6 ($P_{int} < 10^{-3}$), which require external replication. MKLN1, RGL3/ELAVL3, and C2orf43 have been reported associated with BP previously. Among those, MKLN1 regulates the internalization and transport of the $GABA_A$ receptor [71, 72] and ELAVL3

encodes a neural-specific RNA-binding protein involved in neuronal differentiation and maintenance [73]. We did not observe marginal main effects for those loci among normal sleepers (Supplementary Fig. 14), perhaps because of the small sample size of those variants ($N \leq 10{,}038$; Supplementary Table 10). Our findings suggest that their effects on BP may be amplified in the setting of long sleep due to disrupted circadian rhythm regulation when these effects were not detectable in small samples.

In this study we defined short and long sleep duration using self-reported questionnaires, which can result in misclassification [74], potentially reducing statistical power. Although we used a within cohort approach for harmonizing sleep duration that accounted for age and sex differences across cohorts, there may be systematic residual differences in sleep assessments that resulted in heterogeneity across our samples. Future work using objective measurements (e.g., polysomnography and actigraphy data) may provide further insight into sleep-related BP mechanisms.

Some of our most interesting findings, and ones with high potential public health impact due to the burden of extreme sleep duration and HTN in AFR group. Unfortunately, limited samples of AFR were available for replication. We identified 1976 variants with significant association effect in gene–sleep interaction analyses in stage 1. However, only 1081 of those variants were available in stage 2 analyses. Most of the unavailable variants in stage 2 had been identified in non-EUR cohorts and were rare in EUR populations (MAF < 1%). Future studies following-up these "missing" variants in diverse groups and additional studies of minority populations are needed to further understand mechanisms for BP regulation that are modulated by sleep. In addition, some of our findings were mapped to large genomic regions covering many genes. Further fine-mapping analyses using sequencing data or biochemistry experiments may further clarify the causal variants.

In summary, we performed a large-scale gene–sleep interaction meta-analyses in multi-ancestry groups. This study advances our knowledge on the interactions between genetic risk factors, sleep duration, and blood pressure. This work extends prior research that has reported that extreme sleep duration (short or long) is associated with increased blood pressure as well as cardiovascular morbidity [19], and provides evidence that sleep duration may modify genetic risk for hypertension through pathways that influence photoreception, metabolism, adiposity, renal function, and chronotype. These findings also suggest that sleep duration may modify the effects of antihypertensives that target certain genes or pathways—an area that should be further investigated using pharmacogenetics and pathway-level approaches. Finally, the observation of multiple genetic effects only in individuals with extreme sleep duration supports the general guidance for the public to follow published sleep duration recommendations (7–9 h) [75]—potentially reducing cardiovascular diseases in the population, especially for individuals with genetic predispositions.

Code availability

The URLs of genetic software and database used in this study are provided as follows: ProbABEL, https://github.com/GenABEL-Project/ProbABEL; MMAP, https://mmap.github.io; sandwich, https://github.com/cran/sandwich; METAL, http://csg.sph.umich.edu/abecasis/metal/; EasyQC, http://www.genepi-regensburg.de/easyqc; varExp, https://github.com/vincela/VarExp; HaploReg, https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php; RegulomeDB, http://www.regulomedb.org/; GTEx, https://gtexportal.org/home/; PLINK 2.0, https://www.cog-genomics.org/plink/2.0/; SNPsea, http://pubs.broadinstitute.org/mpg/snpsea/; PheGeni, https://www.ncbi.nlm.nih.gov/gap/phegeni; OMIM, https://www.omim.org; DGIdb, https://www.dgidb.org; FUMA, https://fuma.ctglab.nl. The detailed settings are described in Supplementary Methods.

Acknowledgements This project was supported by the US National Heart, Lung, and Blood Institute (NHLBI) R01HL118305. HW and SR were supported by NHLBI R35HL135818. BEC was supported by NHLBI K01HL135405. ARB was supported by the Intramural Research Program of the National Institutes of Health in the Center for Research on Genomics and Global Health (CRGGH). The CRGGH is supported by the National Human Genome Research Institute, the National Institute of Diabetes and Digestive and Kidney Diseases, the Center for Information Technology, and the Office of the Director at the National Institutes of Health (1ZIAHG200362). D. v.H. was supported by the European Commission funded project HUMAN (Health-2013-INNOVATION-1-602757). The CHARGE cohorts were supported in part by NHLBI infrastructure grant HL105756. Study-specific acknowledgments can be found in the Supplementary Notes.

Author contributions HW, BEC, and JL conducted the centralized data analyses, including quality controls, meta-analyses, and post association lookups and bioinformatics. HW, RN, BEC, KS, TWW, JL, YJS, ARB, DCR, SR, and DvanH were part of the writing group and participated in study design, interpreting the data, and drafting the manuscript. All other co-authors were responsible for cohort-level data collection, cohort-level data analysis, and critical reviews of the draft paper. All authors approved the final version of the paper that was submitted to the journal.

Compliance with ethical standards

Conflict of interest DOMK is a part-time research consultant at Metabolon, Inc. BMP serves on the DSMB of a clinical trial funded by the manufacturer (Zoll LifeCor) and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. HJG has received travel grants and speakers honoraria from Fresenius Medical Care, Neuraxpharm, Servier and Janssen Cilag as well as research funding from Fresenius Medical Care. The remaining authors declare no competing interests.

References

1. Berry JD, Dyer A, Cai X, Garside DB, Ning H, Thomas A, et al. Lifetime risks of cardiovascular disease. N Engl J Med. 2012; 366:321–9.
2. Cooper RS, Luke A, Zhu X, Kan D, Adeyemo A, Rotimi C, et al. Genome scan among Nigerians linking blood pressure to chromosomes 2, 3, and 19. Hypertension. 2002;40:629–33.
3. Levy D, DeStefano AL, Larson MG, O'Donnell CJ, Lifton RP, Gavras H, et al. Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the framingham heart study. Hypertension. 2000;36:477–83.
4. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, et al. Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet. 2009;41:666–76.
5. Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, et al. Genome-wide association study of blood pressure and hypertension. Nat Genet. 2009;41:677–87.
6. International Consortium for Blood Pressure Genome-Wide Association S, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. 2011; 478:103–9.
7. Ehret GB, Ferreira T, Chasman DI, Jackson AU, Schmidt EM, Johnson T, et al. The genetics of blood pressure regulation and its target organs from association studies in 342,415 individuals. Nat Genet. 2016;48:1171–84.
8. Liu C, Kraja AT, Smith JA, Brody JA, Franceschini N, Bis JC, et al. Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. Nat Genet. 2016;48:1162–70.
9. Surendran P, Drenos F, Young R, Warren H, Cook JP, Manning AK, et al. Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. Nat Genet. 2016;48:1151–61.
10. Hoffmann TJ, Ehret GB, Nandakumar P, Ranatunga D, Schaefer C, Kwok PY, et al. Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. Nat Genet. 2017;49:54–64.
11. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. Nat Genet. 2017;49:403–15.
12. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nat Genet. 2018;50:1412–25.
13. Giri A, Hellwege JN, Keaton JM, Park J, Qiu C, Warren HR, et al. Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. Nat Genet. 2019;51:51–62.
14. Franceschini N, Fox E, Zhang Z, Edwards TL, Nalls MA, Sung YJ, et al. Genome-wide association analysis of blood-pressure traits in African-ancestry individuals reveals common associated genes in African and non-African populations. Am J Hum Genet. 2013;93:545–54.
15. Zhu X, Feng T, Tayo BO, Liang J, Young JH, Franceschini N, et al. Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. Am J Hum Genet. 2015;96:21–36.
16. Liang J, Le TH, Edwards DRV, Tayo BO, Gaulton KJ, Smith JA, et al. Single-trait and multi-trait genome-wide association analyses identify novel loci for blood pressure in African-ancestry populations. PLoS Genet. 2017;13:e1006728.
17. Sung YJ, Winkler TW, de Las Fuentes L, Bentley AR, Brown MR, Kraja AT, et al. A large-scale multi-ancestry genome-wide study accounting for smoking behavior identifies multiple significant loci for blood pressure. Am J Hum Genet. 2018;102: 375–400.
18. Noordam R, Bos MM, Wang H, Winkler TW, Bentley AR, Kilpelainen TO, et al. Multi-ancestry sleep-by-SNP interaction analysis in 126,926 individuals reveals lipid loci stratified by sleep duration. Nat Commun. 2019;10:5121.
19. Gangwisch JE. A review of evidence for the link between sleep duration and hypertension. Am J Hypertens. 2014;27:1235–42.
20. Rao DC, Sung YJ, Winkler TW, Schwander K, Borecki I, Cupples LA, et al. Multiancestry study of gene-lifestyle interactions for cardiovascular traits in 610 475 Individuals from 124 cohorts: design and rationale. Circ Cardiovasc Genet. 2017;10:e001649.
21. Manning AK, LaValley M, Liu CT, Rice K, An P, Liu Y, et al. Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP x environment regression coefficients. Genet Epidemiol. 2011;35:11–18.
22. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinforma. 2010;11:134.
23. Zeileis A Object-oriented computation of sandwich estimators. 2006.
24. Grandner MA, Schopfer EA, Sands-Lincoln M, Jackson N, Malhotra A. Relationship between sleep duration and body mass index depends on age. Obes (Silver Spring). 2015;23:2491–8.
25. Martins D, Tareen N, Pan D, Norris K. The relationship between body mass index, blood pressure and pulse rate among normotensive and hypertensive participants in the third National Health and Nutrition Examination Survey (NHANES). Cell Mol Biol (Noisy-le-Gd). 2003;49:1305–9.
26. Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Magi R, et al. Quality control and conduct of genome-wide association meta-analyses. Nat Protoc. 2014;9:1192–212.
27. Laville V, Bentley AR, Prive F, Zhu X, Gauderman J, Winkler TW, et al. VarExp: estimating variance explained by genome-wide GxE summary statistics. Bioinformatics. 2018;34:3412–4.
28. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012;40: D930–934.
29. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012;22:1790–7.
30. Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. 2015;348:648–60.
31. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.
32. Slowikowski K, Hu X, Raychaudhuri S. SNPsea: an algorithm to identify cell types, tissues and pathways affected by risk loci. Bioinformatics. 2014;30:2496–7.
33. Ramos EM, Hoffman D, Junkins HA, Maglott D, Phan L, Sherry ST, et al. Phenotype–Genotype Integrator (PheGenI): synthesizing genome-wide association study (GWAS) data with existing genomic resources. Eur J Hum Genet. 2014;22:144–7.
34. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res. 2005;33:D514–517.

35. Cotto KC, Wagner AH, Feng YY, Kiwala S, Coffman AC, Spies G, et al. DGIdb 3.0: a redesign and expansion of the drug-gene interaction database. Nucleic Acids Res. 2018;46:D1068–D1073.
36. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8:1826.
37. Dashti HS, Jones SE, Wood AR, Lane JM, van Hees VT, Wang H, et al. Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. Nat Commun. 2019;10:1100.
38. Sanuki R, Omori Y, Koike C, Sato S, Furukawa T. Panky, a novel photoreceptor-specific ankyrin repeat protein, is a transcriptional cofactor that suppresses CRX-regulated photoreceptor genes. FEBS Lett. 2010;584:753–8.
39. Medzikovic L, de Vries CJM, de Waard V. NR4A nuclear receptors in cardiac remodeling and neurohormonal regulation. Trends Cardiovasc Med. 2019;29:429–37.
40. Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA, et al. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. PLoS One. 2012;7:e51954.
41. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009;41:25–34.
42. Teng X, Aouacheria A, Lionnard L, Metz KA, Soane L, Kamiya A, et al. KCTD: A new gene family involved in neurodevelopmental and neuropsychiatric disorders. CNS Neurosci Ther. 2019;25:887–902.
43. Eder P, Probst D, Rosker C, Poteser M, Wolinski H, Kohlwein SD, et al. Phospholipase C-dependent control of cardiac calcium homeostasis involves a TRPC3-NCX1 signaling complex. Cardiovasc Res. 2007;73:111–9.
44. Dabertrand F, Nelson MT, Brayden JE. Ryanodine receptors, calcium signaling, and regulation of vascular tone in the cerebral parenchymal microcirculation. Microcirculation. 2013;20:307–16.
45. Kashef F, Li J, Wright P, Snyder J, Suliman F, Kilic A, et al. Ankyrin-B protein in heart failure: identification of a new component of metazoan cardioprotection. J Biol Chem. 2012;287:30268–81.
46. Vicario N, Zappala A, Calabrese G, Gulino R, Parenti C, Gulisano M, et al. Connexins in the central nervous system: physiological traits and neuroprotective targets. Front Physiol. 2017;8:1060.
47. Andersen JL, Schroder TJ, Christensen S, Strandbygard D, Pallesen LT, Garcia-Alai MM, et al. Identification of the first small-molecule ligand of the neuronal receptor sortilin and structure determination of the receptor-ligand complex. Acta Crystallogr D Biol Crystallogr. 2014;70:451–60.
48. Zheng JS, Arnett DK, Lee YC, Shen J, Parnell LD, Smith CE, et al. Genome-wide contribution of genotype by environment interaction to variation of diabetes-related traits. PLoS One. 2013;8:e77442.
49. Sandhu MS, Waterworth DM, Debenham SL, Wheeler E, Papadakis K, Zhao JH, et al. LDL-cholesterol concentrations: a genome-wide association study. Lancet. 2008;371:483–91.
50. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013;45:1274–83.
51. Pickrell JK, Berisa T, Liu JZ, Segurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet. 2016;48:709–17.
52. Lambert JC, Grenier-Boley B, Harold D, Zelenika D, Chouraki V, Kamatani Y, et al. Genome-wide haplotype association study identifies the FRMD4A gene as a risk locus for Alzheimer's disease. Mol Psychiatry. 2013;18:461–70.
53. Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511:421–7.
54. Wang Q, Xi B, Liu M, Zhang Y, Fu M. Short sleep duration is associated with hypertension risk among adults: a systematic review and meta-analysis. Hypertens Res. 2012;35:1012–8.
55. Baron KG, Reid KJ. Circadian misalignment and health. Int Rev Psychiatry. 2014;26:139–54.
56. Douma LG, Gumz ML. Circadian clock-mediated regulation of blood pressure. Free Radic Biol Med. 2018;119:108–14.
57. Nikolaeva S, Pradervand S, Centeno G, Zavadova V, Tokonami N, Maillard M, et al. The circadian clock modulates renal sodium handling. J Am Soc Nephrol. 2012;23:1019–26.
58. Paillasse MR, de Medina P. The NR4A nuclear receptors as potential targets for anti-aging interventions. Med Hypotheses. 2015;84:135–40.
59. Xi B, He D, Zhang M, Xue J, Zhou D. Short sleep duration predicts risk of metabolic syndrome: a systematic review and meta-analysis. Sleep Med Rev. 2014;18:293–7.
60. Chambers BE, Clark EG, Gatz AE, Wingert RA. Kctd15 regulates nephron segment development by repressing Tfap2a activity. Development. 2020;147:dev191973.
61. Adeva-Andany MM, Perez-Felpete N, Fernandez-Fernandez C, Donapetry-Garcia C, Pazos-Garcia C. Liver glucose metabolism in humans. Biosci Rep. 2016;36:e00416.
62. Cascone T, McKenzie JA, Mbofung RM, Punt S, Wang Z, Xu C, et al. Increased Tumor Glycolysis Characterizes Immune Resistance to Adoptive T Cell Therapy. Cell Metab. 2018;27:977–87. e974
63. Segovia J, Zarco N. Gas1 is a pleiotropic regulator of cellular functions: from embryonic development to molecular actions in cancer gene therapy. Mini Rev Med Chem. 2014;14:1139–47.
64. Zarco N, Bautista E, Cuellar M, Vergara P, Flores-Rodriguez P, Aguilar-Roblero R, et al. Growth arrest specific 1 (GAS1) is abundantly expressed in the adult mouse central nervous system. J Histochem Cytochem. 2013;61:731–48.
65. Jones SE, Lane JM, Wood AR, van Hees VT, Tyrrell J, Beaumont RN, et al. Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. Nat Commun. 2019;10:343.
66. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, et al. Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association. Circulation. 2019;139:e56–e528.
67. Nunes J, Jean-Louis G, Zizi F, Casimir GJ, von Gizycki H, Brown CD, et al. Sleep duration among black and white Americans: results of the National Health Interview Survey. J Natl Med Assoc. 2008;100:317–22.
68. Hale L, Do DP. Racial differences in self-reports of sleep duration in a population-based study. Sleep. 2007;30:1096–103.
69. Barfield R, Wang H, Liu Y, Brody JA, Swenson B, Li R, et al. Epigenome-wide association analysis of daytime sleepiness in the Multi-Ethnic Study of Atherosclerosis reveals African-American-specific associations. Sleep. 2019;42:zsz101.
70. Lindhorst J, Alexander N, Blignaut J, Rayner B. Differences in hypertension between blacks and whites: an overview. Cardiovasc J Afr. 2007;18:241–7.
71. Delto CF, Heisler FF, Kuper J, Sander B, Kneussel M, Schindelin H. The LisH motif of muskelin is crucial for oligomerization and governs intracellular localization. Structure. 2015;23:364–73.
72. Heisler FF, Loebrich S, Pechmann Y, Maier N, Zivkovic AR, Tokito M, et al. Muskelin regulates actin filament- and microtubule-based GABA(A) receptor transport in neurons. Neuron. 2011;70:66–81.

73. Ogawa Y, Kakumoto K, Yoshida T, Kuwako KI, Miyazaki T, Yamaguchi J, et al. Elavl3 is essential for the maintenance of Purkinje neuron axons. Sci Rep. 2018;8:2722.
74. Jackson CL, Patel SR, Jackson WB, 2nd, Lutsey PL, Redline S. Agreement between self-reported and objectively measured sleep duration among white, black, Hispanic, and Chinese adults in the United States: multi-ethnic study of atherosclerosis. Sleep. 2018; 41:zsy057.
75. Watson NF, Badr MS, Belenky G, Bliwise DL, Buxton OM, Buysse D, et al. Recommended amount of sleep for a healthy adult: a joint consensus statement of the american academy of sleep medicine and sleep research society. Sleep. 2015;38:843–4.

Affiliations

Heming Wang[1,2] · Raymond Noordam[3] · Brian E. Cade[1,2] · Karen Schwander[4,5] · Thomas W. Winkler[6] · Jiwon Lee[1,7,8,9] · Yun Ju Sung[4] · Amy R. Bentley[10] · Alisa K. Manning[2,11] · Hugues Aschard[12,13] · Tuomas O. Kilpeläinen[14,15] · Marjan Ilkov[16] · Michael R. Brown[17] · Andrea R. Horimoto[18] · Melissa Richard[19] · Traci M. Bartz[20,21] · Dina Vojinovic[22,23] · Elise Lim[24] · Jovia L. Nierenberg[25] · Yongmei Liu[26] · Kumaraswamynaidu Chitrala[27] · Tuomo Rankinen[28] · Solomon K. Musani[29] · Nora Franceschini[30] · Rainer Rauramaa[31] · Maris Alver[32,33] · Phyllis C. Zee[34] · Sarah E. Harris[35] · Peter J. van der Most[36] · Ilja M. Nolte[36] · Patricia B. Munroe[37,38] · Nicholette D. Palmer[39] · Brigitte Kühnel[40,41] · Stefan Weiss[42,43] · Wanqing Wen[44] · Kelly A. Hall[45] · Leo-Pekka Lyytikäinen[46,47] · Jeff O'Connell[48,49] · Gudny Eiriksdottir[16] · Lenore J. Launer[27] · Paul S. de Vries[17] · Dan E. Arking[50] · Han Chen[17,51] · Eric Boerwinkle[17,52] · Jose E. Krieger[18] · Pamela J. Schreiner[53] · Stephen Sidney[54] · James M. Shikany[55] · Kenneth Rice[21] · Yii-Der Ida Chen[56] · Sina A. Gharib[57] · Joshua C. Bis[20] · Annemarie I. Luik[22] · M. Arfan Ikram[22,58] · André G. Uitterlinden[22,59] · Najaf Amin[22] · Hanfei Xu[24] · Daniel Levy[24,60] · Jiang He[25] · Kurt K. Lohman[26] · Alan B. Zonderman[27] · Treva K. Rice[4] · Mario Sims[29] · Gregory Wilson[61] · Tamar Sofer[1,2] · Stephen S. Rich[62] · Walter Palmas[63] · Jie Yao[56] · Xiuqing Guo[56] · Jerome I. Rotter[56] · Nienke R. Biermasz[64] · Dennis O. Mook-Kanamori[65,66] · Lisa W. Martin[67] · Ana Barac[68] · Robert B. Wallace[69] · Daniel J. Gottlieb[1,70] · Pirjo Komulainen[31] · Sami Heikkinen[71,72] · Reedik Mägi[32] · Lili Milani[32] · Andres Metspalu[32] · John M. Starr[73] · Yuri Milaneschi[74] · R. J. Waken[75] · Chuan Gao[76] · Melanie Waldenberger[40,41] · Annette Peters[41,77] · Konstantin Strauch[78,79,80] · Thomas Meitinger[81] · Till Roenneberg[82] · Uwe Völker[42,43] · Marcus Dörr[43,83] · Xiao-Ou Shu[44] · Sutapa Mukherjee[84,85] · David R. Hillman[86] · Mika Kähönen[87,88] · Lynne E. Wagenknecht[89] · Christian Gieger[40,90] · Hans J. Grabe[91] · Wei Zheng[44] · Lyle J. Palmer[45] · Terho Lehtimäki[46,47] · Vilmundur Gudnason[16,92] · Alanna C. Morrison[17] · Alexandre C. Pereira[18,93] · Myriam Fornage[17,19] · Bruce M. Psaty[20,94] · Cornelia M. van Duijn[22,95] · Ching-Ti Liu[24] · Tanika N. Kelly[25] · Michele K. Evans[27] · Claude Bouchard[28] · Ervin R. Fox[29] · Charles Kooperberg[96] · Xiaofeng Zhu[97] · Timo A. Lakka[31,71,98] · Tõnu Esko[32] · Kari E. North[30] · Ian J. Deary[35] · Harold Snieder[36] · Brenda W. J. H. Penninx[74] · W. James Gauderman[99] · Dabeeru C. Rao[4] · Susan Redline[1,100] · Diana van Heemst[3]

[1] Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

[2] Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA

[3] Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

[4] Division of Biostatistics, Washington University School of Medicine, St. Louis, MO, USA

[5] Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA

[6] Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany

[7] Joint Carnegie Mellon University-University of Pittsburgh PhD Program in Computational Biology, Pittsburgh, PA, USA

[8] Department of Computational and Systems Biology, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA

[9] Pittsburgh Center for Evolutionary Biology and Medicine, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA

[10] Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

[11] Clinical and Translational Epidemiology Unit, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

[12] Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA

[13] Centre de Bioinformatique, Biostatistique et Biologie Intégrative (C3BI), Institut Pasteur, Paris, France

[14] Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[15] Department of Environmental Medicine and Public Health, The Icahn School of Medicine at Mount Sinai, New York, NY, USA

[16] Icelandic Heart Association, Kopavogur, Iceland

[17] Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA

[18] Laboratory of Genetics and Molecular Cardiology, Heart Institute (InCor), University of São Paulo Medical School, São Paulo, Brazil

[19] Brown Foundation Institute of Molecular Medicine, McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX, USA

[20] Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA

[21] Department of Biostatistics, University of Washington, Seattle, WA, USA

[22] Department of Epidemiology, Erasmus MC, University Medical Center, Rotterdam, The Netherlands

[23] Molecular Epidemiology, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands

[24] Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

[25] Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA

[26] Division of Cardiology, Department of Medicine, Duke Molecular Physiology Institute Duke University School of Medicine, Durham, NC, USA

[27] Laboratory of Epidemiology and Population Sciences, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA

[28] Human Genomics Laboratory, Pennington Biomedical Research Center, Baton Rouge, LA, USA

[29] Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA

[30] Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

[31] Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio, Finland

[32] Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia

[33] Department of Genetic Medicine and Development, University of Geneva, Geneva, Switzerland

[34] Division of Sleep Medicine, Department of Neurology, Northwestern University, Chicago, IL, USA

[35] Lothian Birth Cohorts group, Department of Psychology, University of Edinburgh, Edinburgh, UK

[36] Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

[37] Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

[38] National Institute for Health Research Barts Cardiovascular Biomedical Research Unit, Queen Mary University of London, London, London, UK

[39] Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC, USA

[40] Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[41] Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[42] Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, University Medicine Greifswald, Greifswald, Germany

[43] German Center for Cardiovascular Research (DZHK), partner site Greifswald, Greifswald, Germany

[44] Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA

[45] School of Public Health, The University of Adelaide, Adelaide, SA, Australia

[46] Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland

[47] Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

[48] Division of Endocrinology, Diabetes, and Nutrition, University of Maryland School of Medicine, Baltimore, MD, USA

[49] Program for Personalized and Genomic Medicine, University of Maryland School of Medicine, Baltimore, MD, USA

[50] McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[51] Center for Precision Health, School of Public Health & School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, USA

[52] Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA

[53] Division of Epidemiology & Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA

[54] Kaiser Permanente Northern California, Oakland, CA, USA

[55] Division of Preventive Medicine, School of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA

[56] The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA, USA

[57] Computational Medicine Core, Center for Lung Biology, UW Medicine Sleep Center, Department of Medicine, University of Washington, Seattle, WA, USA

[58] Department of Radiology and Nuclear Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands

[59] Department of Internal Medicine, Erasmus MC, University Medical Center, Rotterdam, the Netherlands

[60] Population Sciences Branch, National Heart, Lung, and Blood Institute Framingham Heart Study, Framingham, MA, USA

[61] JHS Graduate Training and Education Center, Jackson State University, Jackson, MS, USA

[62] Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

[63] Division of General Medicine, Department of Medicine, Columbia University, New York, NY, USA

[64] Division of Endocrinology, Department of Internal Medicine, Leiden University Medical Center, Leiden, The Netherlands

[65] Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, Netherlands

[66] Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, The Netherlands

[67] George Washington University School of Medicine and Health Sciences, Washington, DC, USA

[68] MedStar Heart and Vascular Institute, Washington, DC, USA

[69] Department of Epidemiology, University of Iowa College of Public Health, Iowa City, IA, USA

[70] VA Boston Healthcare System, Boston, MA, USA

[71] Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio Campus, Finland

[72] Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, Kuopio, Finland

[73] Alzheimer Scotland Dementia Research Centre, The University of Edinburgh, Edinburgh, UK

[74] Department of Psychiatry, Amsterdam Neuroscience and Amsterdam Public Health Research Institute, Amsterdam UMC, Vrije Universiteit, Amsterdam, HJ, The Netherlands

[75] Division of Cardiology, Department of Medicine, Washington University in St. Louis, St. Louis, MO, USA

[76] Molecular Genetics and Genomics Program, Wake Forest School of Medicine, Winston-Salem, NC, USA

[77] German Centre for Cardiovascular Research (DZHK), partner site Munich Heart Alliance, Neuherberg, Germany

[78] Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center, Johannes Gutenberg University, Mainz, Germany

[79] Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[80] Chair of Genetic Epidemiology, IBE, Faculty of Medicine, LMU Munich, Munich, Germany

[81] Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[82] Institute and Polyclinic for Occupational-, Social- and Environmental Medicine, LMU Munich, Munich, Germany

[83] Department of Internal Medicine B, University Medicine Greifswald, Greifswald, Germany

[84] Sleep Health Service, Respiratory and Sleep Services, Southern Adelaide Local Health Network, Adelaide, SA, Australia

[85] Adelaide Institute for Sleep Health, Flinders Health and Medical Research Institute, College of Medicine and Public Health, Flinders University, Adelaide, SA, Australia

[86] Department of Pulmonary Physiology and Sleep Medicine, Sir Charles Gairdner Hospital, Perth, WA, Australia

[87] Department of Clinical Physiology, Tampere University Hospital, Tampere, Finland

[88] Department of Clinical Physiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

[89] Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA

[90] German Center for Diabetes Research (DZD e.V.), Neuherberg, Germany

[91] Department Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany

[92] Faculty of Medicine, University of Iceland, Reykjavik, Iceland

[93] Department of Genetics, Harvard Medical School, Boston, MA, USA

[94] Cardiovascular Health Research Unit, Departments of Epidemiology and Health Services, University of Washington, Seattle, WA, USA

[95] Nuffield Department of Population Health, University of Oxford, Oxford, UK

[96] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[97] Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA

[98] Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland

[99] Division of Biostatistics, Department of Preventive Medicine, University of Southern California, Los Angeles, CA, USA

[100] Division of Pulmonary Medicine, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA

Supplementary Methods

Multi-ancestry genome-wide gene-sleep interactions identify novel loci for blood pressure

Wang H, et al.

