Supplementary Information

Combining populations and risk phenotypes greatly expands knowledge of the genetic architecture of susceptibility to melanoma

Inclusion of UK Biobank self-report cases

While the genetic correlation between the UKBB self-report set is lower than 23andMe, the confidence intervals overlapped substantially (UKBB self-report Rg 0.65, 95% CI 0.36-0.95; 23andMe 1.00 95% CI 0.74-1.00). This suggests that the UK Biobank self-report GWAS does not differ in terms of genetic architecture to the other melanoma GWAS. However, even if there are potential uncertainties in the phenotype as discussed in the main text, we show that the UKBB self report is contributing to our discovery power.

Looking specifically at the lead SNPs identified in the confirmed only meta-analysis (**Supplementary Table 6**), the correlation of log(OR) beta estimates (weighted by their inverse variance) from the confirmed only meta-analysis and from the UKBB self-report set is high, r^2 = 0.62 (**Supplementary Figure 9**). As a further sensitivity analysis we excluded the UKBB self-report set from the meta-analysis, retaining all confirmed sets and 23andMe. Three loci found in the total (all confirmed sets plus 23andMe and UKBB self-report) no longer reach genome-wide significance. Heterogeneity was low across these loci (I^2 <= 10%), both with and without the UKBB self-report set, suggesting its presence was improving power without driving the discovery of these loci nor leading to heterogeneity. For each of these three lead SNP the estimate within UKBB self-report was consistent with the wider meta-analysis (**Supplementary Table 18**). For example, rs6908626 in the UKBB self-report has a log(OR) = 0.151, and a 95% CI 0.047 - 0.255. In the meta-analysis without UKBB self-report the log(OR) = 0.078 , 95% CI = 0.049 - 0.107, with overlapping confidence intervals. The consistency of UKBB self-report GWAS at these three loci is further supported by the reported forest plots for these SNPs in **Supplementary Figure 7**. These show that log(OR) estimates from UKBB self-report set are consistent with, and overlap, those of other confirmed case sets.

Thus, although the Rg between UKBB self-report and the confirmed meta-GWAS is not 1 (which may be for a range of reasons e.g., LD score regression power), the inclusion of the UKBB self-report GWAS boosts the power to identify new melanoma GWAS loci without introducing bias.

Replication of previously reported melanoma GWAS loci

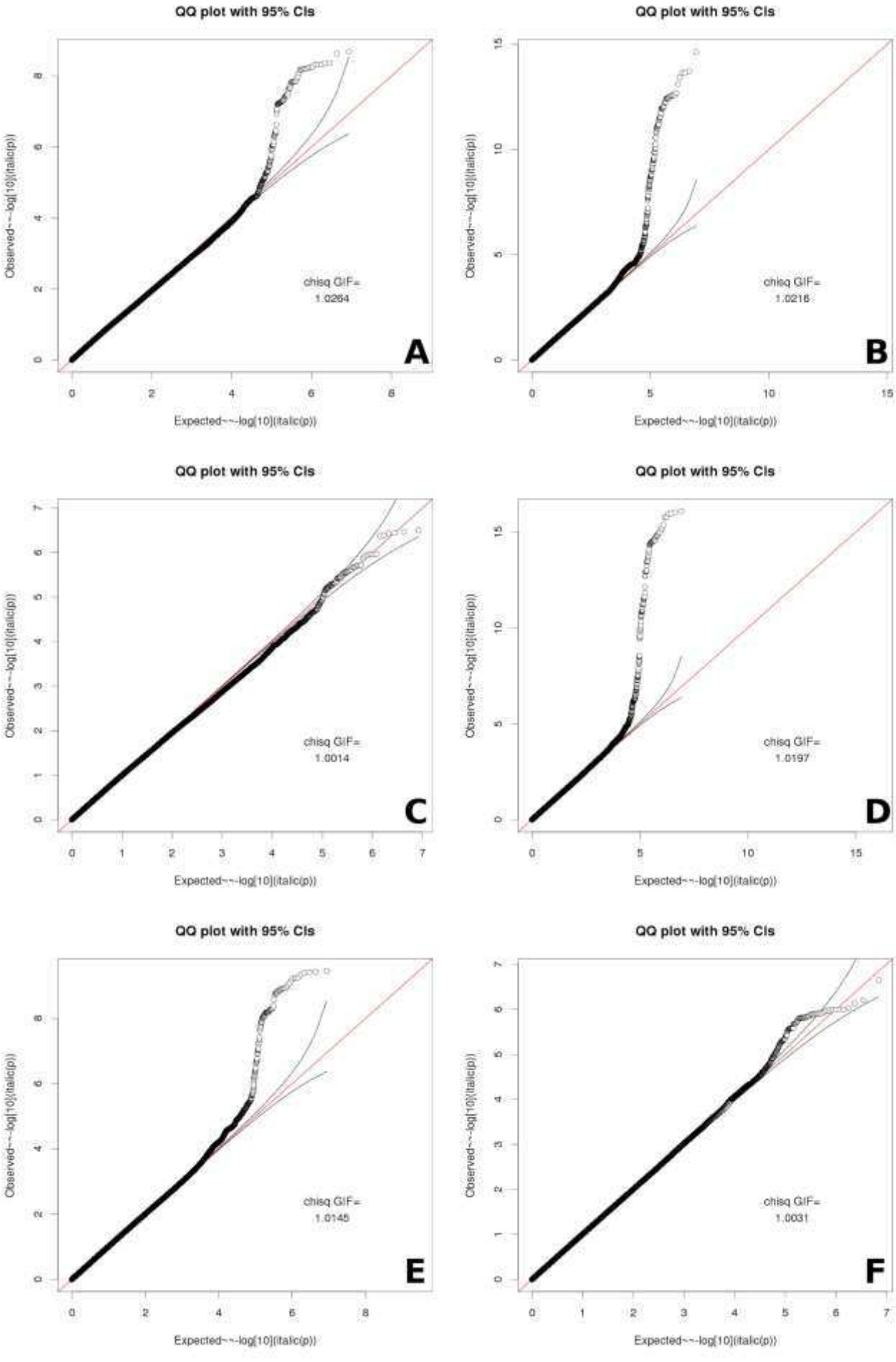
Our analysis directly confirms 19 of the 21 loci reaching genome-wide significance in a previous CM GWAS [1–8] (Table 1). The previously-reported peak variant for the locus in intron 8 of FTO rs16953002 did not formally replicate ($P_{meta\_r} = 1.2 \times 10^{-6}$, I^2 = 35%, $P_{meta} = 3.8 \times 10^{-8}$) nor did any SNPs in LD, with the most strongly associated SNP at this region being rs62034121. However, rs16953002 and rs62034121 are in LD (r^2_{EUR} = 0.96), and these variants are genome-wide significant in the meta-analysis of confirmed-only CM cases, and the combined CM plus nevus analysis (see below and Online Methods). A recent melanoma GWAS [8] reported rs187843643 near BASP1 as a melanoma locus; while we include the contributing GWAS study in our analysis, this locus was no longer genome-wide significant in the meta-analysis (I^2 = 35%, P_{meta_r} = 0.3, P_{meta} = 0.00071).

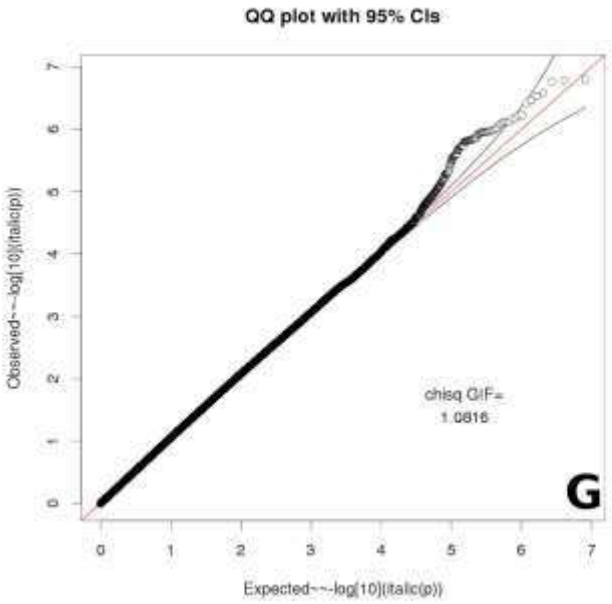
While the loci is still significantly associated with CM, the previously-reported CM susceptibility variant centromeric to AGR3 (rs1636744) on chromosome 7 is no longer the regional peak. While a SNP in LD (r^2_{EUR} = 0.91) with rs1636744, rs73069846, is an independent risk variant for CM (Supplementary Table 3) the peak variant for this locus is now rs117132860 (r^2_{EUR} < 0.03 with previously reported variants) located near the AHR gene (Table 1, Supplementary Table 3, P_{meta}= 3.8×10^{-21}). This region exhibits a complex LD structure and more focused investigations are needed to disentangle the CM association(s) in this region. Similarly, the peak variant near IRF4 is

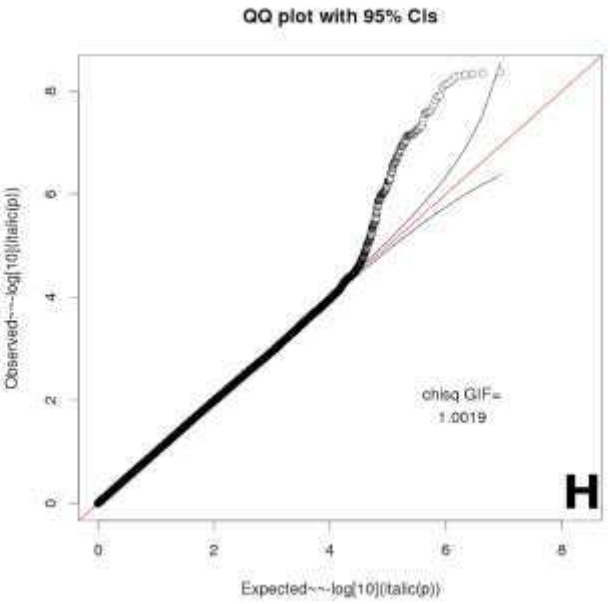
rs12215602 and not the previously reported rs12203592, for which substantial heterogeneity was observed (I^2 = 81%, P_{meta_R} = 0.1, P_{meta} = 1.9×10^{-13}).

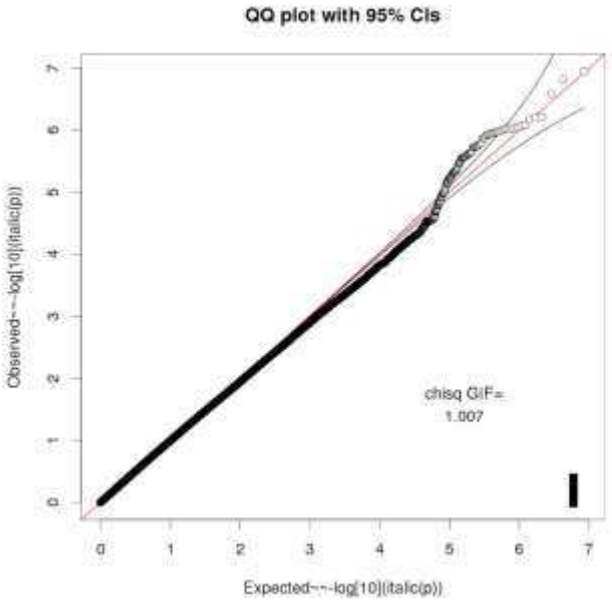
The total CM meta-analysis also confirmed 6 regions previously identified only by combining nevus count and CM GWAS data [9] (**Table 2**). While in our CM meta-analysis the peak variant for another region near MFSD12 and FZR1, rs12984831, is independent from the previously-reported peak at rs34466956 (r^2_{EUR} = 0.001), rs34466956 is also significantly associated in our combined analysis of CM and nevus count described below ($P_{CM+nevus}$ = 2.5×10^{-9}; **Supplementary Table 7**).

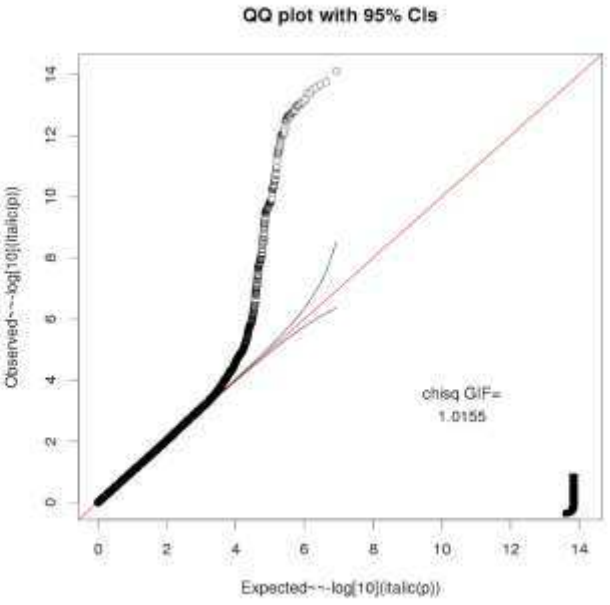
Supplementary Figures

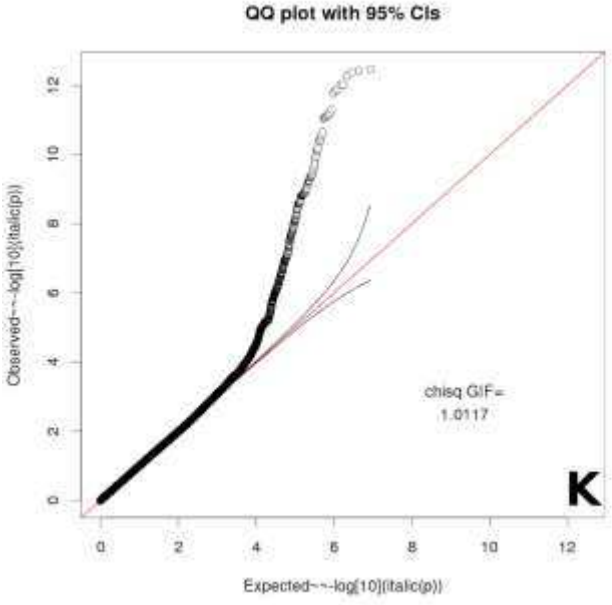

QQ plot with 95% CIs
chisq GIF=
1.0264
A
QQ plot with 95% CIs
chisq GIF=
1.0216
B
QQ plot with 95% CIs
chisq GIF=
1.0014
C
QQ plot with 95% CIs
chisq GIF=
1.0197
D
QQ plot with 95% CIs
chisq GIF=
1.0145
E
QQ plot with 95% CIs
chisq GIF=
1.0031
F
Expected~~-log[10](italic(p))
Observed~~-log[10](italic(p))

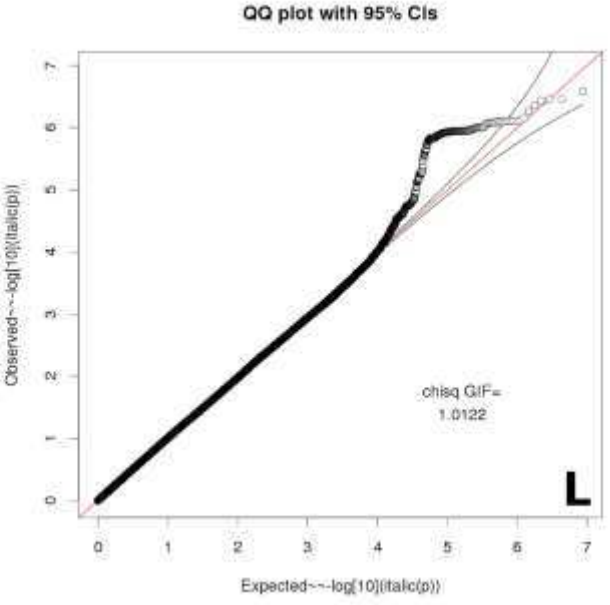

QQ plot with 95% CIs
chisq GIF=
1.0816
G
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))

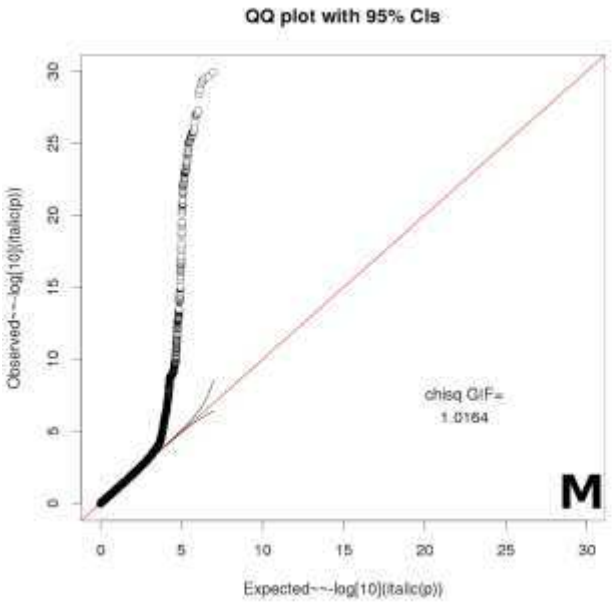

QQ plot with 95% CIs
chisq GIF=
1.0019
H
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))

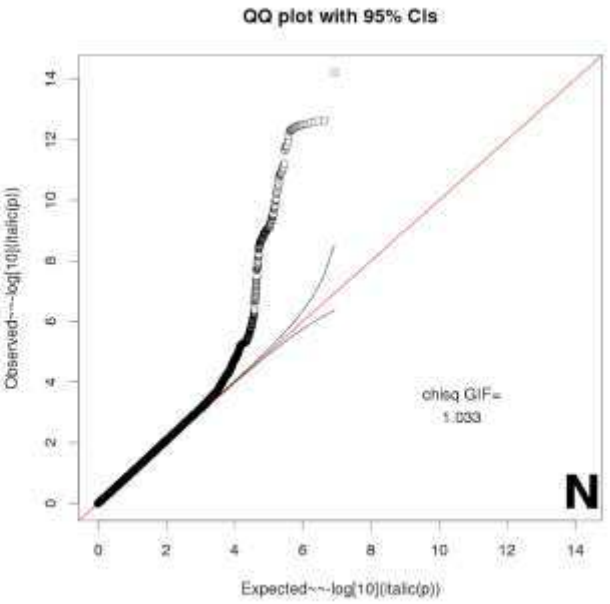

QQ plot with 95% CIs
chisq GIF=
1.007
I
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))

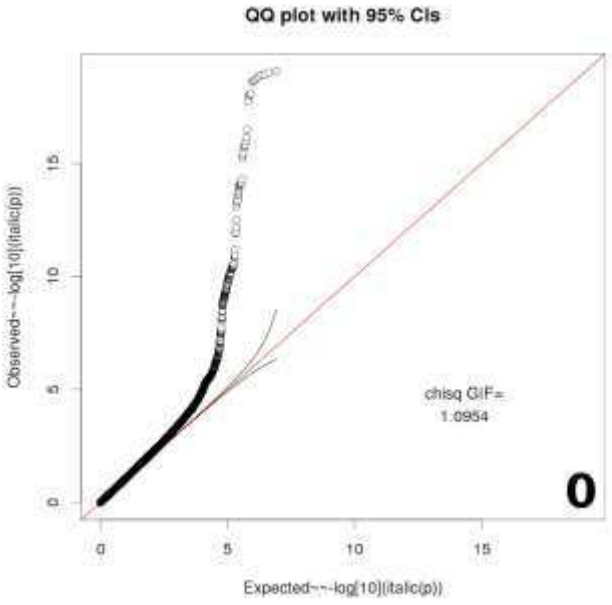

QQ plot with 95% CIs
chisq GIF=
1.0155
J
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))

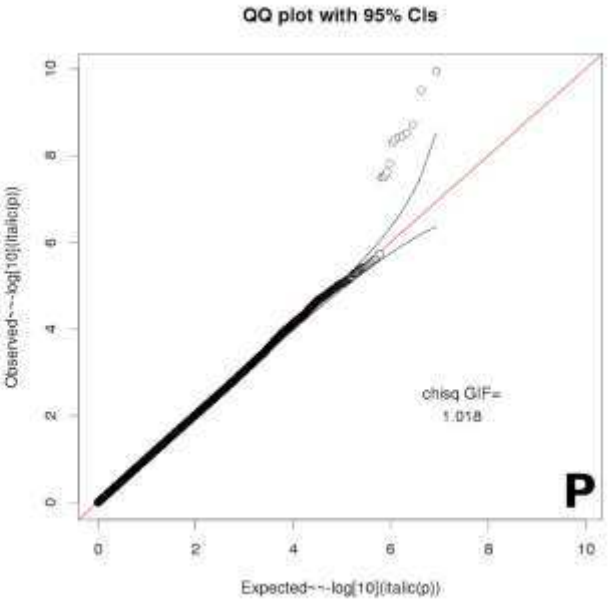

QQ plot with 95% CIs
chisq GIF=
1.0117
K
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))

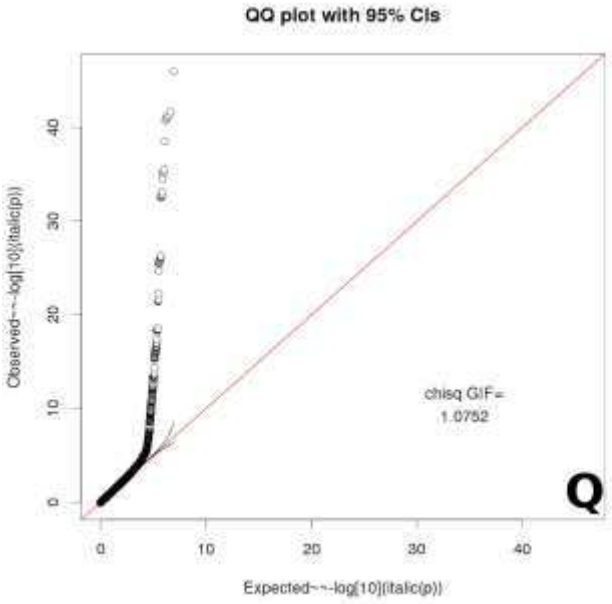

QQ plot with 95% CIs
chisq GIF=
1.0122
L
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))

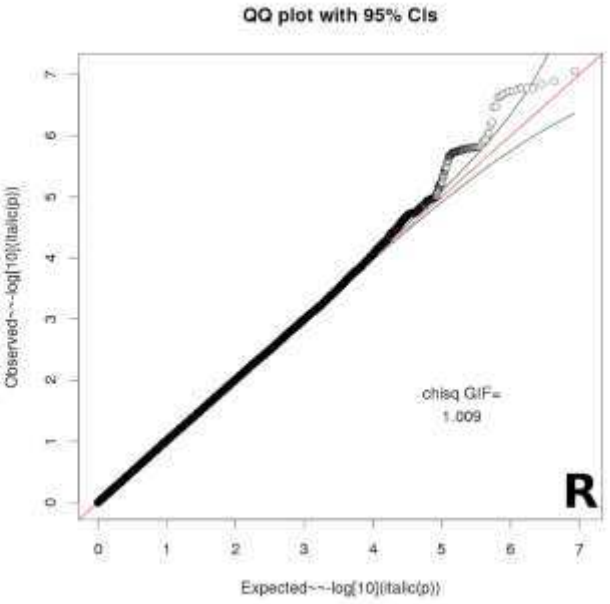

QQ plot with 95% CIs
chisq GIF=
1.0164
M
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))

