scientific reports

Check for updates

OPEN

Multi ancestry genome wide association meta analysis of urinary aMT6s levels

Magdalena Żebrowska[1,6]✉, Ziwei Zhang[3], Gwo-Tsann Chuang[13,14], Daniel S. Evans[9], Jesse Valliere[6], Matthew Maher[6], Jie Hu[3,6,7], Rebecca Richmond[20,21,22], Constance Turman[3], Jaime E. Hart[2,12], Jacqueline Lane[6,18], Loic Le Marchand[23], Lynne Wilkens[23], Matthias Wielscher[1,8], Christopher Haiman[24], Iona Cheng[16,17], A. Heather Eliassen[2,3,25], Katie L. Stone[9,10], Gregory J. Tranah[9], Yi-Cheng Chang[11,14,15], Lorelei Ann Mucci[3,19], Eva S. Schernhammer[1,2,3] & Richa Saxena[4,5,6,7]

Melatonin regulates circadian rhythms, metabolism, and immunity. Its primary metabolite, 6-sulfatoxymelatonin (aMT6s), is a biomarker linked to cancer risk and metabolic disorders. However, genetic determinants of aMT6s remain poorly understood, with only one prior GWAS limited to an East Asian cohort. We conducted the first multi-ancestry genome-wide association meta-analysis of urinary aMT6s, integrating 11,744 participants from five cohorts: East Asians (Taiwan Biobank), European women (Nurses' Health Studies), European men (MrOS), and multiethnic participants (MEC). aMT6s was measured from overnight or first-morning urine samples. Association analyses were conducted using both ancestry-aware meta-regression (MR-MEGA) and fixed-effects meta-analysis (METAL). Polygenic risk scores (PRS) were constructed with PRS-CSx and evaluated in phenome-wide analyses in the Mass General Brigham Biobank and UK Biobank. No genome-wide significant loci were identified, and previously reported East Asian signals were not replicated. At suggestive significance, 23 loci emerged, with eight supported by both MR-MEGA and METAL. Several loci showed ancestry-specific heterogeneity, suggesting that genetic associations with urinary aMT6s may vary by population context, although limited power and cohort heterogeneity may also contribute. PRS analyses identified associations with sleep duration and metabolic traits, including type 2 diabetes, but these findings require cautious interpretation. Overall, our results suggest that urinary aMT6s is influenced by a polygenic and potentially population-dependent genetic architecture. This study provides a multi-ancestry framework for investigating melatonin-related biomarkers and highlights the importance of careful interpretation across diverse populations.

Keywords Melatonin, Sulphatoxymelatonin, GWAS, Meta-analysis, Multi-ancestry

[1]Department of Epidemiology, Center for Public Health, Medical University of Vienna, Kinderspitalgasse 15, Vienna 1090, Austria. [2]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [3]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [4]Division of Sleep Medicine, Harvard Medical School, Boston, MA 02115, USA. [5]Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Boston, MA 02115, USA. [6]Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114, USA. [7]Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA. [8]Department of Dermatology, Medical University of Vienna, Vienna, Austria. [9]Research Institute, California Pacific Medical Center, San Francisco, CA, USA. [10]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA. [11]Department of Endocrinology and Metabolism, National Taiwan University, Taipei, Taiwan. [12]Department of Environmental Health, Harvard T. H. Chan School of Public Health, Boston, MA, USA. [13]Department of Pediatrics, College of Medicine, National Taiwan University Hospital, National Taiwan University, Taipei, Taiwan, ROC. [14]Graduate Institute of Medical Genomics and Proteomics, National Taiwan University, Taipei 100, Taiwan, ROC. [15]Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan. [16]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, USA. [17]San Francisco Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA, USA. [18]Broad Institute, Cambridge, MA 02142, USA. [19]Discovery Science, American Cancer Society, Atlanta, Georgia. [20]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK. [21]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. [22]NIHR Oxford Health Biomedical Research Centre, University of Oxford,

Oxford, UK. [23]University of Hawaii Cancer Center, Honolulu, HI, USA. [24]Department of Population and Public Health Sciences, University of Southern California, Los Angeles, CA, USA. [25]Department of Nutrition, Harvard T. H. Chan School of Public Health, Boston, MA 02215, USA. ✉email: magdalena.zebrowska@meduniwien.ac.at

Melatonin, a hormone synthesized by the pineal gland, plays a critical role in regulating circadian rhythms, cellular metabolism, and immune responses. Its primary metabolite, 6-sulfatoxymelatonin (aMT6s), measured in overnight or first morning urine, serves as a stable and reliable biomarker for assessing nocturnal melatonin secretion in clinical and research settings[1]. This metabolite exhibits profound physiological and pathological associations. Elevated nocturnal urinary aMT6s levels, for instance, are positively correlated with tumor cell proliferation in gastrointestinal and lung cancer patients[2], while reduced levels in renal transplant recipients are linked to higher mortality, underscoring melatonin's potential therapeutic relevance[3]. Low levels of aMT6s are also associated with increased risks of breast[4–6], oral[7], gastric[8] and prostate cancer[9,10]. Additionally, variations in melatonin patterns, such as the timing of dim light melatonin onset (DLMO), offer valuable insights into its temporal dynamics and their connections to chronotype and aging[11]. In people with Rare Genetic Neurodevelopmental Disorders (RGND), irregularities in melatonin production play a role in circadian rhythm disruptions and associated sleep difficulties[12]. Melatonin metabolism is influenced by both environmental and genetic factors. Night-shift work, for example, can suppress melatonin production[13] while genetic variations, such as the rs10830963 G allele in the *MTNR1B* gene, have been associated with altered melatonin signaling and an increased risk of type 2 diabetes mellitus[14] and gestational diabetes mellitus (GDM)[15,16]. Disruptions in melatonin rhythms, such as those observed during night-shift work, can lead to secretion patterns that worsen metabolic dysregulation[17].

Despite its importance, our understanding of genetic determinants of melatonin secretion remains limited. A recent study in 2,373 East Asian ancestry individuals from the Taiwan Biobank did not identify any genome-wide significant loci, and reported five suggestive ($p < 5{\cdot}10^{-6}$) loci associated with urinary aMT6s[18]. Genetic variation, however, exhibits significant differences across ethnic groups, often resulting in population-specific effects. Therefore, the absence of large-scale genome-wide association studies (GWAS) of the melatonin metabolite together with limited racial and ethnic diversity in existing research represent critical limitations. This gap restricts the generalizability of findings and impedes a comprehensive understanding of melatonin's role in health across diverse populations. Furthermore, to the best of our knowledge, currently available datasets that simultaneously include comprehensive genetic information and aMT6s measurements lack sufficient sample sizes to achieve the statistical power necessary for robust genetic studies, thereby hindering advancements in this area of research.

To address these limitations, we aggregated data from the Taiwan Biobank (TWB), the Nurses' Health Study I (NHS1), the Nurses' Health Study II (NHS2), Osteoporotic Fractures in Men[19,20] (MrOS), and the Multiethnic Cohort (MEC). Beyond increasing sample size and ancestry diversity, this study extends prior work in several ways. First, we apply complementary meta-analytic approaches—fixed-effects meta-analysis and ancestry-aware meta-regression—to evaluate whether genetic associations with urinary aMT6s are broadly shared across cohorts or differ by population context. Second, we systematically compare our findings with the only prior aMT6s GWAS to assess whether previously reported suggestive loci show consistent effects across cohorts and whether evidence of heterogeneity is present. Third, we extend beyond variant discovery by constructing multi-ancestry polygenic risk scores using PRS-CSx and evaluating their phenotypic correlates in large biobank-based analyses, including PheWAS as well as targeted analyses of sleep traits. Through this framework, we aim to characterize the polygenic basis of urinary aMT6s, while also assessing the extent to which its genetic associations may be shared or population-dependent.

Results

Among the 11,744 individuals (54.5% men), genetically determined ancestry assignments identified 6,925 (59%) with European, 2,373 (20.2%) with East Asian, 1,494 (12.7%) with Japanese, 426 (3.7%) with African, 290 (2.5%) with Native Hawaiian and 226 (1.9%) with Latino ancestry. The average age of participants was 62.3 years (SD_{age} = 10 years) with MrOS cohort containing the oldest ($mean_{age}$ = 73.1 (SD_{age} = 5.6)) and Taiwan Biobank representing the youngest cohort ($mean_{age}$ = 50.8 (SD_{age} = 10.8); Table 1).

Cohort specific heritabilities and genetic correlations

We estimated cohort-specific heritabilities from cohorts' GWAS summary statistics, filtered to include only high-quality SNPs from the HapMap 3 reference panel[21]. Heritabilities ranged from low for the NHS cohort (h^2_{NHS} = 0.1582 (SE_{NHS} = 0.1683)) to moderately high for MrOS (h^2_{MrOS} = 0.4182 (SE_{MrOS} = 0.2666)) (Supplementary Table A1). Genomic inflation factors were close to 1 across all cohorts (λ_{GC} = 1.0046–1.0165), with mean chi-square statistics also close to 1 and LDSC intercepts near unity, suggesting little evidence of substantial residual inflation (Supplementary Table A1). Together, these findings indicate that while major confounding or technical inflation was unlikely, phenotype measurement heterogeneity across cohorts may still have reduced power by attenuating genetic effects. LDSC regression analysis indicated positive genetic correlation between NHS and MrOS (r_g = 0.8073), moderate positive correlation between NHS and MEC (r_g = 0.5316) and weak negative correlation between MEC and MrOS (r_g = -0.0085), but all with large standard errors relative to estimates themselves, and insignificant p-values ($p > 0.17$) (Supplementary Table A2).

Meta analysis

After cohort-specific quality controls (Supplementary Methods) and data harmonization, 27,217,987 variants met the inclusion criteria, and 2,970,850 were present in all cohorts. These common variants were used for further analyses. Both METAL- and MR-MEGA-based observed scale LD score regression heritabilities

indicated low percentage of genetically explained variance ($h_{METAL}^2 = 0.1184$; $h_{MR-MEGA}^2 = 0.085$), with relatively high standard errors ($SE_{METAL} = 0.051$; $SE_{MR-MEGA} = 0.057$; Supplementary Table A3). Neither METAL nor MR-MEGA identified genome wide significant ($p < 5 \cdot 10^{-8}$) variants (Figs. 1(a)-(b), 2(a)-(b)). Genomic control inflation was minimal ($\lambda_{GC} = 1.0165$ for METAL and 1.0255 for MR-MEGA; Table A3), consistent with the QQ plots and LD score regression intercepts near 1 (METAL: 0.9867 (SE 0.0103); MR-MEGA: 1.0055 (SE 0.0108)).