This study included 30 cohorts comprising of five ancestry groups including European (EUR), African (AFR), Asian (ASN), Hispanic (HIS), and Brazilian (BRZ) groups. Gene-sleep interaction analyses were restricted to men and women between 18-80 years old and with available sleep, blood pressure (BP), and genotype data imputed to 1000 genome reference panel. Analyses of each ancestry group were performed locally by individual cohorts following a universal analytic protocol. Descriptions of each cohort are available in Supplementary Notes.

Phenotypes, lifestyle exposures, and covariates

We focused on two lifestyle variables (short total sleep time [STST] and long total sleep time [LTST]) and investigated their interaction with genetic components on four BP traits (systolic BP [SBP], diastolic BP [DBP], mean artery pressure [MAP], and pulse pressure [PP]).

Resting/sitting SBP and DBP were recorded in mmHg and averaged among multiple readings at the same visit. Effects of anti-hypertensive (BP lowering) medications were addressed by adding 15 mmHg and 10mmHg to SBP and DBP, respectively. We then derived MAP and PP from the medication adjusted SBP and DBP as MAP=DBP + (SBP-DBP)/3 and PP=SBP-DBP. For each of the four BP variables, we winsorized the very extreme BP values to 6 standard deviations (SD)[1].

Total sleep time (TST) was harmonized from questions similar to "On an average day, how long do you sleep?". Questions were either asked as "open (free text)" or "multiple choice". Individuals reported shorter than 3 or longer than 14 hours of TST were further excluded in this study. For most of the cohorts, we performed linear regression on TST adjusting for age and sex in each ancestry group of each cohort and obtained age- and sex-adjusted residuals. Short total sleep time (STST) and long total sleep time (LTST) were then defined as the lowest 20^{th} and highest 80^{th} percentiles of the residual in each ancestry of each cohort. In some cases, cohort-specific definitions (e.g. "multiple choice") were used to defined STST and LTST.

We adjusted for age, sex, and when appropriate for field center, family and population structures, and other study specific covariates (e.g., sampling weights and census block in HCHS/SOL[2]). Body mass index (BMI; calculated as weight / height^2 [kilograms/meters^2]) has been suggested to influence BP and sleep duration[3, 4], and may play a potential role in gene-sleep interactions on BP. Therefore, we performed additional analyses adjusting for BMI.

Genotypes

Each study performed genotyping and imputation locally using various platforms and software as described in Supplementary Tables 3 and 6. The cosmopolitan reference panels from the 1000 Genomes Project Phase I Integrated Release Version 3 Haplotypes (2010-11 data freeze, 2012-03-14 haplotypes) was specified for imputation. Variants with minor allele frequency <1% were excluded by each study.

Upon completion of the analyses by each local institution, all summary data were stored centrally for further processing and meta-analyses. We performed study level quality control (QC) on study specific results using the R package EasyQC[5] (www.genepi-regensburg.de/easyqc). We performed harmonization of alleles, comparison of allele frequencies with ancestry-appropriate 1000 Genomes reference data, and harmonization of all single nucleotide polymorphism (SNP) IDs to a standardized nomenclature according to chromosome and position. We further excluded SNPs with the product of minor allele count (MAC) in the exposed group (LTST/STST=1) and imputation quality (MAC_1×imputation quality) less than 20.

Discovery analysis in stage 1

The discovery analysis includes 40,385 EUR, 15,020 AFR, 2,465 ASN, 4,436 HIS, and 663 BRZ from 16 cohorts (Supplementary Tables 1-3). Each cohort performed genome-wide interaction analyses using the joint effect model

$$Y = \beta_0 + \beta_E E + \beta_G SNP + \beta_{GE} E \times SNP + \beta_C C, \quad (1)$$

where Y is one of the four BP phenotypes, E is STST or LTST, G is the dosage of the imputed genetic variant coded additively, and C is the vector of covariates (with or without adjustment of BMI). Therefore, in total 16 analyses in each ancestry group of each cohort were performed. The number of principal components included in the model was chosen according to cohort-specific preferences (ranging from 0 to 10). Various software such as ProbABEL[6], MMAP and R package sandwich[7] were used to estimate the association effects for each study (Supplementary Table 3). Summary statistics of the estimated genetic main effect (β_G), interaction effect (β_{GE}), and a robust estimation of the covariance between β_G and β_{GE} were provided by each cohort. We plotted the summary statistics of all effect estimates, standard errors (SE) and p-values across studies to visually compare discrepancies. We also generated SE vs sample and QQ plots to identify analytical problems.

We performed inverse-variance weighted meta-analysis for the 1 degree of freedom (df) interaction term, and 2 df joint fixed-effects meta-analysis of the combined main and interaction effects using Manning *et al.*'s method implemented in the METAL software[8]. Genomic control correlation was performed before and after meta-analyses. Variants present in less than 3 cohorts or 5000 individuals were further excluded. Ancestry-specific meta-analyses were performed with the same QC criteria in EUR and AFR but not in other groups. Variants with 2df $P_{joint} < 10^{-6}$ were followed up in phase 2 replication analysis.

Replication analysis in stage 2 cohort

Stage 2 includes 36,695 EUR, 6,981 AFR, 3,287 ASN, and 12,333 HIS individuals from 14 independent studies (Supplementary Tables 4-6). Seven cohorts performed genome-wide interaction analyses and other cohorts performed focused variant lookup upon the convenience of

each analysis team (Supplementary Table 6). Most of the cohorts performed the same joint analysis model (1) as described in the stage 1 analyses, while stratified association analyses of main effects (2) among exposed group (E=1) and unexposed group (E=0) were performed in the UKB AFR individuals.

$$Y = \gamma_0 + \gamma_G \, SNP + \gamma_C \, C, \quad (2)$$

We then estimated the main and interaction effects for the joint regression model as $\beta_G = \gamma_{G|E=0}$ and $\beta_{GE} = \gamma_{G|E=1} - \gamma_{G|E=0}$ using EasyStrata [9]. Study level QC was performed similarly as described for stage 1 cohorts. Meta-level QC and genomic control were not performed in stage 2 meta-analyses. Stage 1 and stage 2 summary statistics were also meta-analyzed. The replication threshold was then defined as stage 2 $P_{joint} < 0.05$ and combined stage 1 + 2 $P_{joint} < 5\times10^{-8}$, with consistent directions of association effects.

To maximize the statistical power, we also performed genome-wide combined stage 1 and 2 meta-analyses using all stage 1 cohorts and seven stage 2 cohorts (Supplementary Table 6) with genome-wide results in multi-ancestry and EUR groups. Genomic control correlation was performed before and after meta-analyses. Significant loci were identified at a stricter threshold ($P_{joint} < 3.125\times10^{-9}$), accounting for 2 independent BP traits, 2 exposures, with and without BMI adjustment, in two groups.

Effects on BP among long, normal, and short sleepers

We investigate the interaction effect with sleep for the significant novel ($r^2<0.1$ and >1Mb from any previously identified BP locus) and known BP loci (≤1Mb) using 1df test (P_{int}). Loci significantly interact with sleep duration were identified with stage 1+2 $P_{int} < 10^{-3}$ accounting for the number of independent loci.

We compared their risk effects on BP among short, normal, and long sleepers (20%, 60%, and 20% of the sample, respectively) without performing additional stratified analyses. We first estimated the main effects in individuals with STST ($\gamma_{G|STST}$), LTST ($\gamma_{G|LTST}$) and their opposite groups ($\gamma_{G|LTST+normal}$ and $\gamma_{G|STST+normal}$) from β_G and β_{GE} in the joint regression model (1) in a reverse approach described for UKB AFR analysis. We then estimate the effect in normal sleepers ($\gamma_{G|normal}$) weighted by their proportion in our sample. I.e., $60\% \times \gamma_{G|normal} = 80\% \times \gamma_{G|LTST+normal} - 20\% \times \gamma_{G|LTST} = 80\% \times \gamma_{G|STST+normal} - 20\% \times \gamma_{G|STST}$.

Bioinformatics analyses

We pruned independently associated loci using the PLINK prune function with $r^2 \geq 0.1$ from the lead variant. Loci with physical distance >1MB from any known BP loci were considered as novel. We annotated functional effects for novel loci using HaploReg[10], Regulome[11], and GTex (v8)[12] database. In addition to genes mapped by PLINK pruning, we also annotated the genes under the association regions using SNPsea[13] with $r^2 \geq 0.5$ from the lead variant using 1000 genome European

reference panel. Genes under the association regions were further interrogated for associated phenotypes, Mendelian diseases, and druggable targets using PheGeni[14], OMIM[15], and DGIdb[16]. Tissue and pathway enrichment analyses were performed using online software FUMA[17] MAGMA[18] and MsigDB[19].

References

1. Rao DC, Sung YJ, Winkler TW, Schwander K, Borecki I, Cupples LA *et al.* Multiancestry Study of Gene-Lifestyle Interactions for Cardiovascular Traits in 610 475 Individuals From 124 Cohorts: Design and Rationale. *Circ Cardiovasc Genet* 2017; **10**(3).

2. Lavange LM, Kalsbeek WD, Sorlie PD, Aviles-Santa LM, Kaplan RC, Barnhart J *et al.* Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol* 2010; **20**(8)**:** 642-649.

3. Grandner MA, Schopfer EA, Sands-Lincoln M, Jackson N, Malhotra A. Relationship between sleep duration and body mass index depends on age. *Obesity (Silver Spring)* 2015; **23**(12)**:** 2491-2498.

4. Martins D, Tareen N, Pan D, Norris K. The relationship between body mass index, blood pressure and pulse rate among normotensive and hypertensive participants in the third National Health and Nutrition Examination Survey (NHANES). *Cell Mol Biol (Noisy-le-grand)* 2003; **49**(8)**:** 1305-1309.

5. Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Magi R *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* 2014; **9**(5)**:** 1192-1212.

6. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* 2010; **11:** 134.

7. Zeileis A. Object-oriented computation of sandwich estimators. 2006.

8. Manning AK, LaValley M, Liu CT, Rice K, An P, Liu Y *et al.* Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP x environment regression coefficients. *Genet Epidemiol* 2011; **35**(1)**:** 11-18.

9. Winkler TW, Kutalik Z, Gorski M, Lottaz C, Kronenberg F, Heid IM. EasyStrata: evaluation and visualization of stratified genome-wide association meta-analysis data. *Bioinformatics* 2015; **31**(2)**:** 259-261.

10. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012; **40**(Database issue)**:** D930-934.

11. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; **22**(9)**:** 1790-1797.

12. Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015; **348**(6235)**:** 648-660.

13. Slowikowski K, Hu X, Raychaudhuri S. SNPsea: an algorithm to identify cell types, tissues and pathways affected by risk loci. *Bioinformatics* 2014; **30**(17)**:** 2496-2497.

14. Ramos EM, Hoffman D, Junkins HA, Maglott D, Phan L, Sherry ST *et al.* Phenotype–Genotype Integrator (PheGenI): synthesizing genome-wide association study (GWAS) data with existing genomic resources. *European Journal of Human Genetics* 2014; **22**(1)**:** 144-147.

15. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* 2005; **33**(Database issue)**:** D514-517.

16. Cotto KC, Wagner AH, Feng YY, Kiwala S, Coffman AC, Spies G *et al.* DGIdb 3.0: a redesign and expansion of the drug-gene interaction database. *Nucleic Acids Res* 2018; **46**(D1)**:** D1068-D1073.

17. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017; **8**(1)**:** 1826.

18. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* 2015; **11**(4)**:** e1004219.

19. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 2011; **27**(12)**:** 1739-1740.

Supplementary Notes

Multi-ancestry genome-wide gene-sleep interactions identify novel loci for blood pressure

Wang H, et al.

Table of Contents

Stage 1 Study Descriptions

Brief descriptions are provided below for each of the discovery studies some of which are based outside the United States:

AGES (Age Gene/Environment Susceptibility Reykjavik Study): The AGES Reykjavik study originally comprised a random sample of 30,795 men and women born in 1907-1935 and living in Reykjavik in 1967. A total of 19,381 people attended, resulting in a 71% recruitment rate. The study sample was divided into six groups by birth year and birth date within month. One group was designated for longitudinal follow up and was examined in all stages; another was designated as a control group and was not included in examinations until 1991. Other groups were invited to participate in specific stages of the study. Between 2002 and 2006, the AGES Reykjavik study re-examined 5,764 survivors of the original cohort who had participated before in the Reykjavik Study. The midlife data blood pressure measurement was taken from stage 3 of the Reykjavik Study (1974-1979), if available. Half of the cohort attended during this period. Otherwise an observation was selected closest in time to the stage 3 visit. The supine blood pressure was measured twice by a nurse using a mercury sphygmomanometer after 5 minutes rest following World Health Organization recommendations.

ARIC (Atherosclerosis Risk in Communities): The ARIC study is a population-based prospective cohort study of cardiovascular disease sponsored by the National Heart, Lung, and Blood Institute (NHLBI). ARIC included 15,792 individuals, predominantly European American and African American, aged 45-64 years at baseline (1987-89), chosen by probability sampling from four US communities. Cohort members completed three additional triennial follow-up examinations, a fifth exam in 2011-2013, a sixth exam in 2016-2017, and a seventh exam in 2018-2019. The ARIC study has been described in detail previously (The ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) study: Design and objectives. Am J Epidemiol. 1989;129:687-702). Blood pressure was measured during the fifth exam using a standardized Hawksley random-zero mercury column sphygmomanometer with participants in a sitting position after a resting period of 5 minutes. The size of the cuff was chosen according to the arm circumference. Three sequential recordings for systolic and diastolic blood pressure were obtained; the mean of the last two measurements was used in this analysis, discarding the first reading. Blood pressure lowering medication use was recorded from the medication history. Sleep duration variables were also determined during the fifth exam.

Baependi Heart Study (Brazil): The Baependi Heart Study, is an ongoing family-based cohort conducted in a rural town of the state of Minas Gerais. The study has enrolled approximate 2,200 individuals (over 10% of the town's adult population) and 10-year follow up period of longitudinal data. Briefly, probands were selected at random across 11 out of the 12 census districts in Baependi. After enrolment, the proband's first-degree (parents, siblings, and offspring), second-degree (half-siblings, grandparents/grandchildren, uncles/aunts, nephews/nieces, and double cousins), and third-degree (first cousins, great uncles/aunts, and great nephews/nieces) relatives, and his/her respective spouse's relatives resident both within Baependi (municipal and rural area) and surrounding towns were invited to participate. Only individuals age 18 and older were eligible to participate in the study. The study is conducted from a clinic/office in an easily accessible sector

of the town, where the questionnaires were completed. A broad range of phenotypes ranging from cardiovascular, neurocognitive, psychiatric, imaging, physiologic and several layers of endophenotypes like metabolomics and lipidomics have been collected throughout the years Details about follow-up visits and available data can be found in the cohort profile paper (PMID: 18430212). DNA samples were genotyped using the Affymetrix 6.0 genechip. After quality control, the data were prephased using SHAPEIT and imputed using IMPUTE2 based on 1000 Genomes haplotypes.

CARDIA (Coronary Artery Risk Development in Young Adults): CARDIA is an observational, prospective, multicenter study with 5,115 Black and White participants, 18-30 years old at baseline, recruited from four centers at the baseline examination in 1985-1986. The recruitment was done from the total community in Birmingham, AL; from selected census tracts in Chicago, IL and Minneapolis, MN; and from the Kaiser Permanente health plan membership in Oakland, CA. The details of the study design for the CARDIA study have been previously published[1]. Nine examinations have been completed since initiation of the study, respectively in the years 0, 2, 5, 7, 10, 15, 20, 25 and 30. Written informed consent was obtained from participants at each examination, and all study protocols were approved by the institutional review boards of the participating institutions. Systolic and diastolic blood pressure was measured in triplicate on the right arm using a random-zero sphygmomanometer through year 15 and an automated blood pressure monitor (Omron) at years 20, 25, and 30, with the participant seated and following a 5-min. rest. The average of the second and third measurements was taken as the blood pressure value. Blood pressure medication use was obtained by questionnaire.

1. Friedman GD, Cutter GR, Donahue RP, Hughes GH, Hully SB, Jacobs DR Jr., Liu K, Savage PJ. CARDIA: study design, recruitment, and some characteristics of the examined subjects. J Clin Epidemiol. 1988;41:1105–1116.

CHS (Cardiovascular Health Study): CHS is a population-based cohort study of risk factors for cardiovascular disease in adults 65 years of age or older conducted across four field centers[1]. The original predominantly European ancestry cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists and an additional predominately African-American cohort of 687 persons was enrolled in 1992-93 for a total sample of 5,888. Research staff with central training in blood pressure measurement assessed repeated right-arm seated systolic and diastolic blood pressure levels at baseline with a Hawksley random-zero sphygmomanometer. Blood samples were drawn from all participants at their baseline examination and DNA was subsequently extracted from available samples. European ancestry participants were excluded from the GWAS study sample due to prevalent coronary heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke, or transient ischemic attack at baseline. After QC, genotyping was successful for 3271 European ancestry and 823 African-American participants. CHS was approved by institutional review committees at each site and individuals in the present analysis gave informed consent including consent to use of genetic information for the study of cardiovascular disease.

1. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, et al. The Cardiovascular Health Study: design and rationale. Ann Epidemiol 1991; 1:263-76.

ERF (Erasmus Rucphen Family study): Erasmus Rucphen Family is a family based study that includes inhabitants of a genetically isolated community in the South-West of the Netherlands,

studied as part of the Genetic Research in Isolated Population (GRIP) program[1,2]. The goal of the study is to identify the risk factors in the development of complex disorders. Study population includes approximately 3,000 individuals who are living descendants of 22 couples who lived in the isolate between 1850 and 1900 and had at least six children baptized in the community church. All data were collected between 2002 and 2005. All participants gave informed consent, and the Medical Ethics Committee of the Erasmus University Medical Centre approved the study.

1. Pardo, Luba M., et al. "The effect of genetic drift in a young genetically isolated population." Annals of human genetics 69.3 (2005): 288-295.
2. Aulchenko, Yurii S., et al. "Linkage disequilibrium in young genetically isolated Dutch population." European Journal of Human Genetics 12.7 (2004): 527.

FHS (Framingham Heart Study): FHS began in 1948 with the recruitment of an original cohort of 5,209 men and women (mean age 44 years; 55 percent women). In 1971 a second generation of study participants was enrolled; this cohort (mean age 37 years; 52% women) consisted of 5,124 children and spouses of children of the original cohort. A third generation cohort of 4,095 children of offspring cohort participants (mean age 40 years; 53 percent women) was enrolled in 2002-2005 and are seen every 4 to 8 years. Details of study designs for the three cohorts are summarized elsewhere. At each clinic visit, a medical history was obtained with a focus on cardiovascular content, and participants underwent a physical examination including measurement of height and weight from which BMI was calculated. Systolic and diastolic blood pressures were measured twice by a physician on the left arm of the resting and seated participant using a mercury column sphygmomanometer. Blood pressures were recorded to the nearest even number. The means of two separate systolic and diastolic blood pressure readings at each clinic examination were used for statistical analyses.

GenSalt (Genetic Epidemiology Network of Salt Sensitivity): GenSalt is a multi-center, family based study designed to identify, through dietary sodium and potassium intervention, salt-sensitivity susceptibility genes which may underlie essential hypertension in rural Han Chinese families. Approximately 629 families with at least one ‘proband’ with high blood pressure were recruited and tested for a wide variety of physiological, metabolic and biochemical measures at baseline and at multiple times during the 3-week intervention. The intervention consisted of one week on a low sodium diet, followed by one week on a high sodium diet, and finally one week on a high sodium diet with a potassium supplement.

HANDLS (Healthy Aging in Neighborhoods of Diversity across the Life Span): HANDLS is a community-based, longitudinal epidemiologic study examining the influences of race and socioeconomic status (SES) on the development of age-related health disparities among a sample of socioeconomically diverse African Americans and whites. This unique study will assess over a 20-year period physical parameters and also evaluate genetic, biologic, demographic, and psychosocial, parameters of African American and white participants in higher and lower SES to understand the driving factors behind persistent black-white health disparities in overall longevity, cardiovascular disease, and cognitive decline. The study recruited 3,720 participants from Baltimore, MD with a mean age of 47.7 years, 2,198 African Americans and 1,522 whites, with 41% reporting household incomes below the 125% poverty delimiter.

Genotyping was done on a subset of self-reporting African American participants by the Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health (NIH). A larger genotyping effort included a small subset of self-reporting European ancestry samples. This research was supported by the Intramural Research Program of the NIH, NIA.

Health ABC (Health, Aging, and Body Composition): Cohort description: The Health ABC study is a prospective cohort study investigating the associations between body composition, weight-related health conditions, and incident functional limitation in older adults. Health ABC enrolled well-functioning, community-dwelling black (n=1281) and white (n=1794) men and women aged 70-79 years between April 1997 and June 1998. Participants were recruited from a random sample of white and all black Medicare eligible residents in the Pittsburgh, PA, and Memphis, TN, metropolitan areas. Participants have undergone annual exams and semi-annual phone interviews. The current study sample consists of 1559 white participants who attended the second exam in 1998-1999 with available genotyping data.
Genotyping: Genotyping was performed by the Center for Inherited Disease Research (CIDR) using the Illumina Human1M-Duo BeadChip system. Samples were excluded from the dataset for the reasons of sample failure, genotypic sex mismatch, and first-degree relative of an included individual based on genotype data. Genotyping was successful in 1663 Caucasians. Analysis was restricted to SNPs with minor allele frequency ≥ 1%, call rate ≥97% and HWE p≥10-6. Genotypes were available on 914,263 high quality SNPs for imputation based on the HapMap CEU (release 22, build 36) using the MACH software (version 1.0.16). A total of 2,543,888 imputed SNPs were analyzed for association with vitamin D levels.
Association analysis: Linear regression models were used to generate cohort-specific residuals of naturally log transformed vitamin D levels adjusted for age, sex, BMI and season defined as summer (June-August), fall (September-November), winter (December to February) and spring (March to May) standardized to have mean 0 and variance of 1. Association between the additively coded SNP genotypes and the vitamin D residuals standardized was assessed using linear regression models. For imputed SNPs, expected number of minor alleles (i.e. dosage) was used in assessing association with the vitamin D residuals.

HERITAGE (Health, Risk Factors, Exercise Training and Genetics): The HERITAGE is the only known family-based study of exercise intervention to evaluate the role of genes and sequence variants involved in the response to a physically active lifestyle. The current study is based on the data collected at baseline of the study from 99 White families (244 males, 255 females). All subjects were required to be sedentary and free of chronic diseases at baseline. There are over 18 trait domains (e.g. dietary, lipids and lipoproteins, glucose and insulin metabolism [fasting and IVGTT], steroids, body composition and body fat distribution, cardiorespiratory fitness), for a grand total of over one thousand variables. Moreover, most of the outcome traits were measured twice on two separate days both at baseline and after exercise training was completed. Marker data include a genome-wide linkage scan and GWAS, in addition to a large number of candidate genes.

JHS (Jackson Heart Study): The Jackson Heart Study is a longitudinal, community-based observational cohort study investigating the role of environmental and genetic factors in the development of cardiovascular disease in African Americans. Between 2000 and 2004, a total of 5,306 participants were recruited from a tri-county area (Hinds, Madison, and Rankin Counties) that encompasses Jackson, MS. Details of the design and recruitment for the Jackson Heart Study

cohort has been previously published.1-3 Briefly, approximately 30% of participants were former members of the Atherosclerosis Risk in Communities (ARIC) study. The remainder were recruited by either 1) random selection from the Accudata list, 2) commercial listing, 3) a constrained volunteer sample, in which recruitment was distributed among defined demographic cells in proportions designed to mirror those in the overall population, or through the Jackson Heart Study Family Study.