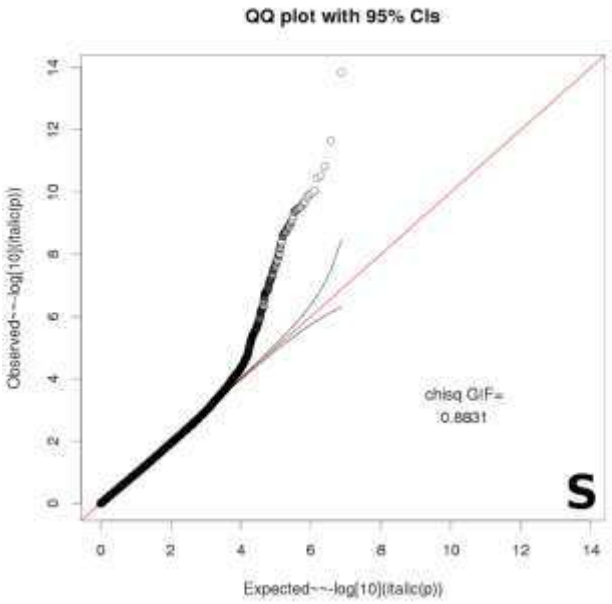

QQ plot with 95% CIs
chisq GIF=
1.033
N
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))

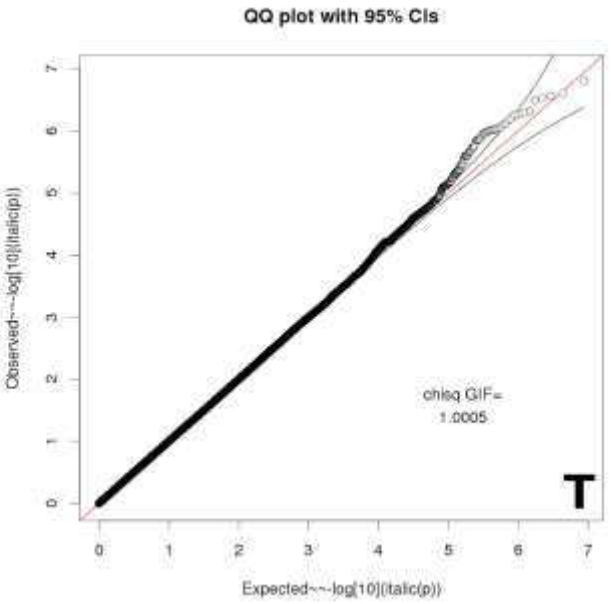

QQ plot with 95% CIs
chisq GIF=
1.0954
O
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))

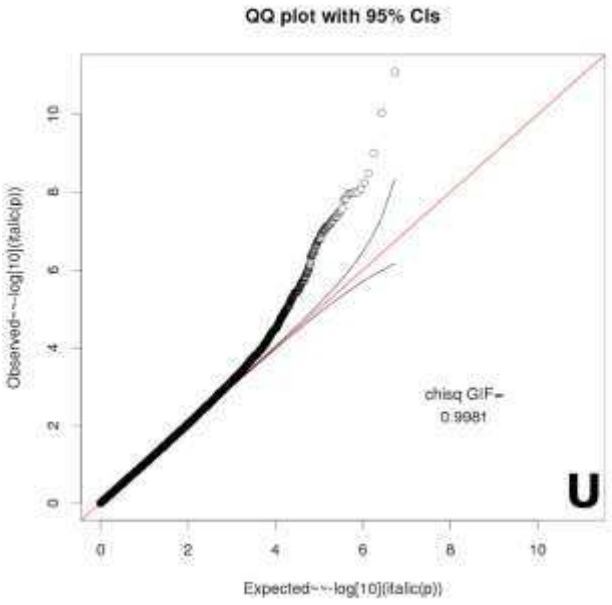

QQ plot with 95% CIs
chisq GIF=
1.018
P
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))

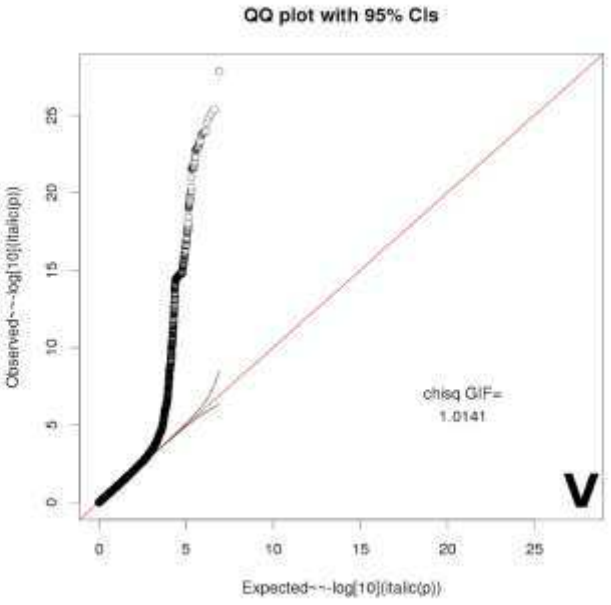

QQ plot with 95% CIs
chisq GIF=
1.0752
Q
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))

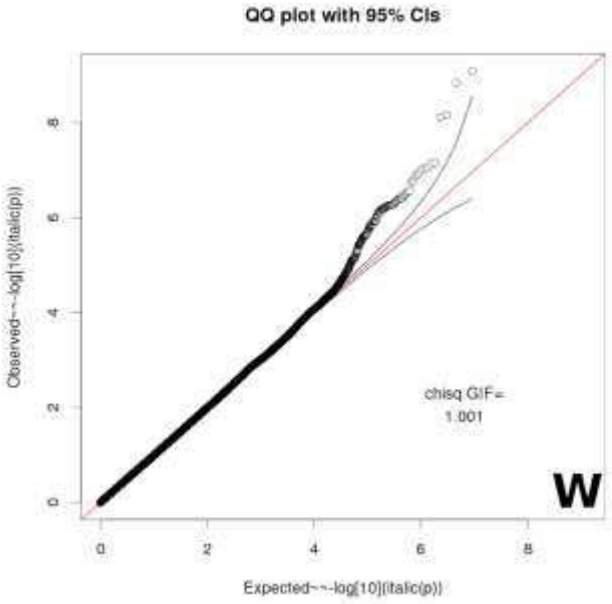

QQ plot with 95% CIs
chisq GIF=
1.009
R
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))


QQ plot with 95% CIs
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))
chisq GIF=
0.8831
S


QQ plot with 95% CIs
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))
chisq GIF=
1.0005
T

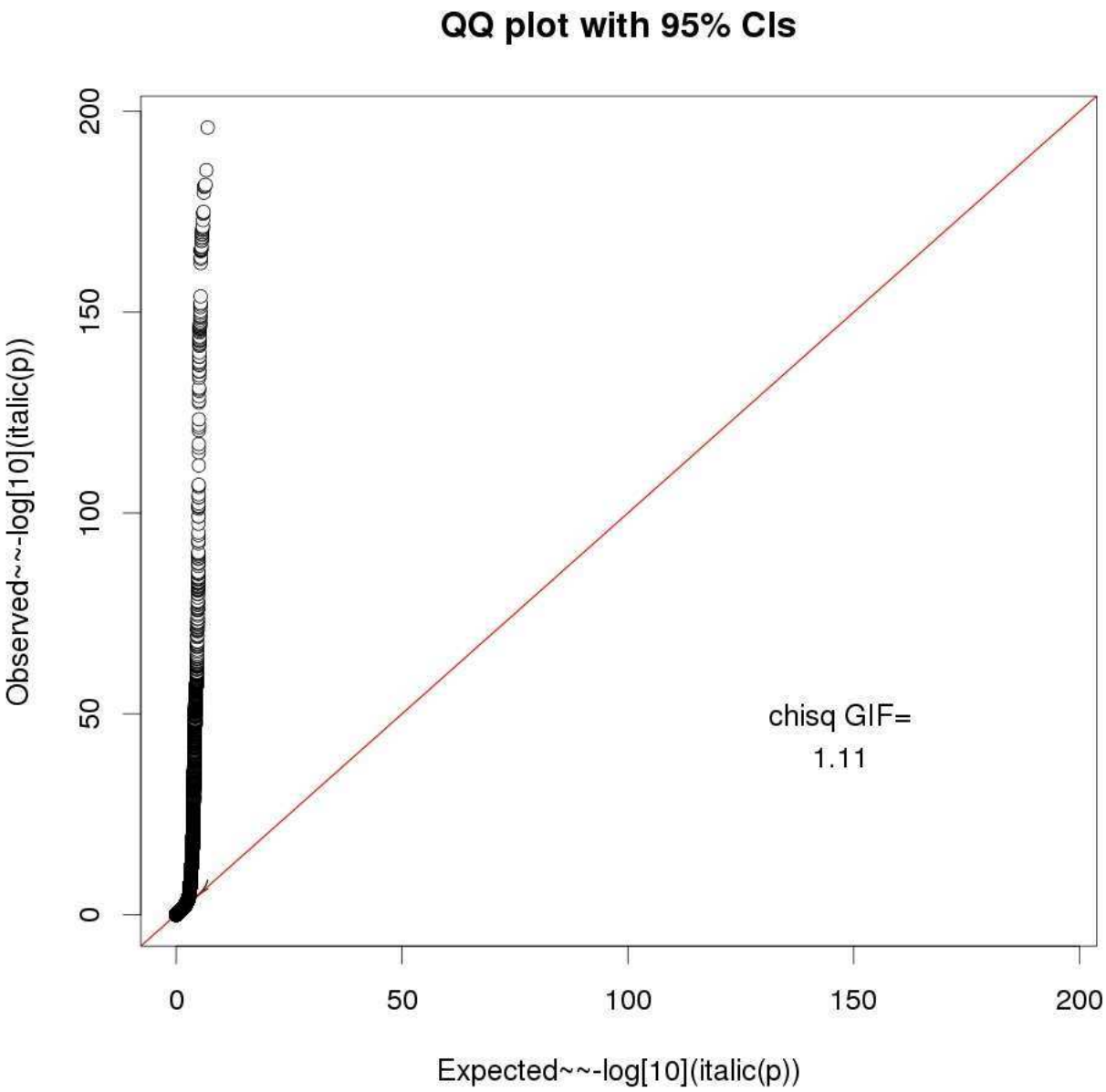

QQ plot with 95% CIs
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))
chisq GIF=
0.9981
U


QQ plot with 95% CIs
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))
chisq GIF=
1.0141
V

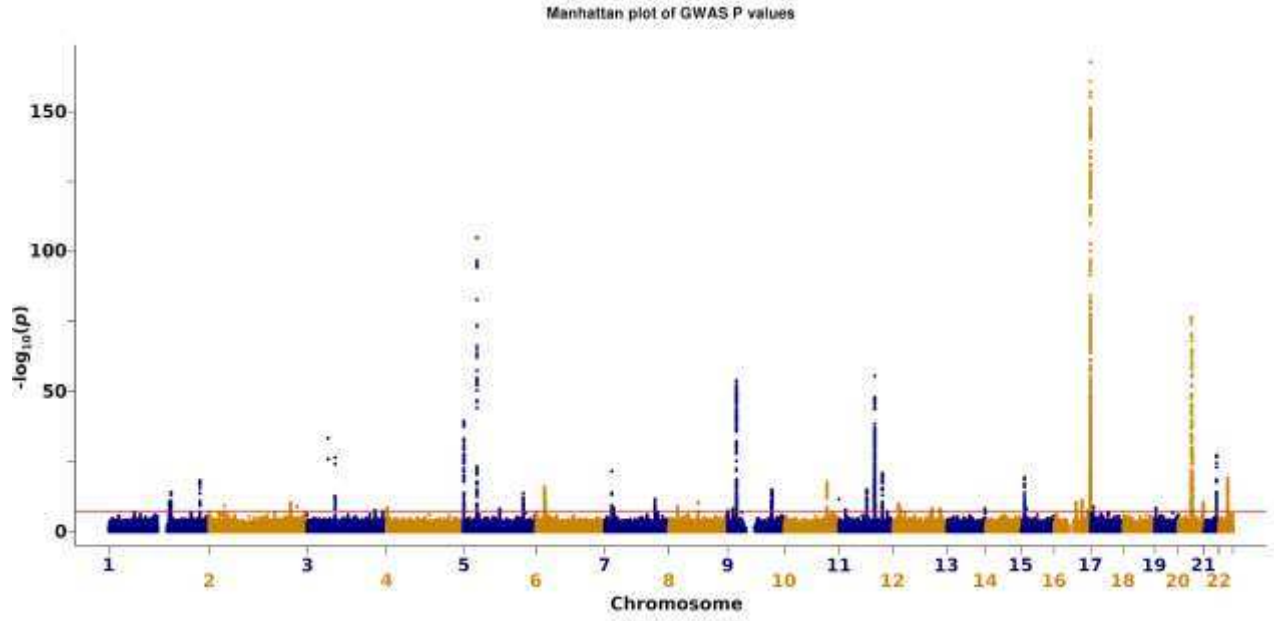

QQ plot with 95% CIs
Observed~~-log[10](italic(p))
Expected~~-log[10](italic(p))
chisq GIF=
1.001
W

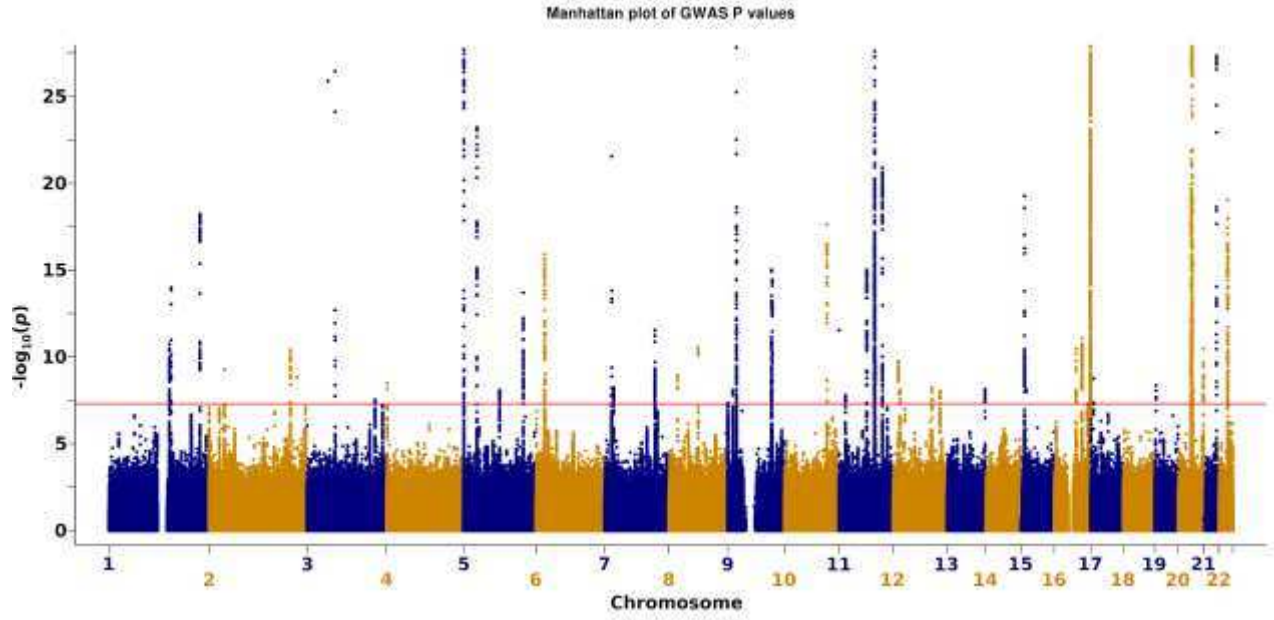
Supplementary Figure 1: Quantile-Quantile plots.

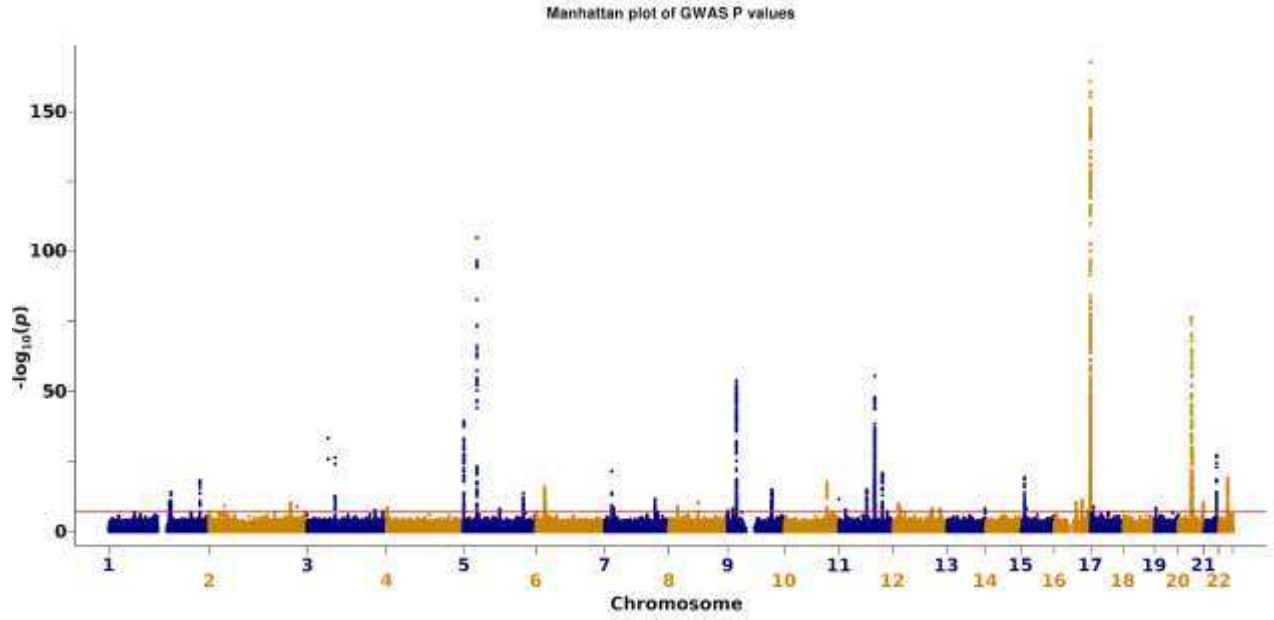
Quantile-quantile plots for all post QC imputed SNPs with a MAF > 0.005. A: AMFS; B: CIDR_WAMHS; C: EPIGENE; D: GenoMEL_P1; E: GenoMEL_P2; F: HARVARD; G: HEIDELBERG; H: MDACC; I: MELARISK; J: MIA_PAH; K: QMEGA_610k; L: QMEGA_omni; M: UKBB500K_pathconf; N: GSEdinCIDRchr; O: italy; P: greek; Q: spain; R: PLCO; S: MICHIGAN; T: QSkin; U: BNMS; V: 23andMe; W: UKBB500K_selfrep.



Supplementary Figure 2: Quantile-Quantile plot of total CM meta-analysis.

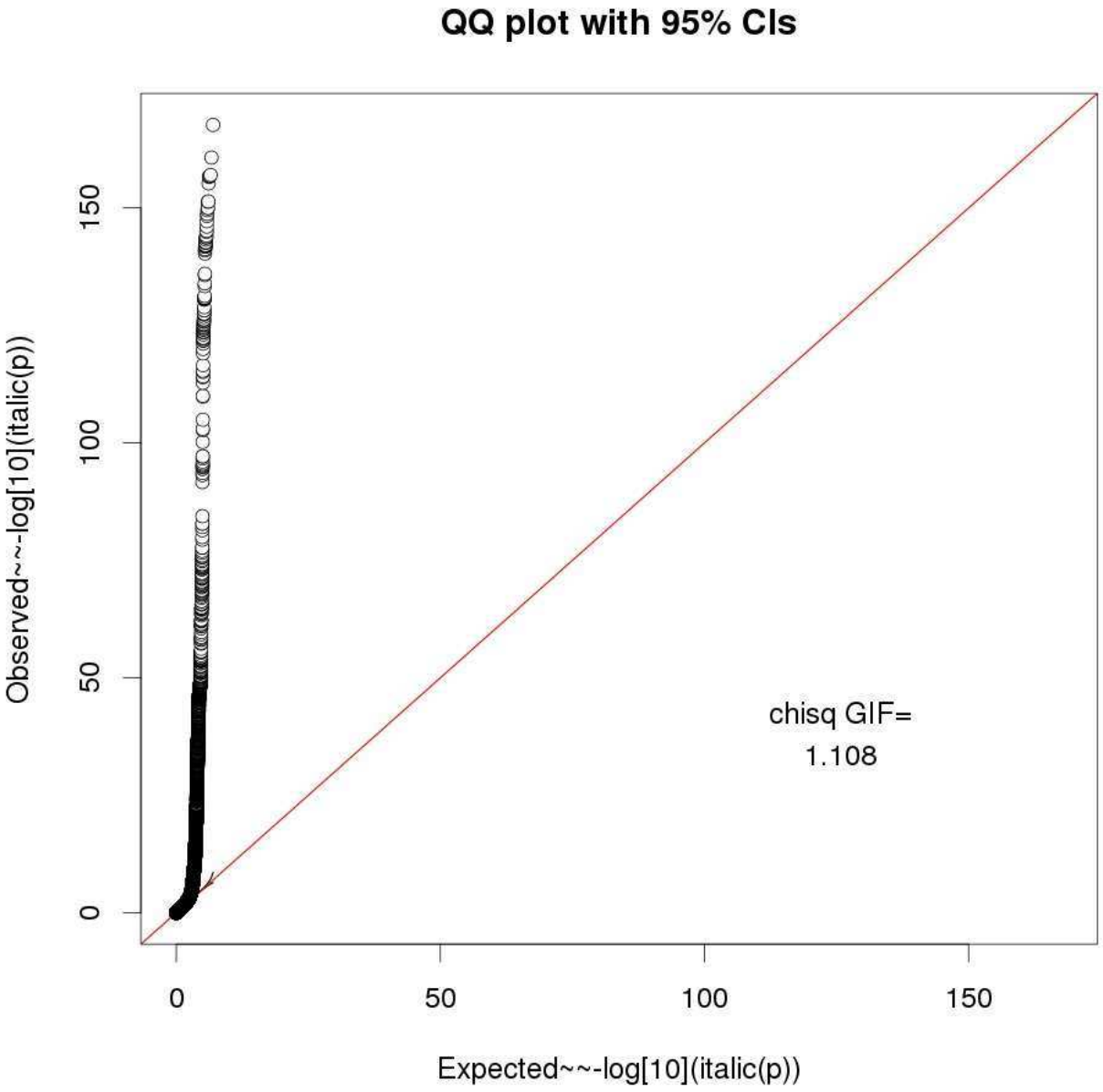

Manhattan plot of GWAS P values
-log10(p)
150
100
50
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Chromosome


Manhattan plot of GWAS P values
-log10(p)
25
20
15
10
5
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Chromosome

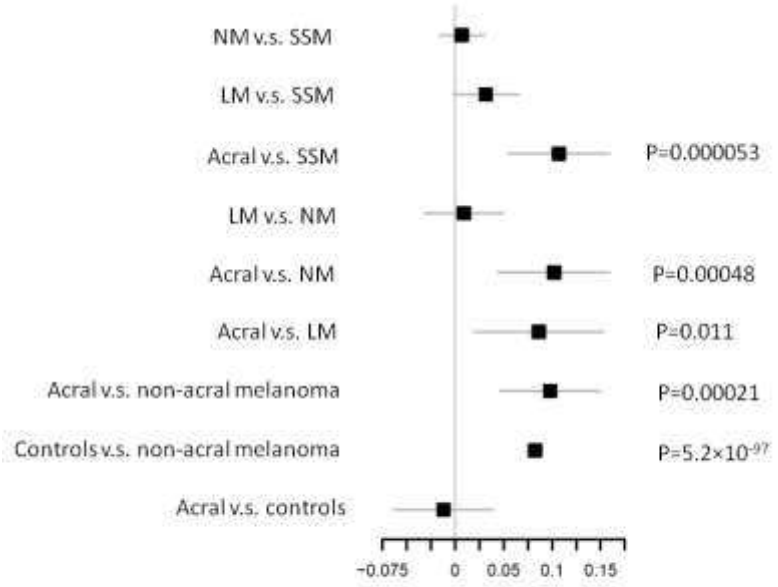

Manhattan plot of GWAS P values
-log10(p)
150
100
50
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Chromosome

Supplementary Figure 3: Manhattan plots of total and confirmed only GWAS-meta-analyses.