Formal power calculations based on the total meta-analysis sample size ($N = 11,744$) indicated limited power to detect small-effect loci. At the conventional genome-wide significance threshold ($p < 5 \times 10^{-8}$), power was 2.2% for variants explaining 0.10% of phenotypic variance and 27.4% for variants explaining 0.20%, while 80% power was achieved only for variants explaining approximately 0.34% of the phenotypic variance, indicating that the study was primarily powered to detect relatively large-effect variants.

When applying our less stringent suggestive level, $p < 1 \cdot 10^{-5}$, power increased but remained limited for modest-effect variants, with 16.1% power for variants explaining 0.10% of phenotypic variance, 66.8% power for variants explaining 0.20% of phenotypic variance, and 80% power reached only for variants explaining approximately 0.23% of phenotypic variance.

Combining results for METAL and MR-MEGA meta-analyses, we identified 23 suggestive genomic loci (Table 2). METAL identified 15 (Supplementary Table B1), while MR-MEGA 16 genomic loci (Supplementary Table B2), with eight genomic loci identified by both methods (Table 3). These eight suggestive genomic loci were located on *RBM6:RBM5* (rs2013208), *SOX5* (rs77480549), *FAM110B* (rs75065017), *ZIC1* (rs9990273), *PIK3CG* (rs185087), *SLIT3* (rs1875972), *PLD1* (rs13083025) and *C12orf55* (rs7137724) genes (Table 3, Supplementary Fig. B1(a)-(p)). Six of them showed homogeneous effects across participating cohorts ($I^2 \in [0,58.2]$; Table 3; Fig. 3, Supplementary Fig. B2). The other two - rs1875972 and rs7137724 - showed ancestrally heterogeneous effects, with significant ancestral heterogeneity ($P_{anc.het} \leq 0.0459$), but not significant residual heterogeneity ($P_{res.het} \geq 0.2747$, Table 3), suggesting population specific differences rather than other confounding factors.

Among eight other genomic loci identified exclusively by MR-MEGA, six (Table 2) showed significant ancestry-related heterogeneity but not residual heterogeneity (Table 2; Fig. 4) with opposite effect directions in the TWB cohort versus all other cohorts (Supplementary Fig. A3(a)-(d)) or effect directions in TWB and MEC opposite to NHS and MrOS (Supplementary Fig. B3(e)-(g)). As opposed to this, genomic loci identified exclusively with METAL exhibited more homogeneous effects across study cohorts (Supplementary Table B3, Fig. 5, Fig. B4).

Functional characterization of suggestive genomic loci

Two of our 23 identified genomic loci - rs2013208 and rs185087 - were previously reported in the literature. The one with the lowest p-values in both meta-analyses, rs2013208 associated with lower aMT6s, was previously associated with high HDL levels and coronary artery disease (CAD)[22], as well as with sex differences in serum lipid profiles[23]. The other one, rs185087 (*PIK3CG*) also associated with lower aMT6s, was previously identified as increasing CAD risk[24]. When looking at the nearest protein coding gene to each suggestive lead SNP, genes with low probability of being loss-of-function intolerant (pLI) were found both in loci with (*C12orf55*, *NEK6*) and without (*PLD1*, *COL25A1*, *MYH13*, *FAM110*) significant ancestral heterogeneity (Table 2). We observed

	TWB	NHS*	MrOS	MEC	Total
Sample size (*n*)	2,373	3,861	2,175	3,335	11,744
Men; N (%)	890 (37.51)	0 (0)	2,175 (100)	3,335 (100)	6,400 (54.5)
Women; N (%)	1,483 (62.5)	3,861 (100)	0 (0)	0 (0)	5,344 (45.5)
Age in years, mean (SD)	50.8 (10.8)	57.6 (12.8)	73.1 (5.6)	68.9 (7.7)	62.3 (10.0)
BMI, mean (SD)	24.7 (5.8)	26.61(5.6)	27.4 (3.7)	26.5 (4.0)	26.3 (4.9)
Ancestry/race					
Black; N (%)	0 (0)	0 (0)	0 (0)	436 (13.1)	436 (3.7)
East Asian; N (%)	2,373 (100)	0 (0)	0 (0)	0 (0)	2,373 (20.2)
Japanese American; N (%)	0 (0)	0 (0)	0 (0)	1,494 (44.8)	1,494 (12.7)
Latino; N (%)	0 (0)	0 (0)	0 (0)	226 (6.8)	226 (1.9)
Native Hawaiian; N (%)	0 (0)	0 (0)	0 (0)	290 (8.7)	290 (2.5)
White; N (%)	0 (0)	3,861 (100)	2,175 (100)	889 (26.7)	6,925 (59)
Outcome measures characteristics					
aMT6s (ng/mL)					
Median (IQR)	20.41 (11.88; 30.19)	24.99	7.58	16.9	
Mean (SD)		44.3(5.65)	9.83 (7.98)	23.5 (23.9)	
log(aMT6s/Cr)					
Median (IQR)	2.83 (2.33; 3.31)		2.14	-1.45	
Mean (SD)		3.11 (1.16)	2.06 (0.85)	-1.59 (1.05)	

Table 1. Descriptive characteristics of study participants with urinary 6-sulfatoxy melatonin and genetic data available in Taiwan Biobank (TWB), the Nurses' Health Study I (NHS1), the Nurses' Health Study II (NHS2), MrOS and Multiethnic Cohort (MEC). *NHS : pooled NHS1 and NHS2 cohorts.

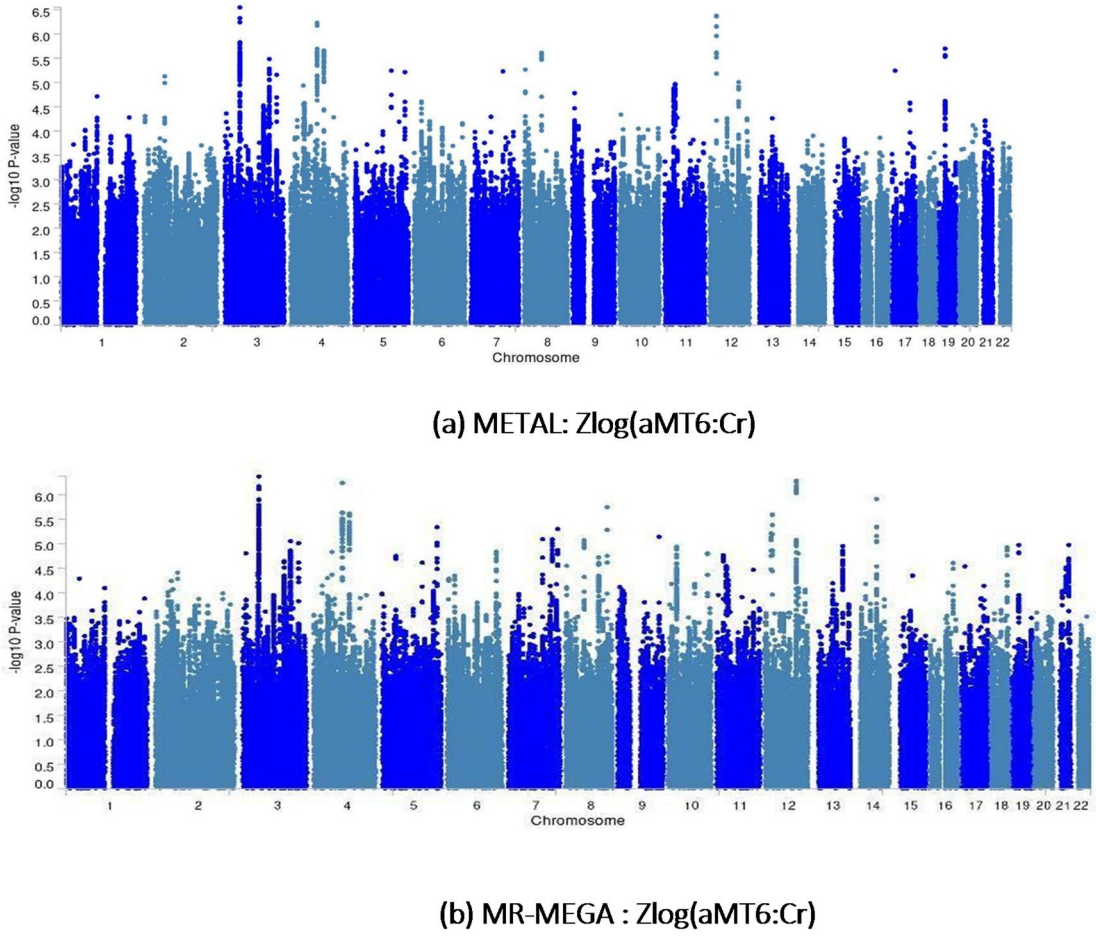

Fig. 1. Manhattan plots of the GWAS meta-analysis results across 11,744 study participants obtained using METAL inverse variance fixed-effects model with genomic control correction or MR-MEGA meta-regression test for Zlog(aMT6:Cr). The x-axis represents chromosomes and base pair positions of variants tested in the meta-analysis, while the y-axis shows -$\log_{10}$ p-values from the two-sided variant association test.

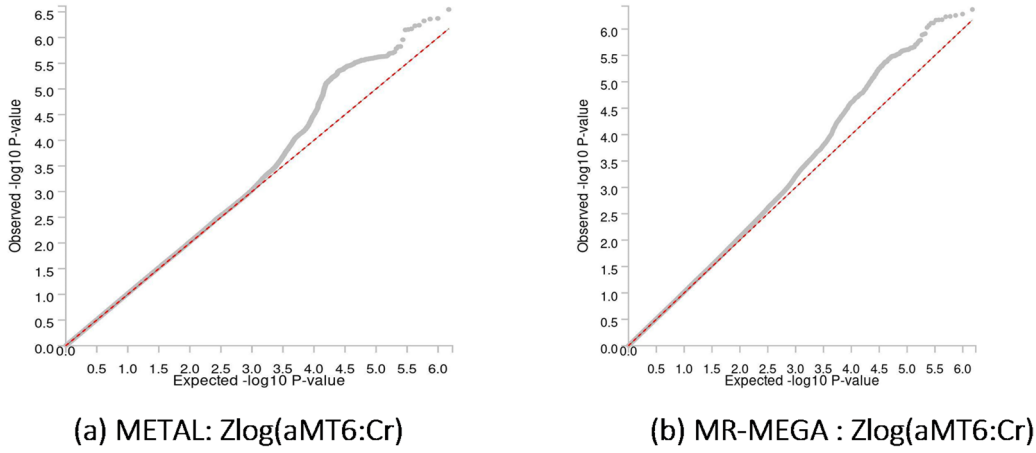

Fig. 2. QQ plots of the GWAS meta-analysis results for Zlog(aMT6:Cr) across 11,744 study participants (a) METAL inverse variance fixed-effects model with genomic control correction; (b) MR-MEGA meta-regression test.

four genes with the pLI > 0.99 (*RBM6: RBM5*, *ACTN1*, *SOX5* and *SLIT3*) and one (*ZIC1*) with pLI = 0.825 (Table 2). Genomic loci placed on these five genes with high pLI were identified as suggestive by both METAL and MR-MEGA.