1. Wyatt SB, Diekelmann N, Henderson F, Andrew ME, Billingsley G, Felder SH et al. A community-driven model of research participation: the Jackson Heart Study Participant Recruitment and Retention Study. Ethn Dis 2003; 13(4):438-455.
2. Taylor HA, Jr., Wilson JG, Jones DW, et al. Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. Ethn Dis 2005; 15:S6-17.
3. Fuqua SR, Wyatt SB, Andrew ME, et al. Recruiting African-American research participation in the Jackson Heart Study: methods, response rates, and sample description. Ethn Dis 2005; 15:S6-29.

MESA (Multi-Ethnic Study of Atherosclerosis): The MESA is a study of the characteristics of subclinical cardiovascular disease and the risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease. MESA consisted of a diverse, population-based sample of an initial 6,814 asymptomatic men and women aged 45-84. 38 percent of the recruited participants were white, 28 percent African American, 22 percent Hispanic, and 12 percent Asian, predominantly of Chinese descent. Participants were recruited from six field centers across the United States: Wake Forest University, Columbia University, Johns Hopkins University, University of Minnesota, Northwestern University and University of California - Los Angeles. Participants are being followed for identification and characterization of cardiovascular disease events, including acute myocardial infarction and other forms of coronary heart disease (CHD), stroke, and congestive heart failure; for cardiovascular disease interventions; and for mortality. The first examination took place over two years, from July 2000 - July 2002. It was followed by five examination periods that were 17-20 months in length. Participants have been contacted every 9 to 12 months throughout the study to assess clinical morbidity and mortality.

1. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR Jr, Kronmal R, Liu K, Nelson JC, O'Leary D, Saad MF, Shea S, Szklo M, Tracy RP. Multi-ethnic study of atherosclerosis: objectives and design. Am J Epidemiol. 2002 Nov 1;156(9):871-81. PubMed PMID: 12397006.

NEO (The Netherlands Epidemiology of Obesity study): The NEO was designed for extensive phenotyping to investigate pathways that lead to obesity-related diseases. The NEO study is a population-based, prospective cohort study that includes 6,671 individuals aged 45–65 years, with an oversampling of individuals with overweight or obesity. At baseline, information on demography, lifestyle, and medical history have been collected by questionnaires. In addition, samples of 24-h urine, fasting and postprandial blood plasma and serum, and DNA were collected. Genotyping was performed using the Illumina HumanCoreExome chip, which was subsequently imputed to the 1000 genome reference panel. Participants underwent an extensive physical examination, including anthropometry, electrocardiography, spirometry, and measurement of the carotid artery intima-media thickness by ultrasonography. In random subsamples of participants, magnetic resonance imaging of abdominal fat, pulse wave velocity of the aorta, heart, and brain,

magnetic resonance spectroscopy of the liver, indirect calorimetry, dual energy X-ray absorptiometry, or accelerometry measurements were performed. The collection of data started in September 2008 and completed at the end of September 2012. Participants are currently being followed for the incidence of obesity-related diseases and mortality.

RS (Rotterdam Study): The Rotterdam Study is a prospective, population-based cohort study among individuals living in the well-defined Ommoord district in the city of Rotterdam in The Netherlands.[1] The aim of the study is to determine the occurrence of cardiovascular, neurological, ophthalmic, endocrine, hepatic, respiratory, and psychiatric diseases in elderly people. The cohort was initially defined in 1990 among 7,983 persons, aged 55 years and older, who underwent a home interview and extensive physical examination at the baseline and during follow-up rounds every 3-4 years (RS-I). Cohort was extended in 2000/2001 (RS-II, 3,011 individuals aged 55 years and older) and 2006/2008 (RS-III, 3,932 subjects, aged 45 and older). Written informed consent was obtained from all participants and the institutional review board (Medical Ethics Committee) of the Erasmus Medical Center and by the review board of The Netherlands Ministry of Health, Welfare and Sports, approved the study.

1. Ikram, M. Arfan, et al. "The Rotterdam Study: 2018 update on objectives, design and main results." European journal of epidemiology 32.9 (2017): 807-850.

WHI (Women's Health Initiative): WHI is a long-term national health study that focuses on strategies for preventing common diseases such as heart disease, cancer and fracture in postmenopausal women. A total of 161,838 women aged 50–79 years old were recruited from 40 clinical centers in the US between 1993 and 1998. WHI consists of an observational study, two clinical trials of postmenopausal hormone therapy (HT, estrogen alone or estrogen plus progestin), a calcium and vitamin D supplement trial, and a dietary modification trial [1]. Study recruitment and exclusion criteria have been described previously [1]. Recruitment was done through mass mailing to age-eligible women obtained from voter registration, driver's license and Health Care Financing Administration or other insurance list, with emphasis on recruitment of minorities and older women [2]. Exclusions included participation in other randomized trials, predicted survival < 3 years, alcoholism, drug dependency, mental illness and dementia. For the CT, women were ineligible if they had a systolic BP > 200 mm Hg or diastolic BP > 105 mm Hg, a history of hypertriglyceridemia or breast cancer. Study protocols and consent forms were approved by the IRB at all participating institutions. Socio-demographic characteristics, lifestyle, medical history and self-reported medications were collected using standardized questionnaires at the screening visit. Physical measures of height, weight and blood pressure were measured at a baseline clinical visit [2]. BP was measured by certified staff using standardized procedures and instruments [3]. Two BP measures were recorded after 5 minutes rest using a mercury sphygmomanometer. Appropriate cuff bladder size was determined at each visit based on arm circumference. Diastolic BP was taken from the phase V Korotkoff measures. The average of the two measurements, obtained 30 seconds apart, was used in analyses. The genome wide association study (GWAS) non-overlapping samples are composed of a case-control study (WHI Genomics and Randomized Trials Network – GARNET, which included all coronary heart disease, stroke, venous thromboembolic events and selected diabetes cases that happened during the active intervention phase in the WHI HT clinical trials and aged matched controls), women selected to be "representative" of the HT trial (mostly

younger white HT subjects that were also enrolled in the WHI memory study - WHIMS) and the WHI SNP Health Association Resource (WHI SHARe), a randomly selected sample of 8,515 African American and 3,642 Hispanic women from WHI. GWAS was performed using Affymetrix 6.0 (WHI-SHARe), HumanOmniExpressExome-8v1_B (WHIMS), Illumina HumanOmni1-Quad v1-0 B (GARNET). Extensive quality control (QC) of the GWAS data included alignment ("flipping") to the same reference panel, imputation to the 1000G data (using the recent reference panel - v3.20101123), identification of genetically related individuals, and computations of principal components (PCs) using methods developed by Price et al. (using EIGENSOFT software 53), and finally the comparison with self-reported ethnicity. After QC and exclusions from analysis protocol, the number of women included in analysis is 4,423 whites for GARNET, 5,202 white for WHIMS, 7,919 for SHARe African American and 3,377 for SHARe Hispanics.

1. Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, Allen C, Rossouw JE. The women's health initiative recruitment methods and results. Ann Epidemiol. 2003;13:S18-77
2. Design of the women's health initiative clinical trial and observational study. The women's health initiative study group. Control Clin Trials. 1998;19:61-109
3. Hsia J, Margolis KL, Eaton CB, Wenger NK, Allison M, Wu L, LaCroix AZ, Black HR. Prehypertension and cardiovascular disease risk in the women's health initiative. Circulation. 2007;115:855-860

Stage 2 Study Descriptions

Brief descriptions are provided below for each of the replication studies/cohorts:

CFS (Cleveland Family Study): The CFS is a family-based, longitudinal study designed to characterize the genetic and non-genetic risk factors for sleep apnea. In total, 2534 individuals (46% African American) from 352 families were studied on up to 4 occasions over a period of 16 years (1990-2006). The initial aim of the study was to quantify the familial aggregation of sleep apnea. 632 African Americans were genotyped on the Affymetrix array 6.0 platform through the CARe Consortium with suitable genotying quality control. A further 122 African-Americans had genotyping based on the Illumina OmniExpress + Exome platform. Genomes were imputed separately for each chip based on a 1000 Genomes Project Phase 3 Version 5 cosmopolitan template using SHAPEIT and IMPUTE2. Participants had three supine BP measurements each performed after lying quietly for 10 minutes, before bed (10:00 P.M.) and upon awakening (7:00 A.M.), and another three sitting at 11 am, following standardized guidelines using a calibrated sphygmomanometer. Cuff size was determined by the circumference of the upper arm and the appropriate bladder size from a standard chart. BP phenotypes were determined from the average of the nine measurements.

DR's EXTRA (Dose-Responses to Exercise Training): The Dose-Responses to Exercise Training (DR's EXTRA) Study is a 4-year RCT on the effects of regular physical exercise and healthy diet on endothelial function, atherosclerosis and cognition in a randomly selected population sample (n=3000) of Eastern Finnish men and women, identified from the national population register, aged 55-74 years. Of the eligible sample, 1410 individuals were randomized into one of the 6 groups: aerobic exercise, resistance exercise, diet, combined aerobic exercise and diet, combined resistance exercise and diet, or reference group following baseline assessments. During the four year intervention the drop-out rate was 15%.

EGCUT (Estonian Genome Center - University of Tartu (Estonian Biobank)): The Estonian Biobank is the population-based biobank of the Estonian Genome Center at the University of Tartu (www.biobank.ee; EGCUT). The entire project is conducted according to the Estonian Gene Research Act and all of the participants have signed the broad informed consent. The cohort size is up to 51535 individuals from 18 years of age and up, which closely reflects the age, sex and geographical distribution of the Estonian population. All of the subjects are recruited randomly by general practitioners and physicians in hospitals. A Computer Assisted Personal interview is filled within 1-2 hours at a doctor's office, which includes personal, genealogical, educational, occupational history and lifestyle data. Anthropometric measurements, blood pressure and resting heart rate are measured and venous blood taken during the visit. Medical history and current health status is recorded according to ICD-10 codes.

HCHS/SOL (Hispanic Community Health Study/ Study of Latinos): The HCHS/SOL is a community-based cohort study of 16,415 self-identified Hispanic/Latino persons aged 18–74 years and selected from households in predefined census-block groups across four US field centers (in Chicago, Miami, the Bronx, and San Diego). The census-block groups were chosen to provide diversity among cohort participants with regard to socioeconomic status and national origin or

background. The HCHS/SOL cohort includes participants who self-identified as having a Hispanic/Latino background; the largest groups are Central American (n = 1,730), Cuban (n = 2,348), Dominican (n = 1,460), Mexican (n = 6,471), Puerto Rican (n = 2,728), and South American (n = 1,068). The HCHS/SOL baseline clinical examination occurred between 2008 and 2011 and included comprehensive biological, behavioral, and sociodemographic assessments. Consenting HCHS/SOL subjects were genotyped at Illumina on the HCHS/SOL custom 15041502 B3 array. The custom array comprised the Illumina Omni 2.5M array (HumanOmni2.5-8v.1-1) ancestry-informative markers, known GWAS hits and drug absorption, distribution, metabolism, and excretion (ADME) markers, and additional custom content including ~150,000 SNPs selected from the CLM (Colombian in Medellin, Colombia), MXL (Mexican Ancestry in Los Angeles, California), and PUR (Puerto Rican in Puerto Rico) samples in the 1000Genomes phase 1 data to capture a greater amount of Amerindian genetic variation. QA/QC procedures yielded a total of 12,803 unique study participants for imputation and downstream association analyses.

IRASFS (Insulin Resistance Atherosclerosis Family Study): The IRASFS was a family study designed to examine the genetic and epidemiologic basis of glucose homeostasis traits and abdominal adiposity. Briefly, self-reported Mexican American pedigrees were recruited in San Antonio, TX and San Luis Valley, CO. Probands with large families were recruited from the initial non-family-based IRAS, which was modestly enriched for impaired glucose tolerance and T2D. Inclusion of IRASFS data is limited to 1040 normoglycemic individuals in 88 pedigrees with genotype data from the Illumina OmniExpress and Omni 1S arrays and imputation to the 1000 Genome Integrated Reference Panel (phase I).

KORA (Cooperative Health Research in the Augsburg Region): The KORA study is a series of independent population-based epidemiological surveys of participants living in the region of Augsburg, Southern Germany. All survey participants are residents of German nationality identified through the registration office and were examined in 1994/95 (KORA S3) and 1999/2001 (KORA S4). In the KORA S3 and S4 studies 4,856 and 4,261 subjects have been examined implying response rates of 75% and 67%, respectively. 3,006 subjects participated in a 10-year follow-up examination of S3 in 2004/05 (KORA F3), and 3080 of S4 in 2006/2008 (KORA F4). The age range of the participants was 25 to 74 years at recruitment. Informed consent has been given by all participants. The study has been approved by the local ethics committee. Individuals for genotyping in KORA F3 and KORA F4 were randomly selected and these genotypes are taken for the analysis of the phenotypes in KORA S3 and KORA S4.

LBC1936 (Lothian Birth Cohort 1936): LBC1936 consists of 1091 (548 male) relatively healthy individuals who underwent cognitive and medical testing at a mean age of 69.6 years (SD = 0.8). They were born in 1936, most took part in the Scottish Mental Survey of 1947, and almost all lived independently in the Lothian region of Scotland[1,2].

1. Deary IJ, Gow AJ, Pattie A, Starr JM. Cohort profile: the Lothian Birth Cohorts of 1921 and 1936. Int J Epidemiol 2012;41:1576-1584.
2. Taylor AM, Pattie A, Deary, IJ Cohort Profile Update: The Lothian Birth Cohorts of 1921 and 1936. Int. J. Epidemiol. 2018 47, 1042-1042.

Lifelines (Netherlands Biobank): Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related

behaviours of 167,729 persons living in the North of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioural, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics. In addition, the Lifelines project comprises a number of cross-sectional sub-studies which investigate specific age-related conditions. These include investigations into metabolic and hormonal diseases, including obesity, cardiovascular and renal diseases, pulmonary diseases and allergy, cognitive function and depression, and musculoskeletal conditions. All survey participants are between 18 and 90 years old at the time of enrollment. Recruitment has been going on since the end of 2006. Baseline data were collected for 167 729 participants, aged from 6 months to 93 years. Follow-up visits are scheduled every 5 years, and in between participants receive follow-up questionnaires. At the baseline examination, the participants in the study were asked to fill in a questionnaire (on paper or online) before the first visit. During the first and second visit, the first or second part of the questionnaire, respectively, are checked for completeness, a number of investigations are conducted, and blood and urine samples are taken. Lifelines is a facility that is open for all researchers. Information on application and data access procedure is summarized on www.lifelines.nl.

1. Scholtens S, Smidt N, Swertz MA, Bakker SJ, Dotinga A, Vonk JM, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. Int J Epidemiol. 2014 Dec 14.

NESDA (Netherlands Study of Depression and Anxiety): NESDA is a multi-center study designed to examine the long-term course and consequences of depressive and anxiety disorders (http://www.nesda.nl)[1]. NESDA included both individuals with depressive and/or anxiety disorders and controls without psychiatric conditions. Inclusion criteria were age 18-65 years and self-reported western European ancestry while exclusion criteria were not being fluent in Dutch and having a primary diagnosis of another psychiatric condition (psychotic disorder, obsessive compulsive disorder, bipolar disorder, or severe substance use disorder).

1. Penninx, B.W. et al. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. Int J Methods Psychiatr Res 17, 121-40 (2008).

SHIP (Study of Health in Pomerania): The Study of Health In Pomerania (SHIP) is a prospective longitudinal population-based cohort study in Mecklenburg-West Pomerania assessing the prevalence and incidence of common diseases and their risk factors (PMID: 20167617). SHIP encompasses the two independent cohorts SHIP and SHIP-TREND. Participants aged 20 to 79 with German citizenship and principal residency in the study area were recruited from a random sample of residents living in the three local cities, 12 towns as well as 17 randomly selected smaller towns. Individuals were randomly selected stratified by age and sex in proportion to population size of the city, town or small towns, respectively. A total of 4,308 participants were recruited between 1997 and 2001 in the SHIP cohort. Between 2008 and 2012 a total of 4,420 participants were recruited in the SHIP-TREND cohort. Individuals were invited to the SHIP study centre for a computer-assisted personal interviews and extensive physical examinations. The study protocol was approved by the medical ethics committee of the University of Greifswald. Oral and written informed consents was obtained from each of the study participants.
Genome-wide SNP-typing was performed using the Affymetrix Genome-Wide Human SNP Array 6.0 or the Illumina Human Omni 2.5 array (SHIP-TREND samples). Array processing was carried out in accordance with the manufacturer's standard recommendations. Genotypes were determined

using GenomeStudio Genotyping Module v1.0 (GenCall) for SHIP-TREND and the Birdseed2 clustering algorithm for SHIP. Imputation of genotypes in SHIP and SHIP-TREND was performed with the software IMPUTE v2.2.2 based on 1000 Genomes release March 2012.

SWHS/SMHS (Shanghai Women's Health Study/ Shanghai Men's Health Study): The Shanghai Women's Health Study (SWHS) is an ongoing population-based cohort study of approximately 75,000 women who were aged 40-70 years at study enrollment and resided in in urban Shanghai, China; 56,832 (75.8%) provided a blood samples. Recruitment for the SWHS was initiated in 1997 and completed in 2000. The self-administered questionnaire includes information on demographic characteristics, disease and surgery histories, personal habits (such as cigarette smoking, alcohol consumption, tea drinking, and ginseng use), menstrual history, residential history, occupational history, and family history of cancer.
The blood pressure were measured by trained interviewers (retired nurses) with a conventional mercury sphygmomanometer according to a standard protocol, after the participants sat quietly for 5 min at the study recruitment. Included in the current project were 2970 women who had GWAS data and blood pressure measurements at the baseline interview.
The Shanghai Men's Health Study (SMHS) is an ongoing population-based cohort study of 61,480 Chinese men who were aged between 40 and 74 years, were free of cancer at enrollment, and lived in urban Shanghai, China; 45,766 (74.4%) provided a blood samples. Recruitment for the SMHS was initiated in 2002 and completed in 2006. The self-administered questionnaire includes information on demographic characteristics, disease and surgery histories, personal habits (such as cigarette smoking, alcohol consumption, tea drinking, and ginseng use), residential history, occupational history, and family history of cancer. The blood pressure were measured by trained interviewers (retired nurses) with a conventional mercury sphygmomanometer according to a standard protocol, after the participants sat quietly for 5 min at the study recruitment. Included in the current project were 892 men who had GWAS data and blood pressure measurements at the baseline interview or 298 men who had GWAS data and lipids data.
Genotyping and imputation: Genomic DNA was extracted from buffy coats by using a Qiagen DNA purification kit (Valencia, CA) or Puregene DNA purification kit (Minneapolis, MN) according to the manufacturers' instructions and then used for genotyping assays. The GWAS genotyping was performed using the Affymetrix Genome-Wide Human SNP Array 6.0 (Affy6.0) platform or Illumina 660, following manufacturers' protocols. After sample quality control, we exclude SNPs with 1) MAF <0.01; 2) call rate <95%; 2) bad genotyping cluster; and 3) concordance rate <95% among duplicated QC samples. Genotypes were imputed using the program MACH (http://www.sph.umich.edu/csg/abecasis/MACH/download/), which determines the probable distribution of missing genotypes conditional on a set of known haplotypes, while simultaneously estimating the fine-scale recombination map. Phased autosome SNP data from HapMap Phase II Asians (release 22) were used as the reference. To test for associations between the imputed SNP data with BMI, linear regression (additive model) was used, in which SNPs were represented by the expected allele count, an approach that takes into account the degree of uncertainty of genotype imputation (http://www.sph.umich.edu/csg/abecasis/MACH/download/).
The lipid profiles were measured at Vanderbilt Lipid Laboratory. Total cholesterol, high-density lipoprotein (HDL) cholesterol, and triglycerides (TG) were measured using an ACE Clinical Chemistry System (Alfa Wassermann, Inc, West Caldwell, NJ). Low-density lipoprotein (LDL) cholesterol levels were calculated by using the Friedwald equation. The levels of LDL cholesterol were directly measured using an ACE Clinical Chemistry System for subjects with TG levels ≥

400 mg/dL. Fasting status was defined as an interval between the last meal and blood draw of 8 hours or longer.

UKB (United Kingdom Biobank, www.ukbiobank.ac.uk): UK Biobank is a major national health resource with the aim of improving the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses. UK Biobank includes data from 502,682 individuals (94% of self-reported European ancestry), with extensive health and lifestyle questionnaire data, physical measures and genetic data. This study includes 6,981 individuals of African ancestries had genetic and phenotypic (blood pressure and sleep) data. Central genotyping quality control (QC) had been performed by UK Biobank [The UK Biobank. UK Biobank Genotyping QC documentation. (2015)]. Further QC was also performed locally.

WASHS (Western Australian Sleep Health Study): The Western Australian Sleep Disorders Research Institute (WASDRI) public sleep clinic provided the sampling frame for the WASHS. Between 2005 and 2010, all consenting new patients having diagnostic overnight polysomnography (PSG) at the WASDRI were serially recruited. PSG data were automatically captured and analysed according to internationally accepted standards. In addition to routine clinical data, all participants completed a detailed questionnaire (sleep behaviour; medical, family, and exposure history), performed spirometry, and provided a blood sample for biochemistry (fasting glucose and insulin, fasting lipids, thyroid function tests, C reactive protein, and fibrinogen levels) and DNA/serum/plasma storage. Weight, height, body mass index, neck circumference, a variety of craniofacial indices, and blood pressure were measured. Complete clinical and epidemiological data were collected on a natural cohort of 4,100 subjects with diagnosed OSA; the participation rate in the study was >98%.

YFS (The Cardiovascular Risk in Young Finns Study): The YFS is a population-based follow up-study started in 1980. The main aim of the YFS is to determine the contribution made by childhood lifestyle, biological and psychological measures to the risk of cardiovascular diseases in adulthood. In 1980, over 3,500 children and adolescents all around Finland participated in the baseline study. The follow-up studies have been conducted mainly with 3-year intervals. The latest 30-year follow-up study was conducted in 2010-11 (ages 33-49 years) with 2,063 participants. The study was approved by the local ethics committees (University Hospitals of Helsinki, Turku, Tampere, Kuopio and Oulu) and was conducted following the guidelines of the Declaration of Helsinki. All participants gave their written informed consent.

Stage 1 Study Acknowledgements

Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant R01HL105756. Infrastructure for the Gene-Lifestyle Working Group is supported by the National Heart, Lung, and Blood Institute grant R01HL118305.

AGES (Age Gene/Environment Susceptibility Reykjavik Study): This study has been funded by NIH contract N01-AG012100, HSSN271201200022C, the NIA Intramural Research Program, an Intramural Research Program Award (ZIAEY000401) from the National Eye Institute, an award from the National Institute on Deafness and Other Communication Disorders (NIDCD) Division of Scientific Programs (IAA Y2-DC_1004-02), Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN: 00-063. The researchers are indebted to the participants for their willingness to participate in the study.

ARIC (Atherosclerosis Risk in Communities) Study: The ARIC study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I and HHSN268201700005I), R01HL087641, R01HL059367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. 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.

Baependi Heart Study (Brazil): The Baependi Heart Study was supported by Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) (Grant 2013/17368-0), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Hospital Samaritano Society (Grant 25000.180.664/2011-35), through Ministry of Health to Support Program Institutional Development of the Unified Health System (SUS-PROADI).

CARDIA (Coronary Artery Risk Development in Young Adults): The Coronary Artery Risk Development in Young Adults Study (CARDIA) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (HHSN268201800005I & HHSN268201800007I), Northwestern University (HHSN268201800003I), University of Minnesota (HHSN268201800006I), and Kaiser Foundation Research Institute (HHSN268201800004I). CARDIA was also partially supported by the Intramural Research Program of the National Institute on Aging (NIA) and an intra-agency agreement between NIA and NHLBI (AG0005). Genotyping was funded as part of the NHLBI Candidate-gene Association Resource (N01-HC-65226) and the NHGRI Gene Environment Association Studies (GENEVA) (U01-HG004729, U01-HG04424, and U01-HG004446). This manuscript has been reviewed and approved by CARDIA for scientific content.

CHS (Cardiovascular Health Study): This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268201800001C, N01HC55222,

N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, 75N92021D00006; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393, and U01HL130114 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ERF (Erasmus Rucphen Family study):
The ERF study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP grant number 018947 (LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (no. QLG2-CT-2002-01254).. High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organisation for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). We are grateful to all study participants and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work, P. Snijders for his help in data collection and E.M. van Leeuwen for genetic imputation.

FHS (Framingham Heart Study): This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract Nos. N01-HC-25195 and HHSN268201500001I) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. This research was partially supported by grant R01-DK089256 and R01-DK122503 from the National Institute of Diabetes and Digestive and Kidney.

GenSalt (Genetic Epidemiology Network of Salt Sensitivity): The Genetic Epidemiology Network of Salt Sensitivity is supported by research grants (U01HL072507, R01HL087263, and R01HL090682) from the National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD.

HANDLS (Healthy Aging in Neighborhoods of Diversity across the Life Span): The Healthy Aging in Neighborhoods of Diversity across the Life Span (HANDLS) study was supported by the Intramural Research Program of the NIH, National Institute on Aging and the National Center on

Minority Health and Health Disparities (project # Z01-AG000513 and human subjects protocol number 09-AG-N248). Data analyses for the HANDLS study utilized the high-performance computational resources of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD. (http://biowulf.nih.gov; http://hpc.nih.gov)).

Health ABC (Health, Aging, and Body Composition): Health ABC was funded by the National Institutes of Aging. This research was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106. The GWAS was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging.

HERITAGE (Health, Risk Factors, Exercise Training and Genetics): The HERITAGE Family Study was supported by National Heart, Lung, and Blood Institute grant HL-45670, HL-47323, HL47317, HL- 47327, and HL47321.

JHS (Jackson Heart Study): The Jackson Heart Study is supported by contracts HSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, HHSN268201300050C from the National Heart, Lung, and Blood Institute on Minority Health and Health Disparities. The authors acknowledge the Jackson Heart Study team institutions (University of Mississippi Medical Center, Jackson State University and Tougaloo College) and participants for their long-term commitment that continues to improve our understanding of the genetic epidemiology of cardiovascular and other chronic diseases among African Americans.