We display in order the total CM meta-analysis without limiting the y-axis; the pathologically confirmed CM cases only meta-analysis with the y-axis limited to 1×10^{-25} and without a limit to more clearly display loci other than MC1R.

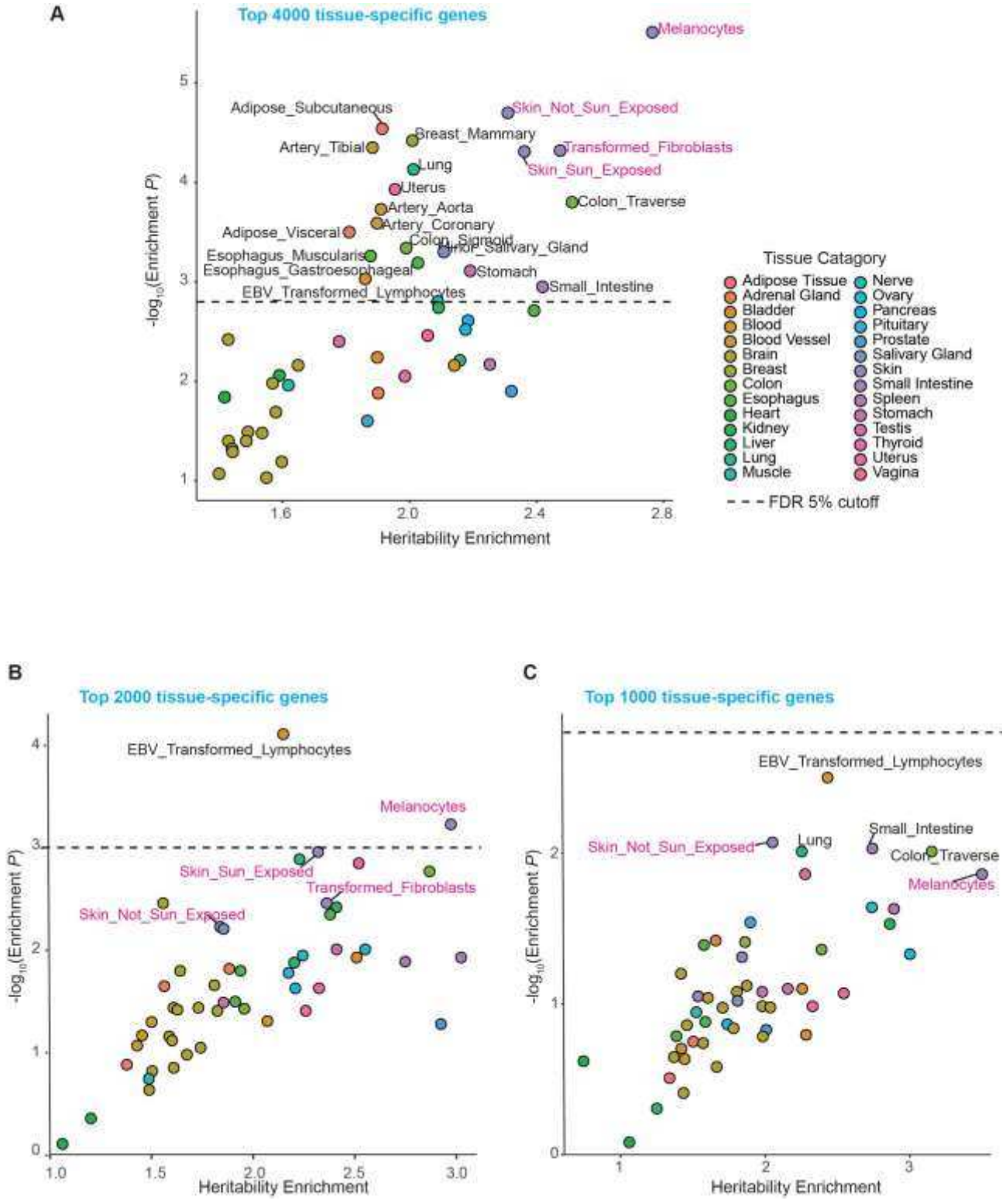


Supplementary Figure 4: Quantile-Quantile plot of confirmed only CM meta-analysis.



Supplementary Figure 5: Distribution of pigmentation polygenic risk scores across melanoma histological subtypes.

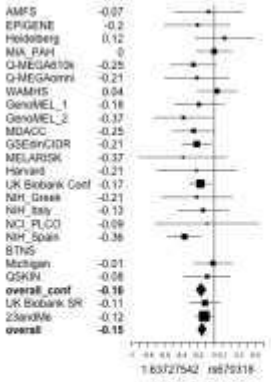
The figure shows whether PRS defined based on SNPs associated with hair colour differ across CM histological types (**Online Methods**; SSM: superficial spreading melanoma; NM: nodular melanoma; LM: lentigo melanoma; Acral: acral lentiginous melanoma). The higher the PRS the lighter the hair colour. When comparing subtype 1 vs. subtype 2, we report the effect size for the linear regression of PRS on subtype 1, including study and principal components as covariates to control for population stratification. The regression coefficient, 95% confidence interval, and statistical significance are shown. The positive beta indicates the PRS is higher in subtype 2 (e.g., non-acral melanomas).

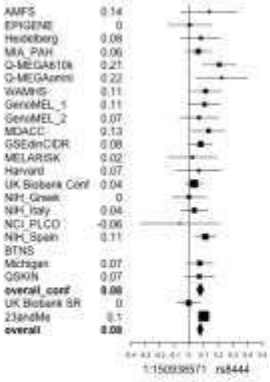


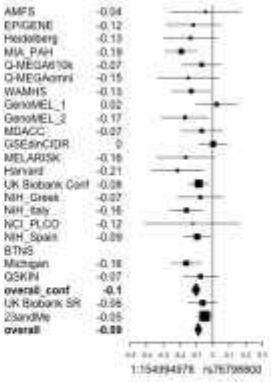
Supplementary Figure 6: LD score regression plots.

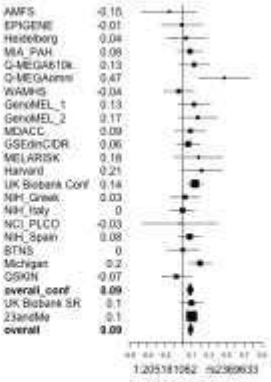
LD score regression was performed for the top 4000 (A) 2000 (B) and 1000 (C) tissue-specific genes from melanocyte and GTEx tissue types, to assess the enrichment of melanoma heritability in these genomic regions. The level of enrichment and P-values are shown, with an FDR = 0.05 cutoff marked as a dashed horizontal line. Tissue categories are color-coded, and a subset of top individual

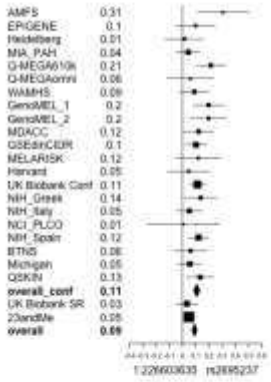
tissue types are shown on the plot. Tissue types from “Skin” category including melanocytes are highlighted in magenta.

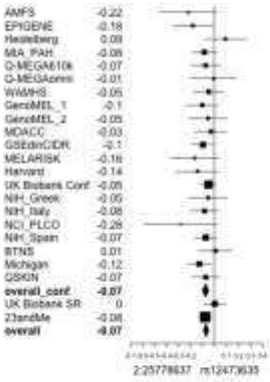

AMFS -0.07
EPIGENE -0.2
Heidelberg 0.12
MIA_PAH 0
Q-MEGA610k -0.25
Q-MEGAomni -0.21
WAMHS 0.04
GenoMEL_1 -0.18
GenoMEL_2 -0.37
MDACC -0.25
GSEdinCIDR -0.21
MELARISK -0.37
Harvard -0.21
UK Biobank Cont -0.17
NIH_Greek -0.21
NIH_Italy -0.13
NCI_PLCO -0.09
NIH_Spain -0.38
BTNS
Michigan -0.21
QSKIN 0.08
overall_cont -0.16
UK Biobank SR -0.11
23andMe -0.12
overall -0.15
1:63727542 rs875318

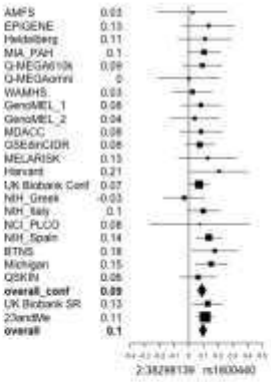

AMFS 0.14
EPIGENE 0
Heidelberg 0.09
MIA_PAH 0.06
Q-MEGA610k 0.21
Q-MEGAomni 0.22
WAMHS 0.11
GenoMEL_1 0.11
GenoMEL_2 0.07
MDACC 0.13
GSEdinCIDR 0.08
MELARISK 0.02
Harvard 0.07
UK Biobank Cont 0.04
NIH_Greek 0
NIH_Italy 0.04
NCI_PLCO -0.06
NIH_Spain 0.11
BTNS
Michigan 0.07
QSKIN 0.07
overall_cont 0.08
UK Biobank SR 0
23andMe 0.1
overall 0.08
1:150938571 rs8444

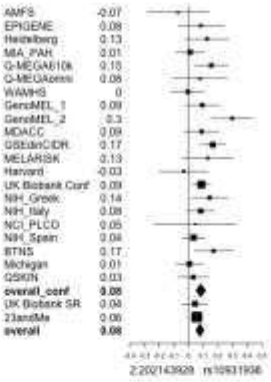

AMFS -0.04
EPIGENE -0.12
Heidelberg -0.13
MIA_PAH -0.19
Q-MEGA610k -0.27
Q-MEGAomni -0.15
WAMHS -0.15
GenoMEL_1 0.02
GenoMEL_2 -0.17
MDACC -0.07
GSEdinCIDR 0
MELARISK -0.16
Harvard -0.21
UK Biobank Cont -0.08
NIH_Greek -0.07
NIH_Italy -0.16
NCI_PLCO -0.12
NIH_Spain -0.09
BTNS
Michigan -0.18
QSKIN -0.07
overall_cont -0.1
UK Biobank SR -0.06
23andMe -0.05
overall -0.09
1:154994978 rs76798800

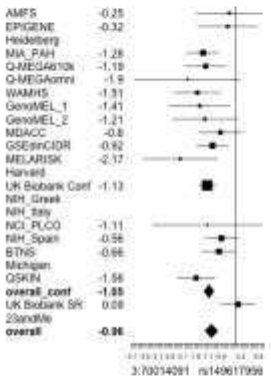

AMFS -0.15
EPIGENE -0.01
Heidelberg 0.04
MIA_PAH 0.08
Q-MEGA610k 0.13
Q-MEGAomni 0.47
WAMHS -0.04
GenoMEL_1 0.13
GenoMEL_2 0.17
MDACC 0.09
GSEdinCIDR 0.06
MELARISK 0.18
Harvard 0.21
UK Biobank Cont 0.14
NIH_Greek 0.03
NIH_Italy 0
NCI_PLCO -0.03
NIH_Spain 0.08
BTNS 0
Michigan -0.2
QSKIN -0.07
overall_cont 0.09
UK Biobank SR 0.1
23andMe 0.1
overall 0.09
1:205161052 rs2369633

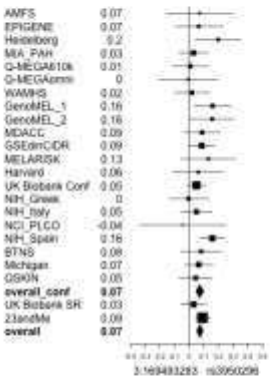

AMFS 0.31
EPIGENE 0.1
Heidelberg 0.01
MIA_PAH 0.04
Q-MEGA610k 0.21
Q-MEGAomni 0.06
WAMHS 0.09
GenoMEL_1 0.2
GenoMEL_2 0.2
MDACC 0.12
GSEdinCIDR 0.1
MELARISK 0.12
Harvard 0.05
UK Biobank Cont 0.11
NIH_Greek 0.14
NIH_Italy 0.05
NCI_PLCO 0.01
NIH_Spain 0.12
BTNS 0.08
Michigan 0.05
QSKIN 0.13
overall_cont 0.11
UK Biobank SR 0.03
23andMe 0.05
overall 0.09
1:229603633 rs2895237

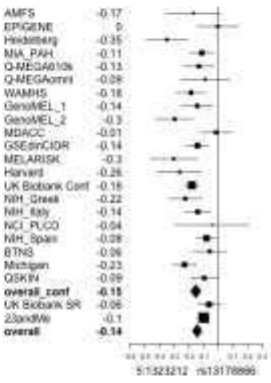

AMFS -0.22
EPIGENE -0.18
Heidelberg 0.09
MIA_PAH -0.08
Q-MEGA610k -0.07
Q-MEGAomni -0.01
WAMHS -0.05
GenoMEL_1 -0.1
GenoMEL_2 -0.05
MDACC -0.03
GSEdinCIDR -0.1
MELARISK -0.16
Harvard -0.14
UK Biobank Cont -0.05
NIH_Greek -0.05
NIH_Italy -0.08
NCI_PLCO -0.28
NIH_Spain -0.07
BTNS 0.01
Michigan -0.12
QSKIN -0.07
overall_cont -0.07
UK Biobank SR 0
23andMe -0.08
overall -0.07
2:25778637 rs12473635

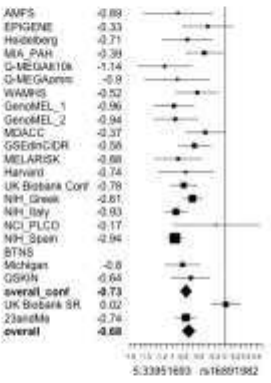

AMFS 0.03
EPIGENE 0.13
Heidelberg 0.11
MIA_PAH 0.1
Q-MEGA610k 0.09
Q-MEGAomni 0
WAMHS 0.03
GenoMEL_1 0.08
GenoMEL_2 0.04
MDACC 0.08
GSEdinCIDR 0.08
MELARISK 0.13
Harvard 0.21
UK Biobank Cont 0.07
NIH_Greek -0.03
NIH_Italy 0.1
NCI_PLCO 0.08
NIH_Spain 0.14
BTNS 0.18
Michigan 0.15
QSKIN 0.08
overall_cont 0.09
UK Biobank SR 0.13
23andMe 0.11
overall 0.1
2:38288139 rs1600460

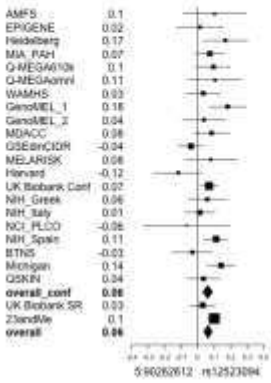

AMFS -0.07
EPIGENE 0.08
Heidelberg 0.13
MIA_PAH 0.01
Q-MEGA610k 0.15
Q-MEGAomni 0.08
WAMHS 0
GenoMEL_1 0.09
GenoMEL_2 0.3
MDACC 0.09
GSEdinCIDR 0.17
MELARISK 0.13
Harvard -0.03
UK Biobank Cont 0.09
NIH_Greek 0.14
NIH_Italy 0.08
NCI_PLCO 0.05
NIH_Spain 0.04
BTNS 0.17
Michigan 0.01
QSKIN 0.03
overall_cont 0.08
UK Biobank SR 0.04
23andMe 0.06
overall 0.08
2:202143929 rs10931936

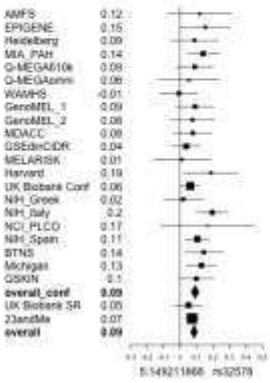

AMFS -0.25
EPIGENE -0.32
Heidelberg
MIA_PAH -1.28
Q-MEGA610k -1.19
Q-MEGAomni -1.9
WAMHS -1.31
GenoMEL_1 -1.41
GenoMEL_2 -1.21
MDACC -0.8
GSEdinCIDR -0.92
MELARISK -2.17
Harvard
UK Biobank Cont -1.13
NIH_Greek
NIH_Italy
NCI_PLCO -1.11
NIH_Spain -0.56
BTNS -0.66
Michigan
QSKIN -1.56
overall_cont -1.05
UK Biobank SR 0.09
23andMe
overall -0.36
3:70014091 rs149617999

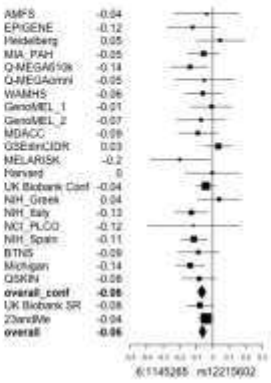

AMFS 0.07
EPIGENE 0.07
Heidelberg 0.2
MIA_PAH 0.03
Q-MEGA610k 0.01
Q-MEGAomni 0
WAMHS 0.02
GenoMEL_1 0.16
GenoMEL_2 0.16
MDACC 0.09
GSEdinCIDR 0.09
MELARISK 0.13
Harvard 0.06
UK Biobank Cont 0.05
NIH_Greek 0
NIH_Italy 0.05
NCI_PLCO -0.04
NIH_Spain 0.16
BTNS 0.08
Michigan 0.07
QSKIN 0.05
overall_cont 0.07
UK Biobank SR 0.03
23andMe 0.09
overall 0.07
3:169483283 rs3950296

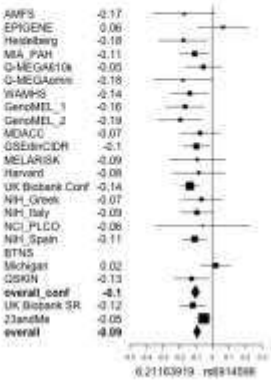

AMFS -0.17
EPIGENE 0
Heidelberg -0.35
MIA_PAH -0.11
Q-MEGA610k -0.13
Q-MEGAomni -0.09
WAMHS -0.16
GenoMEL_1 -0.14
GenoMEL_2 -0.3
MDACC -0.01
GSEdinCIDR -0.14
MELARISK -0.3
Harvard -0.26
UK Biobank Cont -0.16
NIH_Greek -0.22
NIH_Italy -0.14
NCI_PLCO -0.04
NIH_Spain -0.09
BTNS -0.06
Michigan -0.23
QSKIN -0.09
overall_cont -0.15
UK Biobank SR -0.06
23andMe -0.1
overall -0.14
5:1323212 rs13178896

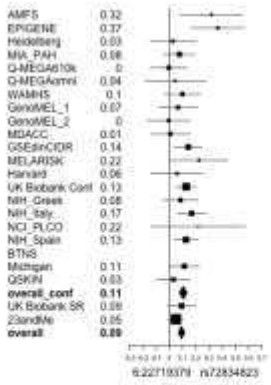

AMFS -0.89
EPIGENE -0.33
Heidelberg -0.71
MIA_PAH -0.39
Q-MEGA610k -1.14
Q-MEGAomni -0.9
WAMHS -0.52
GenoMEL_1 -0.96
GenoMEL_2 -0.94
MDACC -0.37
GSEdinCIDR -0.56
MELARISK -0.68
Harvard -0.74
UK Biobank Cont -0.78
NIH_Greek -0.81
NIH_Italy -0.93
NCI_PLCO -0.17
NIH_Spain -0.94
BTNS
Michigan -0.8
QSKIN -0.64
overall_cont -0.73
UK Biobank SR 0.02
23andMe -0.74
overall -0.68
5:33951693 rs16891982

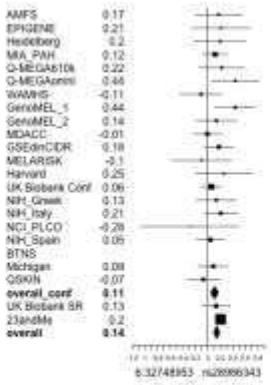

AMFS 0.1
EPIGENE 0.02
Heidelberg 0.17
MIA_PAH 0.07
Q-MEGA610k 0.1
Q-MEGAomni 0.11
WAMHS 0.03
GenoMEL_1 0.18
GenoMEL_2 0.04
MDACC 0.08
GSEdinCIDR -0.04
MELARISK 0.06
Harvard -0.12
UK Biobank Cont 0.07
NIH_Greek 0.06
NIH_Italy 0.01
NCI_PLCO -0.06
NIH_Spain 0.11
BTNS -0.03
Michigan 0.14
QSKIN 0.04
overall_cont 0.06
UK Biobank SR 0.03
23andMe 0.1
overall 0.06
5:80262612 rs12523094