MAGMA gene set and tissue expression analysis

We used Functional Mapping and Annotation (FUMA[25,26] to functionally annotate results of our meta-analyses. We ran multi-marker analysis of genomic annotation (MAGMA[27] for gene ontology, tissue level and single cell expression data. For both METAL and MR-MEGA results, input SNPs were mapped to 15,857 protein coding genes and the genome-wide significance was defined at $p = 3.153 \cdot 10^{-6}$. No gene ontology terms were significant at this level for METAL (Supplementary Table B5), whereas for the MR-MEGA, "gomf_adenyl_nucleotide exchange_factor_activity" gene set was significant (adjusted p_{bon} = 0.021; Supplementary Table B6). Tissue expression analyses did not indicate any significant gene expression, neither with METAL nor with MR-MEGA (Supplementary Figs. B5-B6). Pathway enrichment tests implemented in GENE2FUNC did not indicate any significantly enriched differentially expressed genes ($p_{bon} < 0.05$) in METAL or MR-MEGA (Supplementary Fig. B7). At the gene set level, however, both methods indicated significant enrichment of input genes in positional gene set chr3p21 ($p_{MR-MEGA} = 1.81 \cdot 10^{-19}$; $p_{METAL} = 1.59 \cdot 10^{-18}$, Supplementary Tables B4-B5, Supplementary Fig. B11), canonical pathways (MsigDB c2) in M22069 ("biocarta_msp_pathway" ; $p_{MR-MEGA} = 2.67 \cdot 10^{-2}$; $p_{METAL} = 3.37 \cdot 10^{-2}$, Supplementary Tables B4-B5) and M27744 ("reactome_signaling_ by_mst1"; adjusted $p_{MR-MEGA} = 2.67 \cdot 10^{-2}$; $p_{METAL} = 3.37 \cdot 10^{-2}$, Tables B4-B5) human gene sets[28]. Both METAL and MR-MEGA prioritized genes were enriched in 19 GWAS catalog reported gene sets with a substantial overlap between the two methods, representing gene sets associated with cognition, physical activity, sleep, mental health, brain morphology, socioeconomic factors, and medical conditions (Supplementary Tables B4-B5, Supplementary Fig. B12).

Comparison with previously published results

A previous genome-wide association study (GWAS) analysis conducted on 2,373 individuals from the TWB cohort identified five suggestive loci ($p < 5 \cdot 10^{-6}$) associated with urinary Log(aMT6s: Cr)[18]. However, none of these loci were detected in our meta-analysis (Supplementary Table C1). Among these variants, rs142037747 (*GALNT15*), is exceptionally rare, particularly in non-East Asian populations (EAF = 0.0002 in EUR, EAF = 0.001 in remaining ancestries; Supplementary Table C4). Due to its low minor allele frequency (MAF < 1%), this variant was excluded during quality control from the NHS, MrOS, and MEC cohort data. Another variant, rs7571016 (*GALNT13*) was absent in the MEC cohort GWAS results, likely due to platform-specific genotyping (TaqMan Assay Panel). The remaining three suggestive variants, rs17681554, rs9645614 and rs6451653, exhibited high heterogeneity in effect estimates across the study cohorts (Supplementary Table C5) with the TWB cohort effects differing markedly from those observed in other cohorts (Supplementary Fig.C1(a)-(e)). We additionally performed restricted meta-analyses in the European-ancestry cohorts (NHS + MrOS) and the male-only cohorts (MrOS + MEC). Neither analysis provided evidence that the previously reported Taiwan Biobank suggestive variants generalized to these subsets, supporting limited generalizability across ancestries and providing no strong evidence for sex-specific effects (Supplementary Table C11). Given the substantially smaller sample sizes, these analyses were underpowered for robust locus discovery and were therefore considered exploratory sensitivity analyses rather than primary results. The corresponding comparisons and shareable LocusZoom links are provided in Supplementary Table C11.

A recent meta-analysis[29](with a total sample of 8,011 individuals) of three European cohorts' GWASs of 54 urinary metabolites identified three variants associated with Tryptophan (Supplementary Table C6), playing a key precursor role in melatonin production through the serotonin-melatonin pathway. These three variants were present in three of our cohorts each and were not significant in either METAL (Supplementary Table C7) or MR-MEGA (Supplementary Table C8) meta-analysis ($p > 0.46$). Also, no association was found in our study for any other of the 54 metabolites analyzed in Valo et al.[29] ($p > 0.06$; Supplementary Table C9-C10).

Polygenic risk score and PheWAS analyses

We used PRS-CSx (Supplementary Methods) on our meta-analysis cohorts' data to estimate SNPs' weights for the aMT6s polygenic risk score (PRS), and these weights were then applied to the UK Biobank and MGBB data to derive aMT6s-PRSs. Ancestry-specific (EUR, EAS) and cross ancestry (META) PRSs were generated using ancestry-matched LD reference panels and SNPs with $p < 0.05$. Scores were calculated using PLINK[30] and normalized (z-scored) within each ancestry group or across all ancestries for META-PRS. PRSs were constructed as genetic proxies for nocturnal melatonin secretion. In the NHS1/NHS2 data, after adjustment for age, principal components, cohort, batch and genotyping platform, the ancestry specific EUR-PRS explained an incremental in-sample R^2 of 12.7% overall (16.7% in NHS1 and 8.5% in NHS2; Supplementary Table D14), whereas the META-PRS explained an incremental in-sample R^2 of 21.6% overall (36.3% in NHS1 and 7.6% in NHS2; Supplementary Table D15). These values represent in-sample explained variance and may therefore overestimate out-of-sample predictive performance. Since nocturnal melatonin secretion was not available in either UK Biobank or MGBB, the corresponding variance-explained metric could not be estimated in those cohorts.

Phenome-wide association studies (PheWAS) of aMT6s-PRSs were conducted in MGBB (EUR ancestry only) and UKBB (EUR, EAS, META; Supplementary Methods).

Our PheWAS results, in which the aMT6s PRSs were used as an exposure, should be interpreted as downstream association analyses rather than as direct predictive validation of the PRS across the tested phenotypes. In MGBB, of 1,657 disease outcomes tested, 101 associations were significant after Bonferroni correction (Supplementary Table D1). These included among others genitourinary (e.g. symptoms involving urinary system ($p\_adj = 3.6 \cdot 10^{-4}$)), dermatologic (e.g. cellulitis ($p\_adj = 6.0 \cdot 10^{-3}$)) endocrine/metabolic (e.g. obesity ($p\_adj = 2.5 \cdot 10^{-3}$), type 2 diabetes ($p\_adj = 0.025$)), circulatory system (e.g. chronic pulmonary heart