MESA (Multi-Ethnic Study of Atherosclerosis): This research was supported by the Multi-Ethnic Study of Atherosclerosis (MESA) contracts 75N92020D00001, HHSN268201500003I, N01-HC-95159, 75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420. Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-64278. Genotyping was performed at Affymetrix (Santa Clara, California, USA) and the Broad Institute of Harvard and MIT (Boston, Massachusetts, USA) using the Affymetrix Genome-Wide Human SNP Array 6.0. The authors thank the participants of the MESA study, the Coordinating Center, MESA investigators, and study staff for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org.

NEO (The Netherlands Epidemiology of Obesity study): The authors of the NEO study thank all individuals who participated in the Netherlands Epidemiology in Obesity study, all participating general practitioners for inviting eligible participants and all research nurses for collection of the data. We thank the NEO study group, Petra Noordijk, Pat van Beelen and Ingeborg de Jonge for the coordination, lab and data management of the NEO study. The genotyping in the NEO study was supported by the Centre National de Génotypage (Paris, France), headed by Jean-Francois Deleuze. The NEO study is supported by the participating Departments, the Division and the Board

of Directors of the Leiden University Medical Center, and by the Leiden University, Research Profile Area Vascular and Regenerative Medicine. Dennis Mook-Kanamori is supported by Dutch Science Organization (ZonMW-VENI Grant 916.14.023).

RS (Rotterdam Study): The Rotterdam Study is supported by the Erasmus MC University Medical Center 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 Sports; the European Commission (DG XII); and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.
The generation and management of GWAS genotype data for the Rotterdam Study was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. The GWAS datasets are supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera, Marjolein Peters and Carolina Medina-Gomez for their help in creating the GWAS database, and Karol Estrada, Yurii Aulchenko and Carolina Medina-Gomez for the creation and analysis of imputed data.

WHI (Women's Health Initiative): The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at:
http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf

Stage 2 Study Acknowledgements

Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant R01HL105756. Infrastructure for the Gene-Lifestyle Working Group is supported by the National Heart, Lung, and Blood Institute grant R01HL118305.

CFS (Cleveland Family Study): The CFS was supported by the National Institutes of Health, the National Heart, Lung, Blood Institute grant HL113338, R01HL098433, HL46380.

DR's EXTRA (Dose-Responses to Exercise Training): The study was supported by grants from Ministry of Education and Culture of Finland (722 and 627; 2004-2010); Academy of Finland (102318, 104943, 123885, 211119); European Commission FP6 Integrated Project (EXGENESIS), LSHM-CT-2004-005272; City of Kuopio; Juho Vainio Foundation; Finnish Diabetes Association; Finnish Foundation for Cardiovascular Research; Kuopio University Hospital; Päivikki and Sakari Sohlberg Foundation; Social Insurance Institution of Finland 4/26/2010.

EGCUT (Estonian Genome Center - University of Tartu (Estonian Biobank)): This study was supported by EU H2020 grants 692145, 676550, 654248, Estonian Research Council Grant IUT20-60 and PUT1660, NIASC, EIT – Health and NIH-BMI Grant No: 2R01DK075787-06A1 and EU through the European Regional Development Fund (Project No. 2014-2020.4.01.15-0012 GENTRANSMED.

HCHS/SOL (Hispanic Community Health Study/ Study of Latinos): The baseline examination of HCHS/SOL was supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (N01-HC65233), University of Miami (N01-HC65234), Albert Einstein College of Medicine (N01-HC65235), Northwestern University (N01-HC65236), and San Diego State University (N01-HC65237). The National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research (NIDCR), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Neurological Disorders and Stroke, and NIH Office of Dietary Supplements additionally contributed funding to HCHS/SOL. The Genetic Analysis Center at the University of Washington was supported by NHLBI and NIDCR contracts (HHSN268201300005C AM03 and MOD03). Additional analysis support was provided by 1R01DK101855-01 and 13GRNT16490017. Genotyping was also supported by National Center for Advancing Translational Sciences UL1TR000124 and NIDDK DK063491 to the Southern California Diabetes Endocrinology Research Center. This research was also supported in part by the Intramural Research Program of the NIDDK, contract no. HHSB268201200054C, and Illumina.

IRASFS (Insulin Resistance Atherosclerosis Family Study): The IRASFS was supported by the National Heart Lung and Blood Institute (HL060944, HL061019, and HL060919). Genotyping and analysis for this study was supported by the GUARDIAN Consortium with grant support from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK; DK085175 and DK118062) and in part by UL1TR000124 (CTSI) and DK063491 (DRC). The authors thank study investigators, staff, and participants for their valuable contributions.

KORA (Cooperative Health Research in the Augsburg Region): 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.

LBC1936 (Lothian Birth Cohort 1936): We thank the LBC1936 cohort participants and team members who contributed to these studies. Phenotype collection was supported by Age UK (The Disconnected Mind project). Genotyping was funded by the BBSRC (BB/F019394/1). The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and Medical Research Council (MRC) is gratefully acknowledged. We gratefully acknowledge the contribution of LBC1936 co-author Professor John M. Starr, who died prior to the publication of this manuscript.

Lifelines (Netherlands Biobank): The Lifelines Biobank initiative has been made possible by subsidy from the Dutch Ministry of Health, Welfare and Sport, the Dutch Ministry of Economic Affairs, the University Medical Center Groningen (UMCG the Netherlands), University Groningen and the Northern Provinces of the Netherlands. We acknowledge the services of the Lifelines Cohort Study, the contributing research centers delivering data to Lifelines, all the study participants.
Lifelines group authors: Behrooz Z Alizadeh[1], H Marike Boezen[1], Lude Franke[2], Pim van der Harst[3], Gerjan Navis[4], Marianne Rots[5], Harold Snieder[1], Morris Swertz[2], Bruce HR Wolffenbuttel[6], Cisca Wijmenga[2]
[1]Department of Epidemiology, University of Groningen, University Medical Center Groningen, The Netherlands;
[2]Department of Genetics, University of Groningen, University Medical Center Groningen, The Netherlands; [3]Department of Cardiology, University of Groningen, University Medical Center Groningen, The Netherlands;
[4]Department of Internal Medicine, Division of Nephrology, University of Groningen, University Medical Center Groningen, The Netherlands; [5]Department of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, The Netherlands; [6]Department of Endocrinology, University of Groningen, University Medical Center Groningen, The Netherlands.

NESDA (Netherlands Study of Depression and Anxiety): Funding was obtained from the Netherlands Organization for Scientific Research (Geestkracht program grant 10-000-1002); the Center for Medical Systems Biology (CSMB, NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL), VU University's Institutes for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam, University Medical Center Groningen, Leiden University Medical Center, National Institutes of Health (NIH, R01D0042157-01A, MH081802, Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health.Computing was supported by BiG Grid, the Dutch e-Science Grid, which is financially supported by NWO.

SHIP (Study of Health in Pomerania): 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 were 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.

SWHS/SMHS (Shanghai Women's Health Study/ Shanghai Men's Health Study): We thank all the individuals who took part in these studies and all the researchers who have enabled this work to be carried out. The Shanghai Women's Health Study and the Shanghai Men's Health Study are supported by research grants UM1CA182910 and UM1CA173640 from the U.S. National Cancer Institute, respectively.

UKB (United Kingdom Biobank): This research has been conducted using the UK Biobank Resource. The UK Biobank data was analysed from the data set corresponding to UK Biobank access application no. 236, application title "Genome-wide association study of blood pressure", with Paul Elliott as the PI/applicant. Central analysts from Queen Mary University of London (QMUL) and Imperial College London are funded by the NIHR Cardiovascular Biomedical Research Unit at Barts and The London School of Medicine, QMUL, and the NIHR Imperial College Health Care NHS Trust and Imperial College London Biomedical Research Centre respectively. This work was supported by the UK Biobank CardioMetabolic Consortium (UKB-CMC) and the BP working group. The UKB-CMC was supported by the British Heart Foundation (grant SP/13/2/30111). The UK CardioMetabolic Consortium (UK-CMC) was supported by British Heart Foundation (SP/13/2/30111 - Largescale comprehensive genotyping of UK Biobank for cardiometabolic traits and diseases: UKCMC).

WASHS (Western Australian Sleep Health Study): Funding for the Western Australian Sleep Health study was obtained from the Sir Charles Gairdner and Hollywood Private Hospital Research Foundations, the State Health Research Advisory Council of Western Australia, the Western Australian Sleep Disorders Research Institute, and the Centre for Genetic Epidemiology and Biostatistics at the University of Western Australia. Biobanking support is received from the WA DNA Bank, a National Health and Medical Research Council of Australia Enabling Facility. The authors gratefully acknowledge all the students and volunteers who have assisted with data collection for the WASHS. The authors thank the sleep clinic patients for their participation in this study. We are also grateful for the support received from WA Sleep Disorders Research Institute, Centre for Genetic Epidemiology and Biostatistics, and PathWest.

YFS (The Cardiovascular Risk in Young Finns Study): The Young Finns Study has been financially supported by the Academy of Finland: grants 322098, 286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Competitive State Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals (grant X51001); Juho Vainio Foundation; Paavo

Nurmi Foundation; Finnish Foundation for Cardiovascular Research ; Finnish Cultural Foundation; The Sigrid Juselius Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; Yrjö Jahnsson Foundation; Signe and Ane Gyllenberg Foundation; Diabetes Research Foundation of Finnish Diabetes Association; EU Horizon 2020 (grant 755320 for TAXINOMISIS and grant 848146 for To Aition); European Research Council (grant 742927 for MULTIEPIGEN project); Tampere University Hospital Supporting Foundation and Finnish Society of Clinical Chemistry. We thank the teams that collected data at all measurement time points; the persons who participated as both children and adults in these longitudinal studies; and biostatisticians Irina Lisinen, Johanna Ikonen, Noora Kartiosuo, Ville Aalto, and Jarno Kankaanranta for data management and statistical advice.

Additional Acknowledgments

T. O. K. was funded by the Novo Nordisk Foundation (NNF18CC0034900, NNF17OC0026848). P.S.d.V. was supported by American Heart Association grant number 18CDA34110116. N.F. was supported by R01-MD012765, R01-DK117445, R21-HL140385. X.Z. was supported by grant HG011052 from the National Human Genome Research Institute (NHGRI). W.J.G. was supported by P01-CA196569.

Supplementary Figures

Multi-ancestry genome-wide gene-sleep interactions identify novel loci for blood pressure

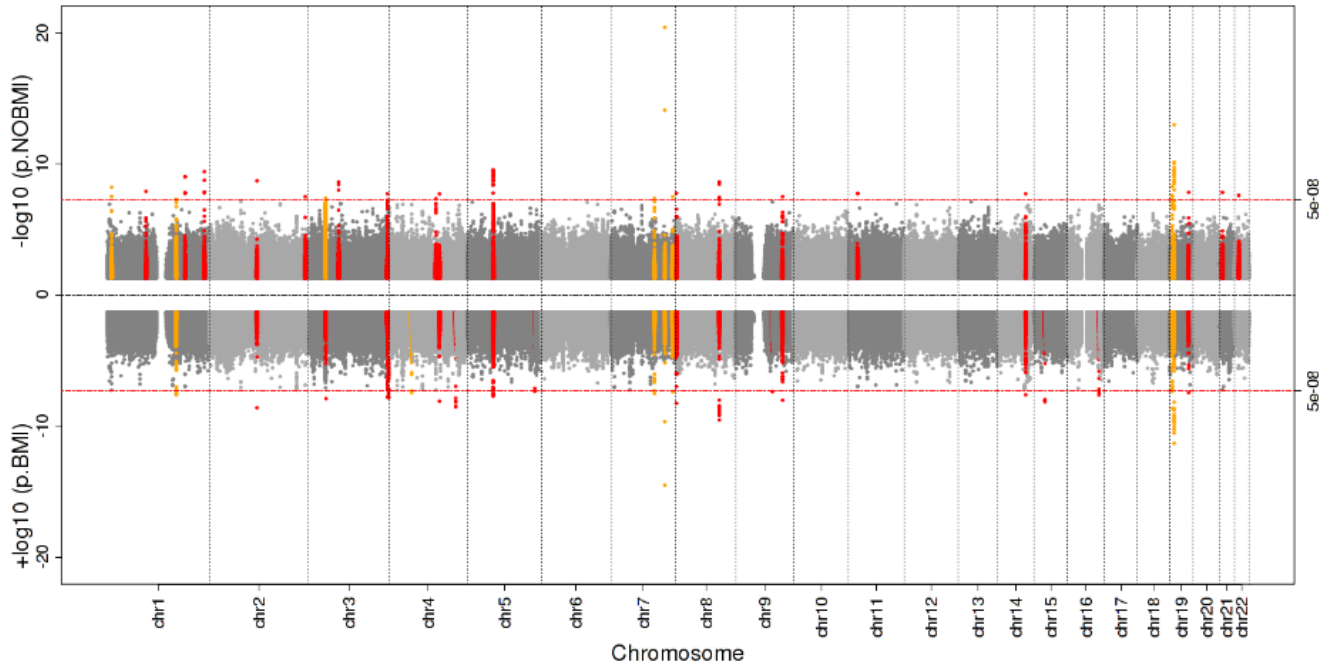
Wang H, et al.

Contents

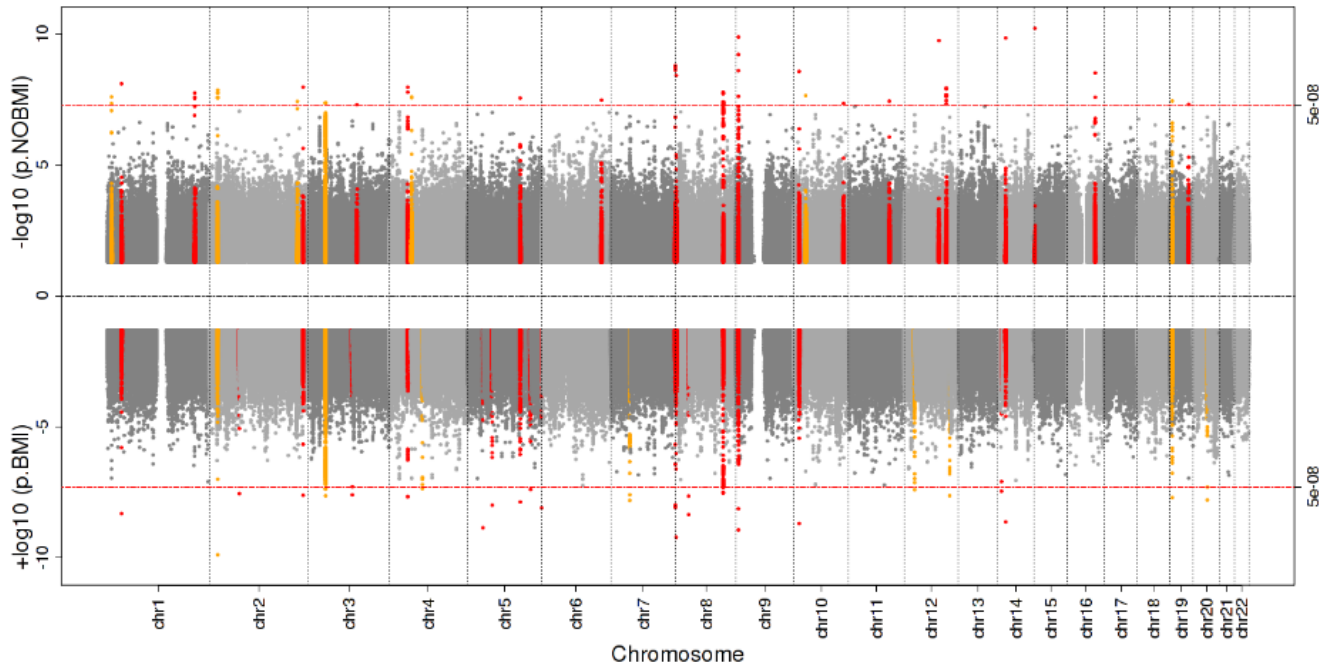
Supplementary Fig. 1. Miami plots of genome-wide gene-sleep interaction analyses in stage 1 MULTI ancestry group.

Upper panel are results without adjusting for BMI, lower panel are results adjusted for BMI. The minimum p-values across four BP traits were used to generate each plot. The orange points correspond to significant loci ($P_{joint}<5\times10^{-8}$) overlapped with previously reported BP loci (distance <1Mb), and the red points correspond to novel significant loci.

LTST

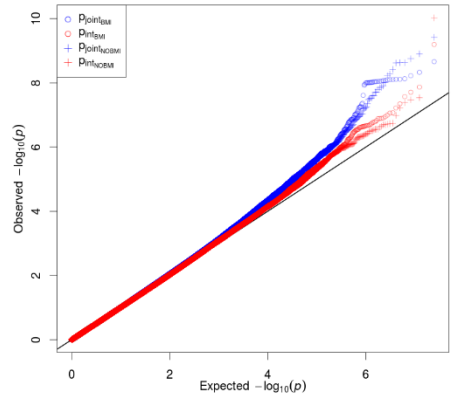


STST

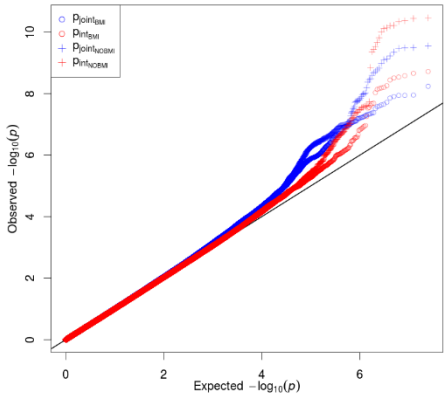


Supplementary Fig. 2. QQ plots of genome-wide gene-sleep interaction analyses in stage 1 MULTI ancestry group.

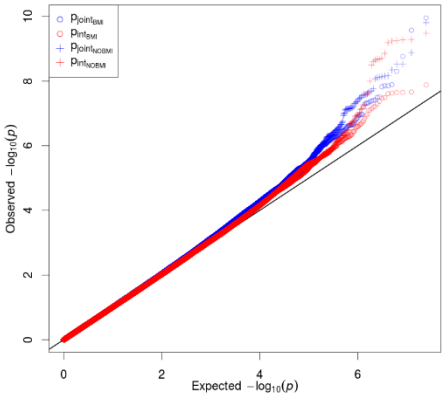
LTST-SBP



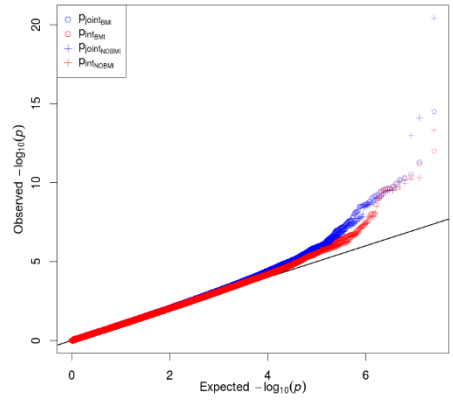
LTST-DBP



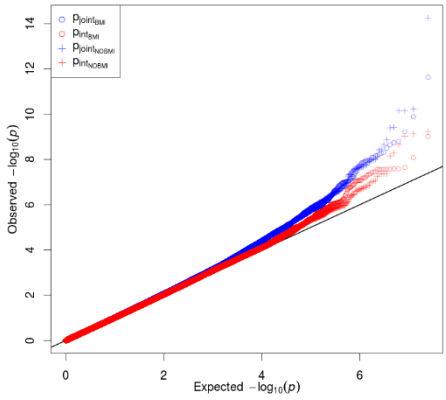
LTST-MAP



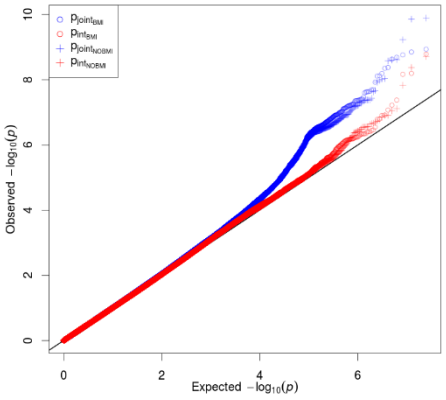
LTST-PP



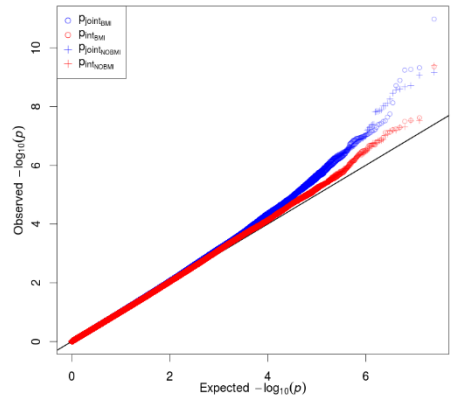
STST-SBP



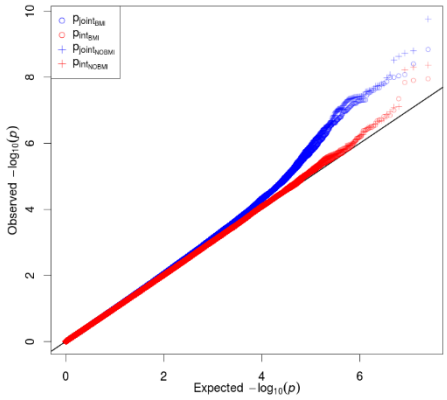
STST-DBP



STST-MAP



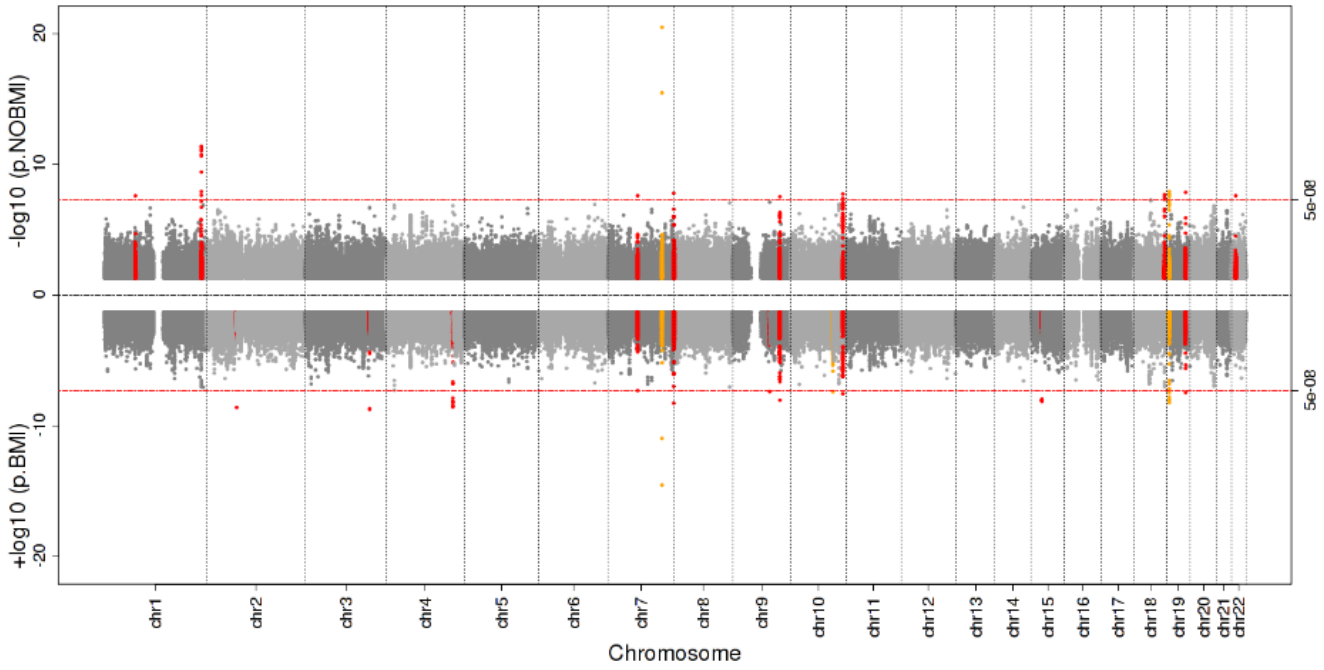
STST-PP



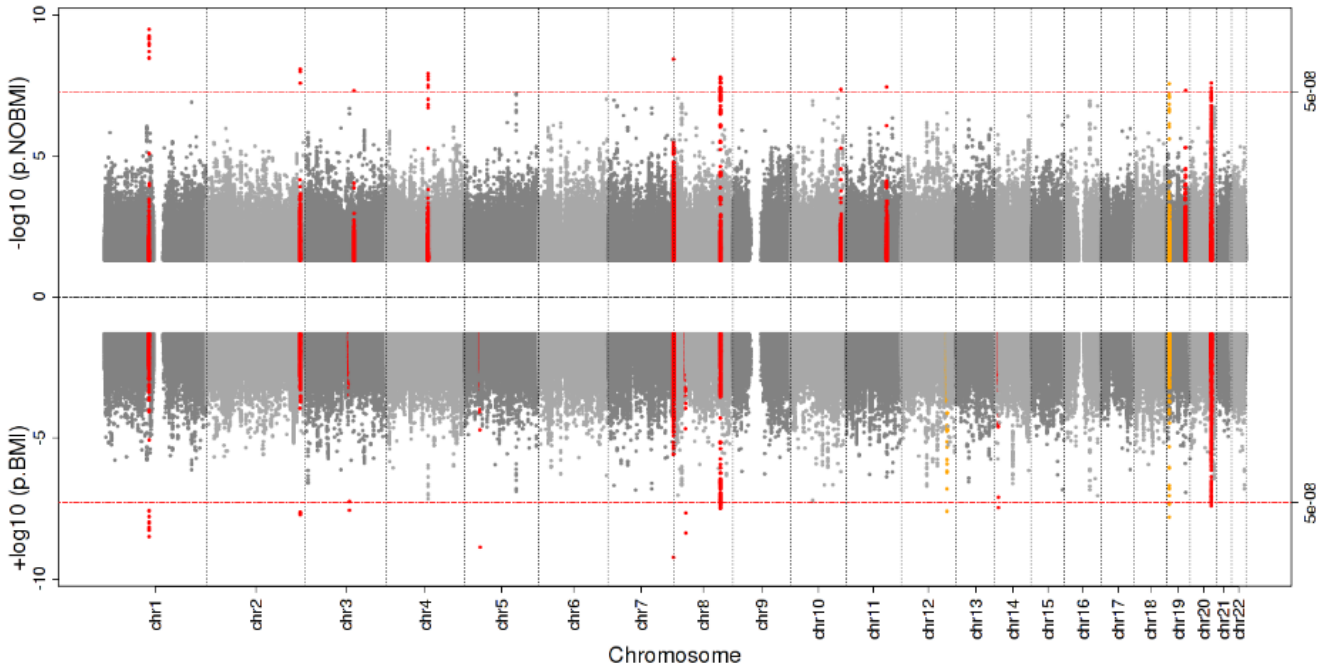
Supplementary Fig. 3. Miami plots of genome-wide gene-sleep interaction analyses in stage 1 EUR group.