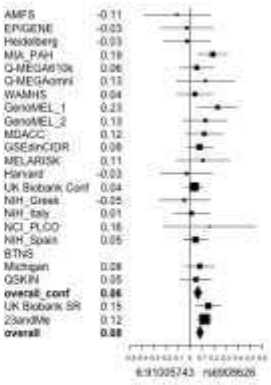

AMFS 0.12
EPIGENE 0.15
Heidelberg 0.09
MIA_PAH 0.14
Q-MEGA610k 0.08
Q-MEGAomni 0.06
WAMHS -0.01
GenoMEL_1 0.09
GenoMEL_2 0.08
MDACC 0.08
GSEdinCIDR 0.04
MELARISK 0.01
Harvard 0.19
UK Biobank Cont 0.06
NIH_Greek 0.02
NIH_Italy 0.2
NCI_PLCO 0.17
NIH_Spain 0.11
BTNS 0.14
Michigan 0.13
QSKIN 0.1
overall_cont 0.09
UK Biobank SR 0.05
23andMe 0.07
overall 0.09
5:149211868 rs32578

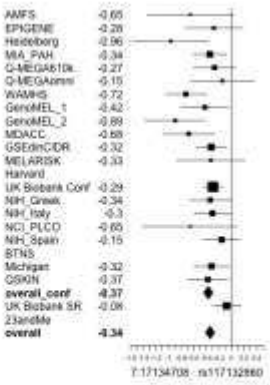

AMFS -0.04
EPIGENE -0.12
Heidelberg 0.05
MIA_PAH -0.05
Q-MEGA610k -0.14
Q-MEGAomni -0.05
WAMHS -0.06
GenoMEL_1 -0.01
GenoMEL_2 -0.07
MDACC -0.09
GSEdinCIDR 0.03
MELARISK -0.2
Harvard 0
UK Biobank Cont -0.04
NIH_Greek 0.04
NIH_Italy -0.13
NCI_PLCO -0.12
NIH_Spain -0.11
BTNS -0.09
Michigan -0.14
QSKIN -0.06
overall_cont -0.06
UK Biobank SR -0.08
23andMe -0.04
overall -0.06
6:1145265 rs12215402

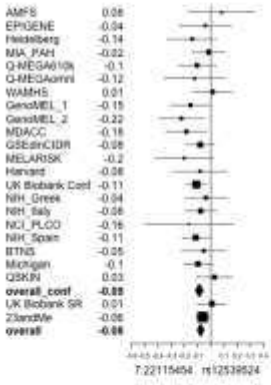

AMFS -0.17
EPIGENE 0.06
Heidelberg -0.18
MIA_PAH -0.11
Q-MEGA610k -0.05
Q-MEGAomni -0.18
WAMHS -0.14
GenoMEL_1 -0.16
GenoMEL_2 -0.19
MDACC -0.07
GSEdinCIDR -0.1
MELARISK -0.09
Harvard -0.08
UK Biobank Cont -0.14
NIH_Greek -0.07
NIH_Italy -0.09
NCI_PLCO -0.06
NIH_Spain -0.11
BTNS
Michigan 0.02
QSKIN -0.13
overall_cont -0.1
UK Biobank SR -0.12
23andMe -0.05
overall -0.09
6:21163919 rs6914588

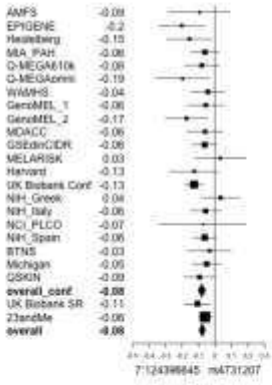

AMFS 0.32
EPIGENE 0.37
Heidelberg 0.03
MIA_PAH 0.08
Q-MEGA610k 0
Q-MEGAomni 0.04
WAMHS 0.1
GenoMEL_1 0.07
GenoMEL_2 0
MDACC 0.01
GSEdinCIDR 0.14
MELARISK 0.22
Harvard 0.06
UK Biobank Cont 0.13
NIH_Greek 0.08
NIH_Italy 0.17
NCI_PLCO 0.22
NIH_Spain 0.13
BTNS
Michigan 0.11
QSKIN 0.03
overall_conf 0.11
UK Biobank SR 0.09
23andMe 0.05
overall 0.09
6:22710378 rs72834823

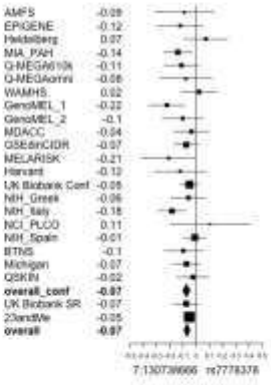

AMFS 0.17
EPIGENE 0.21
Heidelberg 0.2
MIA_PAH 0.12
Q-MEGA610k 0.22
Q-MEGAomni 0.46
WAMHS -0.11
GenoMEL_1 0.44
GenoMEL_2 0.14
MDACC -0.01
GSEdinCIDR 0.18
MELARISK -0.1
Harvard 0.25
UK Biobank Cont 0.06
NIH_Greek 0.13
NIH_Italy 0.21
NCI_PLCO -0.28
NIH_Spain 0.05
BTNS
Michigan 0.08
QSKIN -0.07
overall_conf 0.11
UK Biobank SR 0.13
23andMe 0.2
overall 0.14
6:32748953 rs28986343

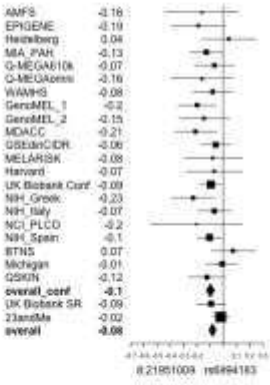

AMFS -0.11
EPIGENE -0.03
Heidelberg -0.03
MIA_PAH 0.19
Q-MEGA610k 0.06
Q-MEGAomni 0.13
WAMHS 0.04
GenoMEL_1 0.23
GenoMEL_2 0.13
MDACC 0.12
GSEdinCIDR 0.08
MELARISK 0.11
Harvard -0.03
UK Biobank Cont 0.04
NIH_Greek -0.05
NIH_Italy 0.01
NCI_PLCO 0.16
NIH_Spain 0.05
BTNS
Michigan 0.08
QSKIN 0.05
overall_conf 0.06
UK Biobank SR 0.15
23andMe 0.12
overall 0.08
6:91005743 rs6908626

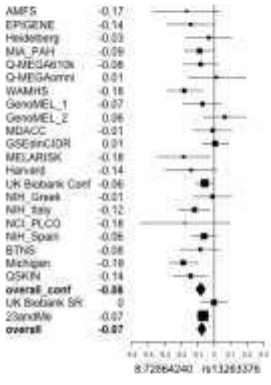

AMFS -0.65
EPIGENE -0.28
Heidelberg -0.06
MIA_PAH -0.34
Q-MEGA610k -0.27
Q-MEGAomni -0.15
WAMHS -0.72
GenoMEL_1 -0.42
GenoMEL_2 -0.89
MDACC -0.68
GSEdinCIDR -0.32
MELARISK -0.33
Harvard
UK Biobank Cont -0.29
NIH_Greek -0.34
NIH_Italy -0.3
NCI_PLCO -0.65
NIH_Spain -0.15
BTNS
Michigan -0.32
QSKIN -0.37
overall_conf -0.37
UK Biobank SR -0.06
23andMe
overall -0.34
7:17134708 rs117132860

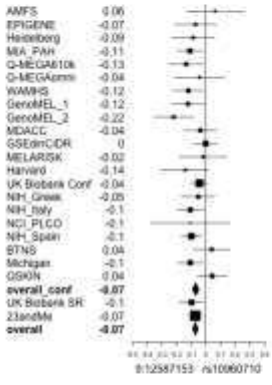

AMFS 0.08
EPIGENE -0.04
Heidelberg -0.14
MIA_PAH -0.02
Q-MEGA610k -0.1
Q-MEGAomni -0.12
WAMHS 0.01
GenoMEL_1 -0.15
GenoMEL_2 -0.22
MDACC -0.18
GSEdinCIDR -0.08
MELARISK -0.2
Harvard -0.08
UK Biobank Cont -0.11
NIH_Greek -0.04
NIH_Italy -0.08
NCI_PLCO -0.16
NIH_Spain -0.11
BTNS -0.05
Michigan -0.1
QSKIN 0.03
overall_conf -0.05
UK Biobank SR 0.01
23andMe -0.06
overall -0.06
7:22115454 rs12539524

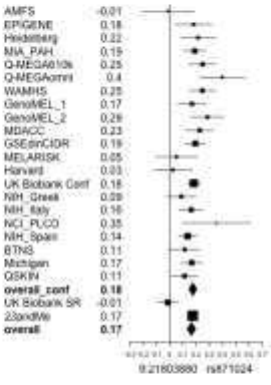

AMFS -0.09
EPIGENE -0.2
Heidelberg -0.15
MIA_PAH -0.06
Q-MEGA610k -0.06
Q-MEGAomni -0.19
WAMHS -0.04
GenoMEL_1 -0.06
GenoMEL_2 -0.17
MDACC -0.06
GSEdinCIDR -0.06
MELARISK 0.03
Harvard -0.13
UK Biobank Cont -0.13
NIH_Greek 0.04
NIH_Italy -0.06
NCI_PLCO -0.07
NIH_Spain -0.06
BTNS -0.03
Michigan -0.05
QSKIN -0.09
overall_conf -0.08
UK Biobank SR -0.11
23andMe -0.06
overall -0.08
7:124396845 rs4731207

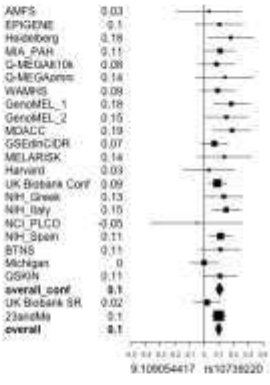

AMFS -0.09
EPIGENE -0.12
Heidelberg 0.07
MIA_PAH -0.14
Q-MEGA610k -0.11
Q-MEGAomni -0.08
WAMHS 0.02
GenoMEL_1 -0.22
GenoMEL_2 -0.1
MDACC -0.04
GSEdinCIDR -0.07
MELARISK -0.21
Harvard -0.12
UK Biobank Cont -0.05
NIH_Greek -0.06
NIH_Italy -0.15
NCI_PLCO 0.11
NIH_Spain -0.01
BTNS -0.1
Michigan -0.07
QSKIN -0.02
overall_conf -0.07
UK Biobank SR -0.07
23andMe -0.05
overall -0.07
7:130736666 rs7778378

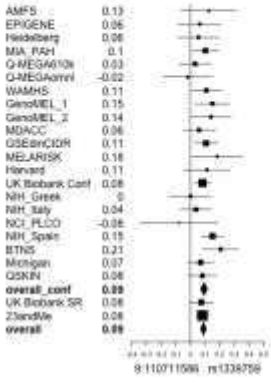

AMFS -0.16
EPIGENE -0.19
Heidelberg 0.04
MIA_PAH -0.13
Q-MEGA610k -0.07
Q-MEGAomni -0.16
WAMHS -0.08
GenoMEL_1 -0.2
GenoMEL_2 -0.15
MDACC -0.21
GSEdinCIDR -0.06
MELARISK -0.08
Harvard -0.07
UK Biobank Cont -0.09
NIH_Greek -0.23
NIH_Italy -0.07
NCI_PLCO -0.2
NIH_Spain -0.1
BTNS 0.07
Michigan -0.01
QSKIN -0.12
overall_conf -0.1
UK Biobank SR -0.09
23andMe -0.02
overall -0.08
8:21851009 rs6994183

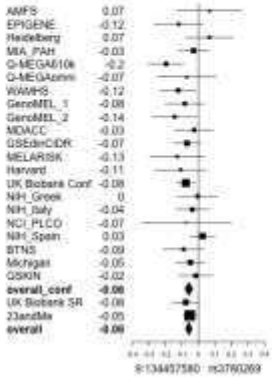

AMFS -0.17
EPIGENE -0.14
Heidelberg -0.03
MIA_PAH -0.09
Q-MEGA610k -0.08
Q-MEGAomni 0.01
WAMHS -0.18
GenoMEL_1 -0.07
GenoMEL_2 0.06
MDACC -0.01
GSEdinCIDR 0.01
MELARISK -0.18
Harvard -0.14
UK Biobank Cont -0.06
NIH_Greek -0.01
NIH_Italy -0.12
NCI_PLCO -0.18
NIH_Spain -0.06
BTNS -0.08
Michigan -0.19
QSKIN -0.14
overall_conf -0.09
UK Biobank SR 0
23andMe -0.07
overall -0.07
8:72864240 rs13263376

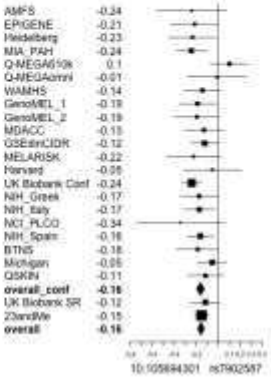

AMFS 0.06
EPIGENE -0.07
Heidelberg -0.09
MIA_PAH -0.11
Q-MEGA610k -0.13
Q-MEGAomni -0.04
WAMHS -0.12
GenoMEL_1 -0.12
GenoMEL_2 -0.22
MDACC -0.04
GSEdinCIDR 0
MELARISK -0.02
Harvard -0.14
UK Biobank Cont -0.04
NIH_Greek -0.05
NIH_Italy -0.1
NCI_PLCO -0.1
NIH_Spain -0.1
BTNS 0.04
Michigan -0.1
QSKIN 0.04
overall_conf -0.07
UK Biobank SR -0.1
23andMe -0.07
overall -0.07
8:12587153 rs10960710

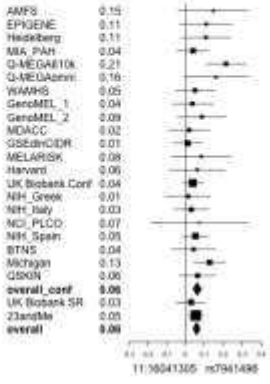

AMFS -0.01
EPIGENE 0.18
Heidelberg 0.22
MIA_PAH 0.19
Q-MEGA610k 0.25
Q-MEGAomni 0.4
WAMHS 0.25
GenoMEL_1 0.17
GenoMEL_2 0.29
MDACC 0.23
GSEdinCIDR 0.19
MELARISK 0.05
Harvard 0.03
UK Biobank Cont 0.18
NIH_Greek 0.09
NIH_Italy 0.16
NCI_PLCO 0.35
NIH_Spain 0.14
BTNS 0.11
Michigan 0.17
QSKIN 0.11
overall_conf 0.18
UK Biobank SR -0.01
23andMe 0.17
overall 0.17
9:21803880 rs871024

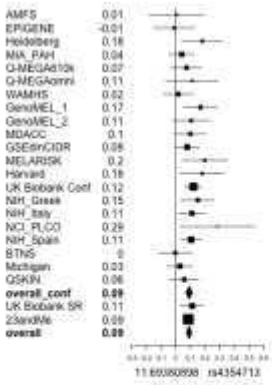

AMFS 0.03
EPIGENE 0.1
Heidelberg 0.18
MIA_PAH 0.11
Q-MEGA610k 0.08
Q-MEGAomni 0.14
WAMHS 0.09
GenoMEL_1 0.18
GenoMEL_2 0.15
MDACC 0.19
GSEdinCIDR 0.07
MELARISK 0.14
Harvard 0.03
UK Biobank Cont 0.09
NIH_Greek 0.13
NIH_Italy 0.15
NCI_PLCO -0.05
NIH_Spain 0.11
BTNS 0.11
Michigan 0
QSKIN 0.11
overall_conf 0.1
UK Biobank SR 0.02
23andMe 0.1
overall 0.1
9:106054417 rs10738220

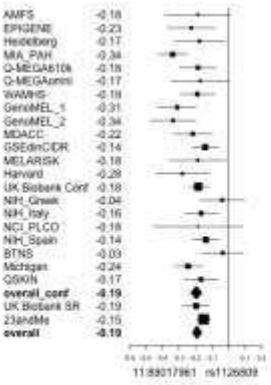

AMFS 0.13
EPIGENE 0.06
Heidelberg 0.06
MIA_PAH 0.1
Q-MEGA610k 0.03
Q-MEGAomni -0.02
WAMHS 0.11
GenoMEL_1 0.15
GenoMEL_2 0.14
MDACC 0.06
GSEdinCIDR 0.11
MELARISK 0.18
Harvard 0.11
UK Biobank Cont 0.08
NIH_Greek 0
NIH_Italy 0.04
NCI_PLCO -0.06
NIH_Spain 0.15
BTNS 0.21
Michigan 0.07
QSKIN 0.06
overall_conf 0.09
UK Biobank SR 0.06
23andMe 0.08
overall 0.09
9:110711586 rs1338759

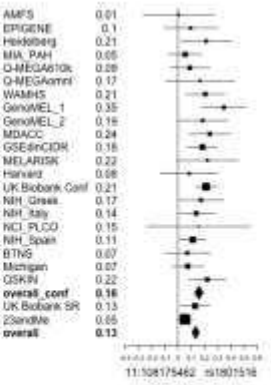

AMFS 0.07
EPIGENE -0.12
Heidelberg 0.07
MIA_PAH -0.03
Q-MEGA610k -0.2
Q-MEGAomni -0.07
WAMHS -0.12
GenoMEL_1 -0.08
GenoMEL_2 -0.14
MDACC -0.03
GSEdinCIDR -0.07
MELARISK -0.13
Harvard -0.11
UK Biobank Cont -0.08
NIH_Greek 0
NIH_Italy -0.04
NCI_PLCO -0.07
NIH_Spain 0.03
BTNS -0.09
Michigan -0.05
QSKIN -0.02
overall_conf -0.06
UK Biobank SR -0.06
23andMe -0.05
overall -0.06
9:134457580 rs3760269

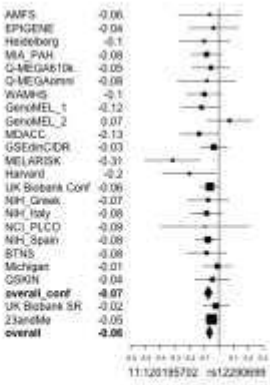

AMFS -0.24
EPIGENE -0.21
Heidelberg -0.23
MIA_PAH -0.24
Q-MEGA610k 0.1
Q-MEGAomni -0.01
WAMHS -0.14
GenoMEL_1 -0.19
GenoMEL_2 -0.19
MDACC -0.13
GSEdinCIDR -0.12
MELARISK -0.22
Harvard -0.05
UK Biobank Cont -0.24
NIH_Greek -0.17
NIH_Italy -0.17
NCI_PLCO -0.34
NIH_Spain -0.16
BTNS -0.16
Michigan -0.05
QSKIN -0.11
overall_conf -0.15
UK Biobank SR -0.12
23andMe -0.15
overall -0.15
10:105894301 rs7902587

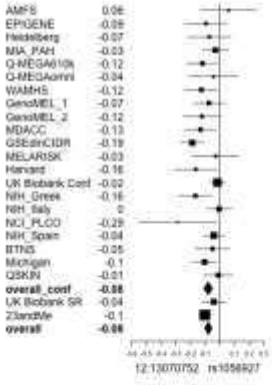

AMFS 0.15
EPIGENE 0.11
Heidelberg 0.11
MIA_PAH 0.04
Q-MEGA610k 0.21
Q-MEGAomni 0.16
WAMHS 0.05
GenoMEL_1 0.04
GenoMEL_2 0.09
MDACC 0.02
GSEdinCIDR 0.01
MELARISK 0.08
Harvard 0.06
UK Biobank Cont 0.04
NIH_Greek 0.01
NIH_Italy 0.03
NCI_PLCO 0.07
NIH_Spain 0.05
BTNS 0.04
Michigan 0.13
QSKIN 0.06
overall_conf 0.06
UK Biobank SR 0.03
23andMe 0.05
overall 0.09
11:16041305 rs7941498

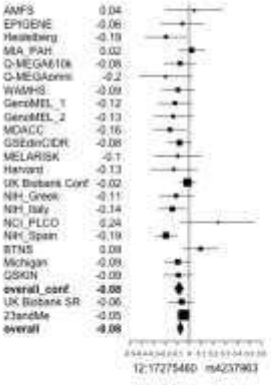

AMFS 0.01
EPIGENE -0.01
Heidelberg 0.18
MIA_PAH 0.04
Q-MEGA610k 0.07
Q-MEGAomni 0.11
WAMHS 0.62
GenoMEL_1 0.17
GenoMEL_2 0.11
MDACC 0.1
GSEdinCIDR 0.08
MELARISK 0.2
Harvard 0.18
UK Biobank Conf 0.12
NIH_Greek 0.15
NIH_Italy 0.11
NCI_PLCO 0.29
NIH_Spain 0.11
BTNS 0
Michigan 0.03
QSKIN 0.06
overall_conf 0.09
UK Biobank SR 0.11
23andMe 0.09
overall 0.09
11:66950938 rs4354713

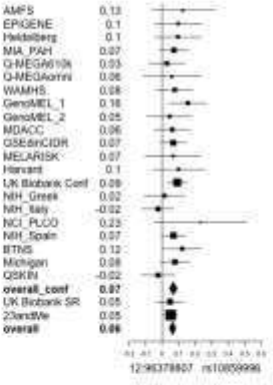

AMFS -0.18
EPIGENE -0.23
Heidelberg -0.17
MIA_PAH -0.34
Q-MEGA610k -0.18
Q-MEGAomni -0.17
WAMHS -0.19
GenoMEL_1 -0.31
GenoMEL_2 -0.34
MDACC -0.22
GSEdinCIDR -0.14
MELARISK -0.18
Harvard -0.28
UK Biobank Conf -0.18
NIH_Greek -0.04
NIH_Italy -0.16
NCI_PLCO -0.18
NIH_Spain -0.14
BTNS -0.03
Michigan -0.24
QSKIN -0.17
overall_conf -0.19
UK Biobank SR -0.19
23andMe -0.15
overall -0.19
11:69017861 rs1126809

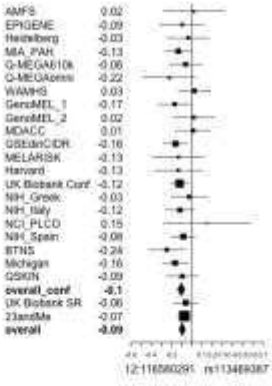

AMFS 0.01
EPIGENE 0.1
Heidelberg 0.21
MIA_PAH 0.05
Q-MEGA610k 0.09
Q-MEGAomni 0.17
WAMHS 0.21
GenoMEL_1 0.35
GenoMEL_2 0.19
MDACC 0.24
GSEdinCIDR 0.18
MELARISK 0.22
Harvard 0.08
UK Biobank Conf 0.21
NIH_Greek 0.17
NIH_Italy 0.14
NCI_PLCO 0.15
NIH_Spain 0.11
BTNS 0.07
Michigan 0.07
QSKIN 0.22
overall_conf 0.16
UK Biobank SR 0.13
23andMe 0.05
overall 0.13
11:108175462 rs1801516