Lead SNP							MRMEGA					METAL					EAF				Direction TWB/ NHS/ MrOS/ MEC
GL	rsID	Nearest Gene	CHR	BP	NEA	EA	Effect	StdErr	*P*.value	Panc.het	Pres.het	Effect	StdErr	*P*.value	*P*.Het	pLI	TWB	NHS	MrOS	MEC	
1	**rs2013208**	*RBM6:RBM5*	3	50,129,399	C	T	-0.065	0.020	**4.26E-07**	9.54E-02	0.181	-0.074	0.014	**2.84E-07**	1.04E-01	**1.000**	**0.858**	0.496	0.514	0.679	-/-/-/-
2	**rs7137724**	*C12orf55*	12	97,265,256	C	T	0.057	0.010	**5.20E-07**	**2.44E-03**	0.676	0.070	0.016	**9.87E-06**	**1.93E-02**	0.000	**0.849**	0.717	0.722	0.811	-/+/+/+
3	rs342460	*AFF1*	4	88,059,242	A	G	-0.078	0.010	**5.76E-07**	5.24E-02	0.697	-0.083	0.017	**6.76E-07**	2.15E-01	0.477	0.181	0.201	**0.210**	0.173	-/-/-/-
4	**rs8017601**	*ACTN1*	14	69,380,814	A	G	-0.040	0.037	**1.22E-06**	**2.12E-07**	0.225	-0.016	0.030	5.91E-01	**1.56E-06**	**1.000**	0.051	0.037	0.034	**0.084**	+/-/-/-
5	rs13268851	*PVT1*	8	129,100,295	G	A	0.026	0.003	**1.79E-06**	**4.97E-07**	0.952	0.015	0.014	2.77E-01	**1.37E-05**	*	0.288	**0.406**	**0.406**	0.326	+/-/-/+
6	**rs2526452**	*COL25A1*	4	110,103,895	C	G	-0.065	0.024	**2.45E-06**	7.33E-02	0.065	-0.067	0.014	**2.37E-06**	**3.45E-02**	0.000	**0.738**	0.675	0.683	0.681	+/-/-/-
7	**rs77480549**	*SOX5*	12	24,107,330	T	C	-0.123	0.031	**2.53E-06**	8.90E-01	0.460	-0.119	0.024	**4.24E-07**	6.69E-01	**0.999**	**0.286**	0.013	0.014	0.146	-/-/-/-
4	rs372896	*RP11-723P16.3*	14	69,324,446	A	C	0.029	0.025	**4.52E-06**	**3.97E-06**	0.457	0.051	0.028	6.84E-02	**4.57E-05**	*	0.321	0.324	0.333	**0.373**	+/-/-/-
8	**rs1875972**	*SLIT3*	5	168,288,992	G	A	-0.060	0.016	**4.62E-06**	4.59E-02	0.275	-0.063	0.014	**6.11E-06**	9.05E-02	**0.992**	0.127	0.127	0.133	**0.139**	-/-/-/-
9	rs13236792	*XRCC2*	7	152,394,108	T	C	-0.016	0.007	**5.04E-06**	**8.49E-07**	0.871	-0.008	0.020	6.97E-01	**2.10E-05**	*	**0.396**	0.184	0.180	0.288	-/+/+/-
10	rs7315106	*RP11-405A12.2*	12	20,025,120	G	A	0.037	0.013	**6.19E-06**	**5.76E-06**	0.478	0.029	0.016	6.57E-02	**6.89E-05**	*	**0.366**	0.258	0.258	0.249	-/+/+/-
11	rs10986249	*NEK6*	9	126,918,060	C	G	-0.054	0.010	**7.20E-06**	**1.81E-03**	0.644	-0.055	0.015	2.07E-04	1.42E-02	0.039	0.066	0.257	**0.264**	0.177	+/-/-/-
12	**rs185087**	*PIK3CG*	7	106,489,559	T	C	-0.062	0.029	**8.12E-06**	9.56E-02	0.104	-0.077	0.017	**5.91E-06**	6.62E-02	0.000	**0.184**	0.140	0.146	0.160	-/-/-/-
13	rs2044251	*AC091736.1*	7	135,507,742	A	T	0.046	0.013	**8.18E-06**	**4.77E-05**	0.606	0.048	0.018	8.99E-03	**5.65E-04**	*	0.049	**0.080**	0.078	0.068	-/+/+/+
14	**rs75065017**	*FAM110B*	8	59,081,401	A	G	-0.130	0.019	**8.57E-06**	3.48E-01	0.608	-0.123	0.026	**2.46E-06**	6.01E-01	0.173	**0.500**	0.412	0.412	0.360	-/-/-/-
15	**rs9990273**	*ZIC1*	3	147,267,582	G	C	0.061	0.014	**8.95E-06**	2.46E-01	0.357	0.062	0.013	**3.28E-06**	3.37E-01	0.825	0.248	**0.437**	**0.437**	0.314	+/+/+/+
16	**rs13083025**	*PLD1*	3	171,487,632	G	C	0.059	0.014	**9.79E-06**	1.08E-01	0.363	0.063	0.014	**6.95E-06**	2.05E-01	0.000	**0.858**	0.496	0.514	0.679	-/+/+/+
17	rs342462	*AFF1*	4	88,057,481	G	A	-0.101	0.010	**5.85E-07**	8.23E-01	0.886	-0.083	0.017	**5.90E-07**	2.93E-01	0.477	0.174	0.200	0.209	0.178	-/-/-/-
18	rs56169609	*AC007204.2*	19	20,091,038	A	C	-0.079	0.006	1.06E-05	6.18E-02	0.884	0.072	0.015	**2.01E-06**	3.04E-01	*	0.136	0.358	**0.375**	0.224	+/+/+/+
19	rs2704102	*COL25A1*	4	110,103,485	C	T	-0.065	0.024	**2.50E-06**	8.26E-02	0.056	-0.067	0.014	**2.19E-06**	**3.31E-02**	0.000	**0.738**	0.683	0.683	0.682	+/-/-/-
20	rs2725643	*RP11-281H11.1*	8	5,552,149	A	G	-0.062	0.029	2.85E-05	4.99E-01	**0.012**	-0.067	0.015	**5.42E-06**	9.29E-01	*	0.240	**0.316**	0.312	0.304	-/-/-/-
21	rs10040924	*HMGB1P29*	5	123,466,039	C	T	-0.067	0.007	2.44E-05	1.00E+00	0.795	-0.061	0.014	**5.70E-06**	**2.62E-02**	*	**0.628**	0.500	0.521	0.604	-/-/+/-
22	rs12948227	*MYH13*	17	10,265,366	C	T	-0.080	0.017	2.89E-05	7.23E-01	0.429	-0.079	0.017	**5.71E-06**	6.16E-01	0.000	0.117	0.224	**0.225**	0.144	-/-/-/-
23	rs73934335	*AC127383.1*	2	68,657,017	G	A	0.075	0.021	3.88E-05	6.64E-01	0.176	-0.104	0.023	**7.45E-06**	9.62E-01	*	0.026	**0.130**	0.120	0.077	-/-/-/-

Table 2. Meta analyses results for lead SNPs. Bolded are (i) lead variants with p values < 10^{-5} (suggestive) in meta-analyses by both METAL and MRMEGA (ii) p values < 10^{-5} (suggestive) from association tests; (iii) p-values for heterogeneity tests <0.05 (significant); (iv) pLI scores >0.99 (putative causal) and (v) effect allele frequencies (EAF) in the cohort with the highest EAF. GL- genomic loci, CHR-Chromosome, BP- base pair, NEA- non-effect allele, EA- effect allele, Effect- association test effect size estimate, StdErr- standard error of the effect size estimate, P.value- p value for the two sided association test, Panc.het- p value for the two sided ancestral heterogeneity test (chi-square test with 1df), Pres.het- p value for the two sided residual heterogeneity test (chi-square test with 2df), P.Het - p value from the two sided test for heterogeneity (chi-square test with 3df), pLI - probability of being loss-of-function-intolerant for the nearest coding gene, EAF - effect allele frequency, TWB- Taiwan Biobank, NHS - merged NHS1 and NHS2 cohorts, MEC - Multiethnic cohort; * pLI score was not available for this gene (gnomAD v2.1.1).

disease (p_adj = 0.013)) or respiratory (e.g. septal deviations/turbinate hypertrophy (p_adj = 0.044)) diseases. Further, in the cross-ancestry UKB sample, both the EUR-PRS (Supplementary Table D2) and META-PRS (Supplementary Table D3) showed significant associations (Bonferroni corrected $p < 0.05/1{,}509 = 3.3{\cdot}10^{-5}$) with phenotypes including hereditary hemolytic anemias (p_adj = $1.3{\cdot}10^{-29}$), sickle cell anemia (p_adj = $9.3{\cdot}10^{-22}$), vitamin D levels (p_adj = $2.7{\cdot}10^{-44}$), type 2 diabetes (p_adj = $9.8{\cdot}10^{-8}$) and sleep duration (p_adj = $8.0{\cdot}10^{-9}$). When restricted to European-ancestry UKB sample, only the META-PRS (Supplementary Table D4) was significantly associated with psoriasis and related disorders (p_adj < 0.017), while the EUR-PRS (Supplementary Table D5) and EAS-PRS (Supplementary Table D6) showed no significant associations.

Except from the disease outcomes included in the pheWAS analysis, we further examined associations of ancestry-specific (East Asian (EAS) and European (EUR)) and cross-ancestry (META) polygenic risk scores (PRSs) with four self-reported sleep related traits (sleep duration (in hours), short sleep duration, morningness and blood vitamin D levels; see Supplementary Tables D7-D13 for details). When all UK Biobank participants were included in the sample, both the z-scored East Asian-specific PRS (EAS-zPRS) and the European-specific PRS (EUR-zPRS) showed significant (Bonferroni adjusted for 3 PRSs·4 sleep traits; $p < 0.05/12 = 0.004$) associations with sleep duration ($p < 2.81 \times 10^{-3}$) short sleep duration ($p < 7.97 \times 10^{-6}$), blood vitamin D levels ($p < 2.40 \times 10^{-6}$) and morningness ($p < 7.32 \times 10^{-7}$) (Supplementary Table D7). The cross-ancestry PRS (META-zPRS) was significantly associated with all these traits except morningness ($p = 0.716$) (Supplementary Table D7). When we restricted to UK Biobank participants of European ancestry, the associations for the EUR-zPRS remained significant. In contrast, for the EAS-zPRS, only the association with morningness remained significant. For the META-zPRS, all associations except that with overall sleep duration remained significant (Supplementary Table D8). None of these associations remained significant when East Asian (EAS, Supplementary Table D9), Admixed American (AMR, Supplementary Table D11), Central/South Asian (CSA, Supplementary Table D12), or African (AFR, Supplementary Table D13) ancestry groups were used as the validation set. However, in the Middle Eastern (MID) group (Supplementary Table D10), the EAS-zPRS showed a significant association with blood vitamin D levels ($p = 0.0012$).

Genetic correlation with sleep related and cardiometabolic traits

We used linkage disequilibrium score regression (LDSC) analysis to assess genetic correlations between urinary aMT6s and both sleep-related and cardiometabolic traits. Since LDSC estimates can be biased when applied across ancestries with different LD patterns and allele frequencies, we restricted the present analyses to European-ancestry GWAS datasets and used European reference LD scores. Sleep analyses included self-reported and actigraphy-derived sleep traits in the UK Biobank, including relative amplitude[31]. Type 2 diabetes summary statistics (BMI-adjusted) were obtained from OpenGWAS (ebi-a-GCST007516[32–34]), for essential hypertension we used IEU OpenGWAS (ukb-a-531[34,35]) and for GDM, the NHGRI-EBI GWAS Catalog[36] study GCST90044484[37].

No strong genetic correlations were observed for the sleep traits (r_g ranging from −0.3 to 0.2), and none reached statistical significance (smallest $p = 0.09$). (Table E1, Figure E1).

Similarly, there was no evidence of significant genetic correlation with type 2 diabetes (r_g=0.0665, SE = 0.1542, $p = 0.666$) or hypertension (r_g = -0.2711, SE = 0.3396, $p = 0.4248$). For GDM, LDSC did not yield an interpretable genetic correlation estimate, as the GDM summary statistics showed an out-of-bounds heritability estimate (h2 = -0.0024, SE = 0.0018) with mean χ^2 below 1 (0.9893). Overall, our LDSC analyses for sleep related traits, T2D, hypertension, and GDM do not support significant genome-wide genetic correlation with urinary aMT6s.

Discussion

In this study, we conducted the largest to date GWAS meta-analysis of urinary 6-sulfatoxymelatonin (aMT6s) levels across five cohorts, including participants from diverse populations. Despite aggregating data from over 11,000 individuals, no genome-wide significant variants ($p < 5 \times 10^{-8}$) were detected, indicating that melatonin secretion is likely influenced by a complex polygenic architecture with small effect sizes. We identified 23 suggestive genomic loci, with 8 loci detected consistently across both METAL and MR-MEGA approaches. Two suggestive loci were previously implicated in coronary artery disease (rs2013208, rs185087) and one in lipid metabolism (rs2013208), suggesting potential pleiotropic effects linking melatonin metabolism to cardiometabolic pathways.