Upper panel are results without adjusting for BMI, lower panel are results adjusted for BMI. The minimum p-values across four BP traits were used to generate each plot. The orange points correspond to significant loci ($P_{joint}<5\times10^{-8}$) overlapped with previously reported BP loci (distance <1Mb), and the red points correspond to novel significant loci.

LTST : P_{joint}

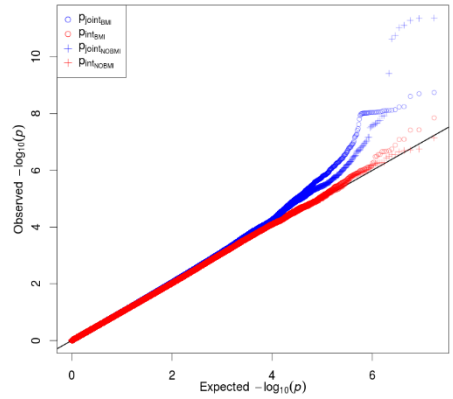


STST : P_{joint}

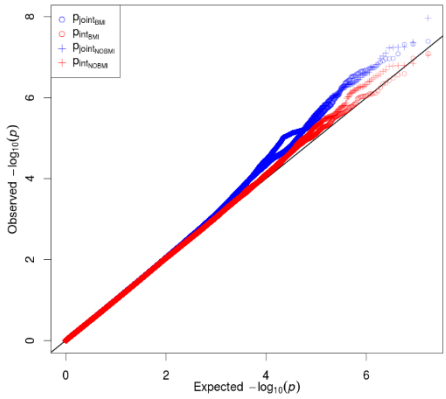


Suppementary Fig. 4. QQ plots of genome-wide gene-sleep interaction analyses in stage 1 EUR group.

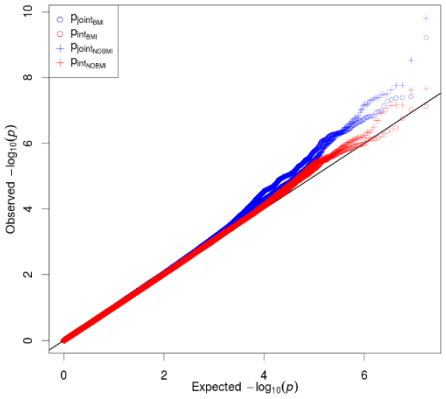
LTST-SBP



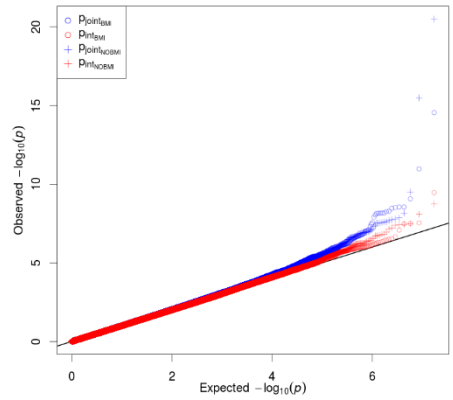
LTST-DBP



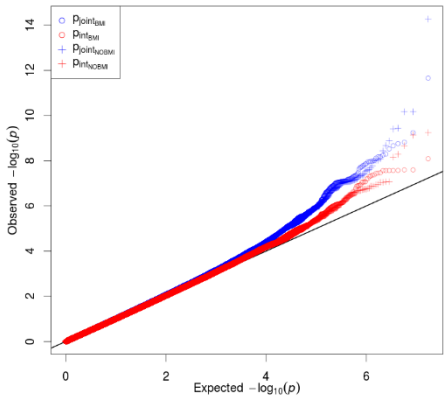
LTST-MAP



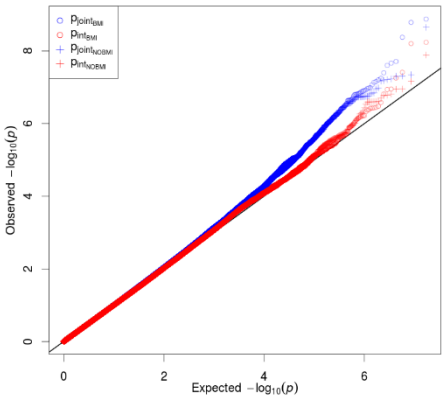
LTST-PP



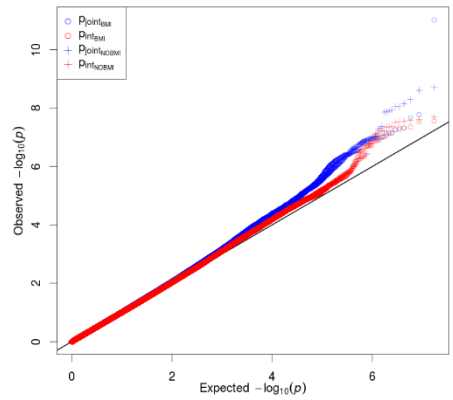
STST-SBP



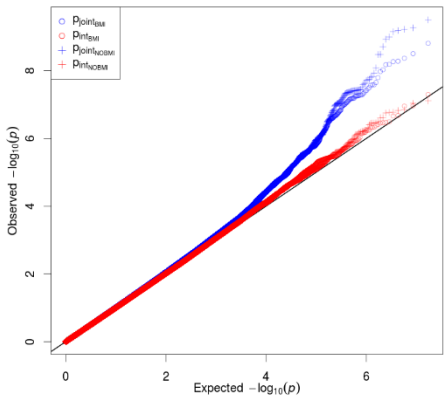
STST-DBP



STST-MAP



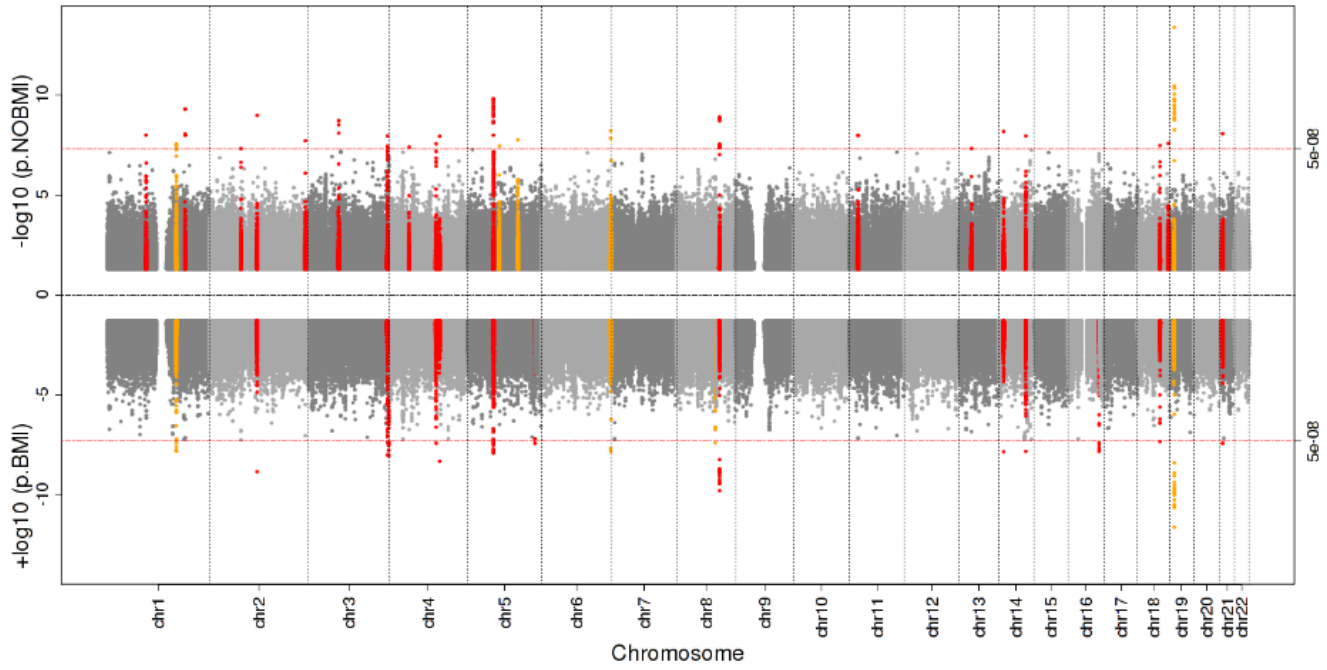
STST-PP



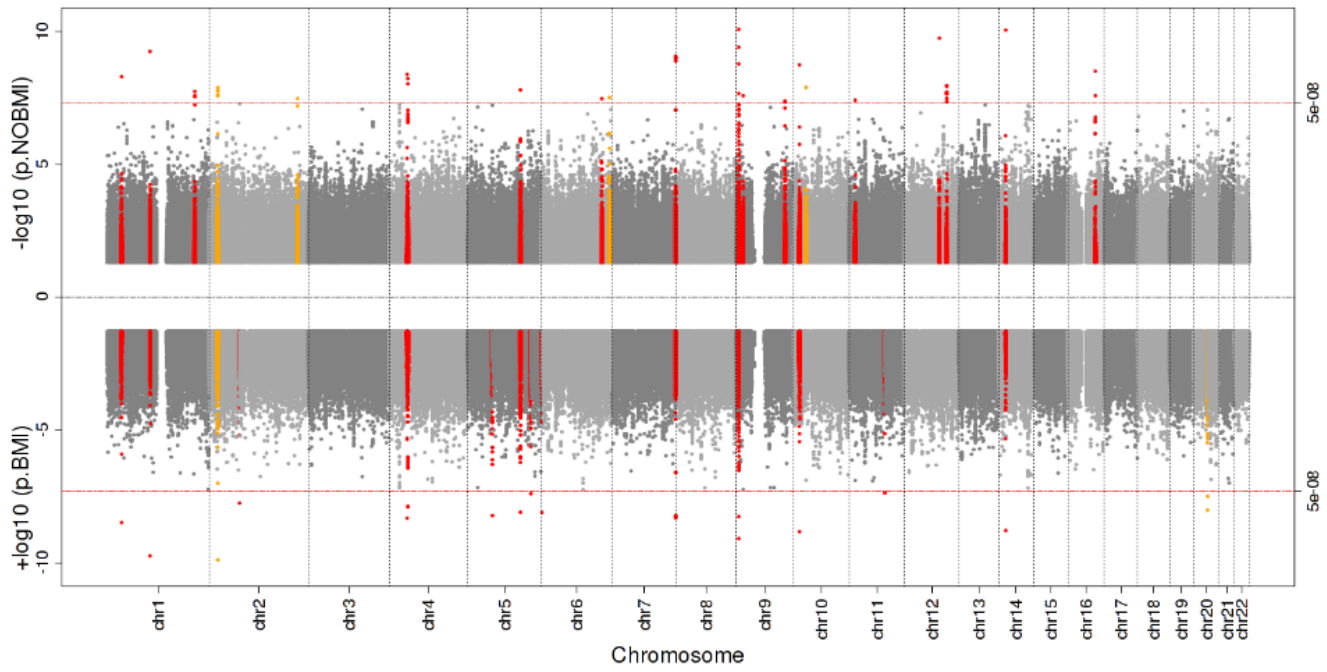
Supplementary Fig. 5. Miami plots of genome-wide gene-sleep interaction analyses in stage 1 AFR group.

Upper panel are results without adjusting for BMI, lower panel are results adjusted for BMI. The minimum p-values across four BP traits were used to generate each plot. The orange points correspond to significant loci ($P_{joint}<5\times10^{-8}$) overlapped with previously reported BP loci (distance <1Mb), and the red points correspond to novel significant loci.

LTST : P_{joint}

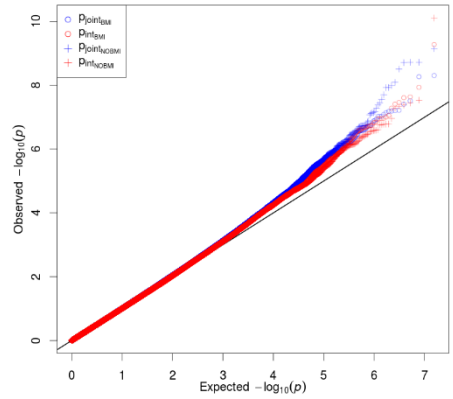


STST : P_{joint}

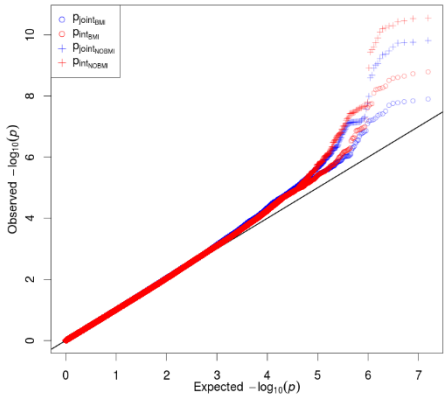


Supplementary Fig. 6. QQ plots of genome-wide gene-sleep interaction analyses in stage 1 AFR group.

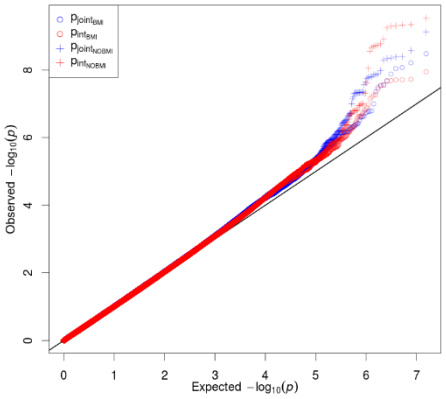
LTST-SBP



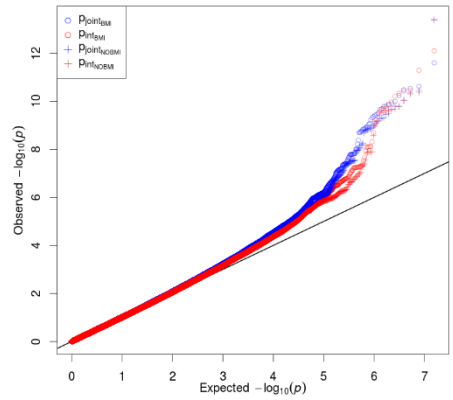
LTST-DBP



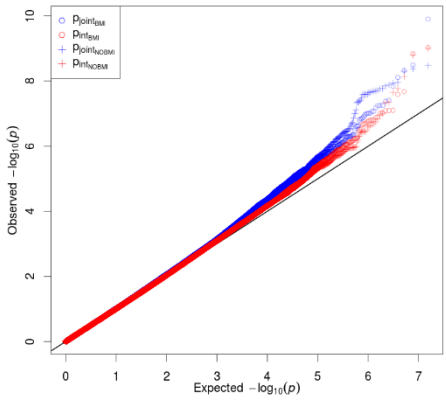
LTST-MAP



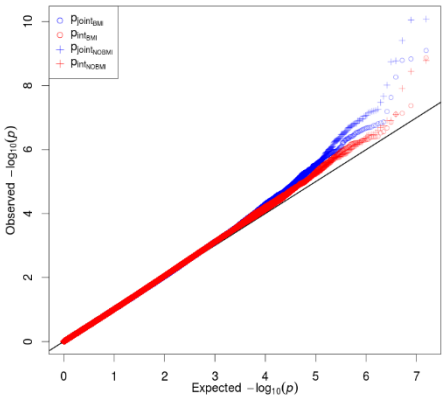
LTST-PP



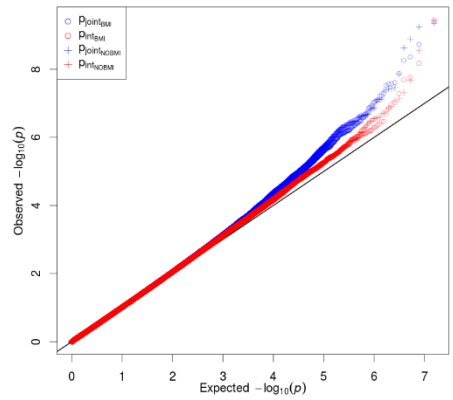
STST-SBP



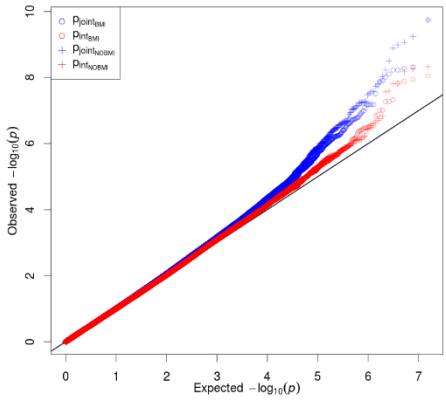
STST-DBP



STST-MAP

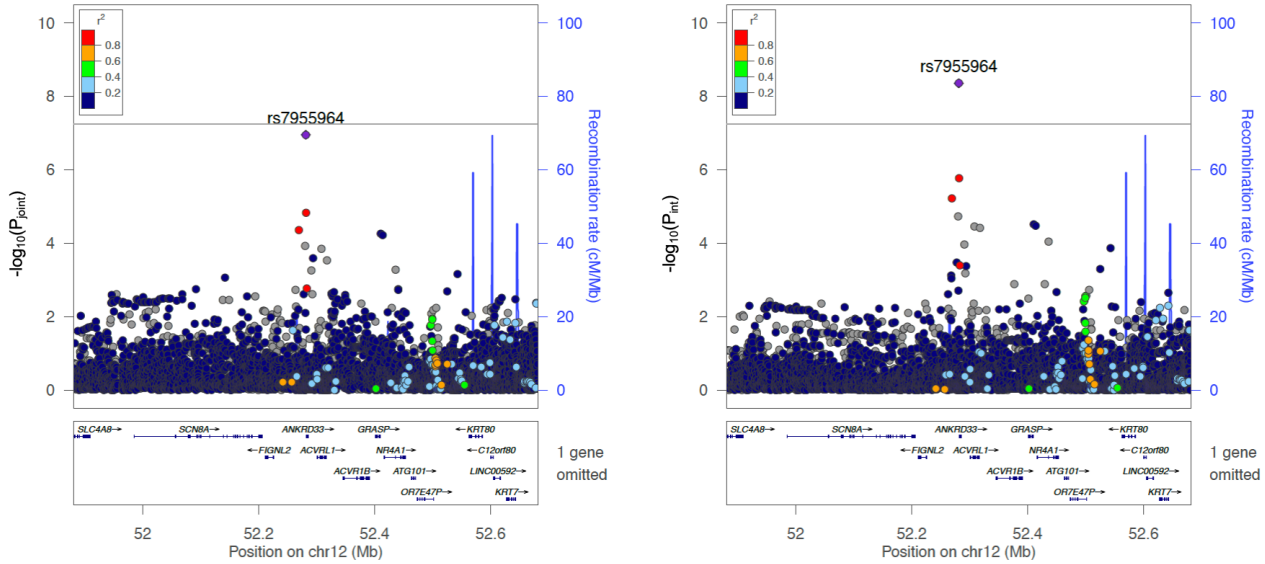


STST-PP

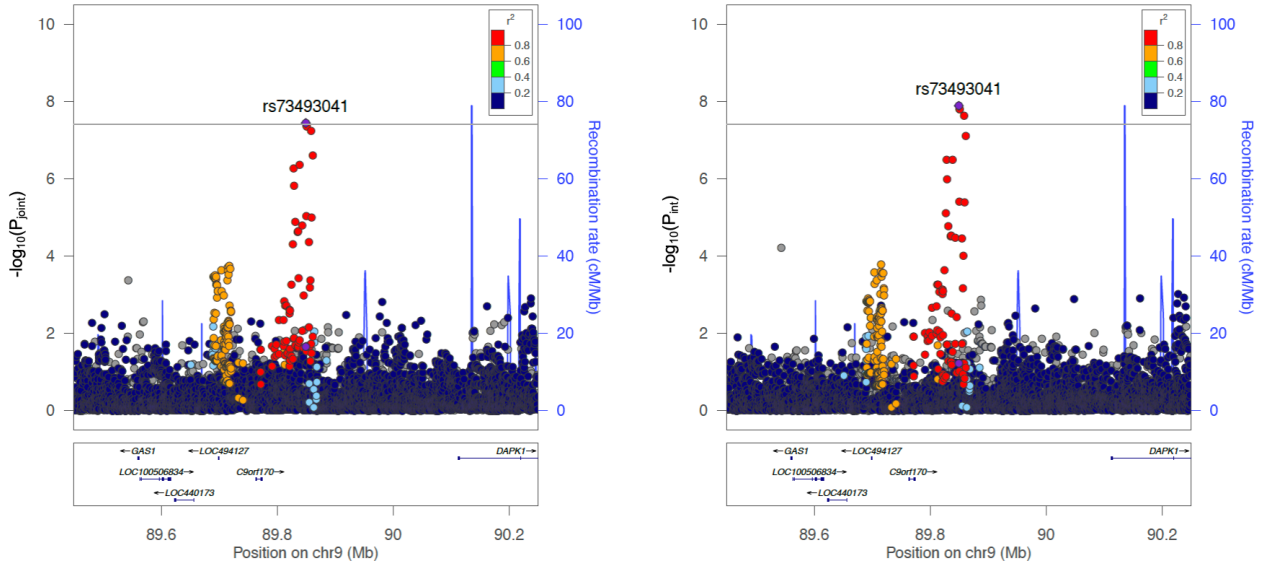


Supplementary Fig. 7. Regional association plots of P_{joint} and P_{int} for 3 replicated novel loci in stage 1+2 MULTI ancestry analyses (Table 1).