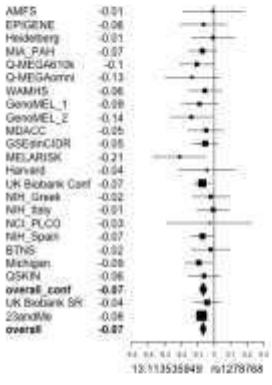

AMFS -0.06
EPIGENE -0.04
Heidelberg -0.1
MIA_PAH -0.08
Q-MEGA610k -0.05
Q-MEGAomni -0.08
WAMHS -0.1
GenoMEL_1 -0.12
GenoMEL_2 0.07
MDACC -0.13
GSEdinCIDR -0.03
MELARISK -0.31
Harvard -0.2
UK Biobank Conf -0.06
NIH_Greek -0.07
NIH_Italy -0.08
NCI_PLCO -0.09
NIH_Spain -0.08
BTNS -0.08
Michigan -0.01
QSKIN -0.08
overall_conf -0.07
UK Biobank SR -0.02
23andMe -0.05
overall -0.06
11:120195702 rs12290699

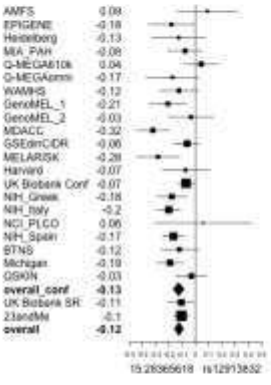

AMFS 0.06
EPIGENE -0.09
Heidelberg -0.07
MIA_PAH -0.03
Q-MEGA610k -0.12
Q-MEGAomni -0.04
WAMHS -0.12
GenoMEL_1 -0.07
GenoMEL_2 -0.12
MDACC -0.13
GSEdinCIDR -0.19
MELARISK -0.03
Harvard -0.16
UK Biobank Conf -0.02
NIH_Greek -0.16
NIH_Italy 0
NCI_PLCO -0.29
NIH_Spain -0.04
BTNS -0.05
Michigan -0.1
QSKIN -0.01
overall_conf -0.08
UK Biobank SR -0.04
23andMe -0.1
overall -0.08
12:13070752 rs1058827

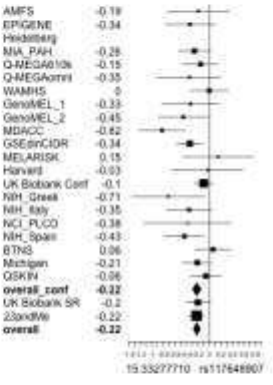

AMFS 0.04
EPIGENE -0.06
Heidelberg -0.19
MIA_PAH 0.02
Q-MEGA610k -0.08
Q-MEGAomni -0.2
WAMHS -0.09
GenoMEL_1 -0.12
GenoMEL_2 -0.13
MDACC -0.16
GSEdinCIDR -0.08
MELARISK -0.1
Harvard -0.13
UK Biobank Conf -0.02
NIH_Greek -0.11
NIH_Italy -0.14
NCI_PLCO 0.24
NIH_Spain -0.19
BTNS 0.08
Michigan -0.09
QSKIN -0.09
overall_conf -0.08
UK Biobank SR -0.06
23andMe -0.05
overall -0.08
12:17275460 rs4237903

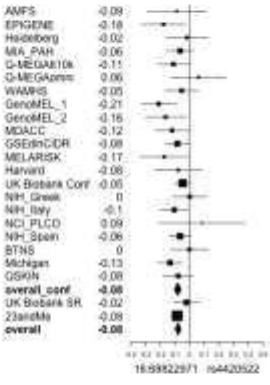

AMFS 0.13
EPIGENE 0.1
Heidelberg 0.1
MIA_PAH 0.07
Q-MEGA610k 0.03
Q-MEGAomni 0.06
WAMHS 0.08
GenoMEL_1 0.18
GenoMEL_2 0.05
MDACC 0.06
GSEdinCIDR 0.07
MELARISK 0.07
Harvard 0.1
UK Biobank Conf 0.09
NIH_Greek 0.02
NIH_Italy -0.02
NCI_PLCO 0.23
NIH_Spain 0.07
BTNS 0.12
Michigan 0.08
QSKIN -0.02
overall_conf 0.07
UK Biobank SR 0.05
23andMe 0.05
overall 0.06
12:96378807 rs10859996

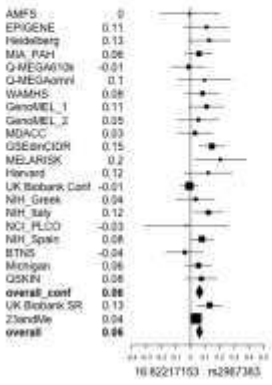

AMFS 0.02
EPIGENE -0.09
Heidelberg -0.03
MIA_PAH -0.13
Q-MEGA610k -0.06
Q-MEGAomni -0.22
WAMHS 0.03
GenoMEL_1 -0.17
GenoMEL_2 0.02
MDACC 0.01
GSEdinCIDR -0.16
MELARISK -0.13
Harvard -0.13
UK Biobank Conf -0.12
NIH_Greek -0.03
NIH_Italy -0.12
NCI_PLCO 0.15
NIH_Spain -0.08
BTNS -0.24
Michigan -0.16
QSKIN -0.09
overall_conf -0.1
UK Biobank SR -0.06
23andMe -0.07
overall -0.09
12:116580291 rs113469387

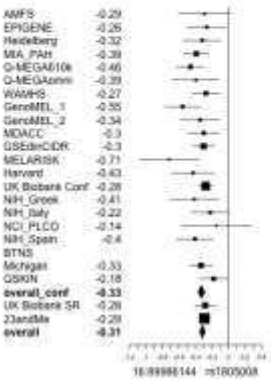

AMFS -0.01
EPIGENE -0.06
Heidelberg -0.01
MIA_PAH -0.07
Q-MEGA610k -0.1
Q-MEGAomni -0.13
WAMHS -0.06
GenoMEL_1 -0.09
GenoMEL_2 -0.14
MDACC -0.05
GSEdinCIDR -0.05
MELARISK -0.21
Harvard -0.04
UK Biobank Conf -0.07
NIH_Greek -0.02
NIH_Italy -0.01
NCI_PLCO -0.03
NIH_Spain -0.07
BTNS -0.02
Michigan -0.09
QSKIN -0.06
overall_conf -0.07
UK Biobank SR -0.04
23andMe -0.08
overall -0.07
13:113535849 rs1278788

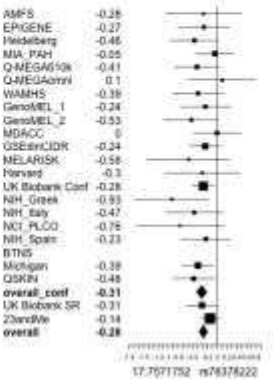

AMFS 0.08
EPIGENE -0.18
Heidelberg -0.13
MIA_PAH -0.08
Q-MEGA610k 0.04
Q-MEGAomni -0.17
WAMHS -0.12
GenoMEL_1 -0.21
GenoMEL_2 -0.03
MDACC -0.32
GSEdinCIDR -0.06
MELARISK -0.28
Harvard -0.07
UK Biobank Conf -0.07
NIH_Greek -0.18
NIH_Italy -0.2
NCI_PLCO 0.06
NIH_Spain -0.17
BTNS -0.12
Michigan -0.19
QSKIN -0.03
overall_conf -0.13
UK Biobank SR -0.11
23andMe -0.1
overall -0.12
15:28365618 rs12913832

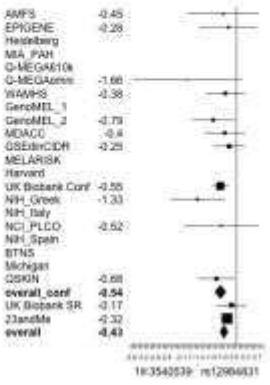

AMFS -0.19
EPIGENE -0.34
Heidelberg
MIA_PAH -0.28
Q-MEGA610k -0.15
Q-MEGAomni -0.35
WAMHS 0
GenoMEL_1 -0.33
GenoMEL_2 -0.45
MDACC -0.62
GSEdinCIDR -0.34
MELARISK 0.15
Harvard -0.03
UK Biobank Conf -0.1
NIH_Greek -0.71
NIH_Italy -0.35
NCI_PLCO -0.38
NIH_Spain -0.43
BTNS 0.06
Michigan -0.21
QSKIN -0.06
overall_conf -0.22
UK Biobank SR -0.2
23andMe -0.22
overall -0.22
15:33277710 rs117648807

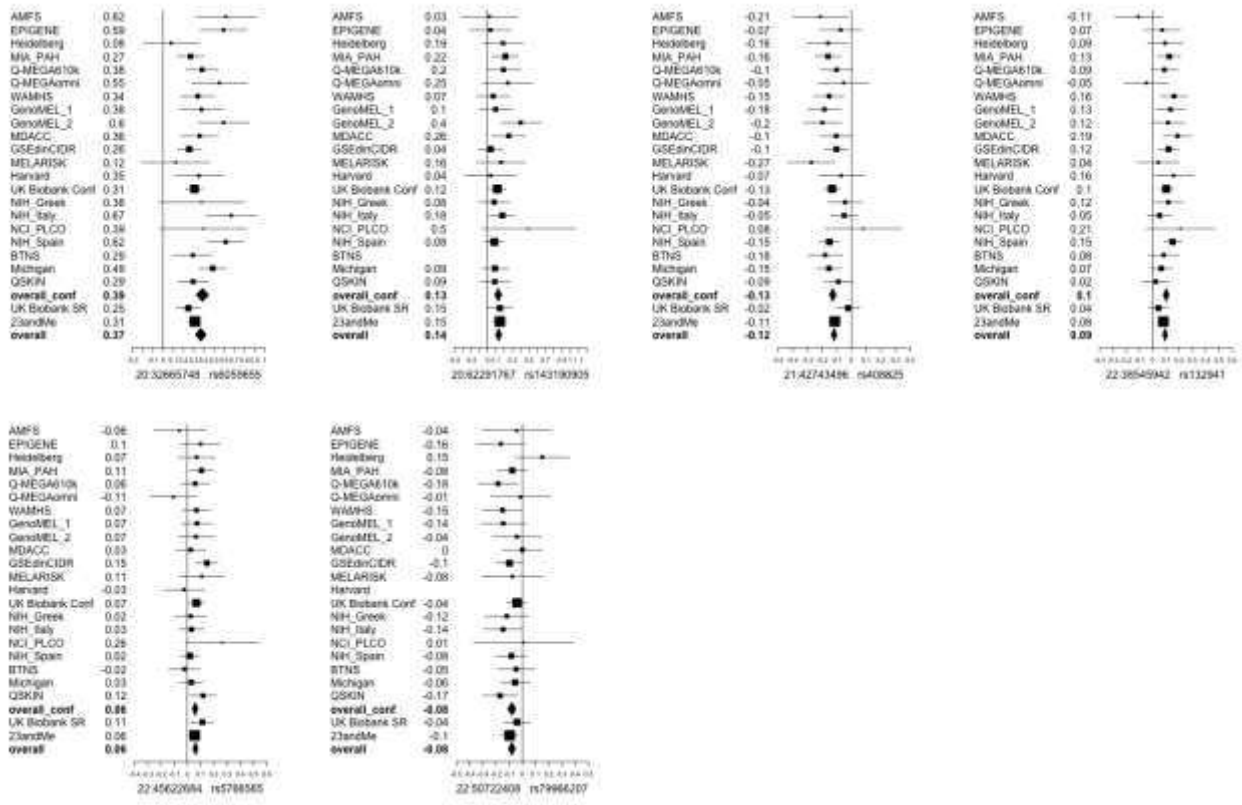

AMFS -0.09
EPIGENE -0.18
Heidelberg -0.02
MIA_PAH -0.06
Q-MEGA610k -0.11
Q-MEGAomni 0.06
WAMHS -0.05
GenoMEL_1 -0.21
GenoMEL_2 -0.16
MDACC -0.12
GSEdinCIDR -0.08
MELARISK -0.17
Harvard -0.08
UK Biobank Conf -0.05
NIH_Greek 0
NIH_Italy -0.1
NCI_PLCO 0.09
NIH_Spain -0.06
BTNS 0
Michigan -0.13
QSKIN -0.08
overall_conf -0.08
UK Biobank SR -0.02
23andMe -0.09
overall -0.08
16:89822971 rs4420522


AMFS 0
EPIGENE 0.11
Heidelberg 0.13
MIA_PAH 0.06
Q-MEGA610k -0.01
Q-MEGAomni 0.1
WAMHS 0.08
GenoMEL_1 0.11
GenoMEL_2 0.05
MDACC 0.03
GSEdinCIDR 0.15
MELARISK 0.2
Harvard 0.12
UK Biobank Conf -0.01
NIH_Greek 0.04
NIH_Italy 0.12
NCI_PLCO -0.03
NIH_Spain 0.08
BTNS -0.04
Michigan 0.06
QSKIN 0.08
overall_conf 0.06
UK Biobank SR 0.13
23andMe 0.04
overall 0.06
16:82217153 rs2987383


AMFS -0.29
EPIGENE -0.26
Heidelberg -0.32
MIA_PAH -0.39
Q-MEGA610k -0.46
Q-MEGAomni -0.39
WAMHS -0.27
GenoMEL_1 -0.55
GenoMEL_2 -0.34
MDACC -0.3
GSEdinCIDR -0.3
MELARISK -0.71
Harvard -0.43
UK Biobank Conf -0.28
NIH_Greek -0.41
NIH_Italy -0.22
NCI_PLCO -0.14
NIH_Spain -0.4
BTNS
Michigan -0.33
QSKIN -0.18
overall_conf -0.33
UK Biobank SR -0.28
23andMe -0.29
overall -0.31
16:89986144 rs1805008

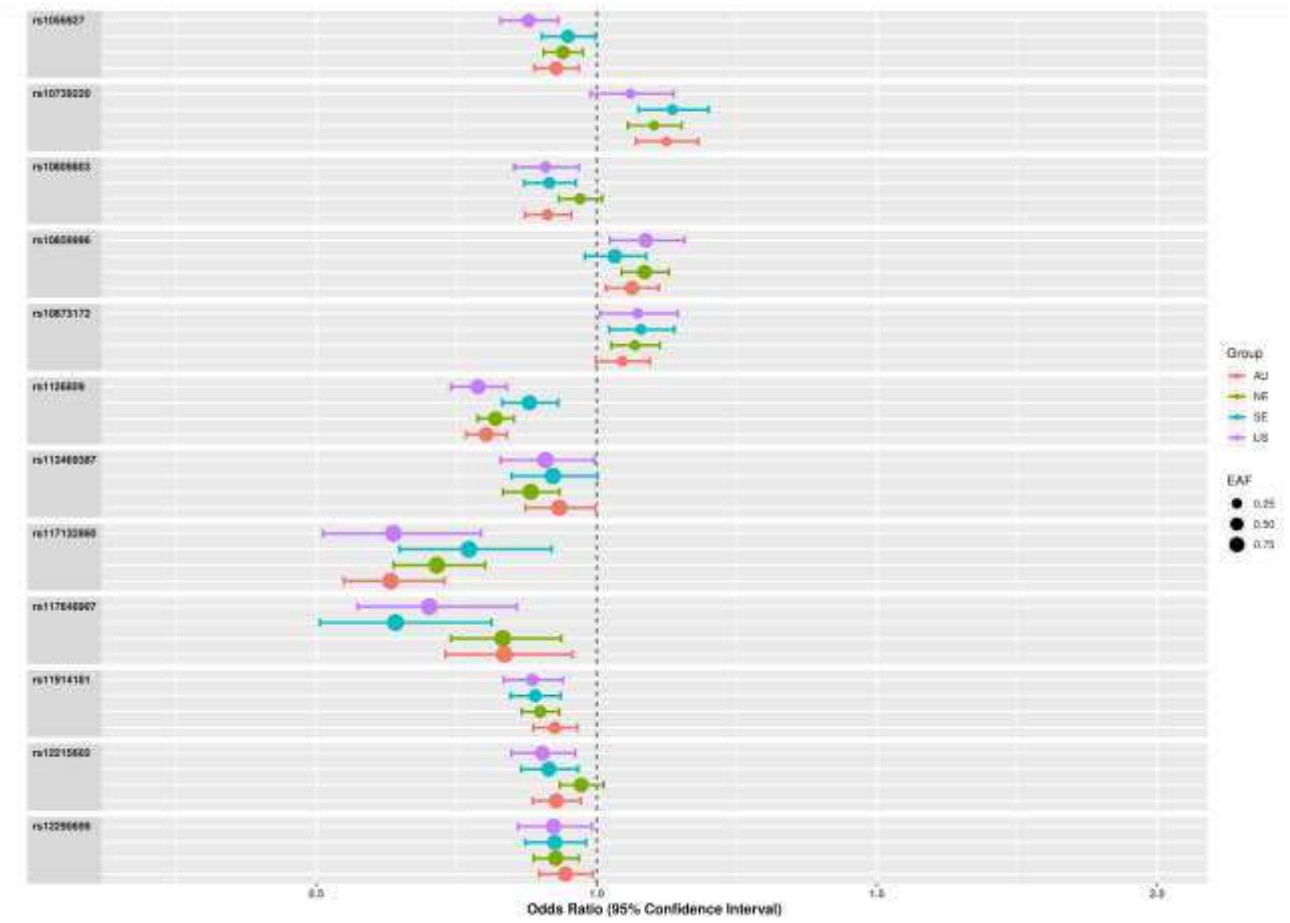

AMFS -0.28
EPIGENE -0.27
Heidelberg -0.46
MIA_PAH -0.05
Q-MEGA610k -0.41
Q-MEGAomni 0.1
WAMHS -0.39
GenoMEL_1 -0.24
GenoMEL_2 -0.53
MDACC 0
GSEdinCIDR -0.24
MELARISK -0.58
Harvard -0.3
UK Biobank Conf -0.28
NIH_Greek -0.93
NIH_Italy -0.47
NCI_PLCO -0.76
NIH_Spain -0.23
BTNS
Michigan -0.39
QSKIN -0.48
overall_conf -0.31
UK Biobank SR -0.31
23andMe -0.14
overall -0.28
17:7577752 rs76376222

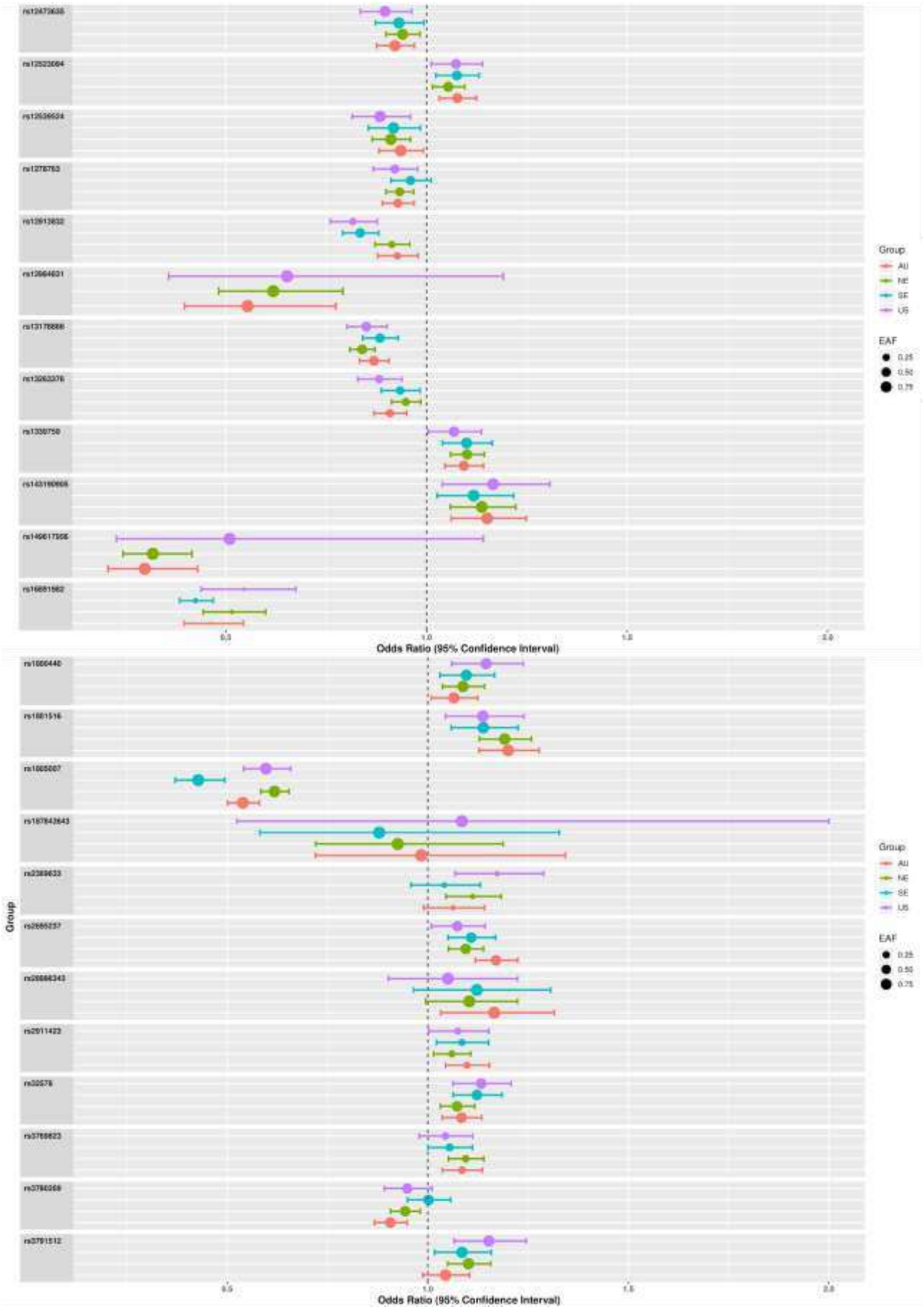

AMFS -0.45
EPIGENE -0.28
Heidelberg
MIA_PAH
Q-MEGA610k
Q-MEGAomni -1.66
WAMHS -0.38
GenoMEL_1
GenoMEL_2 -0.79
MDACC -0.4
GSEdinCIDR -0.29
MELARISK
Harvard
UK Biobank Conf -0.55
NIH_Greek -1.33
NIH_Italy
NCI_PLCO -0.62
NIH_Spain
BTNS
Michigan
QSKIN -0.68
overall_conf -0.54
UK Biobank SR -0.17
23andMe -0.32
overall -0.43
19:3540539 rs12984631

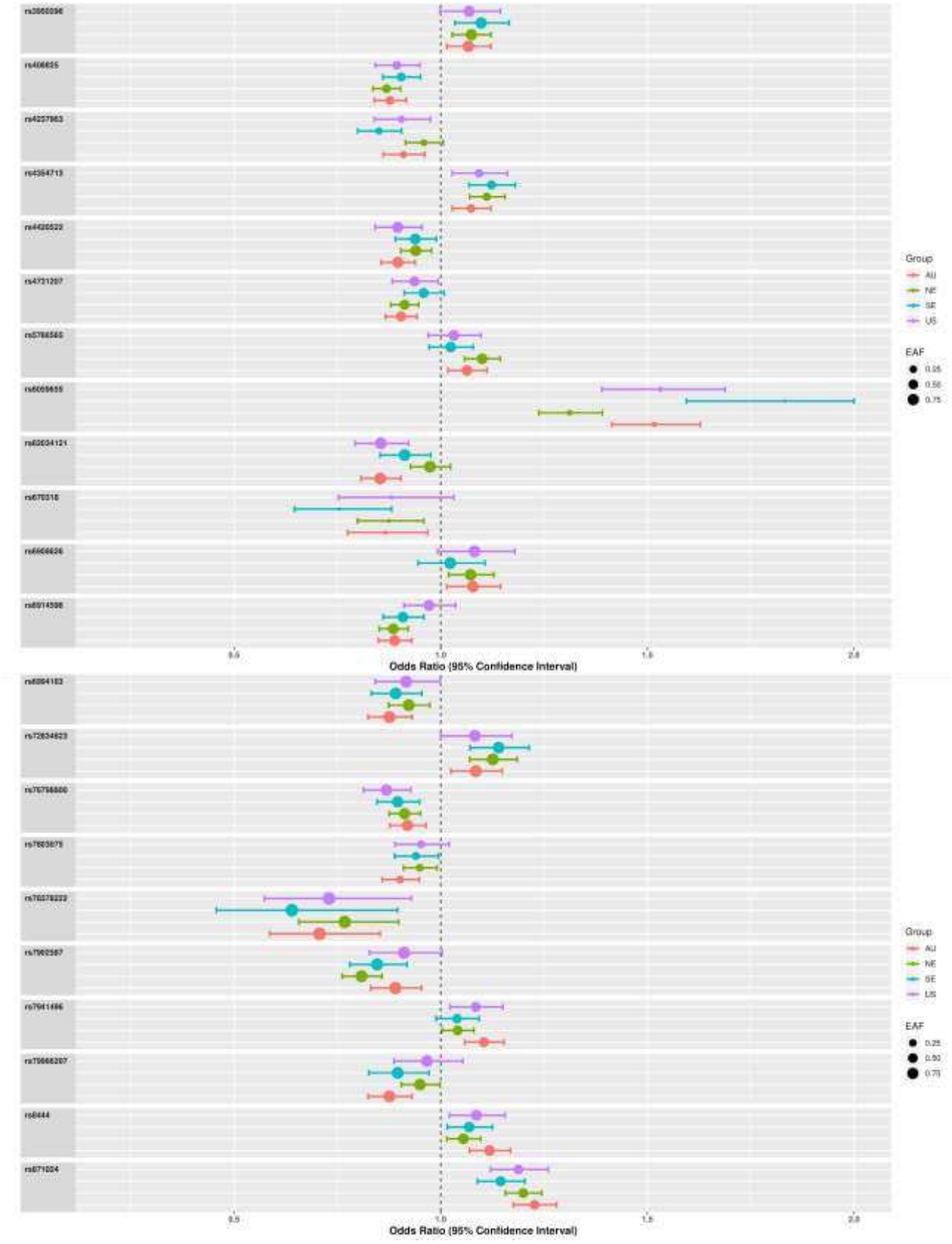


Supplementary Figure 7: **Effect size estimates across contributing GWAS.**

Forest plots showing results across the studies for the most significant SNP at each of the 54 genome-wide significant loci. Results shown for meta-analysis with only clinically-confirmed cases ("overall_conf") and including self-reported cases ("overall"). Estimated log(OR) and 95% confidence intervals are plotted.