A recent large-scale GWAS of pineal gland volume[38] identified 34 genome-wide significant loci and highlighted robust genetic contributions to melatonin-related neuroanatomy. Across both our melatonin metabolite GWAS and the pineal gland volume GWAS, several genes of interest overlap, pointing to shared biological mechanisms. The *RBM6:RBM5* locus, which encodes RNA-binding proteins central to splicing regulation, emerges in both studies and highlights RNA processing pathways relevant to melatonin metabolism and pineal morphology. *COL25A1*, implicated in amyloid plaque formation and neurodegeneration, links melatonin biology to brain aging and disease vulnerability. Finally, signals in the *ZIC* gene family (*ZIC1* in our study, *ZIC4* in the pineal gland study) underscore the contribution of neurodevelopmental transcription factors to melatonin-related traits. Taken together, these convergences suggest common genetic pathways influencing both structural and metabolic aspects of melatonin biology.

A central finding of our study is the presence of significant ancestry-related heterogeneity in genetic associations. Several loci exhibited differing, and in some cases opposing effect estimates across populations, particularly between the Taiwan Biobank (EAS) and the other cohorts. However, these patterns should be interpreted cautiously. Although they may reflect population-specific genetic influences on melatonin metabolism, they may also arise from ancestry-related differences in allele frequency and LD structure, unequal sample sizes across ancestry groups, and limited precision of effect estimates in the smaller ancestry strata. In the present study, the majority of participants were of European or East Asian ancestry, whereas other ancestry groups were represented by much smaller numbers, making heterogeneity signals in those strata less stable

	Suggestive genomic loci				MRMEGA					METAL					
	rsID	chr	pos	nearest Gene	*P*.value	beta	se	Panc.het	*P* res.het	*P*.value	beta	se	I2	*P*. Het.	Direction TWB/ NHS/ MrOS/ MEC
1	rs2013208	3	50,129,399	*RBM6:RBM5*	**4.26E-07**	-0.0654	0.0199	0.0954	0.1805	**2.84E-07**	-0.0739	0.0144	51.4	0.1036	-/-/-/-
2	rs77480549	12	24,107,330	*SOX5*	**2.53E-06**	-0.1226	0.0308	0.8902	0.4600	**4.24E-07**	-0.1192	0.0236	0	0.6694	-/-/-/-
3	rs75065017	8	59,081,401	*FAM110B*	**8.57E-06**	-0.1296	0.0190	0.3481	0.6082	**2.46E-06**	-0.1230	0.0261	0	0.6010	-/-/-/-
4	rs9990273	3	147,267,582	*ZIC1*	**8.95E-06**	0.0613	0.0136	0.2456	0.3565	**3.28E-06**	0.0623	0.0134	11.3	0.3365	+/+/+/+
5	rs185087	7	106,489,559	*PIK3CG*	**8.12E-06**	-0.0623	0.0289	0.0956	0.1044	**5.91E-06**	-0.0773	0.0171	58.2	0.0662	-/-/-/-
6	rs1875972	5	168,288,992	*SLIT3*	**4.62E-06**	-0.0604	0.0159	**0.0459**	0.2747	**6.11E-06**	-0.0632	0.0140	53.7	0.0905	-/-/-/-
7	rs13083025	3	171,487,632	*PLD1*	**9.79E-06**	0.0585	0.0143	0.1080	0.3629	**6.95E-06**	0.0630	0.0140	34.6	0.2045	-/+/+/+
8	rs7137724	12	97,265,256	*C12orf55*	**5.20E-07**	0.0568	0.0102	**0.0024**	0.6764	**9.87E-06**	0.0698	0.0158	69.8	**0.0193**	-/+/+/+

Table 3. Common suggestive genomic loci identified in meta-analysis of **Zlog(aMT6s: Cr)** by both METAL and MR MEGA, with association and heterogeneity statistics. Significant values are in [bold]. P.value – p value for a SNP from the association test; beta – SNP's effect size estimate; se – standard error of the SNP's effect size estimate; I2 -I^2 Statistics: percentage of variation in the SNP's effect size across studies due to heterogeneity; P.Het. – p value for Cochrane heterogeneity test; P anc.het- p value for the ancestry heterogeneity test; P res.het – p value for the residual heterogeneity test.

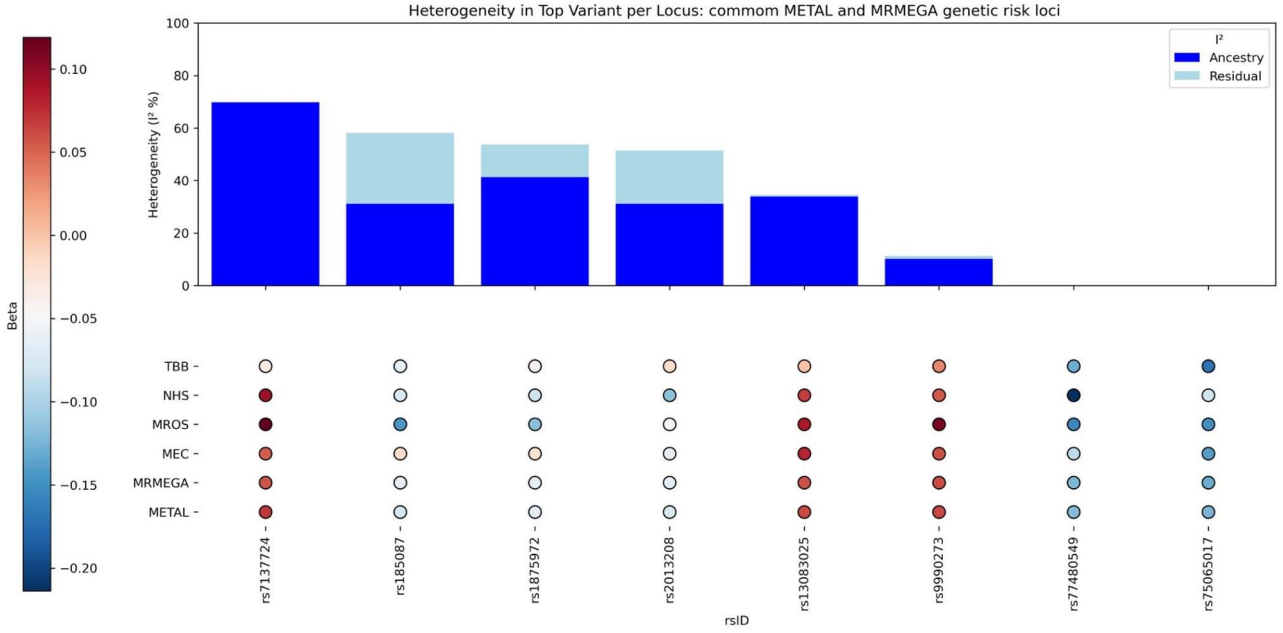

Fig. 3. Heterogeneity in lead suggestive SNPs common in METAL and MRMEGA meta-analyses.

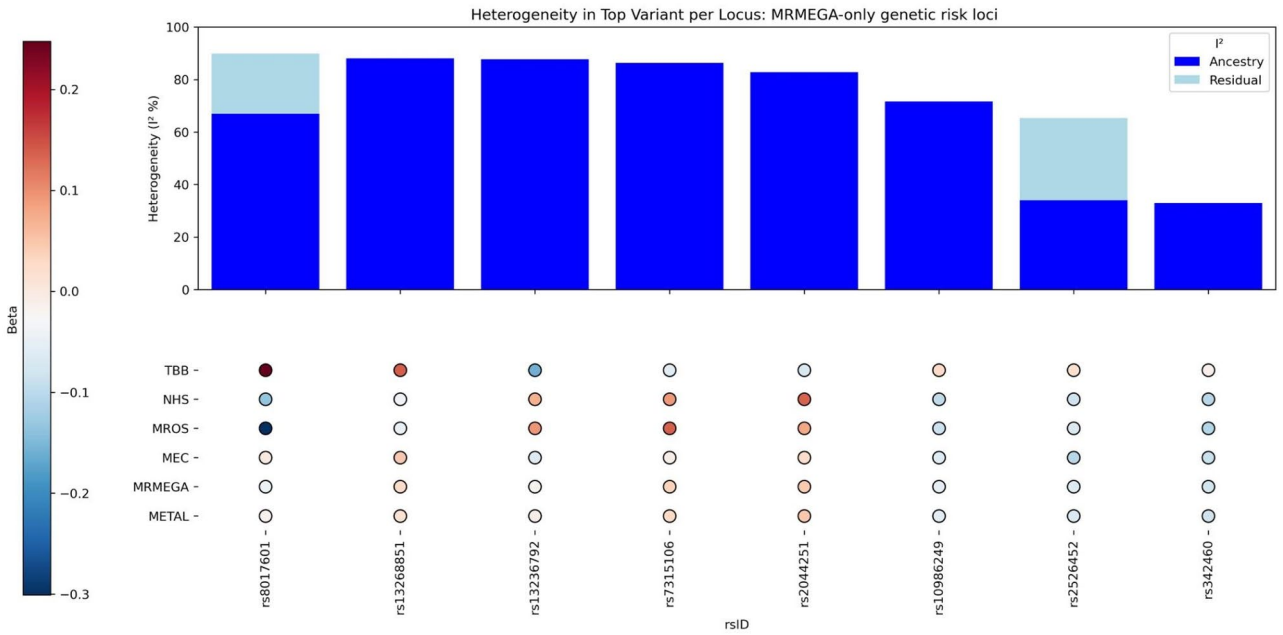

Fig. 4. Heterogeneity in lead suggestive SNPs identified exclusively in MRMEGA meta-analyses.

and more susceptible to sampling variability. Consistent with this, the previously reported Taiwan Biobank suggestive variants were not replicated in our study, including in restricted subset analyses, suggesting limited transferability of these signals across populations. Overall, our findings support the need for larger ancestry-specific GWASs and more balanced multi-ancestry datasets to distinguish true biological heterogeneity from power-related instability.