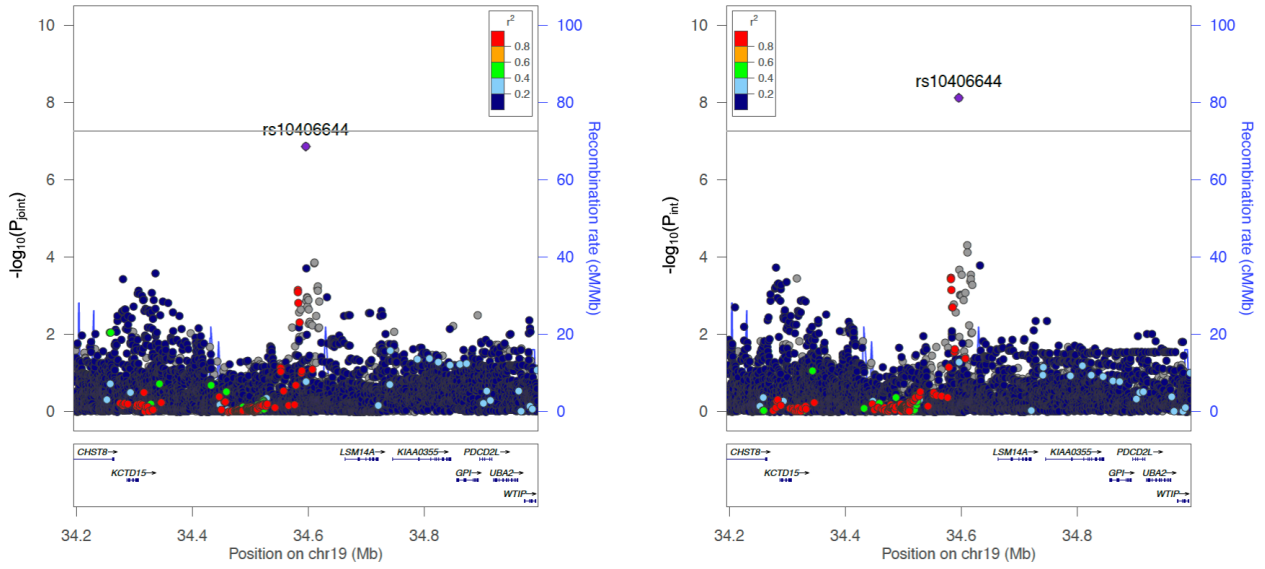
Locus 1. LTST- MAP



Locus 2. STST- DBP

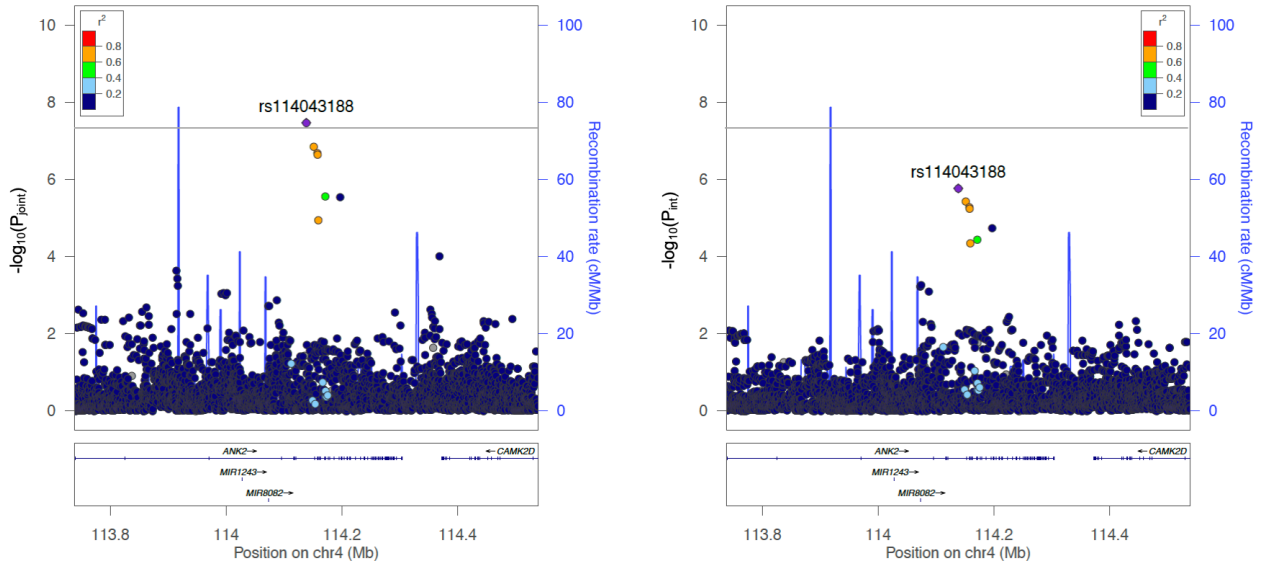


Locus 3. STST- PP

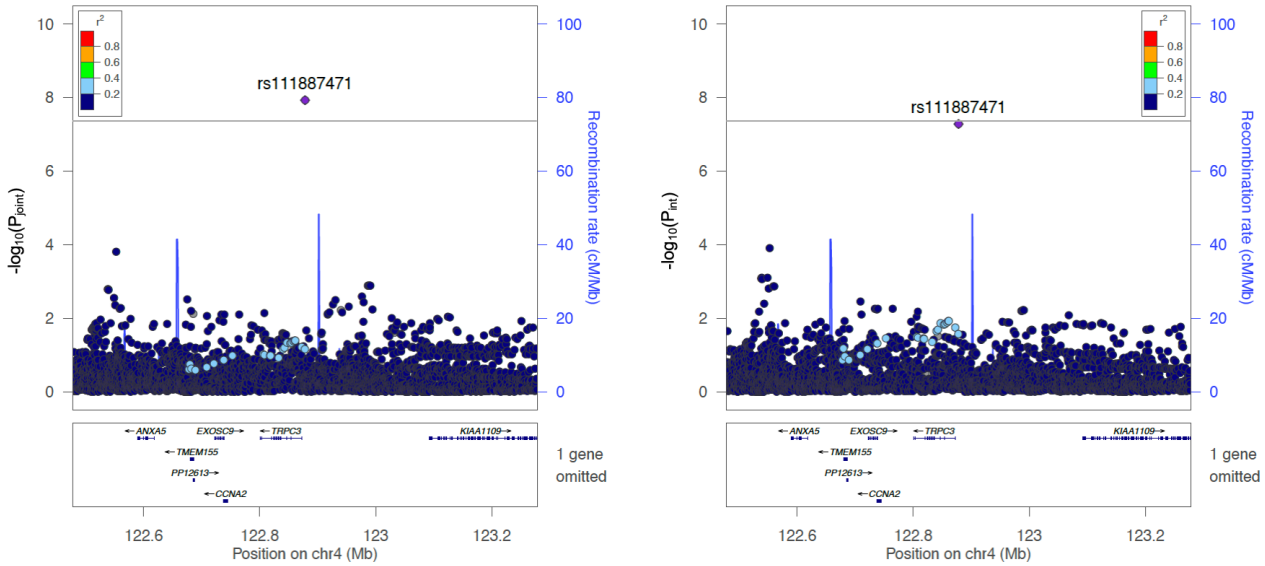


Supplementary Fig. 8. Regional association plots of P_{joint} and P_{int} for 3 novel loci in stage 1 AFR analyses (Supplementary Table 8).

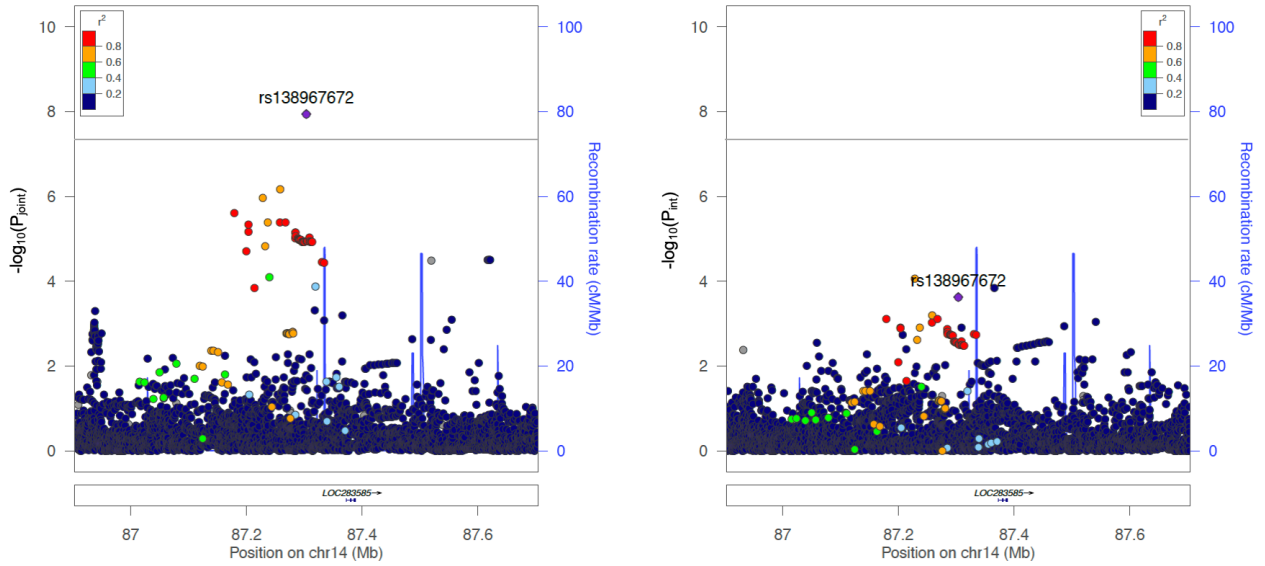
Locus 1. LTST- PP



Locus 2. LTST-SBP



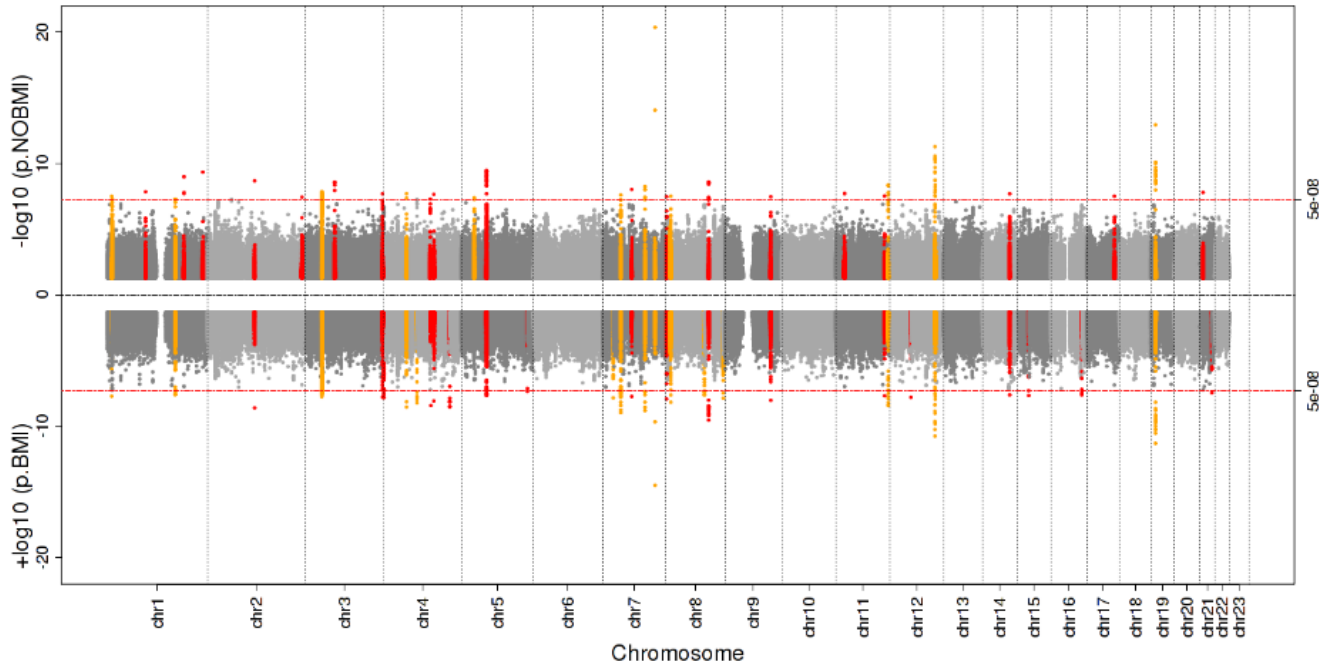
Locus 3. LTST-PP



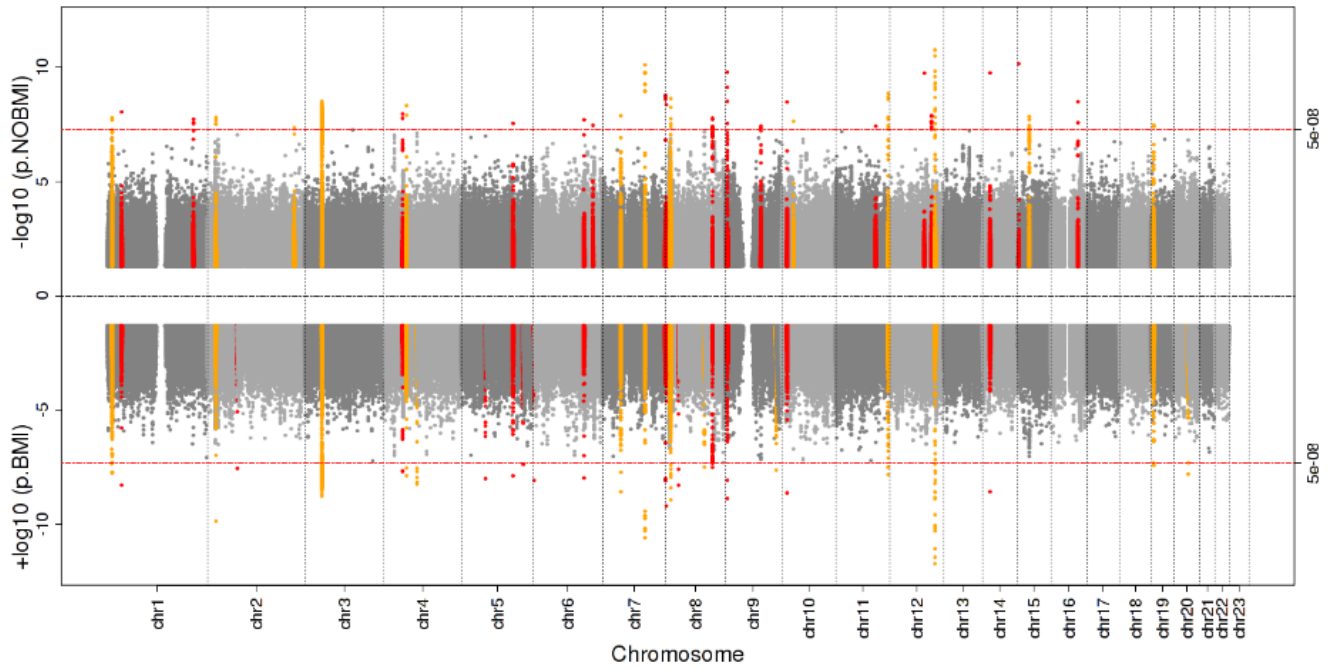
Supplementary Fig. 9. Miami plots of genome-wide gene-sleep interaction analyses in stage 1+2 MULTI ancestry group.

Upper panel are results without adjusting for BMI, lower panel are results adjusted for BMI. The minimum p-values across four BP traits were used to generate each plot. The orange points correspond to significant loci ($P_{joint}<5\times10^{-8}$) overlapped with previously reported BP loci (distance <1Mb), and the red points correspond to novel significant loci.

LTST : P_{joint}

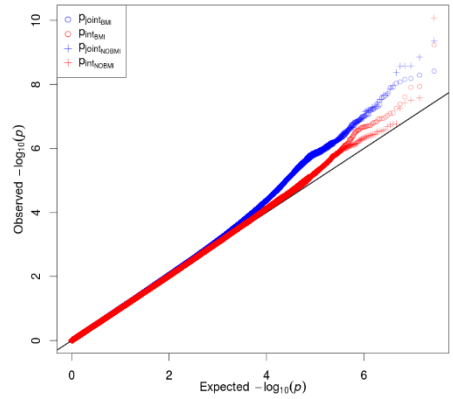


STST : P_{joint}

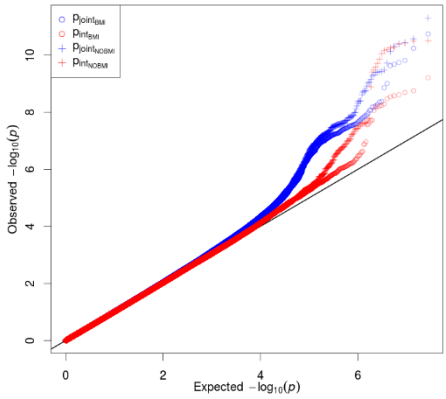


Supplementary Fig. 10. QQ plots of genome-wide gene-sleep interaction analyses in stage 1+2 MULTI ancestry group.

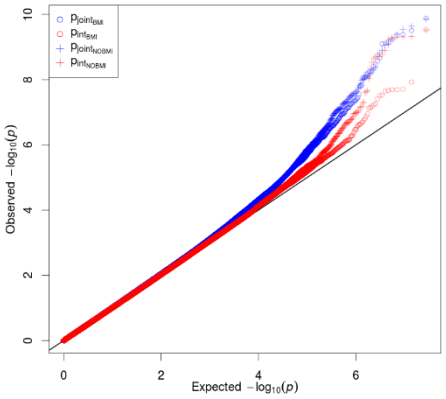
LTST-SBP



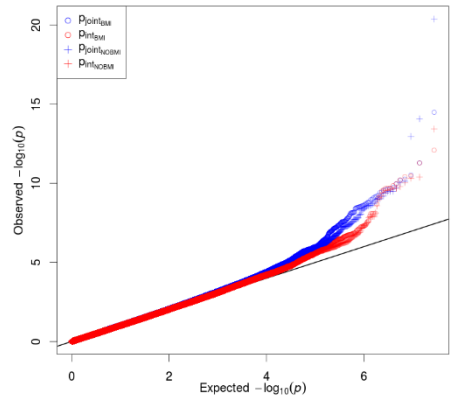
LTST-DBP



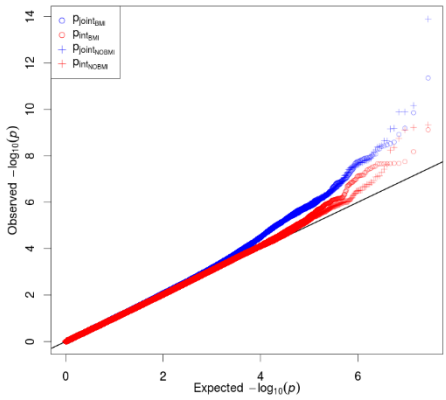
LTST-MAP



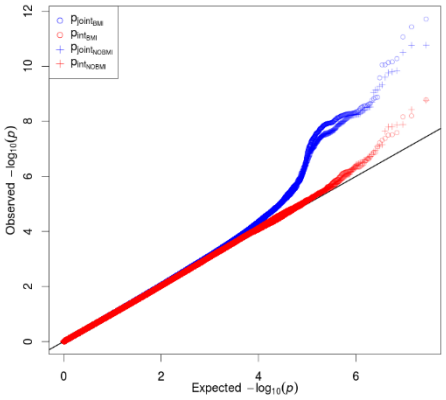
LTST-PP



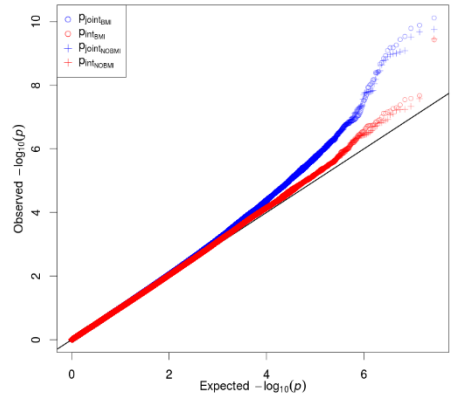
STST-SBP



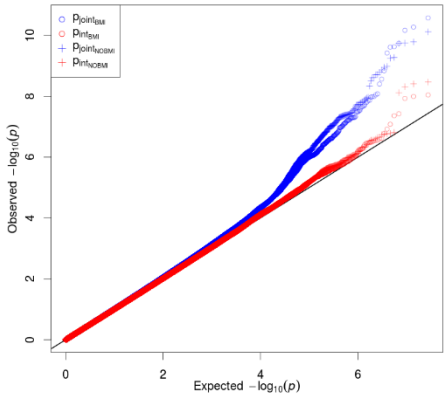
STST-DBP



STST-MAP



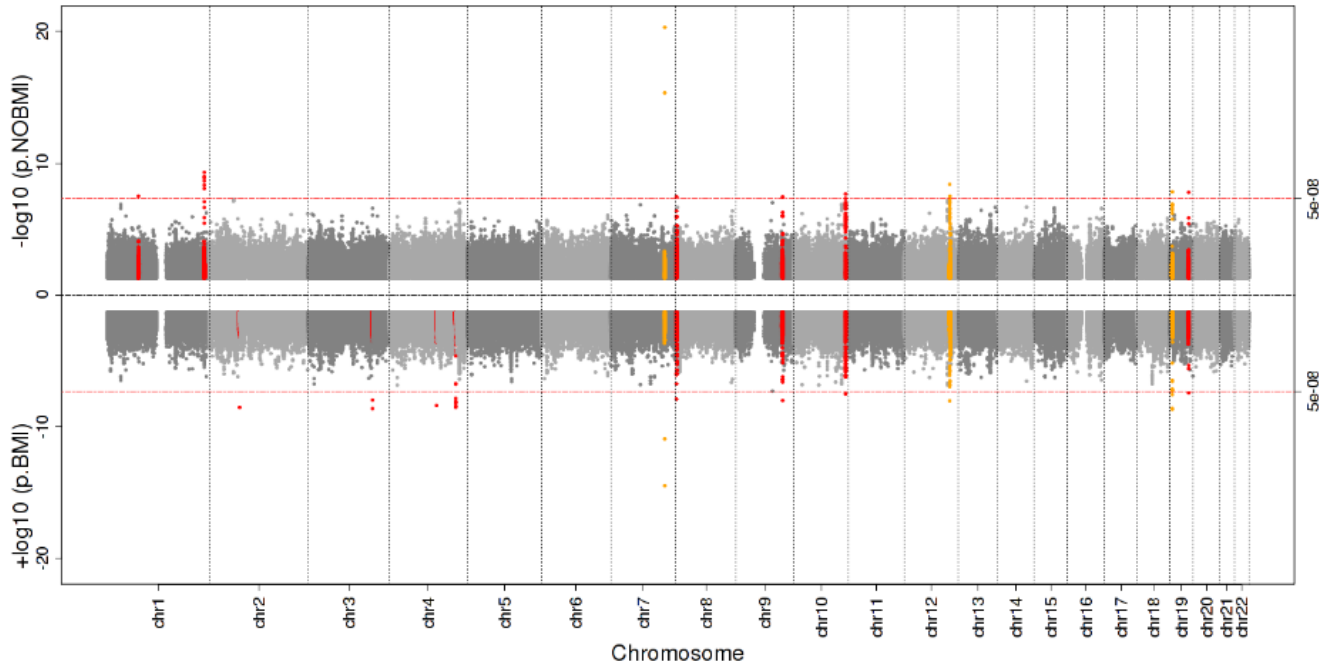
STST-PP



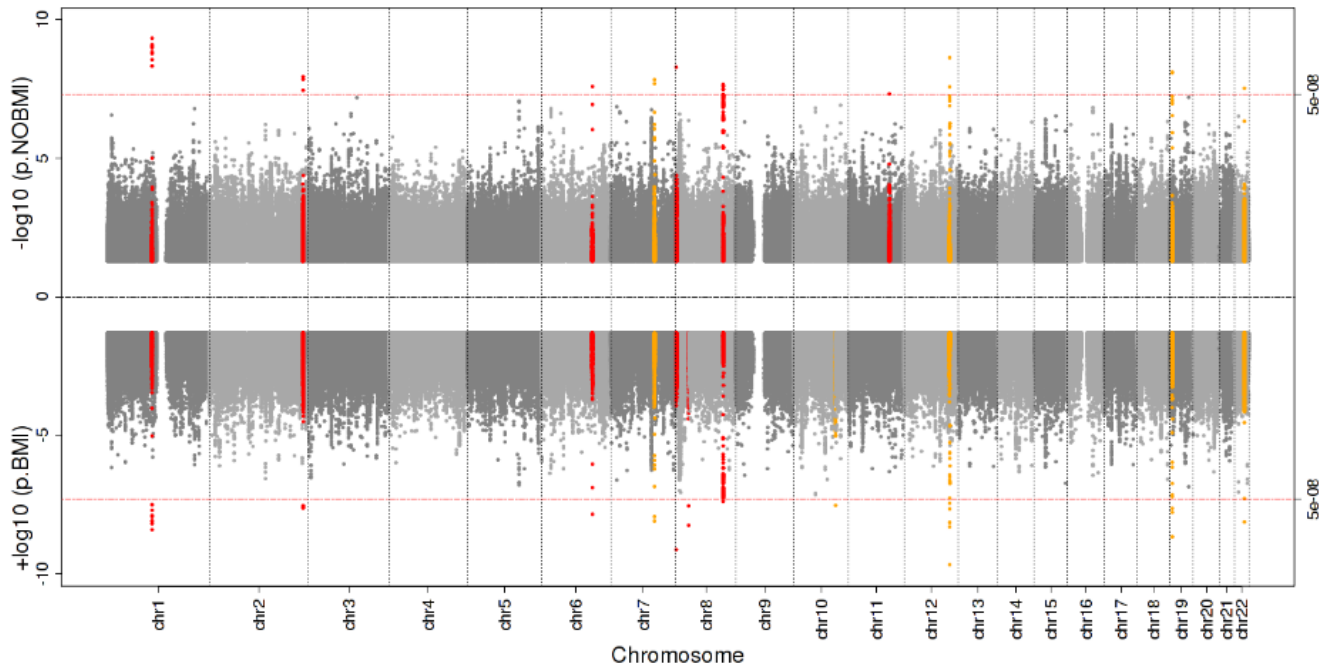
Supplementary Fig. 11. Miami plots of genome-wide gene-sleep interaction analyses in stage 1+2 EUR group.

Upper panel are results without adjusting for BMI, lower panel are results adjusted for BMI. The minimum p-values across four BP traits were used to generate each plot. The orange points correspond to significant loci ($P_{joint}<5\times10^{-8}$) overlapped with previously reported BP loci (distance <1Mb), and the red points correspond to novel significant loci.

LTST : P_{joint}

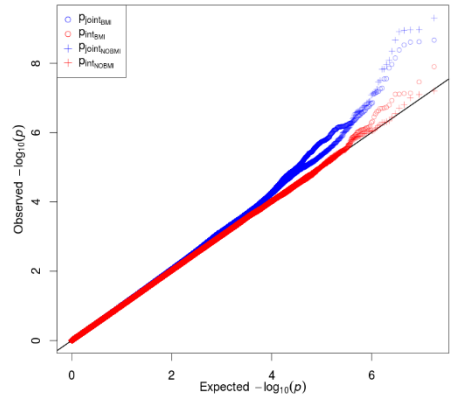


STST : P_{joint}

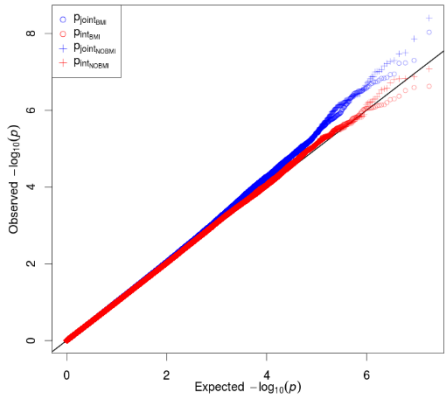


Supplementary Fig. 12. QQ plots of genome-wide gene-sleep interaction analyses in stage 1+2 EUR group.