Group
AU
NR
SE
US
EAF
0.25
0.50
0.75
0.5
1.0
1.5
2.0
Odds Ratio (95% Confidence Interval)

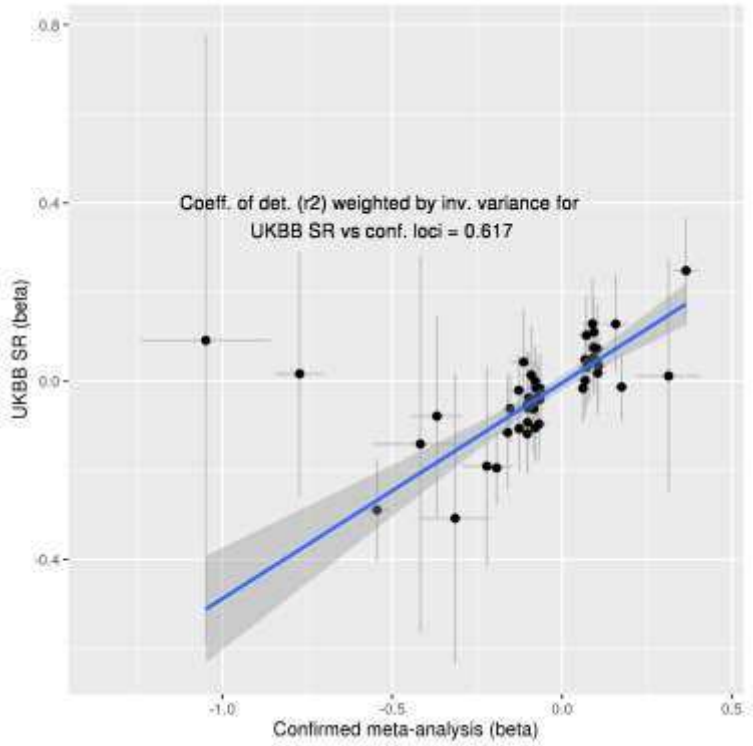

Group
AU
NE
SE
US
EAF
0.25
0.50
0.75
Odds Ratio (95% Confidence Interval)
Group
AU
NE
SE
US
EAF
0.25
0.50
0.75
Odds Ratio (95% Confidence Interval)



Supplement Figure 8: Effect size and minor allele frequency by geographic region.

Forest plots showing the effect size estimates, 95% confidence intervals and allele frequency for each independent genome-wide significant SNP grouped by geographic region (AU: Australia, NE:

Northern European; SE: Southern European; US: United States). Allele frequency is indicated by point size.



Supplementary Figure 9: Effect sizes for confirmed only meta-analysis versus UKBB self-report set

UK Biobank self-report GWAS (UKBB SR); plotted effect sizes are the log(OR) beta.

Supplementary Table Legends

Supplementary Table 1. Summary of GWAS sets.

Type is clinically confirmed cutaneous melanoma (CM) or self report. **Population** is the broad population of origin for samples. **European** is multiple European countries; **US** United States, **UK** United Kingdom**; European-derived** is combined populations drawn from Australian and European countries. MIA_PAH includes Australian case and control samples supplemented with additional UK controls. The Brisbane Nevus Morphology Study (**BNMS**) includes a combined set of Australian and German cases and controls. Principal component analysis was used to ensure homogeneity within GWAS; see GWAS descriptions for details. **Study** is the GWAS name. In addition to **case** and **control** number, the **effective case and control** number to better display discovery power was calculated as
$2/\sigma^2$, where $\sigma^2 = \sqrt{((1/Case_N) + (1/Control_N))}$

GWAS Array is the array used for genotyping. When arrays were combined SNPs were filtered to the minimum overlapping set (see Methods). The Harvard, BNMS, and 23andMe GWAS were imputed to 1000 genomes phase 1 v3; all other data sets were imputed to HRC v1.1 using the Michigan Imputation Server. The genomic inflation (**Lambda**) and LD Score Regression (**LDSC**) **Intercept** and its standard error (**SE**) [10] are both reported as the intercept better reflects test statistic inflation while lambda is also influenced by true polygenic signal and sample size.

The field "**In previous meta-analyses?**" indicates if a GWAS was included in [7] *This GWAS was included in the previous meta-analysis, but has now been supplemented with additional cases and controls such that 39% of cases and 92% of controls are new.

Supplementary Table 2. Inflation, LD Score intercept and genetic correlation between confirmed and self report sets.

Dataset is the **confirmed** meta-analysis, the total confirmed plus self report (**Confirmed plus SR**)**,** the UK Biobank self report GWAS (**UK Biobank SR**), and the **23andMe** self report set. Array heritability ($\mathbf{h^2}$), genomic inflation factor (**Lambda**), **LDSC intercept** and its **SE**, and the ratio between them are reported [10]. The genetic correlation ($\mathbf{R_g}$) and its **SE** as calculated by LDSC indicate a high degree of genetic overlap between self report and confirmed GWAS datasets.

Supplementary Table 3. Annotation and detailed results for independent SNPs associated with CM in the total confirmed plus self report meta-analysis.

Locus: Independent SNP within 1 MB are combined into a single locus. **SNP**: Independent SNP number. **CHR** and **BP** are provided as per hg 19 positions. **rsID**: dbSNP142 rs number.

We list the **Nearest gene** and **Nearest protein coding genes** as reported by the open targets genetics platform [11]. Where loci overlap we summarise the TWAS results using melanocyte-specific expression (Online Methods, Supplementary Table 11) or gene expression from the skin representing sun-exposed skin, non-sun-exposed skin, and transformed skin fibroblasts from GTEx (**Online Methods, Supplementary Table 12**). If known we list the functional gene target based on presence of a missense protein change, or if identified by a previous publication the gene name and PUBMED ID.

CHR:BP:REFALT is a positional ID made by combining HRC position, ref and alt allele. Effect allele (**EA**) and non effect allele (**NEA**) as well as the freq of the effect allele (**EA freq**) are provided. For the two meta-analysis performed using PLINK v1.9 [12] - the total (confirmed and self report combined), and the confirmed cases only - we report summary data including number of contributing GWAS (**N**), Fixed effects P-value and OR (**Pmeta**, **OR**) and random effects (**Pmeta_r**, **OR random**), and measures of heterogeneity (**Q** and **I2**). Finally we report the per study OR for the EA allele for each study.

Supplementary Table 4. SNPs independently associated with CM in the total analysis set.

Locus: Independent SNP within 1 MB are combined into a single locus. **SNP**: Independent SNP number. **CHR** and **BP** are provided as per hg 19 positions. **rsID**: dbSNP142 rs number. **CHR:BP:REFALT** is a positional ID made by combining HRC position, ref and alt allele. Effect allele (**EA**) and non effect allele (**NEA**) as well as the freq of the effect allele (**EA freq**) are provided. **BETA** is the log of the fixed effects OR for the total (confirmed plus self report) meta-analysis and its standard error (**SE**) and P-value (**Pmeta**).

Conditionally independent SNPs were determined using GCTA [13,14] (**Methods**) with the resultant adjust **joint conditional BETA**, **SE** and **P**-value reported. Finally we report the linkage disequilibrium r as estimated in the reference population for SNP i and SNP i + 1 (**Ld_r**). This can be informative when the joint conditional P is more significant than the single SNP analysis (e.g. where two adjacent SNPs are in negative LD).

To limit the chance of false positive claims of novel SNP/loci, we further filtered the list of 75 conditionally independent variants to those (i) genome-wide significant ($P < 5 \times 10^{-8}$) in single SNP and joint conditional analysis, and (ii) as recommended [15] where there was evidence of heterogeneity between studies ($I^2 > 31\%$) the random effect P-value also needed to be $< 5 \times 10^{-8}$. The **Independent SNP** field reports this; if a variant passes all filters its number is recorded; otherwise the reason for filtering is noted; **het** indicates the SNP has $I^2 > 31\%$ and the random effects is not significant; **cojo** indicates the variant is only significant in the cojo analysis and not p 5×10^{-8} in the initial single SNP GWAS so is not counted as a secondary SNP at this locus.

[a]While the lead SNP selected by joint analysis rs12523094 has meta-analysis $I^2 > 31$ and a random effects p-value $< 5 \times 10^{-8}$ there are SNPs in LD that pass this filter. e.g. rs12173258 has the strongest p-value and is passes filters ($P_{meta} = 1.1 \times 10^{-11}$ and $I^2 = 29.6$) and this loci was retained. [b]The HLA region is excluded from LD reference panel used for conditional analysis (See methods) and the lead SNP is reported. [c] All SNPs at this locus near the FTO gene, including the lead SNP, have $^2 > 31$ and a random effects p-value $< 5 \times 10^{-8}$, and as such is not counted as an independent CM risk locus in this analysis.

Supplementary Table 5. Independant CM SNPs and their association with risk phenotypes.

Locus: Independent SNPs within 1 MB are combined into a single locus. **SNP**: Independent SNP number. **CHR** and **BP** are provided as per hg 19 positions. **rsID**: dbSNP142 rs number. Nearest, or if known likely functional **gene** target, is provided. The meta-analysis P-value for the confirmed only sets (**Pmeta confirmed**) and total including self report (**Pmeta confirmed SR**) are listed for each variant, as is their their association P-value from a series of GWAS of melanoma risk phenotypes from the UK Biobank (**Online methods**); Ease of tanning (**P Ease**), Number of Childhood sunburns (**P Childhood Sunburns**), **Skin Colour**, non red hair hair colour (**Hair Colour**) as well as red hair yes/no (**Red Hair**). Each SNP's association with leukocyte telomere length (P Telomere Length) taken from a GWAS of 37,684 individuals [16] and the same nevus count GWAS meta-analysis used for the joint analyses of CM and nevus count and pigmentation (N = 65,597, Methods).

For a number of SNPs, telomere GWAS results were not available for the lead CM SNP; where possible a proxy SNP (**Telomere SNP**) was identified using an expanding LD window in PLINK v1.9 using the same LD reference panel as the GCTA conditional analysis (**Online methods**). The LD window to search for proxies started at 2mb and 100 SNPs wide, with a minimum LD r^2_{EUR} > 0.9, with the threshold relaxed until a SNP was identified; if no SNP was found with a window of 1000 SNPs, 4 Mb and LD r^2_{EUR} > 0.5 the search was stopped. LD r^2 between the **Telomere SNP** and the lead SNP is listed (**r^2**)

aFor two SNPs, rs7705526 and rs143190905 data for telomere association was not available for the lead SNP nor a proxy with LD r^2_{EUR} > 0.5. For these two SNPs the p-value is derived from a separate telomere GWAS [17] identified via the open targets genetic platform. **b**SNP is not in the Nevus GWAS meta-analysis data, nor is there a proxy SNP ith a window of 1000 SNPs, 4 Mb and LD r^2_{EUR} > 0.5. **c**P value for a proxy SNP for rs116927526 is reported; rs193247693 LD r^2_{EUR} 0.70

Supplementary Table 6. SNPs independently associated with CM in the confirmed cases only analysis set.

Locus: Independent SNPs from the confirmed cases meta-analysis within 1 MB are combined into a single locus. **SNP**: Independent SNP number. **SNP formally significant and independent** discounts SNPs that, while selected in the joint conditional analysis (**Online Methods**), are not significant in the single SNP GWAS (e.g. $P_{meta} > 5 \times 10^{-8}$ or $P_{meta\_r} > 5 \times 10^{-8}$ where heterogeneity $I^2 > 31$). See footnotes below for further detail.

CHR and **BP** are provided as per hg 19 positions. **rsID**: dbSNP142 rs number. **CHR:BP:REFALT** is a positional ID made by combining HRC position, ref and alt allele. Effect allele (**EA**) and non effect allele (**NEA**) as well as the freq of the effect allele (**EA freq**) are provided. **BETA** is the log of the fixed effects OR for the confirmed cases only meta-analysis and its standard error (**SE**) and P-value (**Pmeta**). Conditionally independent SNPs were determined using GCTA (**Methods**) with the resultant adjust **joint conditional BETA**, **SE** and **P**-value reported. Finally we report the linkage disequilibrium r as estimated in the reference population for SNP i and SNP i + 1 (**Ld_r**). This can be informative when the joint conditional P is more significant than the single SNP analysis (e.g. where two adjacent SNPs are in negative LD).

For reference we append the relevant summary data including number of contributing GWAS (**N**), Fixed effects P-value and OR (**Pmeta**, **OR**) and random effects (**Pmeta_r**, **OR random**), and measures of heterogeneity (**Q** and **I2**) from the confirmed cases only meta-analysis results performed using PLINK v1.9 [12].

We also report the per study OR for the EA allele for each study.

Results are compared to the "**Total Meta-analysis lead SNP**" with the overlapping locus number from **Supplementary Table 3** provided.

A While the lead SNP in this locus selected by conditional analysis has fixed $P > 5 \times 10^{-8}$ I^2 heterogeneity measure > 31% and a random meta-analysis P-value $> 5 \times 10^{-8}$ other SNPs exist that qualify e.g. chromosome 2 rs3769823 $r^2_{EUR} = 0.75$ with lead SNP rs10931936; chromosome 5 rs12173258 $r^2_{EUR} = 0.82$ with lead SNP rs11747245; chromosome 8 rs1481853 ($r^2_{EUR} = 0.99$ with lead SNP rs13263376); chromosome 12 rs4257028 ($r^2_{EUR} = 0.99$ with lead SNP rs4763456); chromosome 16 rs16953002 ($r^2_{EUR} = 0.85$ with lead SNP rs62034139); chromosome 16 rs2967361 (r^2_{EUR} 0.85 with lead SNP rs2911423). **B**While the lead SNP in the total meta-analysis is rs2301293, that SNP is P > 5e-8 in the total meta-analysis and not formally a CM locus. **C** I^2 heterogeneity measure > 31% and a random meta-analysis P-value $> 5 \times 10^{-8}$. D Indicates the variant is only significant in the cojo analysis and not p 5×10^{-8} in the initial single SNP GWAS so is not counted as a secondary SNP at this locus.

Supplementary Table 7. Results from total confirmed plus self report meta-analysis for SNPs with meta-analysis P < 5 x 10^{-8}.

Displayed results are with respect to I^2 < 31% (**Online Methods**).

Supplementary Table 8. Sample sizes for sub phenotype GWAS analysis.

Numbers of cases and controls available across contributing GWAS data for the histology, age and sex stratified GWAS analyses.

Supplementary Table 9. Combined analysis of CM and Nevus count GWAS meta-analyses.

LD calculations performed in PLINK using a reference panel of 10,000 white british UK Biobank individuals as implemented in the FUMA platform was used to identify independent SNPs with $P < 5 \times 10^{-8}$; independent SNPs within 1 Mb were considered to be single loci (**Online Methods**). The lead SNP for each locus is reported. Bolding indicates that the locus is significant at a Bonferroni corrected P-value threshold of $P \leq 1.25 \times 10^{-8}$. Chromosome (**CHR**), base pair (**BP**) and **rsID** are provided for the peak variant for each locus using hg19 positions. Effect allele (**EA**) and non effect allele (**NEA**) as well as the freq of the effect allele (**EA freq**) are provided. We list the **Nearest gene** and **Nearest protein coding genes** as reported by the open targets genetics platform [11]. If the lead SNP is an eQTL for skin tissues in GTEx we list the gene (**GTEX skin eQTL gene**); if the p-value is not significant following bonferroni correction (> 0.05 divided by the number of genes) the actual P-value is also reported. If known we list the functional gene target based on presence of a missense protein change, or if identified by a previous publication the gene name and PUBMED ID.

If the locus overlaps with a locus identified in the total confirmed plus self report CM meta-analysis (**Table 1**, **Table 2**, **Supplementary Table 3**) its number is noted here, as is the peak CM rsID if appropriate, and the LD r^2_{EUR} between the two SNPs (**Overlapping CM locus, rsID and LD r2 between lead SNPs**).

If the locus overlaps with a recent combined analysis of CM and nevus count GWAS meta-analysis [9] this is noted (otherwise novel ones are indicated by new), and the peak SNP there is provided, as is the LD r^2_{EUR} between the new peak and that SNP, as well as the likely model for this locus (CM alone, Nevus alone, or pleiotropic for CM and nevus count (**Overlapping 2018 Nevus+CM locus**; **rsID**, **LD r2 between lead SNPs**, and **2018 Model**).

The **Model** field indicates which GWAS-PW model has a PPA > 0.5 (**Online Methods**); model 3, pleiotropic for CM and Nevus count (CM+Nevus) or Model 1 (CM alone). No loci were associated with model 2, Nevus count alone. Where appropriate we also indicate is this locus is pleiotropically associated with pigmentation (**Online Methods**, **Supplementary Table 10**).

We summarise the stouffer's method meta-analysis and GWAS-PW approach (Online methods). For the peak locus SNP in field **rsID** we provide the Z score and it variance (**V**) for the total CM GWAS meta-analysis (**CM**), the Nevus GWAS meta-analysis (**Nevus**). The P-value from Stouffer meta-analysis of both (**CM + Nevus P**). GWAS-PW assigns SNPs to **LD blocks**, and the hg19 start and end bp for each is provided as well as the maximum Z score for CM or Nevus for SNPs within each block. The PPA for GWAS-PW models for the association signal at each block being CM alone, Nevus alone, or CM + Nevus are provided. GWAS-PW also reports a 4th model, where the signal is co-located but independant. This is unlikely for nevus count and melanoma, where nevus count is a direct risk for melanoma [9] and as such all model 4 PPAs were dropped and scores for the remaining 3 models reweighted to sum to 1. As all PPA values for model 4 were all < 0.01 barring locus 41, peak SNP rs9651783 on chromosome 11 which reach 0.03, again supporting that colocation is an unlikely model for CM and nevus count, this made little different to the final model PPA.

Supplementary Table 10. Combined analysis of CM and hair colour GWAS.

The independent locus, as defined by LD calculations implemented in FUMA platform (**Online Methods**), that is novel to previous reports and has a GWAS-PW model 3, pleiotropy, PPA > 0.5 is reported (**CM + Hair colour Locus**; **Supplementary methods**). Bolding indicates that the locus is significant at a Bonferroni corrected P-value threshold of $P \leq 1.25 \times 10^{-8}$. Chromosome (**CHR**), base pair (**BP**) and **rsID** are provided for the peak variant for each locus using hg19 positions. Effect allele (**EA**) and non-effect allele (**NEA**) as well as the freq of the effect allele (**EA freq**) are provided. [For the peak locus SNP in field **rsID** we provide the **nearest gene** and **nearest protein coding gene**.]. If the lead SNP is an eQTL for skin tissues in GTEx we list the gene (**GTEX skin eQTL gene**); if the p-value is not significant following bonferroni correction (> 0.05 divided by the number of genes) the actual P-value is also reported. If known we list the functional gene target based on presence of a missense protein change, or if identified by a previous publication the gene name and PUBMED ID.

If the Locus overlaps with a locus identified in the total confirmed plus self-report CM meta-analysis (**Table 1**, **Table 2**, **Supplementary Table 3**) its number is noted here (**CM Locus**), as is the peak **CM rsID** if appropriate, and the LD r^2_{EUR} (**LD_r2**) between the two SNPs.

The **Model** field indicates which GWAS-PW model has a PPA > 0.5 (**Online Methods**); model 3, pleiotropic for CM and Pigmentation (CM+Pigment) or Model 1 (CM alone). Where appropriate we also indicate is this locus is pleiotropically associated with nevus count (**Online Methods**, **Supplementary Table 9**).