Gene set enrichment analysis identified significant pathways related to nucleotide exchange factors[39], cellular signaling[40], mitochondrial and energy regulation[41] and cancer[42], including signals in MST1 and MSP pathways with established roles in immune, metabolic and cancer-related processes[43–47]. Our PRS PheWAS analyses demonstrated associations between melatonin PRSs and both blood vitamin D levels and sleep duration, consistent with melatonin's role in circadian regulation and vitamin D metabolism and supporting a potentially shared genetic component between melatonin secretion and sleep-related traits. However,

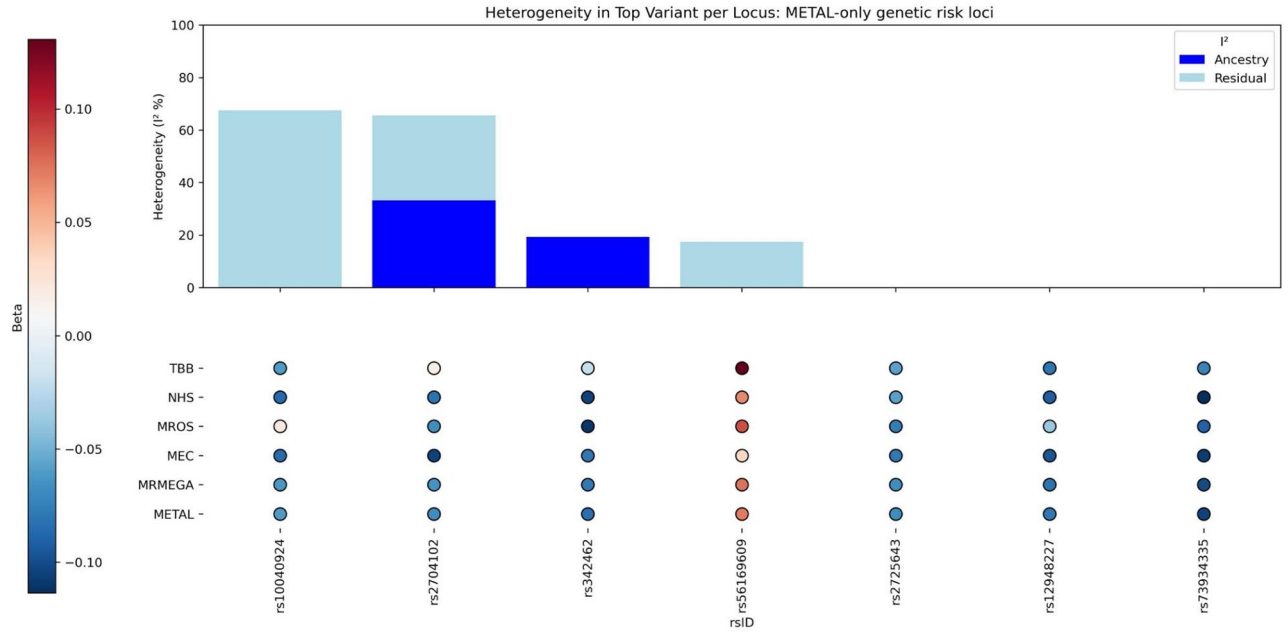

Fig. 5. Heterogeneity in lead suggestive SNPs identified exclusively in METAL meta-analyses.

associations with phenotypes known to vary by ancestry or population structure, including vitamin D-related and hematologic traits, should be interpreted cautiously. Although the PheWAS models included adjustment for age, sex, genotyping batch, and 10 principal components, residual ancestry-related confounding cannot be fully excluded, particularly for phenotypes that are themselves strongly patterned by ancestry or population structure. Accordingly, these signals should not be interpreted as evidence of shared melatonin biology alone, as they may also reflect ancestry-related differences in phenotype prevalence, environmental exposures, allele-frequency distributions, or reduced PRS transferability across populations. These findings are therefore best regarded as hypothesis-generating and in need of replication in ancestry-stratified datasets.

Our study has several limitations. First, the overall heritability estimates suggest that aMT6s levels have a relatively modest genetic component, which may require larger sample sizes to detect genome-wide significant associations. Power calculations indicated limited power to detect small-effect loci. Together with the modest SNP-based heritability estimates, these findings suggest that the genetic architecture of 6-sulfatoxymelatonin is likely polygenic, with individual variants expected to explain only a small fraction of trait variance. In addition, measurement variability in circulating 6-sulfatoxymelatonin may have further reduced locus discovery power. Second, phenotype measurement heterogeneity across cohorts is an important limitation. Although we applied within-cohort z-scoring of log(aMT6s: Cr) to improve harmonization, this procedure standardizes scale rather than fully correcting for systematic differences related to ELISA platform, creatinine assay, laboratory site, and urine collection or processing protocols. These differences may have introduced residual non-biological variation, attenuated genetic effect estimates, and reduced power for locus discovery. In addition, raw melatonin levels varied across cohorts, reflecting cohort-specific characteristics, including particularly low levels among older MrOS men and much higher levels among NHS women. The lack of substantial genomic inflation across the individual cohorts and meta-analyses suggests that such heterogeneity likely acted primarily by adding measurement noise rather than inflating association statistics. Together with the inclusion of multiple ancestries and relatively modest sample sizes, these factors likely contributed to the lack of genome-wide significant associations. Accordingly, the present meta-analysis should be regarded as exploratory and power-limited, and loci reaching only suggestive significance should be interpreted as hypothesis-generating rather than definitive associations. Future studies should focus on expanding sample sizes, particularly in non-European populations, improving phenotype harmonization across cohorts, and applying more standardized biospecimen collection and assay protocols to improve discovery power and clarify ancestry-specific effects.

In conclusion, this study represents one of the largest multi-ancestry GWAS meta-analyses of urinary aMT6s levels to date, providing novel insights into the genetic determinants of melatonin secretion. While no genome-wide significant loci were identified, we uncovered several suggestive genetic signals with potential links to cardiometabolic traits and cancer. The observed ancestry-specific effects highlight the complexity of melatonin metabolism and the importance of diverse study populations. Future research should aim to replicate and extend these findings using larger and more diverse cohorts, ultimately improving our understanding of melatonin's role in human health.

	TWB	NHS*	MrOS	MEC
Sample size (*n*)	2,373	3,861	2,175	3,335
aMT6s concentration measuring kit and manufacturer's protocol used	Human Melatonin Sulfate ELISA kit, Lab science	ELISA; ALPCO, Windham, NH	Buhlmann 6-sulphatoxymelatonin ELISA; ALPCO Diagnostics, Windham, NH	Melatonin-Sulfate ELISA; IBL International
Cr measuring method	In the same lab but not at the same time; one plate of 96-well ELISA at a time; no precise batch information	ALPCO, Windham, NH; Jaffe method	OHSU clinical laboratory; using the Synchron LX System (Beckman Coulter, Fullerton, CA), which utilizes the Jaffe rate method and has a sensitivity of 10 mg/dL for creatinine in urine	Jaffe method on Hitachi 912 (Roche Diagnostic)
Were there several study sites/labs?	NO; measured in the same lab	YES; 28 nested case-control studies	YES; Birmingham *N* = 471; Minneapolis *N* = 481; Palo Alto *N* = 463; Pittsburgh *N* = 500; Portland *N* = 430; San Diego *N* = 499	NO; measurements done by the same lab, measured in duplicate in 32 batches

Table 4. Urinary 6-sulphatoxymelatonin (aMT6s) measurement characteristics in Taiwan Biobank (TWB), the Nurses' Health Study I (NHS1), the Nurses' Health Study II (NHS2), MrOS and Multiethnic Cohort (MEC). *NHS : pooled NHS1 and NHS2 cohorts.

Methods

Study sample

The analytic sample for our study was made up of 11,744 individuals (Table 1) with 2,373 individuals of East Asian ancestry from the Taiwan Biobank (TWB), 3,861 female nurses of European ancestry from the Nurses' Health Study (NHS1) and Nurses' Health Study II (NHS2), 2,175 men of European ancestry from the Osteoporotic Fractures in Men (MrOS), and 3,335 individuals of diverse ancestries from the Multiethnic Cohort Study (MEC). Further details on each study cohort can be found in Supplementary Methods.

Outcome definition

Table 4 summarizes the aMT6s measurement details across participating cohorts, with additional cohort-specific information provided in the Supplementary Methods. The concentration of melatonin metabolite (6-sulfatoxymelatonin, aMT6s) in the urine —whether measured from a first-morning void or an overnight collection— depends on the sample's volume, with urinary creatinine levels serving as a reliable substitute for this variation[48]. The melatonin-to-creatinine ratio (aMT6s: Cr) is conventionally used to control for differences in sample dilution, with further log-transformation ensuring normality. To address inter-study variability, we additionally normalized (z-scored) the log-transformed aMT6s: Cr values within each study cohort, thereby establishing our primary outcome measure as Zlog(aMT6s: Cr).

Genome-Wide association and meta-analysis

Genome wide association studies (GWASs) of our outcome variable were performed within each study cohort using linear regression models adjusted for age at sample collection, sex (where applicable) and first 10 principal components of ancestry. Further cohort specific details on GWAS analyses are described in Supplementary Methods. GWAS summary statistics were aligned to genome build hg19 and harmonized by excluding variants with poor imputation quality (info score $R^2 < 0.7$), low minor allele frequency (MAF < 1%), significant deviation from the Hardy-Weinberg equilibrium (p-value $< 10^{-8}$) and the low genotype call rate (< 95%). Meta analyses were performed by using meta-regression of multi-ethnic genetic association (MR-MEGA[49] and inverse variance fixed-effects model in METAL[50] with genomic control correction for the individual study level data. Briefly, MR-MEGA conducts a meta-regression analysis by creating axes of genetic variation specific to each cohort, which are next used as covariates in the meta-analysis to adjust for possible differences in population structure. This approach allows to distinguish between ancestral and residual sources of heterogeneity. Although MR-MEGA shows an increased power to detect SNPs associations under both fixed and random effects meta-analysis settings[49], for variants with homogenous effects across populations its power is reduced[51]. To account for these aspects and potentially detect genetic variants with both homogeneous and heterogeneous effects across study cohorts, we applied both METAL and MR-MEGA methods. Genetic ancestry was not assigned using a single cross-cohort clustering pipeline for this meta-analysis. Instead, ancestry definitions were inherited from the original cohort-specific GWAS pipelines and therefore reflected cohort-specific recruitment, eligibility criteria, quality-control procedures, and analytic design. Consequently, no single PCA-projection, ADMIXTURE-based framework, or uniform ancestry-classification threshold was applied across all cohorts. Taiwan Biobank was restricted to participants of East Asian ancestry, MrOS to white/European-ancestry men, and NHS excluded participants self-identifying as non-White or multi-racial, while MEC was analyzed as a multiethnic cohort. Cohort-specific GWAS models included 10 ancestry principal components as covariates. For genotype imputation, cohort-specific reference panels were used, including the 1000 Genomes East Asian

reference panel in Taiwan Biobank, TOPMed in NHS, and the 1000 Genomes Project reference panel in MrOS and MEC. For MR-MEGA, 1 genetic principal component was used to construct the meta-regression axis. To claim significance, we applied the conventional genome-wide threshold of $5{\cdot}10^{-8}$ and less stringent level of $1{\cdot}10^{-5}$ for suggestive associations[52]. The overall SNP-based heritability and genetic correlations were estimated by using Linkage Disequilibrium Score Regression (LDSC)[53].