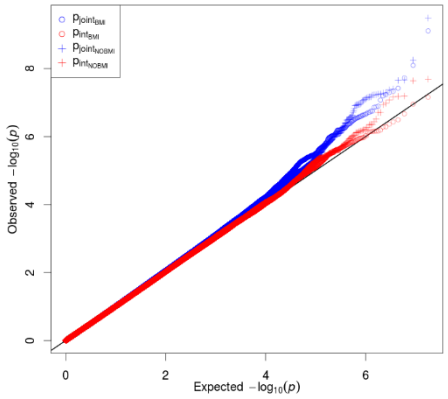
LTST-SBP



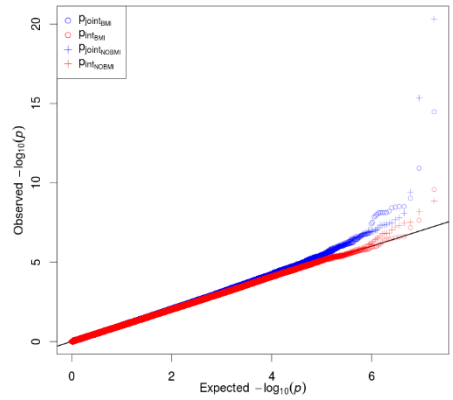
LTST-DBP



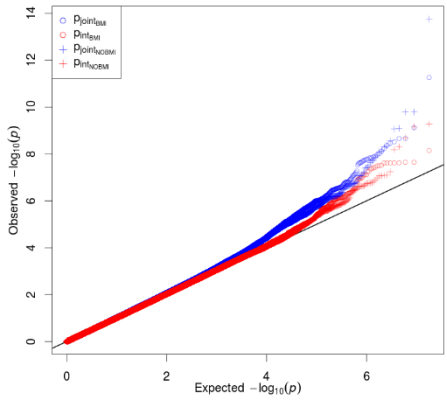
LTST-MAP



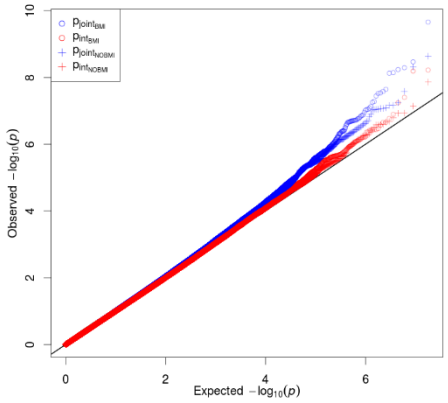
LTST-PP



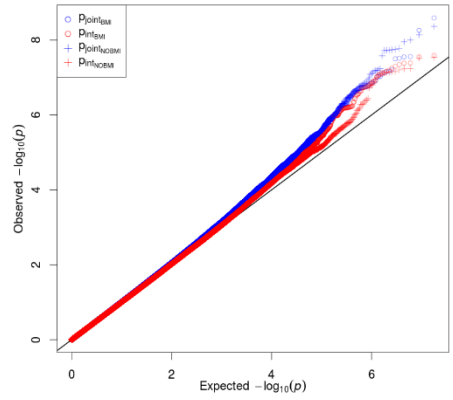
STST-SBP



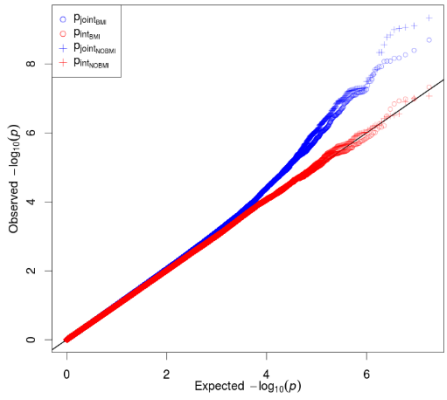
STST-DBP



STST-MAP

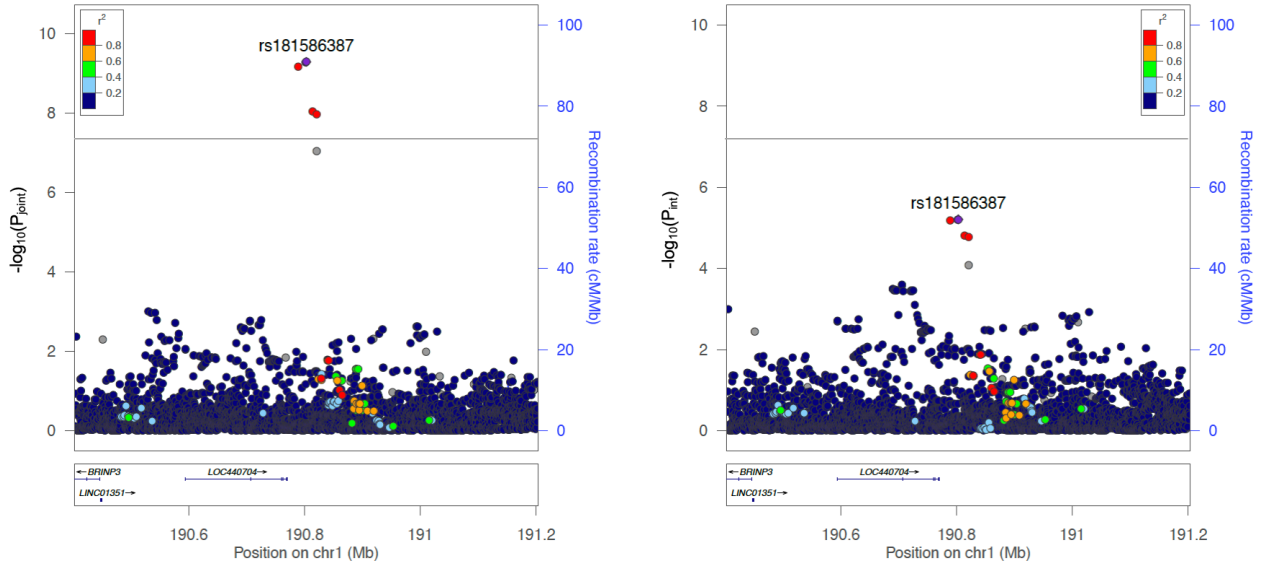


STST-PP

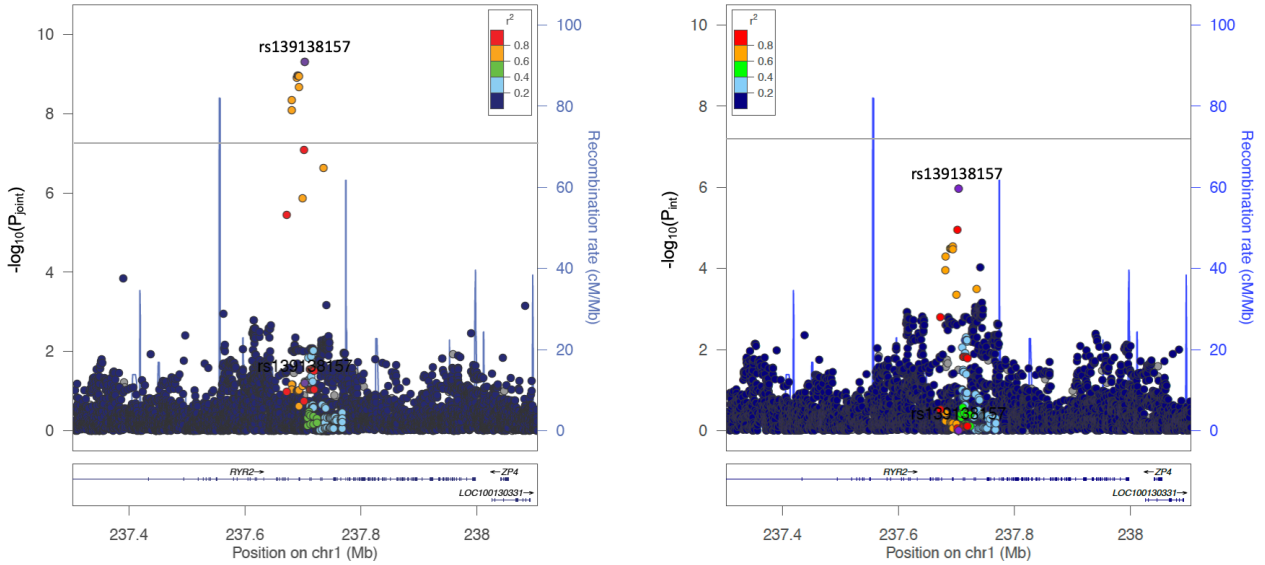


Supplementary Fig. 13. Regional association plots of P_{joint} and P_{int} for 20 novel loci in stage 1+2 analyses (Supplementary Table 9).