For the peak locus SNP in field **rsID** we provide the Z score (**Z**) and its variance (**V**) for the total CM GWAS meta-analysis (**CM**), the Hair colour GWAS (**Hair**), and the P-value from Stouffer meta-analysis of both (**CM + Hair P**). GWAS-PW assigns SNPs to **LD blocks** (50 SNPs wide), and provided the maximum absolute Z score for CM or Hair colour for SNPs within each block. The PPA for GWAS-PW models for the association signal at each block (**PPA**) being CM alone, hair colour alone, or CM + hair colour are provided.

Supplementary Table 11. Top melanoma TWAS genes using melanocyte eQTL data as a reference.

For genes whose expression in melanocytes is predicted to be significantly associated with CM by TWAS (**Online Methods**), the transcript gene name (**Gene**) along with its hg19 chromosome (**CHR**), transcription start (**Start**) and end (**End**) positions are reported. We also provide the peak CM SNP's rsID (**CM rsID**) and Z score for association with CM in the total (confirmed plus self-report) meta-analysis (**CM Z**). In addition we report the rsID of the most significant eQTL variant in the locus from the 1000 genomes reference panel for the named transcript (**eQTL Top Varian**t), that variant's eQTL Z score (**eQTL Z**), and that variant's total GWAS meta-analysis Z-score (**CM Z for eQTL variant**). The number of variants included in the TWAS models for each locus is reported (**Number of variants**), as is the weighted number of variants (**Number of Weight**). For each transcript the best performing TWAS **Model**, and its **TWAS Z** and **TWAS P** are provided. We also provide **TWAS P cutoff** that we applied based on the multiple testing corrections described in **Online Methods**. To assist interpretation we note of if the significant gene maps to a previously or newly identified **CM locus**, and if so report the applicable peak SNP and locus number from Supplementary Table 3 or is otherwise **new** to CM. As done for independent SNPs identified in the total and confirmed CM meta-analysis, TWAS genes identified genes within 1 Mb of each other were combined in a single locus. Finally we reported if this same transcript was reported as significantly associated with CM by Zhang and colleagues [18] which applied melanocyte TWAS to the phase 1 CM data (**Reported by Zhang**; Y=yes, N=no).

Supplementary Table 12. Top melanoma TWAS genes using skin eQTL data as a reference.

For genes whose expression in three skin tissues (sun-exposed, not-sun-exposed, and fibroblasts) within GTEx (V7) is predicted to be significantly associated with CM by TWAS (**Online Methods**), the transcript gene name (**Gene**) along with its hg19 chromosome (**CHR**), transcription start (**Start**) and end (**End**) positions are reported. We also provide the peak CM SNP's rsID (**CM rsID**) and Z score for association with CM in the total (confirmed plus self-report) meta-analysis (**CM Z**). In addition we report the rsID of the most significant eQTL variant in the locus from the 1000 genomes reference panel for the named transcript (**eQTL Top Varian**t), that variant's eQTL Z score (**eQTL Z**), and that variant's total GWAS meta-analysis Z-score (**CM Z for eQTL variant**). The number of variants included in the TWAS models for this locus are reported (**Number of Variants**), as is the weighted number of variants (**Number of Weight**). For each transcript the best performing TWAS **Model**, and its **TWAS Z** and **TWAS P** are provided. We also provide **TWAS P cutoff** that we applied based on the multiple testing corrections described in **Online Methods**. To assist interpretation we note of if the significant gene maps to a previously or newly identified **CM locus**, and if so report the applicable peak SNP and locus number from Supplementary Table 3 or is otherwise **new** to CM. As done for independent SNPs identified in the total and confirmed CM meta-analysis, TWAS identified genes within 1 Mb of each other were combined in a single locus.

Supplementary Table 13. Top melanoma TWAS genes using all tissues eQTL data as a reference.

For genes whose expression in 45 GTEx (V7) tissues (excluding 3 skin tissues shown in **Supplementary Table 12**) is predicted to be significantly associated with CM by TWAS (**Online Methods**), the transcript gene name (**Gene**) along with its hg19 chromosome (**CHR**), transcription start (**Start**) and end (**End**) positions are reported. We also provide the peak CM SNP's rsID (**CM rsID**) and Z score for association with CM in the total (confirmed plus self-report) meta-analysis (**CM Z**). In addition we report the rsID of the most significant eQTL variant in the locus from the 1000 genomes reference panel for the named transcript (**eQTL Top Varian**t), that variant's eQTL Z score (**eQTL Z**), and that variant's total GWAS meta-analysis Z-score (**CM Z for eQTL variant**). The number of variants included in the TWAS models for this locus are reported (**Number of Variants**), as is the weighted number of variants (**Number of Weight**). For each transcript the best performing TWAS **Model**, and its **TWAS Z** and **TWAS P** are provided. **TWAS P cutoff** was set for each tissue based on the number of genes passing a heritability cutoff and being tested within each tissue type (0.05/number of genes). P-values of genes that pass a Bonferroni-corrected cross-tissue wide P-value cutoff (based on the total number of tested genes in 45 tissues, N=211,941) are bolded. To assist interpretation we note of if the significant gene maps to a previously identified **CM locus**, and report the applicable peak SNP and locus number from **Supplementary Table 3**.

Supplementary Table 14: Tabulation of loci associated across primary and secondary analyses

Each **Locus** is numbered in order (see **Online Methods** for how loci are defined), and the positions (**CHR**, **BP**) and **rsID** for the lead SNP are reported for the SNP with the strongest total CM associations signal (totalCM, TWAS results) or for the loci derived from the joint analysis, the SNP with the strongest joint p-value. **Gene** prioritises the functional target if known, followed by melanocyte or skin tissue TWAS data, or finally the closest protein coding gene (**Supplementary Table 3, 11, and 12**). For each of the named analyses we indicate with a "**1**" if the loci passes the significance threshold for that analysis (**Supplementary Table 3, 9, 10, 11, 12).**

Supplementary Table 15. Colocalization of melanoma GWAS signal and eQTLs from melanocytes and skin tissues.

eCAVIAR [19] was used for testing colocalization of melanocyte and GTEx skin tissue eQTLs (listed in **eQTL dataset**) for the reported gene (listed in **Gene** column), and the melanoma GWAS signal (here we report the **CM GWAS locus**, lead independent **SNP**). Fifty SNPs upstream of and downstream from GWAS lead SNP in each locus were chosen to quantify the probability of the variant to be causal in both the GWAS and eQTL studies.

We report the colocalization posterior probability (**CLPP**), which is the probability that the same variant is causal in both GWAS and eQTL. A cutoff of CLPP > 0.01 was applied to the resulting list of SNPs, and SNPs were further filtered using an LD r^2 cutoff of > 0.9 with the GWAS lead SNP (1000 Genomes, EUR population), as well as to those with a significant eQTL P-value (<0.05/number of eGenes tested for each eQTL dataset). We report the SNP with the highest CLPP score for each gene, as well as its LD **r2 with the lead GWAS SNP**, and **eQTL P-value**. Finally we indicate if the gene is also significantly associated with CM by TWAS in melanocyte or skin tissues (**Supplementary Tables 11-12**) .

Supplementary Table 16. Enriched canonical pathways in genes identified through TWAS and colocalization from melanocyte or skin tissues.

All significant TWAS genes from melanocyte (**Supplementary Table 11**) plus skin tissues (**Supplementary Table 12**), as well as all colocalized genes identified by eCAVIAR applied to melanocyte and 3 skin tissue datasets (**Supplementary Table 15**) were queried as a gene list in IPA analysis (Qiagen, version 46901286). Significantly enriched canonical pathways are displayed (**-log(P-value)** > 1.3).

Supplementary Table 17. Top 5 enriched upstream regulators in genes identified through TWAS and colocalization from melanocyte or skin tissues.

All significant TWAS genes from melanocyte (**Supplementary Table 11**) plus skin tissues (**Supplementary Table 12**), as well as all colocalized genes identified by eCAVIAR applied to melanocyte and all 3 skin tissue datasets (**Supplementary Table 15**) were queried as a gene list in IPA analysis (Qiagen, version 46901286). The top 5 enriched **Upstream Regulators** are listed (**P-value of overlap** < 0.05; **Molecule Type** and **Target molecules in dataset** are also reported).

Supplementary Table 18: Results for UKBB self-report set for three loci no longer significant following their removal from the total meta-analysis

For each locus that is no longer significant following the exclusion of the UKBB self-report set we report the position and rsID for the lead SNP. For each SNP we report the results for the total meta-analysis, which includes all pathology-confirmed GWAS together with 23andMe and UKBB self-report. We also report results for the meta-analysis without UKBB self-report, and for UKBB self-report itself. For each GWAS we tabulate the respective log(OR) (**BETA**), standard error (**SE**) and P-value (**P**).

Supplementary Methods

GWAS descriptions

Q-MEGA
1,617 CM cases were drawn from population-based studies undertaken in Queensland were combined as The Queensland study of Melanoma: Environment and Genetic Associations (Q-MEGA) study [20]. CM cases were histologically confirmed using the Queensland Cancer Registry [20]. All contributing studies were approved by the Human Research Ethics Committees of QIMR Berghofer, University of Sydney, University of Melbourne and cancer registries of NSW, Victoria and Queensland. Informed consent was obtained from all participants.

Samples were genotyped in two phase using the Illumina HumanHap610 and Omni1-Quad arrays (San Diego, CA, USA), and to ensure maximum SNPs were retain each phase was cleaned, imputed and analysed as QMEGA_610k and QMEGA_omni respectively.

Q-MEGA_610k: 926 Q-MEGA cases were combined with 1,799 unrelated controls genotyped on the Illumina HumanHap610 from the Brisbane Adolescent Twin Study [21,22]. 20% of the samples were a singleton from twin pairs or sibling sets, with the remainder drawn from parents of twins. Controls were further supplemented by 2,155 endometriosis patients genotyped using the Illumina HumanHap670 array (San Diego, CA, USA) [23]. Controls who self reported melanoma in themselves or a family member were excluded. Following cleaning (see **Online Methods**) there were 912 cases and 3,777 controls and 487,089 autosomal SNPs.

Q-MEGA_omni: 691 Q-MEGA CM samples were with 553 healthy cancer-free controls from the Study of Digestive Health (SDH) [24]. After QC (**Online Methods**) there were 656 cases, 538 controls and 767,979 SNPs available for imputation.

AMFS
549 cases with histopathologically confirmed invasive CM, and 431 city, age and sex-matched controls from Melbourne, Brisbane, and Sydney, Australia were drawn from the Australian Melanoma Family Study (AMFS) [25]. The outlined study was approved by the Human Research Ethics Committees of the University of Sydney, University of Melbourne and cancer registries of NSW, Victoria and Queensland. Informed consent was obtained from all participants. AMFS samples were genotyped using Illumina Omni1-Quad (San Diego, CA, USA) . After QC (see online methods) there were 535 cases, 427 controls and 791,961 SNPs.

WAMHS
DNA samples were available from 1,289 CM cases from the Western Australian Melanoma Health Study (WAMHS) [26]. The WAMHS population based study was collected from Perth, Australia based on histopathology confirmed invasive, cutaneous melanoma cases in the Western Australian Cancer Registry.

WAMHS CM cases were genotyped using Infinium HumanOmniExpressExome, and were combined with two control sets on matching arrays. The first was 960 Australian individuals (including healthy individuals, and a subset with ulcerative colitis) genotyped on the Infinium HumanOmniExpressExome arrays from the Inflammatory Bowel disease (IBD) Clinical and Research Programme at the Royal Brisbane & Women's Hospital Brisbane, Queensland [27]. The second was 1,298 advanced glaucoma cases from the ANZRAG; Australian & New Zealand Registry of Advanced Glaucoma; 651 were genotyped on Omni1M and 647 on the HumanOmniExpress arrays from Illumina (San Diego, CA, USA). Controls were not screened for melanoma. Following GWAS cleaning (**Online methods**), including filtering to a common set of

SNPs that passed QC across all arrays, 1,237 melanoma cases, 1,980 controls, and 567,914 SNPs were available for imputation.

This work was undertaken with ethics approval from The University of Western Australia and the Department of Health Western Australia, with ANZRAG inclusion approved by the Southern Adelaide Clinical Human Research Ethics Committee. Informed consent was obtained from all participants.

UK Biobank

A full description of the UK Biobank dataset and the quality control and imputation approaches applied are described elsewhere [28]. Briefly, UK Biobank is a volunteer cohort of approximately 500,000 persons recruited between 2006 and 2010. Besides answering a broad range of questions, participants gave permission for their health records to be accessed and for linkage to national cancer registries which record pathologically and clinically diagnosed cancers; for melanoma essentially all diagnoses have pathological verification. The Cancer Registry records were searched for diagnoses made under ICD9 (172.x) or ICD10 (C43.x). Overall, 3,659 cases were identified with one or more diagnoses with the date of first diagnosis ranging from 1968 to 2015 at the time of this analysis; about 40% of diagnoses were incident (following recruitment) while the remainder were made prior to recruitment into UK Biobank. UK Biobank participants also self-reported cancer diagnoses at recruitment; overall 1,651 participants had a confirmed diagnosis of melanoma without a self-report of melanoma (the majority of these confirmed diagnoses being after recruitment, the time at which self-reports were made), 1,873 reported melanoma but did not have a confirmed diagnosis and 2,008 had both a self-report of melanoma and a confirmed diagnosis.

Within UK Biobank, biological samples were available for genetic analysis from 488,000 participants. The majority of participants were genotyped using a purpose designed UK Biobank Applied Biosystems Axiom array assessing 826,000 SNPs and indels. The quality control and imputation approaches applied are described elsewhere [28]. UK Biobank provides a lists of participants whose genetic results should be excluded on the basis of poor performance or close relatedness; these persons were excluded in our analysis. Non-European outliers were identified using the same approach that UK Biobank apply to their "Caucasian" definition: the 'aberrant' routine in R [29] was applied to PCs 1&2, 3&4 and 5&6, but anyone declaring themselves to be White was included in this analysis (where UK Biobank automatically exclude "Irish" and "any other white background"); the lambda parameter used was 100.

Following the application of QC and restriction by ancestry as described above, we included 3,499 individuals with confirmed melanoma from the cancer registry and a further 1,802 individuals with self-reported melanoma without cancer registry confirmation. The confirmed and self-reported cases were analysed separately, in a case-control analysis using four times as many controls as cases, randomly selected from UK Biobank.

Analysis was conducted in PLINK2 correcting for population stratification by including PCs 1-15 as covariates.

Michigan

The Michigan Genomics Initiative (MGI) is a repository of patient electronic health and genetic data at Michigan Medicine [30]. MGI participants are recruited primarily through pre-surgical encounters at Michigan Medicine and consent to linking of genetic and clinical data for research purposes. Blood derived DNA was genotyped on Illumina Infinium CoreExome-24 bead arrays (San Diego, CA, USA). Genotype data was imputed to the Haplotype Reference Consortium using the Michigan Imputation Server, providing > 17 million imputed variants after standard quality control and filtering (minimac imputation $R^2 > 0.3$). We restricted our analysis to samples of

European descent based on Principal Component Analysis using the Human Genome Diversity Panel as references [31]. The CM phenotype was derived from ICD-9 codes extracted from patient electronic health records using the PheWAS R package [32]. We identified 1,198 CM cases samples required to have at least two inclusion ICD-9 codes for melanoma and 26,211 controls with zero inclusion or exclusion codes. The association analysis consisted of a Firth bias-reduced logistic regression controlling for sex, age and 10 principal components. The Firth test corrects for imbalance in case-control ratio [33] and was performed using the epacts software (https://github.com/statgen/EPACTS).

GSXP1 & P2

GenoMEL Phase 1 and 2 are updated relative to previous reports [4,7] with CM samples from the Leeds melanoma cohort now included in the GSEdinCIDR dataset. For a more detailed description see [2,4], but briefly phase 1 includes cases with histologically confirmed CM from 7 collections centres from 5 European countries genotyped on the Illumina HumanHap300 BeadChip and HumanCNV370k arrays (San Diego, CA, USA). Controls were collected by the same groups as cases and were melanoma-free. After QC there were 1,075 cases and 2,163 controls. Phase 2 samples includes cases and controls from 9 centres (4 new to phase 1) from 8 Europeans countries plus Israel, totalling 925 cases and 1,128 controls, with genotyping performed on the Illumina HumanHap610 array.

GSEdinCIDR - Leeds/Cambridge/Scotland

Cases are from two sources. 2,949 are from several studies from Leeds and 1,379 from the Cambridge SEARCH study. Both were genotyped on the Illumina Infinium HumanOmniExpressExome (San Diego, CA, USA) array.

The Leeds data are predominantly from the Leeds melanoma cohort, a population-based melanoma study recruited population-based incident melanoma cases diagnosed between September 2000 and December 2012 from a geographically defined area of Yorkshire and the Northern region of the UK [2,34–36]. Cases were identified by clinicians, pathology registers and via the Northern and Yorkshire Cancer Registry and Information Service to ensure overall ascertainment. For all but 18 months of the study period, recruitment was restricted to patients with Breslow thickness of at least 0.75mm. All controls and the first 960 cases recruited were examined by trained interviewers who performed a standardized examination of the skin, recording nevi by anatomical site and size.

CM cases were identified and recruited from the Eastern Cancer Registry and Information Centre (Eastern England) by the SEARCH study, University of Cambridge [37,38]. CM cases were aged 18-70 at diagnosis. 5,304 of the controls are from Generation Scotland, sampled to preferentially select those individuals reporting English ancestry. 476 of the controls are from the Leeds Melanoma Cohort, selected to be matched by age and GP surgery to the cases. The Leeds controls were selected to be melanoma-free while the Generation Scotland controls are simply population samples.

In Leeds participants consented to studies approved by the Northern and Yorkshire Research Ethics Committee. Approval for the SEARCH study was obtained from the Eastern Multicentre Research Ethics Committee, UK. Informed consent was obtained from all participants.

MDACC

The MDACC dataset consists of 931 white (non-Hispanic) individuals with cutaneous melanoma who were recruited together with 1,026 cancer-free controls (friends or acquaintances of patients reporting to other clinics at M.D. Anderson). Controls were frequency matched on age and sex. The The GWAS CM dataset also includes an additional 873 individuals presenting for treatment for cutaneous melanoma at M.D.Anderson. All samples were collected between March 1998 and

August 2008. Samples were genotyped using the Illumina HumanOmni1-Quad_v1-0_B array (San Diego, CA, USA) and called using the BeadStudio algorithm. A more detailed description of the QC procedures applied to these data can be found elsewhere [3] The study protocols were approved by the Institutional Review Board at MD Anderson, and informed consent was obtained from all participants.

Harvard
410 pathologically confirmed invasive CM cases and 2,920 melanoma free controls were drawn from US non-Hispanic Caucasian participants of two cohorts; the Nurses' Health Study and the Health Professionals Follow-up Study who were genotyped on Illumina HumanHap660 arrays (San Diego, CA, USA) [39]). Cohort recruitment required informed consent, and ethical oversight and approval was via the Institutional Review Board at Brigham and Women's Hospital and the Harvard School of Public Health. Following QC, imputation to 1000 genomes phase 1 Integrated Release Version 3 Haplotypes (2010-11 data freeze, 2012-03-14 haplotypes) was conducted using MACH [40,41]. Only SNPs with a MAF > 0.01 were carried forward for analysis using logistic regression; age, sex, family history of melanoma and the first 5 principal components were included as covariates.

EPIGENE
787 CM cases genotyped on the Illumina CoreExome array (San Diego, CA, USA) were available from the EPIGENE cohort, which recruited histopathologically confirmed invasive CM cases (with informed consent) from southern Queensland, Australia [42]. CM cases were aged 18-79 at the time of diagnosis, and were diagnosed between April 1st 2007 and September 30, 2010.
These were combined with 983 healthy controls from the Brisbane Adolescent Twin Study [21,22] genotyped on the Illumina CoreExome array (San Diego, CA, USA); control samples were selected to be unrelated (PLINK IBD PI_HAT < 0.1875) to each other, to EPIGENE samples, and to controls used in the QMEGA_610k GWAS. Following QC as performed for the other GWAS sets (Methods) there were 773 CM cases and 910 controls. Genotyped data was imputed to the HRC as outlined in the **Online Methods**. The study protocol was approved by the HREC of QIMR Berghofer Medical Research Institute, Brisbane, Australia.

MIA_PAH
Clinically confirmed CM cases were were drawn from participants three cohorts genotyped on the Illumina Oncoarray (San Diego, CA, USA). Germline DNA samples and data from the Melanoma Institute of Australia (MIA) cohort (1,745 cases) were obtained from the MIA Biospecimen Bank (protocol HREC/10/RPAH/530), which is based on an Institute-wide prospective protocol of registration and clinical data collection from all melanoma cases (MIA Melanoma Research Database, protocol HREC/11/RPAH/444). All MIA cases were confirmed by histopathology. A further N = 236 were recruited from the Princess Alexandra Hospital and Brisbane treatment centres (PAH). All patients with melanoma recruited into the PAH study were diagnosed with locally invasive primary cutaneous melanoma and referred to the Multidisciplinary Melanoma Clinic at the Princess Alexandra Hospital and related clinics. Patients who proceeded to a sentinel lymph node biopsy procedure between 1994 and 2007 were considered for inclusion. Patient demographics, date of diagnosis and clinicopathological characteristics of the primary melanoma for each patient were prospectively recorded in a hospital database through the PAH surgical clinic. Consenting patients were sent a saliva kit for DNA collection. The third set of samples was participants in the Q-MEGA cohort not previously genotyped an unrelated (IBD < 0.15) with participants in the Q-MEGA omni and Q-MEGA 610k GWAS (N = 60).

Controls were cancer free women genotyped on the Illumina Oncoarray (San Diego, CA, USA) [43] drawn from the population based controls of the Australian Ovarian Cancer Study (N = 1,134) [44,45], combined with N = 1,794 UK cancer free controls (all female) from the SEARCH study [46]. PCA was used to confirm cancer free controls were homogenous with the Australian cases and controls.

GWAS samples were cleaned as per other datasets (Online Methods). In addition to the standard genotype QC (**Online Methods**) the Oncoarray genotype data was first cleaned and aligned as per Oncoarray Consortium protocols to remove poorly performing assays specific to that array [43]. After QC and cleaning there were 1,933 CM cases (N = 1,641 from the MIA set, 232 from the PAH set , and 60 from Q-MEGA). 2,841 controls were available after cleaning (AOC N = 1,113, SEARCH = 1,778).

MELARISK, France

Additional clinically confirmed CM cases and controls were available from two French cohorts contributing to GenoMEL Phase 1 and 2; MELARISK (535 cases) [47] and the Supplementation in Vitamins and Mineral Antioxidants (SU.VI.MAX) study (856 controls - population samples not screened for melanoma) [48] [34]. The project was approved by the Human Research Ethics Committees of PARIS SAINT-LOUIS, PARIS NECKER and Ile de France II for the MELARISK study and from the Human Research Ethics Committee of PARIS Cochin for the SU.VI.MAX study. Informed consent was collected for all individuals taking part. Following genotyping on IHuman660W-Quad array (San Diego, CA, USA) at the Centre National de Genotypage (CNG, Evry, France), QC as per the Online Methods were applied. Post QC, 511 melanoma cases and 815 controls were available for GWAS.