Formal power calculations were performed as an approximate assessment of variant-level discovery power using the observed total meta-analysis sample size ($N = 11{,}744$) and significance thresholds of 5×10^{-8} and 1×10^{-5}. Power was calculated from the non-central chi-square distribution across a range of assumed per-variant proportions of phenotypic variance explained (R^2), and the minimum R^2 required for 80% power was determined. Because the GWAS results were obtained from multi-cohort meta-analysis using both fixed-effects and meta-regression approaches, these calculations should be interpreted as approximations rather than exact estimates of power.

To identify independent significant genomic risk loci we used Functional Mapping and Annotation (FUMA[25,26] v1.5.2. We further derived a polygenic risk score for the aMT6s using PRS-CSx[54], restricting score construction to SNPs nominally associated with aMT6s in our discovery GWAS ($p < 0.05$), and checked its associations with common diseases by running phenome-wide association studies (PheWASs) in the Mass General Brigham Biobank (MGBB) and in the UK Biobank (UKBB). Since harmonized individual-level data were not available across all contributing cohorts, formal PRS predictive performance could not be assessed. PheWAS analyses were performed using PRSs derived from PRS-CSx weights estimated from our meta-analysis, with regression models adjusted for age, sex, genotyping batch, and 10 ancestry principal components. Further details are described in Supplementary Methods.

Ethics statements

All studies were approved by the respective local ethical committees. All methods were carried out in accordance with relevant guidelines and regulations, and informed consent was obtained from all participants. The Taiwan Biobank study protocol was approved by the Institutional Review Board of Chang Gung Medical Foundation and the Institutional Review Board of National Taiwan University Hospital. All subjects have provided written informed consent. The Nurses' Health Study and Nurses' Health Study 2 protocols were approved by the Institutional Review Board of Brigham and Women's Hospital and the Committee on the Use of Human Subjects in Research of Harvard T.H. Chan School of Public Health (Boston, MA, USA). Voluntary return of questionnaires indicates their informed consent. For the MrOS study men were recruited at six US clinical centers in Birmingham, AL; Minneapolis, MN; Palo Alto, CA; the Monongahela Valley near Pittsburgh, PA; Portland, OR; and San Diego, CA and the institutional review boards at each clinic site approved the study. The Institutional Review Boards of the University of Hawaii and the University of Southern California approved the Multiethnic Cohort study.

Data availability

GWAS summary statistics generated in this study have been deposited in the NHGRI-EBI GWAS Catalog under GCP ID: GCP001636 (https://www.ebi.ac.uk/gwas/deposition/bodyofwork/GCP001636). The corresponding accession numbers are: GCST90833020: METAL meta-analysis (https://www.ebi.ac.uk/gwas/studies/GCST90833020), GCST90833021: MRMEGA meta-analysis (https://www.ebi.ac.uk/gwas/studies/GCST90833021), GCST90833022: TWB GWAS (https://www.ebi.ac.uk/gwas/studies/GCST90833022), GCST90833023: NHS GWAS (https://www.ebi.ac.uk/gwas/studies/GCST90833023), GCST90833024: MrOS GWAS (https://www.ebi.ac.uk/gwas/studies/GCST90833024), GCST90833025: MEC GWAS (https://www.ebi.ac.uk/gwas/studies/GCST90833025), GCST90833026: METAL meta-analysis among European ancestry participants (https://www.ebi.ac.uk/gwas/studies/GCST90833026).

Received: 26 September 2025; Accepted: 15 April 2026
Published online: 22 April 2026

References

1. Bojkowski, C. J., Arendt, J., Shih, M. C. & Markey, S. P. Melatonin secretion in humans assessed by measuring its metabolite, 6-sulfatoxymelatonin. *Clin. Chem.* **33** (8), 1343–1348. https://doi.org/10.1093/clinchem/33.8.1343 (1987).
2. Bartsch, C. et al. Nocturnal urinary 6-sulfatoxymelatonin and proliferating cell nuclear antigen-immunopositive tumor cells show strong positive correlations in patients with gastrointestinal and lung cancer. *J. Pineal Res.* **23** (2), 90–96. https://doi.org/10.1111/j.1600-079X.1997.tb00340.x (1997).
3. Van Der Veen, A. et al. Urinary excretion of 6-sulfatoxymelatonin, the main metabolite of melatonin, and mortality in stable outpatient renal transplant recipients. *J. Clin. Med.* **9** (2), 525. https://doi.org/10.3390/jcm9020525 (2020).
4. Schernhammer, E. S. et al. Urinary 6-sulphatoxymelatonin levels and risk of breast cancer in premenopausal women: the ORDET cohort. *Cancer Epidemiol. Biomarkers Prev.* **19** (3), 729–737. https://doi.org/10.1158/1055-9965.EPI-09-1229 (2010).
5. Basler, M. et al. Urinary excretion of melatonin and association with breast cancer: meta-analysis and review of the literature. *Breast Care*. **9** (3), 182–187. https://doi.org/10.1159/000363426 (2014).
6. Wong, A. T. Y. et al. Urinary melatonin in relation to breast cancer risk: nested case–control analysis in the DOM study and meta-analysis of prospective studies. *Cancer Epidemiol. Biomarkers Prev.* **30** (1), 97–103. https://doi.org/10.1158/1055-9965.EPI-20-0822 (2021).
7. Cutando, A., Gómez-Moreno, G., Arana, C., Acuña-Castroviejo, D. & Reiter, R. J. Melatonin: potential functions in the oral cavity. *J. Periodontol.* **78** (6), 1094–1102. https://doi.org/10.1902/jop.2007.060396 (2007).
8. Asghari, M. H., Moloudizargari, M., Ghobadi, E., Fallah, M. & Abdollahi, M. Melatonin as a multifunctional anti-cancer molecule: implications in gastric cancer. *Life Sci.* **185**, 38–45. https://doi.org/10.1016/j.lfs.2017.07.020 (2017).
9. Bazzi, L. A. et al. Exploratory assessment of pineal gland volume, composition, and urinary 6-sulfatoxymelatonin levels on prostate cancer risk. *Prostate* **81** (8), 487–496. https://doi.org/10.1002/pros.24130 (2021).