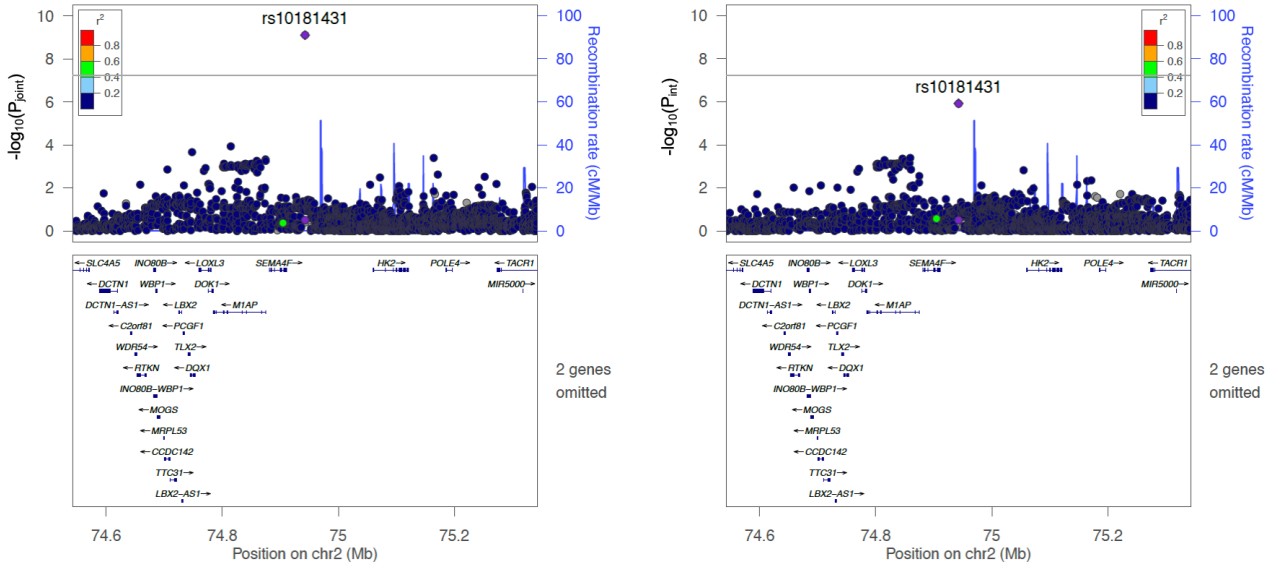
Locus 1. LTST- PP



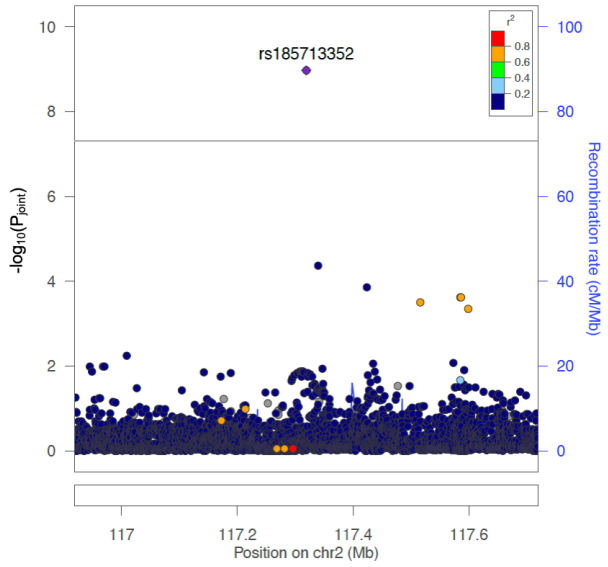
Locus 2. LTST- SBP

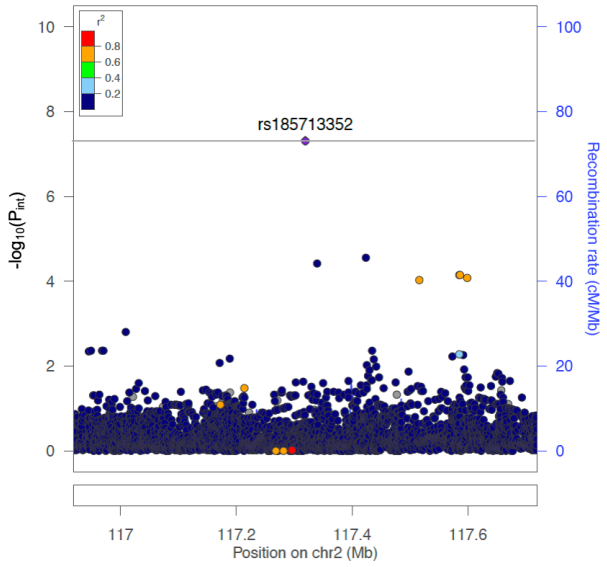


Locus 3. LTST- MAP

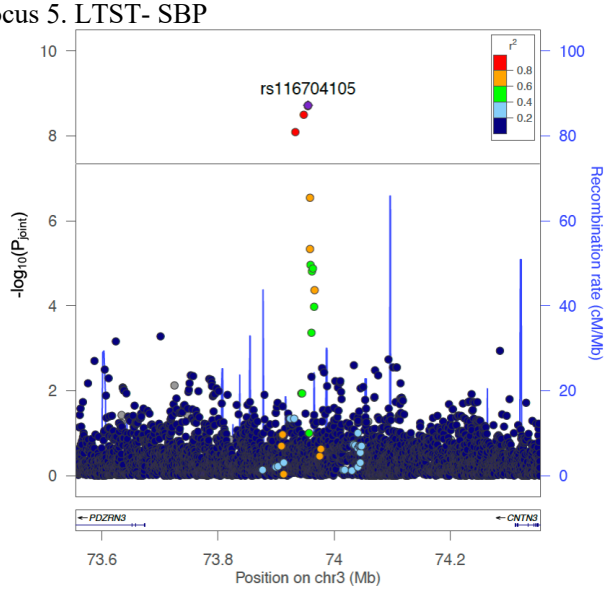


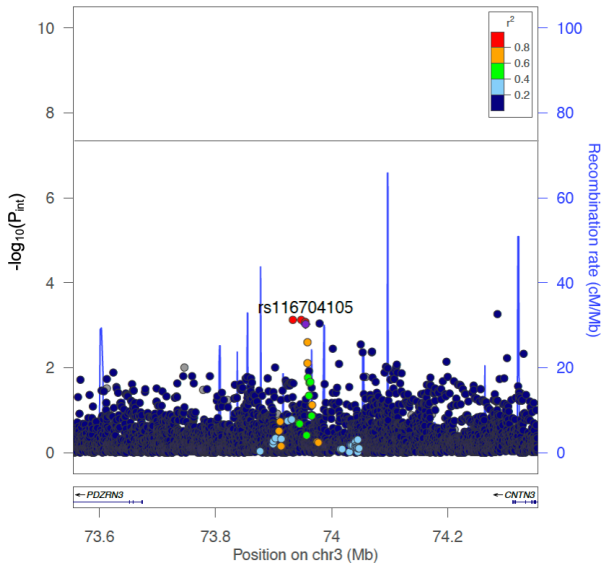
Locus 4. LTST- PP



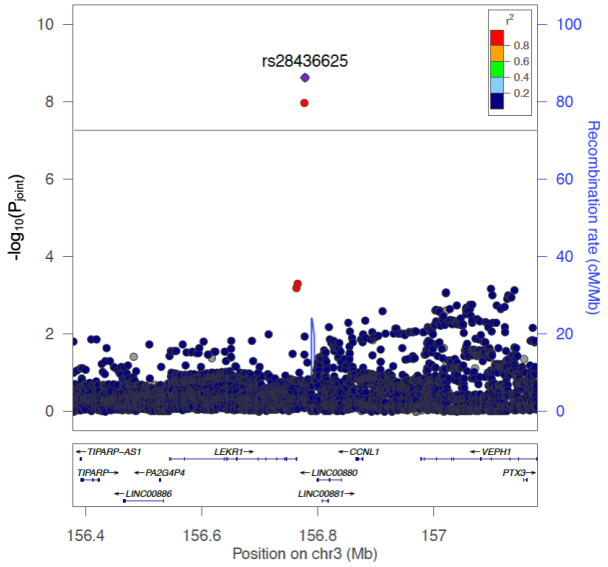


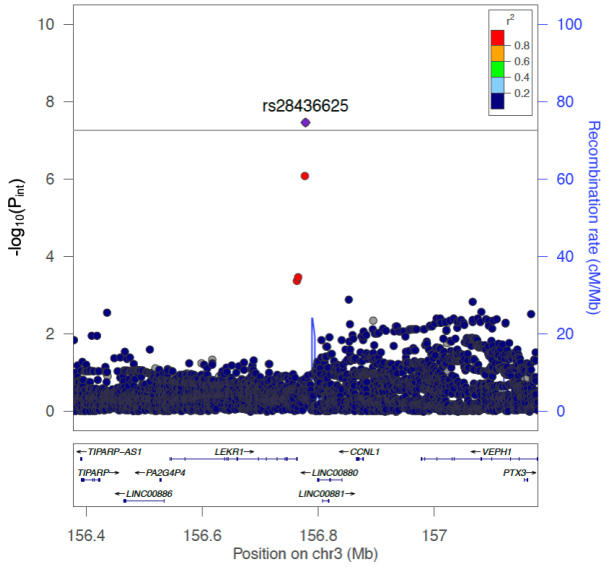
Locus 5. LTST- SBP



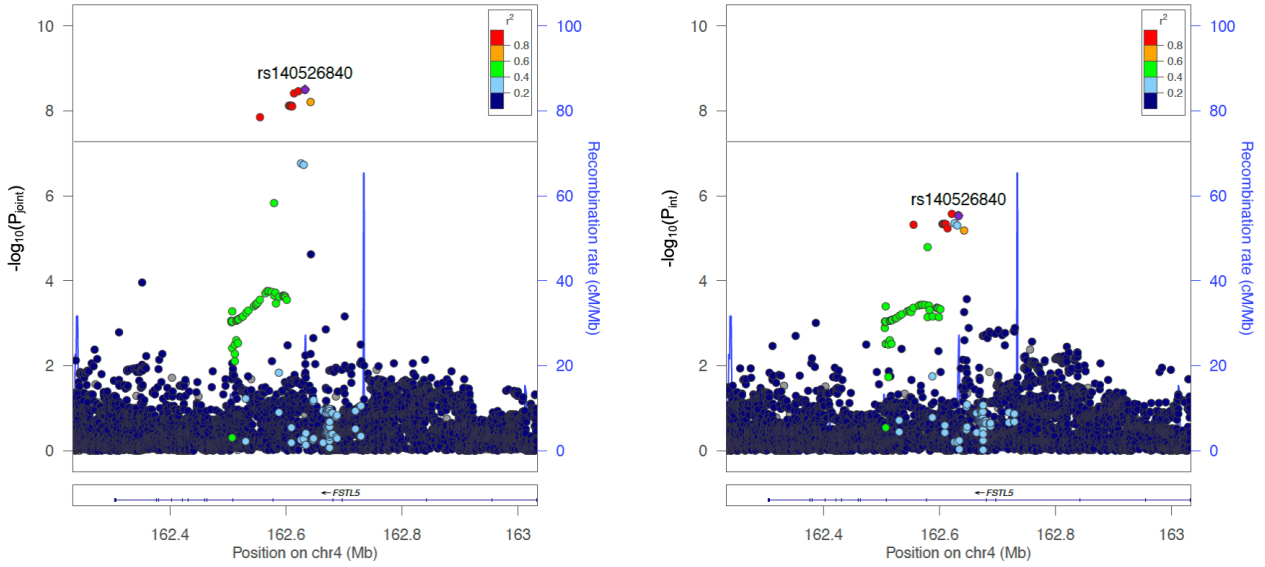


Locus 6. LTST- SBP

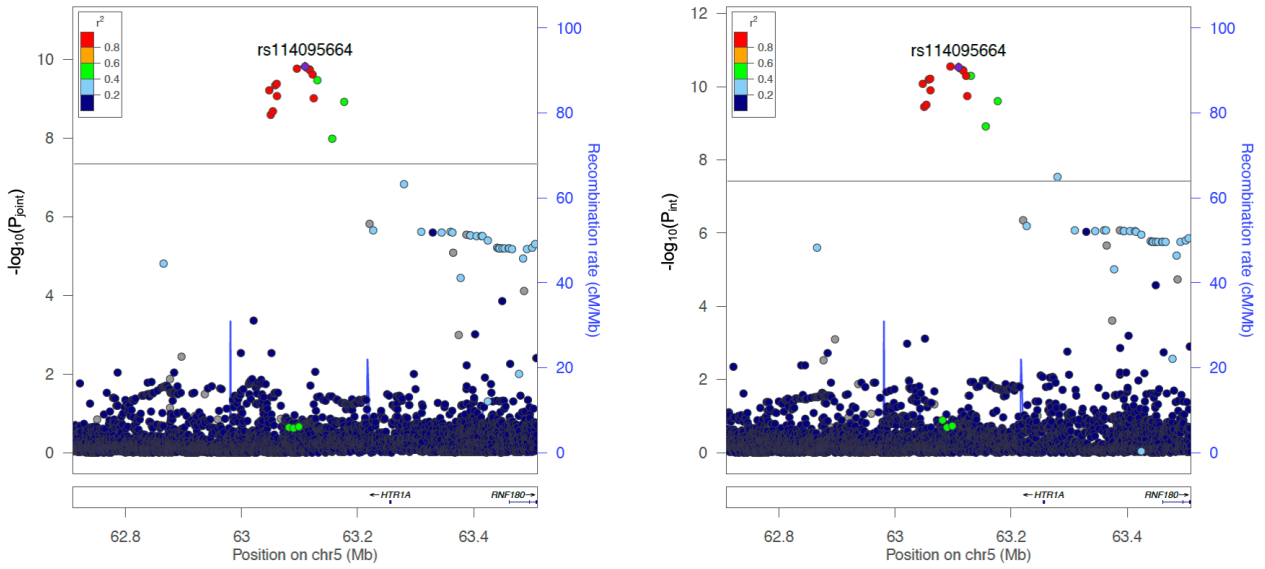




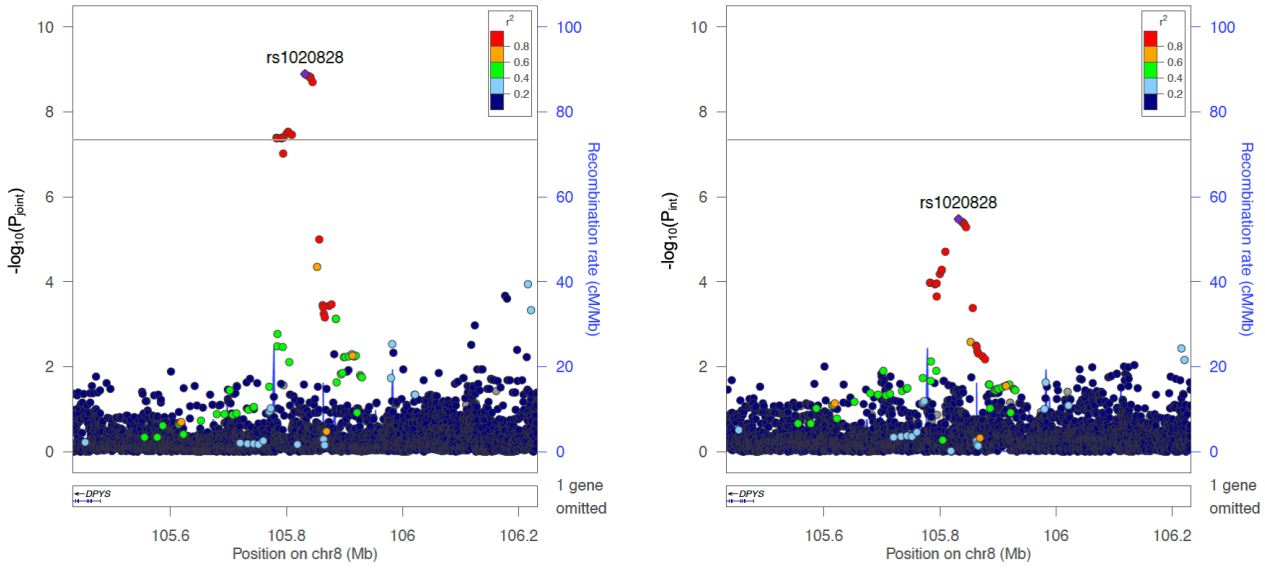
Locus 7. LTST- PP



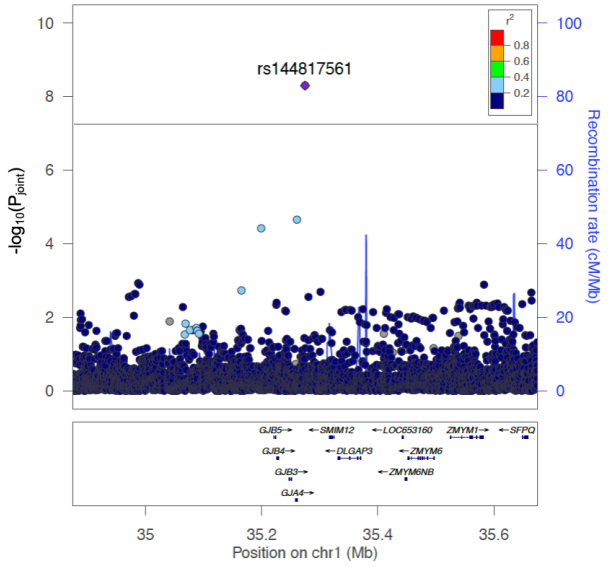
Locus 8. LTST- DBP

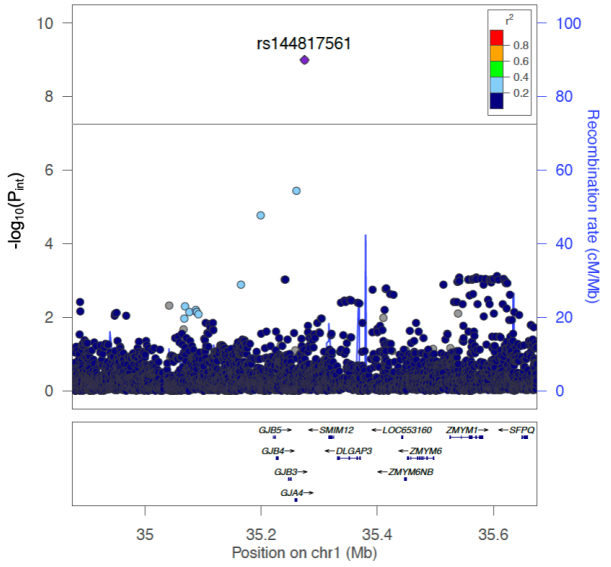


Locus 9. LTST- PP

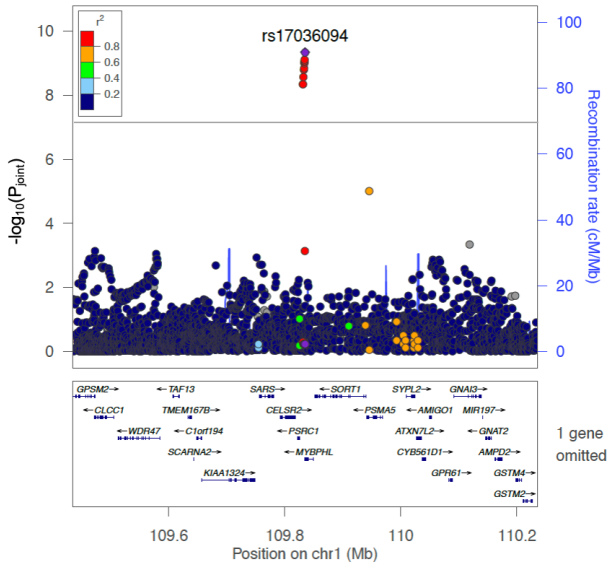


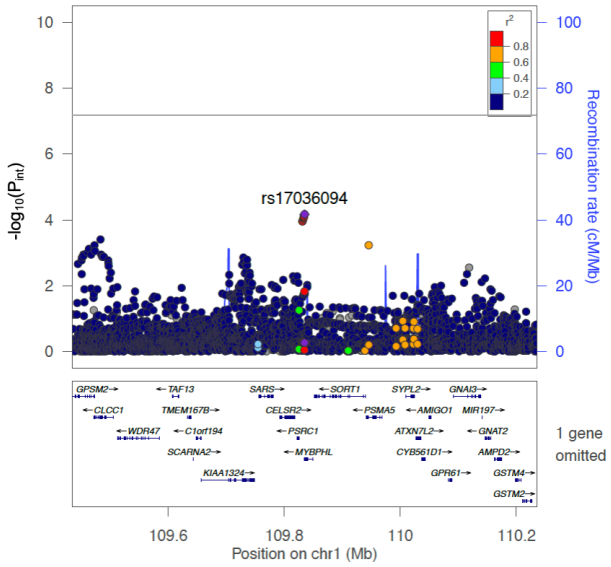
Locus 10. STST- SBP



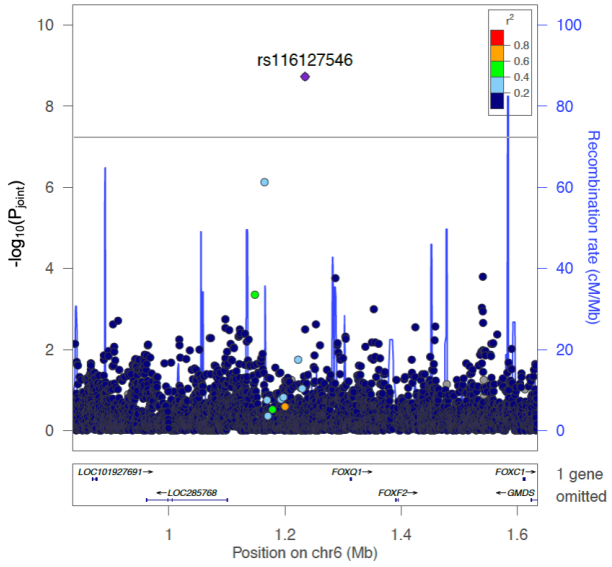


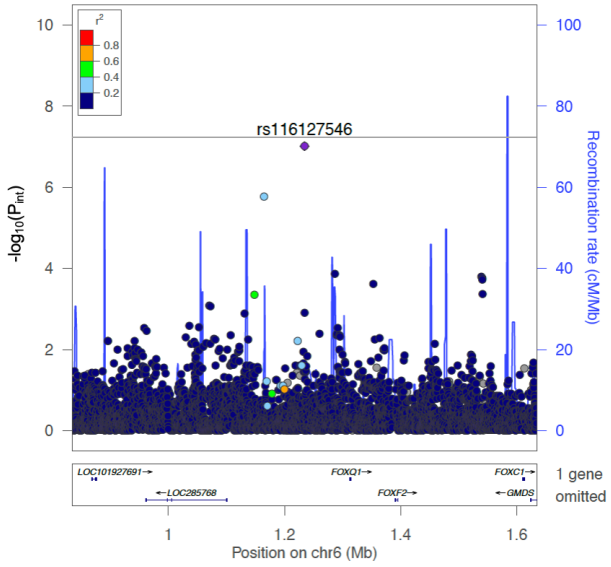
Locus 11. STST- PP



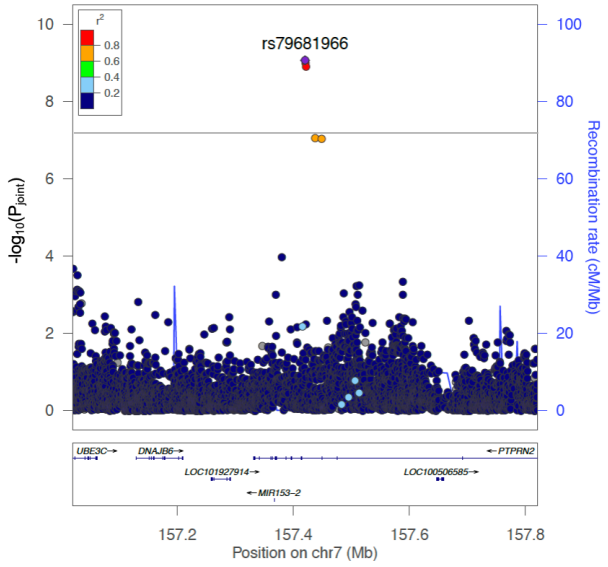


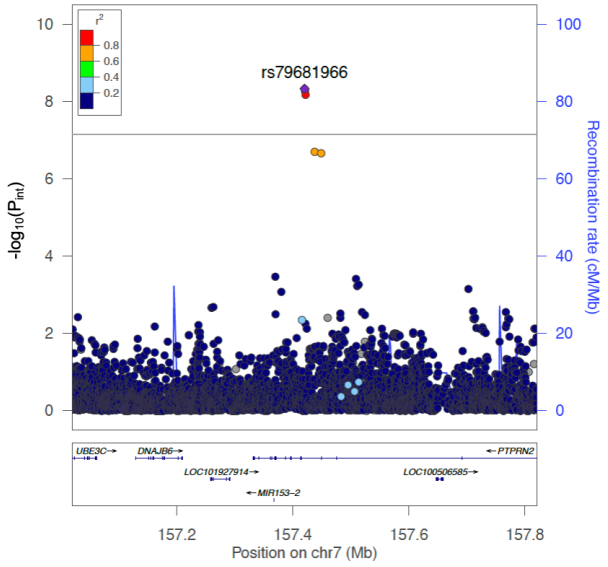
Locus 12. STST- MAP



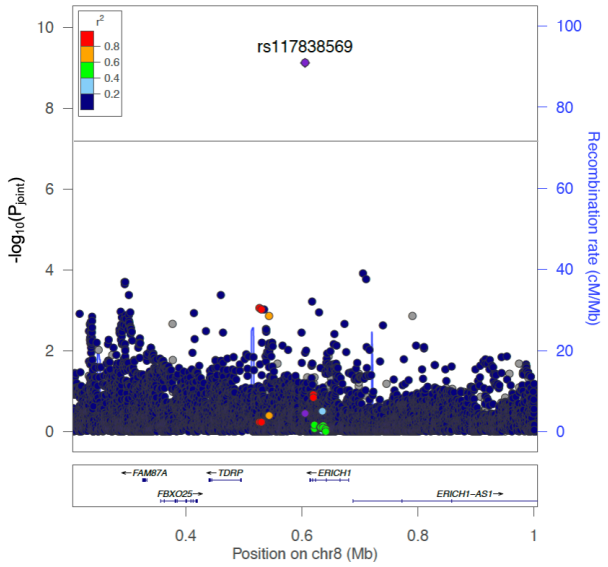


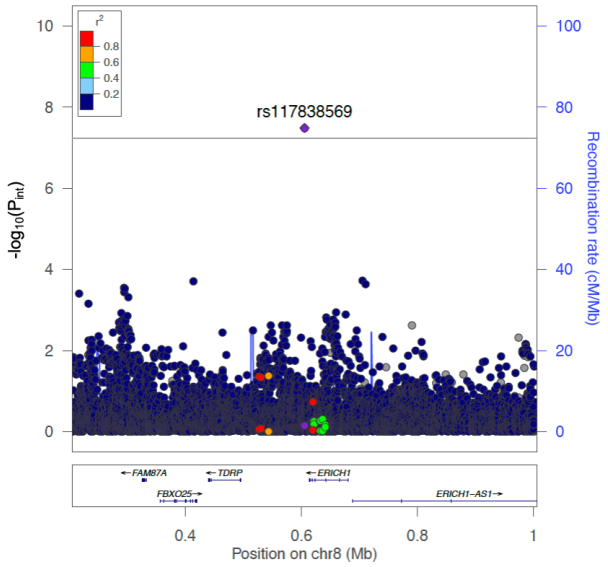
Locus 13. STST- PP



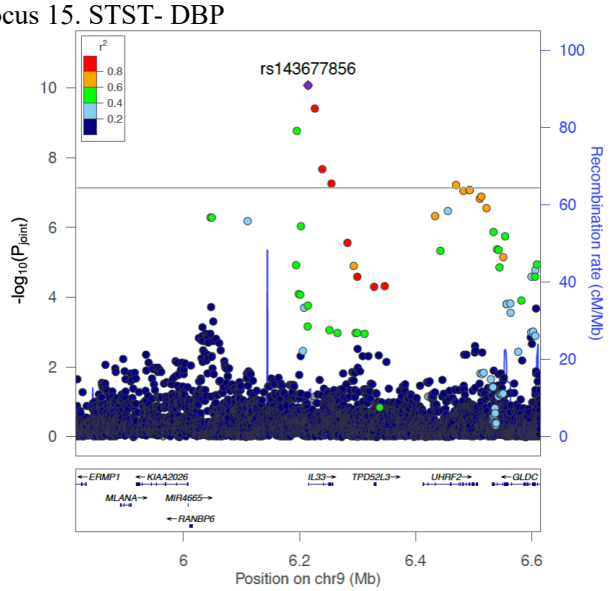


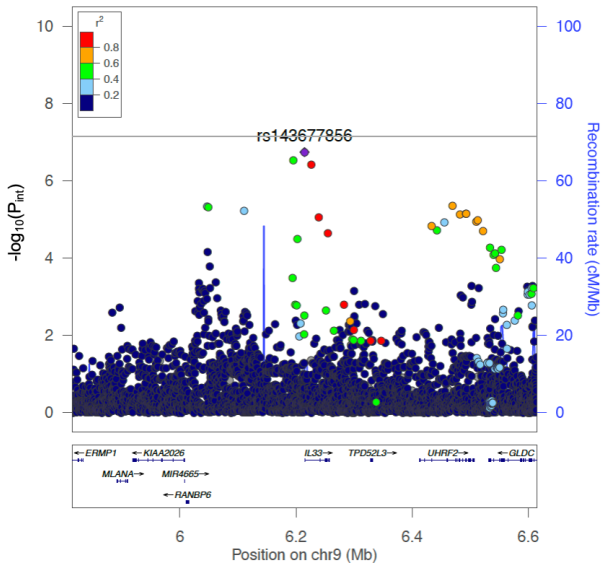
Locus 14. STST- SBP



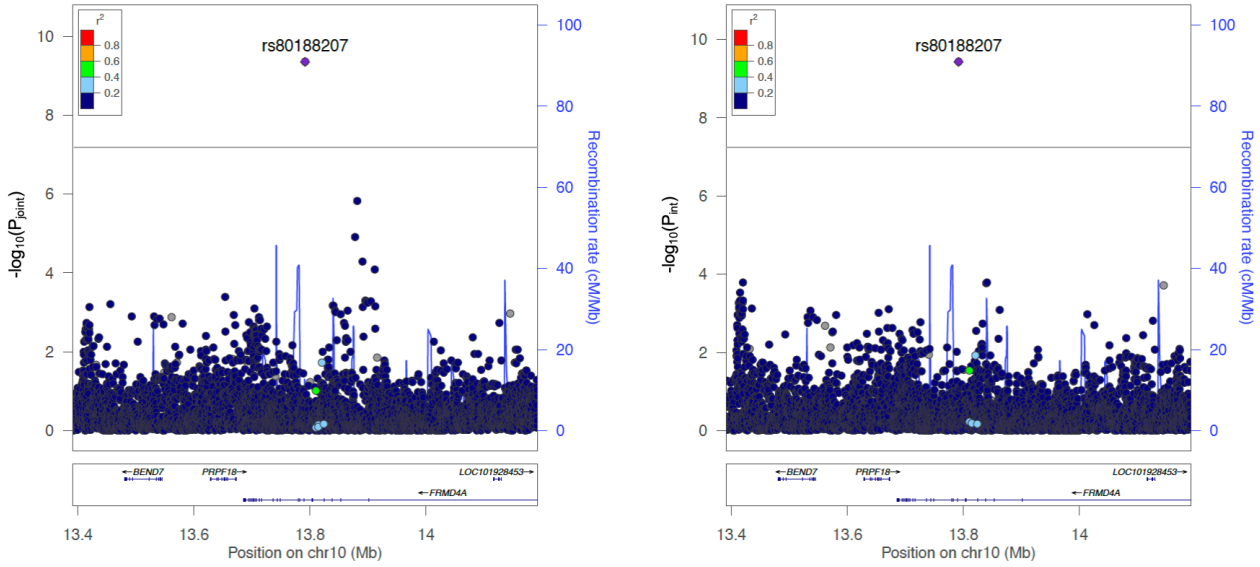


Locus 15. STST- DBP

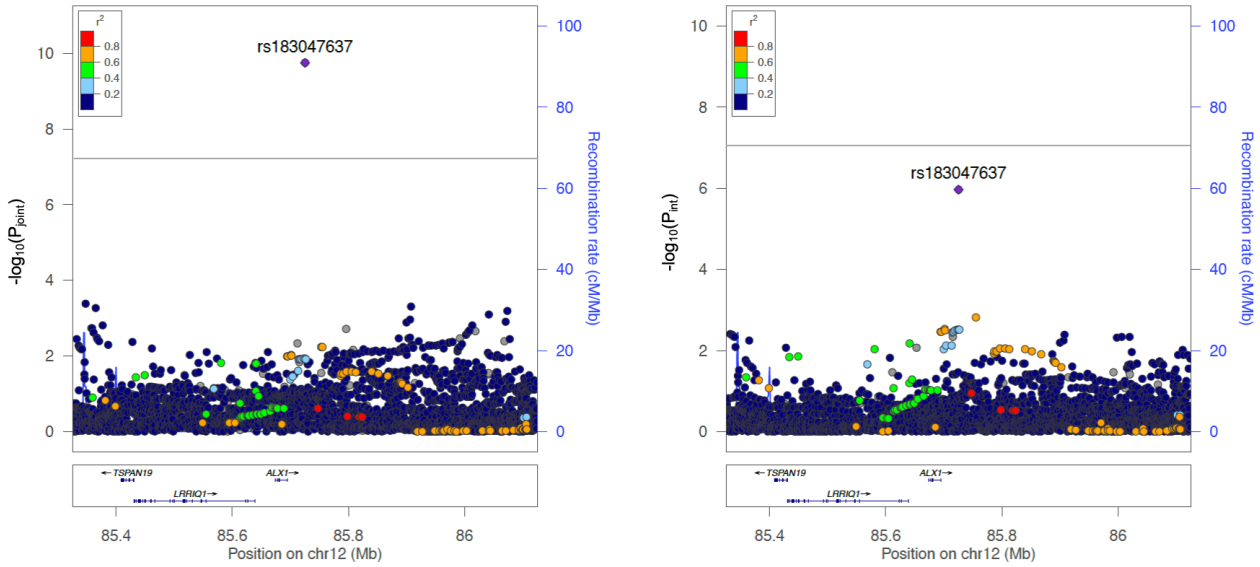




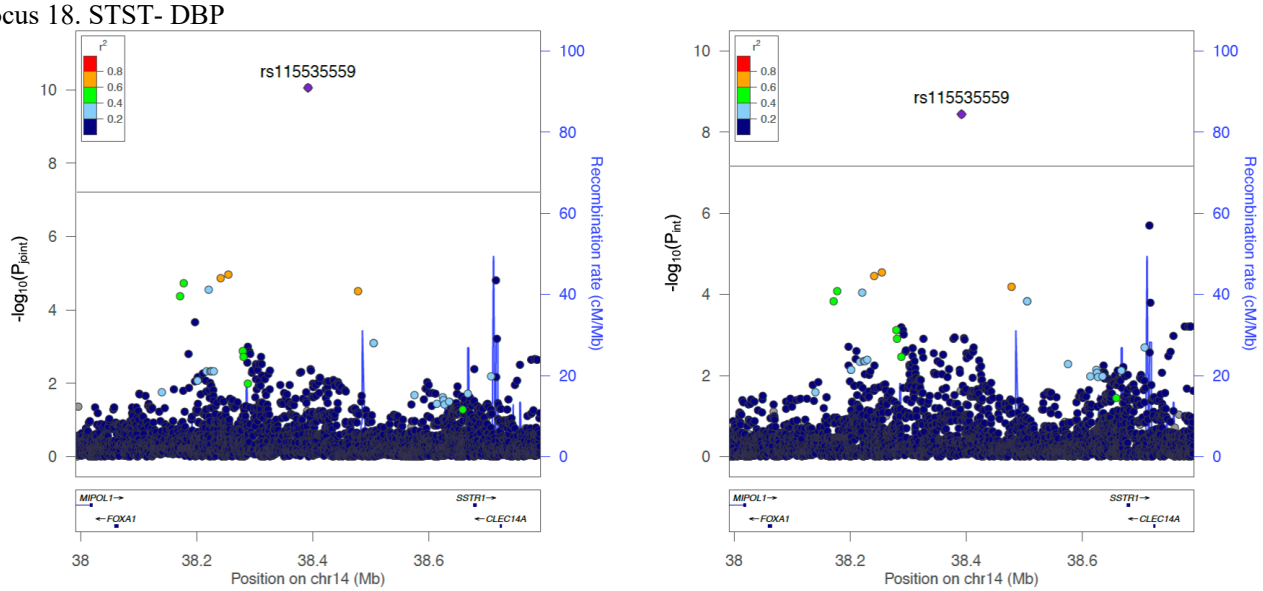
Locus 16. STST- MAP



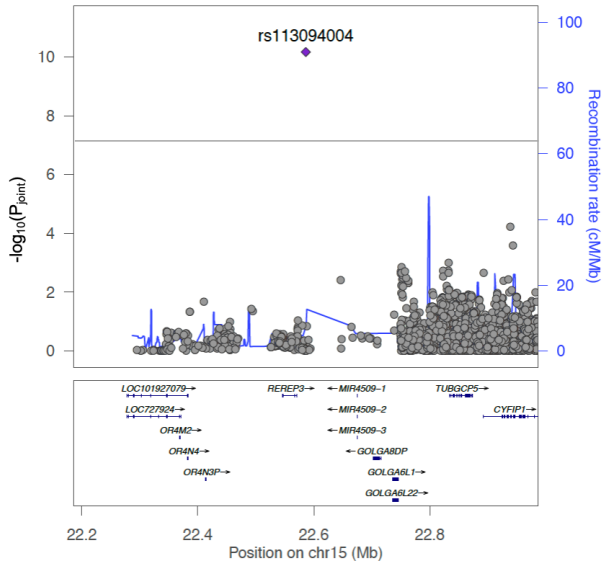
Locus 17. STST- PP

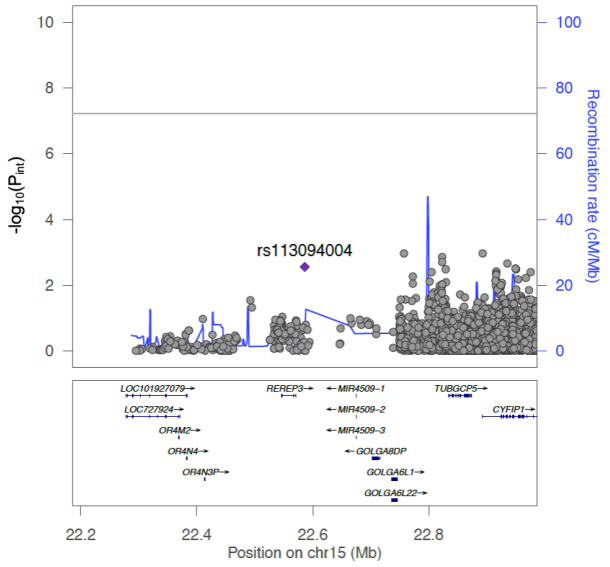


Locus 18. STST- DBP

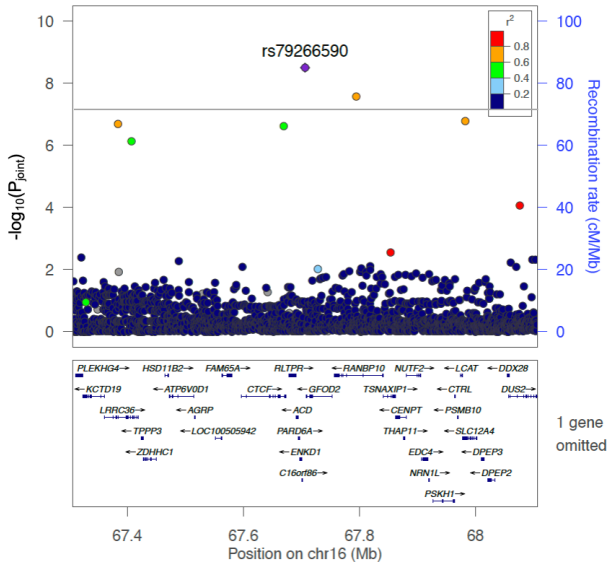


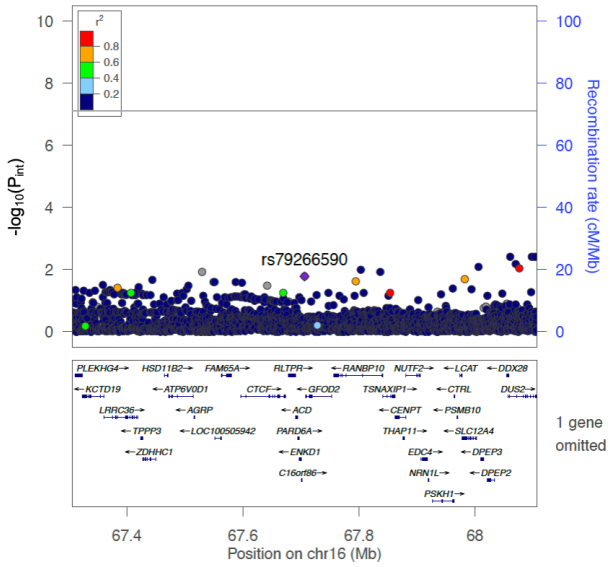
Locus 19. STST- SBP





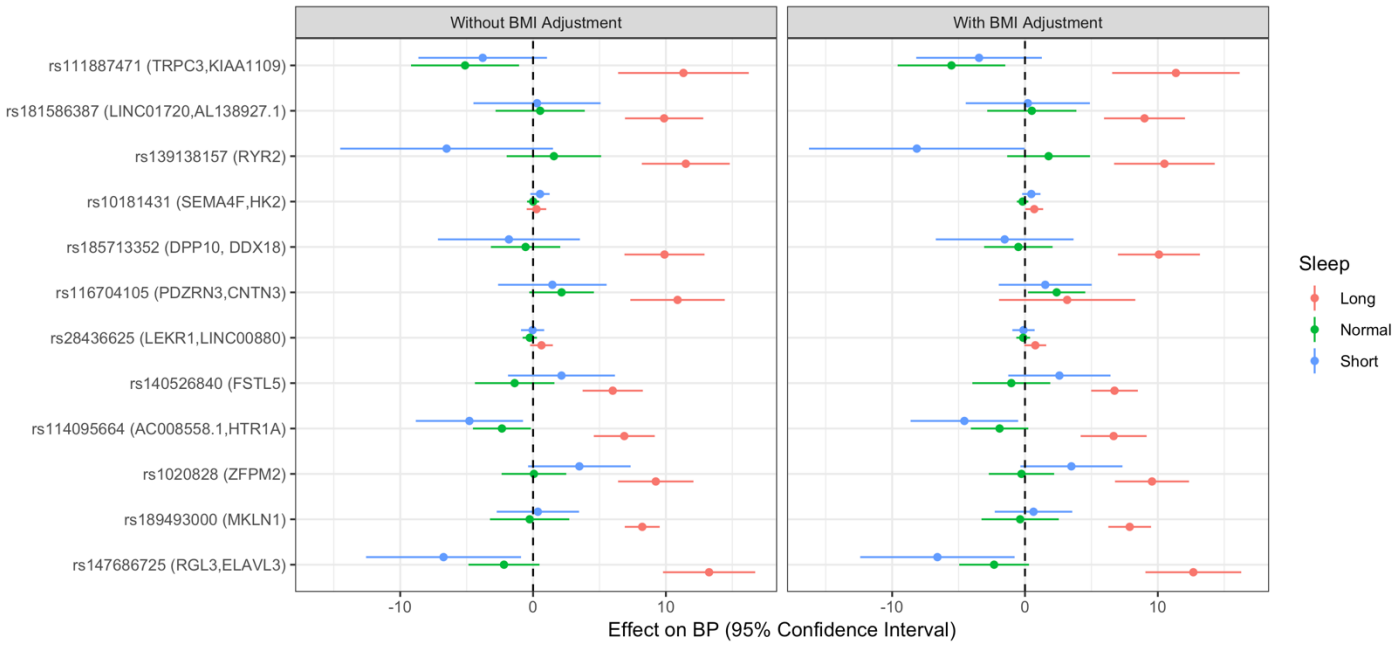
Locus 20. STST- PP





Supplementary Fig. 14. Effects on BP in long, normal and short sleep for loci identified in two-stage ancestry specific analyses and combined stage 1 and 2 analyses.

A. Loci significantly interact with LTST



B. Loci significantly interact with STST