10. Tai, S. Y., Huang, S. P., Bao, B. Y. & Wu, M. T. Urinary melatonin-sulfate/cortisol ratio and the presence of prostate cancer: a case-control study. *Sci. Rep.* **6** (1), 29606. https://doi.org/10.1038/srep29606 (2016).
11. Kennaway, D. J. The dim light melatonin onset across ages, methodologies, and sex and its relationship with morningness/eveningness. *SLEEP* **46** (5), zsad033. https://doi.org/10.1093/sleep/zsad033 (2023).
12. Woodford, E. C. et al. Endogenous melatonin and sleep in individuals with Rare Genetic Neurodevelopmental Disorders (RGND): a systematic review. *Sleep. Med. Rev.* **57**, 101433. https://doi.org/10.1016/j.smrv.2021.101433 (2021).
13. Wei, T. et al. Association between night-shift work and level of melatonin: systematic review and meta-analysis. *Sleep. Med.* **75**, 502–509. https://doi.org/10.1016/j.sleep.2020.09.018 (2020).
14. Lyssenko, V. et al. Common variant in MTNR1B associated with increased risk of type 2 diabetes and impaired early insulin secretion. *Nat. Genet.* **41** (1), 82–88. https://doi.org/10.1038/ng.288 (2009).
15. Brito Nunes, C. et al. Multi-ancestry, trans-generational GWAS meta-analysis of gestational diabetes and glycaemic traits during pregnancy reveals limited evidence of pregnancy-specific genetic effects. *Genetic Genomic Med.* https://doi.org/10.1101/2025.08.19.25333735 (2025).
16. Pervjakova, N. et al. Multi-ancestry genome-wide association study of gestational diabetes mellitus highlights genetic links with type 2 diabetes. *Hum. Mol. Genet.* **31** (19), 3377–3391. https://doi.org/10.1093/hmg/ddac050 (2022).
17. Qian, J. et al. Unanticipated daytime melatonin secretion on a simulated night shift schedule generates a distinctive 24-h melatonin rhythm with antiphasic daytime and nighttime peaks. *J. Pineal Res.* **72** (3), e12791. https://doi.org/10.1111/jpi.12791 (2022).
18. Liu, P. H. et al. A genome-wide association study for melatonin secretion. *Sci. Rep.* **12** (1), 8025. https://doi.org/10.1038/s41598-022-12084-w (2022).
19. Blank, J. B. et al. Overview of recruitment for the osteoporotic fractures in men study (MrOS). *Contemp. Clin. Trials.* **26** (5), 557–568. https://doi.org/10.1016/j.cct.2005.05.005 (2005).
20. Orwoll, E. et al. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study — a large observational study of the determinants of fracture in older men. *Contemp. Clin. Trials.* **26** (5), 569–585. https://doi.org/10.1016/j.cct.2005.05.006 (2005).
21. HapMap 3 reference panel. https://ibg.colorado.edu/cdrom2021/Day06-nivard/GenomicSEM_practical/eur_w_ld_chr/ (2025).
22. Willer, C. J. et al. Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **45** (11), 1274–1283. https://doi.org/10.1038/ng.2797 (2013).
23. Gao, H. et al. Association of RBM5 rs2013208 SNP with serum lipid levels in two Chinese ethnic groups. *Int. J. Clin. Exp. Pathol.* **10** (8), 8520–8534 (2017).
24. Raina, J. K. et al. Association of ESR1 (rs2234693 and rs9340799), CETP (rs708272), MTHFR (rs1801133 and rs2274976) and MS (rs185087) polymorphisms with Coronary Artery Disease (CAD). *BMC Cardiovasc. Disord.* **20** (1), 340. https://doi.org/10.1186/s12872-020-01618-7 (2020).
25. Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8** (1), 1. https://doi.org/10.1038/s41467-017-01261-5 (2017).
26. Watanabe, K., Umićević Mirkov, M., De Leeuw, C. A., Van Den Heuvel, M. P. & Posthuma, D. Genetic mapping of cell type specificity for complex traits. *Nat. Commun.* **10** (1), 1. https://doi.org/10.1038/s41467-019-11181-1 (2019).
27. De Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLOS Comput. Biol.* **11** (4), e1004219. https://doi.org/10.1371/journal.pcbi.1004219 (2015).
28. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell. Syst.* **1** (6), 417–425. https://doi.org/10.1016/j.cels.2015.12.004 (2015).
29. Valo, E. et al. Genome-wide characterization of 54 urinary metabolites reveals molecular impact of kidney function. *Nat. Commun.* **16** (1), 325. https://doi.org/10.1038/s41467-024-55182-1 (2025).
30. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4** (1), 1. https://doi.org/10.1186/s13742-015-0047-8 (2015).
31. Ferguson, A. et al. Genome-wide association study of circadian rhythmicity in 71,500 UK biobank participants and polygenic association with mood instability. *EBioMedicine* **35**, 279–287. https://doi.org/10.1016/j.ebiom.2018.08.004 (2018).
32. Mahajan, A. et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat. Genet.* **50** (11), 1505–1513. https://doi.org/10.1038/s41588-018-0241-6 (2018).
33. Lyon, M. et al. The variant call format provides efficient and robust storage of GWAS summary statistics. *Genetics* **30**, 586. https://doi.org/10.1101/2020.05.29.115824 (2020).
34. Elsworth, B. et al. The MRC IEU OpenGWAS data infrastructure. *Genetics.* **10**, 586. https://doi.org/10.1101/2020.08.10.244293 (2020).
35. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562** (7726), 7726. https://doi.org/10.1038/s41586-018-0579-z (2018).
36. Cerezo, M. et al. The NHGRI-EBI GWAS Catalog: standards for reusability, sustainability and diversity. *Nucleic Acids Res.* **53** (D1), D998–D1005. https://doi.org/10.1093/nar/gkae1070 (2025).
37. Jiang, L., Zheng, Z., Fang, H. & Yang, J. A generalized linear mixed model association tool for biobank-scale data. *Nat. Genet.* **53** (11), 1616–1621. https://doi.org/10.1038/s41588-021-00954-4 (2021).
38. Xu, P. et al. Genome-wide association study of human pineal gland volume as proxy for melatonin secretion. *Genetic Genomic Med.* **11**, 256. https://doi.org/10.1101/2025.03.10.25323654 (2025).
39. Laco, J., Zeman, I., Pevala, V., Polčic, P. & Kolarov, J. Adenine nucleotide transport via Sal1 carrier compensates for the essential function of the mitochondrial ADP/ATP carrier: adenine nucleotide transport via Sal1. *FEMS Yeast Res.* **10** (3), 290–296. https://doi.org/10.1111/j.1567-1364.2010.00606.x (2010).
40. Chen, S. & Qiu, G. Cloning and activity analysis of the promoter of nucleotide exchange factor gene ZjFes1 from the seagrasses Zostera japonica. *Sci. Rep.* **10** (1), 17291. https://doi.org/10.1038/s41598-020-74381-6 (2020).
41. Atlante, A. & Valenti, D. A walk in the memory, from the first functional approach up to its regulatory role of mitochondrial bioenergetic flow in health and disease: focus on the adenine nucleotide translocator. *Int. J. Mol. Sci.* **22** (8), 4164. https://doi.org/10.3390/ijms22084164 (2021).
42. Bartolini, M., Wainer, I. W., Bertucci, C. & Andrisano, V. The rapid and direct determination of ATPase activity by ion exchange chromatography and the application to the activity of heat shock protein-90. *J. Pharm. Biomed. Anal.* **73**, 77–81. https://doi.org/10.1016/j.jpba.2012.03.021 (2013).
43. Dong, F., Zheng, L. & Zhang, X. Alpha-boswellic acid accelerates acute wound healing via NF-κB signaling pathway. *PLOS ONE* **19** (9), e0308028. https://doi.org/10.1371/journal.pone.0308028 (2024).
44. Nakamura, M., Zhai, P., Del Re, D. P., Maejima, Y. & Sadoshima, J. Mst1-mediated phosphorylation of Bcl-xL is required for myocardial reperfusion injury. *JCI Insight*. **1**, 5. https://doi.org/10.1172/jci.insight.86217 (2016).
45. Wang, S. et al. Nicorandil alleviates myocardial injury and post-infarction cardiac remodeling by inhibiting Mst1. *Biochem. Biophys. Res. Commun.* **495** (1), 292–299. https://doi.org/10.1016/j.bbrc.2017.11.041 (2018).
46. Peta, K. T., Ambele, M. A. & Pepper, M. S. Similarities between tumour immune response and chronic wound microenvironment: influence of mesenchymal stromal/stem cells. *J. Immunol. Res.* **2021**, 1–11. https://doi.org/10.1155/2021/6649314 (2021).
47. Azarbarzin, S., Hosseinpour-Feizi, M. A., Banan Khojasteh, S. M., Baradaran, B. & Safaralizadeh, R. MicroRNA – 383-5p restrains the proliferation and migration of breast cancer cells and promotes apoptosis via inhibition of PD-L1. *Life Sci.* **267**, 118939. https://doi.org/10.1016/j.lfs.2020.118939 (2021).

48. Klante, G., Brinschwitz, T., Secci, K., Wollnik, F. & Steinlechner, S. Creatinine is an appropriate reference for urinary sulphatoxymelatonin of laboratory animals and humans. *J. Pineal Res.* **23** (4), 191–197. https://doi.org/10.1111/j.1600-079X.1997.tb00354.x (1997).
49. Mägi, R. et al. Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution. *Hum. Mol. Genet.* **26** (18), 18. https://doi.org/10.1093/hmg/ddx280 (2017).
50. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26** (17), 17. https://doi.org/10.1093/bioinformatics/btq340 (2010).
51. Schormair, B. et al. Genome-wide meta-analyses of restless legs syndrome yield insights into genetic architecture, disease biology and risk prediction. *Nat. Genet.* **56** (6), 1090–1099. https://doi.org/10.1038/s41588-024-01763-1 (2024).
52. Duggal, P., Gillanders, E. M., Holmes, T. N. & Bailey-Wilson, J. E. Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genom.* **9** (1), 516. https://doi.org/10.1186/1471-2164-9-516 (2008).
53. Schizophrenia Working Group of the Psychiatric Genomics Consortium et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47** (3), 291–295. https://doi.org/10.1038/ng.3211 (2015).
54. Ruan, Y. et al. Improving polygenic prediction in ancestrally diverse populations. *Nat. Genet.* **54** (5), 573–580. https://doi.org/10.1038/s41588-022-01054-7 (2022).

Acknowledgements

The authors would like to thank Dr Sarah Coseo Markt for her generous support, guidance and data sharing, which was critical for the completion of this study.

Author contributions

M.Z. and R.S. wrote the main manuscript text; M.Z performed NHS data analyses and meta analysis; Z.Z. analysed MEC data; G-T.C. and Y-C.C. analysed TBB data; D.S.E analysed MrOS data; J.V. performed MGBB pheWAS and UKB analyses; M.W. shared UKB pheWAS pipeline; K.L.S.,G.J.T. : conceptualization, supervision (MrOS); A.H.E.,J.E.H , C.T.: conceptualization, supervision (NHS); L.M.,I.C.,C.H.,L.W.; L.L.M,:conceptualization, supervision (MEC); Y-C.C: conceptualization, supervision (TBB); M.M., J.H.: conceptualization, technical support in data analysis, supervision (general); E.S.S., R.S., J.L., R.R.: conceptualization, supervision (general); All authors read and reviewed the manuscript.

Funding

This study was supported by European Union, European Research Council (ERC) Advanced Grant CLOCK-risk (grant number 101053225), Department of Epidemiology, Medical University of Vienna to PI Eva Schernhammer. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. Richmond is supported by Cancer Research UK (grant number C18281/A29019), the Medical Research Council Integrative Epidemiology Unit (grant number: MC_UU_0032/1) and the NIHR Oxford Health Biomedical Research Centre (grant number: NIHR203316). This research was conducted using the UK Biobank resource under application number 48576. We thank all the participants and staff of the UK Biobank for enabling us to conduct this research. Taiwan Biobank : Grants from the Ministry of Science and Technology in Taiwan (MOST 105-2314-B-182-062, MOST 106-2314-B-182-043), the Translational Medical Research Program of Academia Sinica (ASTM-108-01-04), National Taiwan University Hospital, Yunlin Branch Intramural Grant (NTUHYL106.X003, NTUHYL.107S004, NTUHYL 111.X019, NTUHYL110.X017) and Chang Gung University, Taoyuan, Taiwan (NMRPD1F154, NMRPD1G0711, and BMRPD08). NHS and NHS2: This work was supported by National Institute of Health: P01CA87969, P01CA055075, P01DK070756, U01HG004728, UM1CA186107, U01CA176726, UM1CA176726, U01 CA167552, R01CA49449, R01CA50385, R01CA67262, R01CA131332, R01HL034594, R01HL088521, R01HL35464, R01HL116854, R01EY015473, R01EY022305, P30EY014104, R03DC013373 and R03CA165131. MrOS : The Osteoporotic Fractures in Men (MrOS) Study is supported by National Institutes of Health funding. The following institutes provide support: the National Institute on Aging (NIA), the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Center for Advancing Translational Sciences (NCATS), and NIH Roadmap for Medical Research under the following grant numbers: U01 AG027810, U01 AG042124, U01 AG042139, U01 AG042140, U01 AG042143, U01 AG042145, U01 AG042168, U01 AR066160, R01 AG066671, and UL1 TR002369. The National Heart, Lung, and Blood Institute (NHLBI) provides funding for the MrOS Sleep ancillary study "Outcomes of Sleep Disorders in Older Men" under the following grant numbers: R01 HL071194, R01 HL070848, R01 HL070847, R01 HL070842, R01 HL070841, R01 HL070837, R01 HL070838, and R01 HL070839. MEC : This research was supported by the Public Health Service (National Cancer Institute) grant RO1 CA 5428 and U01 CA164973.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-49491-2.

Correspondence and requests for materials should be addressed to M.Ż.