Essen-Heidelberg, Germany

Histologically confirmed CM cases (N = 1,218) were genotyped using the Illumina HumanCytoSNP-12v1-0_D arrays (San Diego, CA, USA). These were combined with healthy controls (N = 1,223) from Germany genotyped on the Illumina HumanOmniExpress12_v1_C array (San Diego, CA, USA). The combined genotyped data was thinned to the 220,000 SNPs overlapping across the arrays before applying the standard cleaning and quality control (Online Methods). Participation required informed consent, and the study was approved and overseen by the Human Research Ethics Committee of University Hospital Essen.

QSkin

Data for 17,695 individuals genotyped using the Illumina Global Screening Array (San Diego, CA, USA) was available from the QSkin cohort [49]. Following excluding participants with > 5% genotype missingness (n = 322), were related to another participant at PLINK IBD PI_HAT > 0.1875 (N = 400), or were more than 6 sd on PC1 and PC2 compared to a European reference population constructed using 1000 genomes data (N = 378; some individuals overlap across exclusions), 16,708 participants were available. Genotyped SNPs were cleaned for GenTrain score <0.6, Hardy-Weinberg P-value $< 1 \times 10^{-6}$, or a minor allele frequency (MAF) < 1%, leaving 496,695 SNPs for imputation. Imputation was performed using the University of Michigan Imputation Server as per the other GWAS datasets (Online Methods). Cleaned, imputed data were split into a CM GWAS set (N = 1,285 with histopathologically confirmed CM and 2,493 controls without melanoma) and a distinct Nevus GWAS dataset (N = 12,930). The CM GWAS was performed using logistic regression in PLINK v1.9 with the first 10 principal components fitted as covariates. The QSkin nevus phenotype was a self reported 4 point scale (None, some, a few, many based on a reference picture; see https://qskin.qimrberghofer.edu.au/page/About/Baseline_survey). Following rank normalisation in R. this was analysed using PLINK v1.9 as a quantitative phenotype with the orthoganol transformation of age, age^2, sex, sex × age, and sex × age^2 and the first 10 principal components fitted as covariates.

Brisbane Nevus Morphology Study

The ethical aspects of this research project have been given approval by the Human Research Ethics Committee of Princess Alexandra Hospital and The University of Queensland. Patients and controls for this study were all enrolled in the larger Brisbane Naevus Morphology Study, (BNMS) study

run by University of Queensland Dermatology Research Centre. Recruitment of patients was via referral with patient consent from treating doctor (including private and public hospital clinics) or some patients had heard about the trial from friends or family and self referred. Controls were people with no history of melanoma who volunteered for the BNMS, often due to knowledge of it via friends, family, or who were being seen in the dermatology clinic and were offered participation. Once identified, participants are contacted by research assistants and asked if they would like to come for an in-person consultation and more information, and if they elect to proceed, to complete the consent process.

Patients are asked to provide a saliva sample using a self collection kit (Oragene-DNA self collection, DNA Genotec, Ottowa, ON, Canada). A minimum of 2.5ug of DNA was provided to UQ Centre for Clinical Genomics (UQCCG) at the Translational Research Institute (TRI) for high throughput microarray processing. The concentration range for the submitted samples was 100 to 300ng/ul, aliquoted in 96-well semi-skirted PCR Plates (Axygen Scientific) and sealed with Clear Self-Adhesive Topseal (PerkinElmer). These samples were genotyped on the Illumina Infinium Microarray on HumanCoreExome-24.

To account for the family structure in the BNMS sample, dosage data was analysed using the Genome-Wide Efficient Mixed Model Association (GEMMA) algorithm [50]. The ordinary least squares beta output from GEMMA regression was converted to a log odds ratio a using a simplified formula adapted from the first term of the formula provided by [51] of B/(C(1-C)), where B is the ordinary least squares beta and C the proportion of cases.

23andMe

Self reported CM cases were defined as having answered yes, or reported melanoma for at least one of the following question asked to 23andMe, Inc. participants 1) "Have you ever been diagnosed by a doctor with melanoma?"; 2) "What type(s) of cancer were you diagnosed with? Please check all that apply."; 3) "What type(s) of skin cancer did you have? Please check all that apply."; 4) "Have you ever been diagnosed with melanoma?"; 5) "Have you ever been diagnosed with or treated for *any of the following conditions?"* with the list including "Melanoma". Cancer free controls were drawn from those who answered No, or not reporting melanoma, to any of the above questions, and without responding yes to any of these questions. Further, controls answered No to at least one general cancer questions, and Yes to none 1) "Have you ever been diagnosed with cancer, including skin cancer or cancerous moles?"; 2) "Has a doctor ever told you that you have a type of cancer?"; 3) "Have you ever been diagnosed or treated with any of the following conditions? Cancer, any type".

DNA samples were genotyped using one of four genotyping platforms; two variants of the Illumina HumanHap550+ BeadChip (V1, V2), a custom Illumina OmniExpress+ BeadChip (V3) with improved overlap with the V2 array, and a custom array (V4) designed for 23andMe, Inc. from Illumina that overlaps with the V2 and V3 arrays and improved coverage of lower frequency SNPs. Samples were filtered on < 98.5% call rate, < 97% European ancestry when compared to 1000 genomes reference panels, or more closely related by IBD than first cousins. Genotyped SNPs were filtered for being non-polymorphic; failing parent offspring tests; HWE $P < 1 \times 10^{-20}$; genotype missingness >= 10% or exhibiting batch effects across the 4 GWAS arrays used. SNPs were imputed by genotyping batch (V1-V4), with 945446 SNPs passing QC across the batches. Following imputation to the September 2013 release of 1000 Genomes Phase1 reference panel using Minimac2, there were 12833621 SNPs in total with 9,829,439 exceeding imputation quality threshold (INFO) > 0.5. Variants in the HLA region, including imputed HLA alleles for HLA-A, B, C, DPB1, DQA1, DQB1, and DRB1 loci at four-digit resolution, were imputed using HIBAG6. Imputed HLA alleles were otherwise analysed as dosage data as per the rest of the imputed

genotype data. Logistic regression using genotype dosages was performed with Age, sex and PC0-4 (first 5 Pcs) as covariates.

Athens – Andreas Sygros Hospital Study

The Athens sample is a case-control study that collected melanoma cases from Andreas Sygros Hospital and the Medical Oncology Center of Laiko Hospital. These are large referral centers for melanoma and skin cancer that receives the majority of melanoma cases in Athens and central Greece [52]. Melanoma cases were defined as individuals with a histologically confirmed diagnosis of invasive melanoma between 2000 to 2014. Melanoma-free controls were collected from blood donors of an Athens blood donation center or individuals with minor skin diseases and no history of skin malignancy who attended the hospital outpatient service or who were hospital personnel. All participants were older than 18 years. Demographic variables, pigmentation traits (eye, hair, and skin color), skin phototype, tanning ability, and information from clinical examination were obtained through a questionnaire that was filled out by all participants under the supervision of a certified dermatologist who performed the clinical examination. The questionnaire and clinical examination was conducted for both melanoma cases and controls. The participation rate was approximately 95% for patients and 90% for controls. The study protocol was approved by the Scientific and Ethics Committee of Andreas Sygros Hospital, and all participating individuals gave written informed consent before study participation.

Athens – Hellenic Cooperative Oncology Group (HeCOG)

Melanoma cases from this study were previously enrolled in Trial 13A/98, a prospective, multicenter, randomized phase III trial conducted at 13 institutions by the Hellenic Cooperative Oncology Group (HeCOG) [53]. The trial enrolled patients with histologically confirmed primary cutaneous melanoma between 1998 and 2004. Eligible cases were also required to be ≥18 years of age with an ECOG performance status ≤1, adequate organ function and no significant medical or psychiatric comorbidity. Cases were excluded if they had received prior adjuvant chemotherapy, radiotherapy or immunotherapy. Melanoma-free controls were randomly selected healthy unrelated Greek individuals. All participants gave informed consent to participate in the trial. This study was approved by the ethics committees of the Hellenic Cooperative Oncology Group, University Authorities and "G. Gennimatas" Hospital.

Barcelona – Melanoma: Image, Genetics and Immunology Study

The Melanoma Unit of from IDIBAPS (Barcelona, Spain), Hospital Clínic of Barcelona, and CIBERER (Instituto de Salud Carlos III, Spain) started recruiting familial and sporadic melanoma cases and controls in 1992. The collection consists of a series of case-series, case-control and cohort studies from Barcelona, Spain [54–63]. Since 1999, data has been collected on clinical information, follow-up, and biological species. Melanoma cases were defined as individuals with clinically diagnosed melanoma based on histopathological evaluation. Controls were cancer-free individuals, some of which were non-affected family members recruited for familial melanoma studies. In the case-control melanoma studies, the controls used have no personal or family history of melanoma. Matching was performed based on age and gender for some included studies.

Cyprus

The Cypriot cohort consisted of 32 histologically confirmed unrelated cases of malignant melanoma and 201 healthy controls. The control samples were cancer-free volunteers participating in the MASTOS study [64]. All subjects were above 18 years.

Cesena

The Cesena study is a case-control study conducted at M. Bufalini Hospital, Cesena, Italy. Cases are cutaneous melanoma cases with histopathological confirmation which included hematoxylin and eosin staining, HMB-45, S100, and Melan-A (MART-1) makers. Healthy control subjects were

recruited from patient's spouses or friends from in-patients or out-patients at the dermatologic department mainly with diagnoses of streptodermitis, subcutaneous abscesses, lipomas, mild contact dermatititis, hair or nail diseases. Controls were melanoma free at time of recruitment and matched to cases on age, gender and region of birth. All participants donated a sample of peripheral blood for molecular analysis and were examined by a single physician who reported data on pigmentation, nevi and other skin characteristics. All participants provided written informed consent and responded to an interview-based questionnaire on melanoma risk factors.

Genoa – The Genoa Study

The Genoa study is a hospital based case-control study conducted at the Genetics of Rare Cancers Unit at the San Martino University Hospital in Genoa, Italy and IRCCS AOU San Martino-IST Research Hospital [65–68]. Cases were cutaneous melanoma patients recruited over the past 15 years referred for clinical genetic testing. Cases were either sporadic or affected by multiple primary melanoma or were probands from melanoma families. All cases were 18 years of age or older with a histologically confirmed in situ or invasive cutaneous melanoma. Controls were healthy volunteers or spouses 18 years old or older who were recruited in the same period as the cases and matched by gender, age and place of residence. Participants donated a sample of peripheral blood for molecular analyses as well as filled out a detailed questionnaire on hair, eye and skin color, sun exposure, and history of melanoma. A trained interviewer also assessed presence of freckles and nevi. All participants provided written informed consent and the study was approved by the Ethics Committees of the participating Institutions.

L'Aquila

The Department of Dermatology of the University of L'Aquila, Italy recruited familial and sporadic melanoma cases and controls for study participation [69–72]. Familial melanoma cases were defined as patients from melanoma families with at least 2 first- or second-degree relatives or with at least 3 documented cases of melanoma irrespective of the degree of relatedness. Sporadic melanoma cases were defined as patients consecutively presenting with histopathologically confirmed sporadic primary cutaneous melanoma of any stage either as a first diagnosis or during the periodic follow-up. Controls were individuals free of melanoma at time of study entry. Only controls related to sporadic melanoma patients were recruited. Control individuals were matched to cases by gender, age (within 1 year), ethnicity and residential area and were consecutively recruited in the same period from patients with allergic disorders, skin infections, psoriasis, ulcers or autoimmune diseases referred to the same department. All sporadic and familial melanoma patients and control individuals were of Italian origin. Basic demographic information and phenotypic characteristics of patients were collected through a standardized questionnaire. Skin examination was performed by a dermatologist who evaluated skin type, hair color, eye color, number of melanocytic nevi and presence or absence of clinically atypical nevi.

Milano

The Milano study is a case-control study conducted at the Fondazione IRCCS Istituto Nazionale Tumori in Milan, Italy. Melanoma cases were defined as patients hospitalized for surgical treatment of melanoma at the Melanoma and Sarcoma Surgery Unit. All cases of melanoma were confirmed by histopathology. Controls were melanoma free at study entry and recruited from healthy blood donors from the Immunohematology and Transfusion Medicine Unit. No matching of controls to cases was performed. Phototype information was obtained by a self-administered questionnaire. Atypical nevi and phototypes were defined by a single clinician.

Padua

The Padua study is comprised of melanoma-prone families with at least 2 melanoma cases referred to the Veneto Institute of Oncology (IOV) in Padua, Italy [73,74]. All cases of melanoma were confirmed by pathology reports. Family history and phenotypic features were obtained from

personal interviews using a standardized questionnaire. Self-reported hair/eye/skin colour was recorded using reference colour classes. Presence/absence of freckles and nevus count were defined according to physical examination by a trained dermatologist. Control participants were healthy individuals without personal or family history of melanoma who were donors at the Blood Collection Center, Hospital Transfusion Centre of Padua. All melanoma patients and control individuals were of Caucasian origin and living in Veneto, a region in the North-East Italy. Written informed consent was obtained by all participants enrolled in the study and the project was approved by the institutional local ethical committee.

Roma

The Roma study is a series of familial and sporadic melanoma cases recruited from the Institute of Dermatology, Catholic University Rome, Italy since 2014. Familial melanoma cases were defined as patients from melanoma families with at least 2 first- or second-degree relatives or with at least 3 documented cases of melanoma irrespective of the degree of relatedness. Sporadic cases were defined as patients consecutively presenting with histopathologically confirmed sporadic primary cutaneous melanoma of any stage either as a first diagnosis or during the periodic follow-up. All cases were confirmed by histopathological review of melanoma tissue slides. No controls were recruited that were related to sporadic melanoma patients. A standardized questionnaire was administered by physicians trained in data gathering to collect demographic information and phenotypic characteristics. Skin examination was performed by a dermatologist who evaluated skin type, hair color, eye color, number of melanocytic nevi and presence or absence of clinically atypical nevi.

Valencia

The Valencia study consists of sporadic cases, familial cases and melanoma-free controls [75].
The sporadic cases consisted of Spanish cutaneous melanoma patients recruited at the Department of Dermatology, Instituto Valenciano de Oncologia, a referral skin cancer center for the provinces of Valencia, Alicante and Castellon, with a catchment population of approximately 5 million people. All cases are histopathologically confirmed melanoma cases by an expert pathologist devoted to skin cancer pathology and were treated at the center. Clinical and pathological data from the patients were prospectively collected since January 2000 through the review of medical history, personal interview and clinical examination by expert dermatologists. Controls are disease free and ethnically matched blood donors recruited at the Transfusion Center of Valencia between September 2008 and January 2009. All the epidemiological and phenotypic characteristics of skin were obtained from a structured questionnaire performed at the time of blood sampling.

Acknowledgements and Funding

NCI
This study was supported by the Intramural Research program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services.

AOCS/OCAC/SEARCH
AOCS/OCAC/SEARCH is accessible via European Genome-Phenome Archive. We acknowledge their support and data, and the contribution of the study nurses, research assistants and all clinical and scientific collaborators in generation of these data. We also acknowledge their funding sources; OCAC (NIH U19CA148112) , SEARCH team (Cancer Research UK C490/A16561), AOCS (U.S. Army Medical Research and Materiel Command under DAMD17-01-1-0729, The Cancer Council Victoria, Queensland Cancer Fund, The Cancer Council New South Wales, The Cancer Council South Australia, The Cancer Foundation of Western Australia, The Cancer Council Tasmania and the NHMRC (ID400413 and ID400281, as well as support from S. Boldeman, the Agar family, Ovarian Cancer Action (UK), Ovarian Cancer Australia and the Peter MacCallum Foundation).

Endometriosis
Contributors to the Endometriosis collection : Anjali K. Henders, S.H. Kennedy, S. Macgregor, N.G. Martin, S. Missmer, G.W. Montgomery, D.R. Nyholt, J.N. Painter, S.A. Treloar, L. Wallace, K.T. Zondervan. Acknowledgements: We acknowledge all the participants in the QIMR and endometriosis studies. We thank Anjali Henders, Leanne Wallace, and Lisa Bowdler for project management, sample processing and database development. We thank Endometriosis Associations for supporting study recruitment and S. Nicolaides and the Queensland Medical Laboratory for assistance with blood collection including pro bono collection and delivery of blood samples. Funding: This work was supported by the Cooperative Research Centre (CRC) for Discovery of Genes for Common Human Diseases, Cerylid Biosciences (Melbourne), The Wellcome Trust and donations from Neville and Shirley Hawkins. Endometriosis sample genotyping was funded by a grant from the Wellcome Trust (WT084766/Z/08/Z) and NHMRC (496610, GNT1049472, GNT1050208). G.W.M. is supported by the NHMRC Fellowships scheme. D.R.N. was supported by the NHMRC Fellowship (613674) and Australian Research Council (ARC) Future Fellowship (FT0991022) schemes.

QTWIN/QMEGA
The Q-MEGA and QTWIN study would like to recognise and thank A. Baxter, M. de Nooyer, I. Gardner, D. Statham, B. Haddon, M.J. Wright, J. Palmer, J. Symmons, B. Castellano, L. Bowdler, S. Smith, D. Smyth, L. Wallace, M.J. Campbell, A. Caracella, M. Kvaskoff, O. Zheng, B. Chapman and H. Beeby for their input in project management, sample processing and database development. We are also grateful to the many research assistants and interviewers for assistance with the collection of these datasets.

The Q-MEGA/QTWIN study was supported by the Melanoma Research Alliance, the NIH NCI (CA88363, CA83115, CA122838, CA87969, CA055075, CA100264, CA133996 and CA49449), the National Health and Medical Research Council of Australia (NHMRC) (200071, 241944, 339462, 380385, 389927,389875, 389891, 389892,389938, 443036, 442915, 442981, 496610, 496675, 496739, 552485, 552498, APP1049894), the Cancer Councils New South Wales, Victoria and Queensland, the Cancer Institute New South Wales, the Cooperative Research Centre for Discovery of Genes for Common Human Diseases (CRC), Cerylid Biosciences (Melbourne), the Australian Cancer Research Foundation, The Wellcome Trust (WT084766/Z/08/Z) and donations from Neville and Shirley Hawkins. Stuart MacGregor acknowledges fellowship support from the

Australian National Health and Medical Research Council and from the Australian Research Council.

Princess Alexandra Hospital (PAH) samples
We acknowledge the support and assistance of study participants, and Adele Greene (securing funding), Mark Smithers (cohort establishment), as well as Casey Rowe and Maryrose Malt for consenting and collection of patient samples. Funding for the PAH collection was via the University of Queensland Diamantina Institute, the Meehan Foundation, NHMRC CDF (1125290), and Cancer Council Queensland (1125237).

AMFS
This work was supported by the National Health and Medical Research Council of Australia (NHMRC) project grants 566946, 107359, 211172, and program grant number 402761 to GJM and RFK. Work was also funded by the Cancer Councils of Victoria, Queensland, and New South Wales (project grants 77/00, 06/10, 371) and by US NIH RO1 grant CA-83115-01A2 and 2R01CA083115-11A1. Anne E. Cust is supported by fellowships from the Cancer Institute NSW and the NHMRC. We gratefully thank the support and involvement of all participants, research coordinators, interviewers, examiners and data management staff.

Study of Digestive Health
Controls for use with the Q-MEGA_omni dataset were derived from the Study of Digestive Health group (SDH) which was funded by NCI grant 5 RO1 CA 001833-02. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute. We gratefully acknowledge the cooperation of the following institutions: Sullivan and Nicolaides Pathology (Brisbane); Queensland Medical Laboratory (Brisbane); Queensland Health Pathology Services (Brisbane); Institute of Medical and Veterinary Science (Adelaide); SouthPath (Adelaide). We also acknowledge the contribution of the study nurses and research assistants and would like to thank all of the people who participated in the study. DCW was supported by an NHMRC Research Fellowship (APP1058522).

Inflammatory Bowel Disease (IBD)
We gratefully acknowledge the aid in identifying study participants by Sullivan and Nicolaides Pathology, Queensland Medical Laboratories and the Queensland Health Pathology Service. The IBD study would like to recognise and thank Peter Schultz, Lauren Aoude, Loralie Parsonson, Stephen Walsh, Mitchell Stark, John Cardinal and Herlina Handoko for technical support.

This work was funded by US NCI grant CA 001833-03. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute. PMW and DCW are Senior Research Fellows of the National Health and Medical Research Council (NHMRC) of Australia. NP was supported by a NHMRC PhD scholarship. The funding bodies played no role in the design or conduct of the study, the collection, management, analysis, or interpretation of the data, or the preparation, review or approval of the manuscript.

Essen-Heidelberg
The study was supported by a grant from Deutsche Forshungsgemeinschaft (GZ: SCHA 422/11-1).

Western Australian Melanoma Health Study (WAMHS)
The WAMHS, and the salaries of its staff and PhD students, was funded by the Scott Kirkbride Melanoma Research Centre. The Cancer Council Western Australia is also acknowledged for current salary support for Sarah Ward (Capacity Building and Collaboration grant). We acknowledge the donation of time and samples from study participants, the WAMHS study team, and the WAMHS Management Committee. This work could not have occurred without the

assistance and support of the Western Australian (WA) DNA Bank, and the Ark at The University of WA, and the WA Cancer Registry, for which we are grateful.

Australian & New Zealand Registry of Advanced Glaucoma (ANZRAG)
Participant collection was funded by the Royal Australian and New Zealand College of Ophthalmology Eye Foundation. Genotyping was supported by NHMRC grants 535074 and 1023911. ANZRAG also acknowledges funding from the BrightFocus Foundation and a Ramaciotti Establishment Grant. The authors acknowledge the support of Ms. Bronwyn Usher-Ridge in patient recruitment and data collection, and Dr Patrick Danoy and Dr Johanna Hadler for genotyping.

Generations Scotland
Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award "STratifying Resilience and Depression Longitudinally" (STRADL) Reference 104036/Z/14/Z)."

Michigan
The authors acknowledge the University of Michigan Precision Health Initiative and Medical School Central Biorepository for providing biospecimen storage, management, processing and distribution services and the Center for Statistical Genetics in the Department of Biostatistics at the School of Public Health for genotype data curation, imputation, and management in support of this research.

Brisbane Nevus Morphology Study
This work was funded by project grant numbers NHMRC 1004999, 1062935 and the Centre of Research Excellence for the Study of Naevi 1099021. This research was carried out at the Translational Research Institute, Woolloongabba, QLD 4102, Australia.
The Translational Research Institute is supported by a grant from the Australian Government

MIA
This work was supported by research funding from Melanoma Institute Australia, the National Health and Medical Research Council of Australia (NHMRC) through program grants to GJM, RAS & GVL and from Cancer Institute New South Wales and infrastructure grants from Macquarie University and the Australian Cancer Research Foundation. R.A.S. and G.V.L. are supported by NHMRC Fellowships, and G.V.L. is supported by the University of Sydney Medical Foundation.

23andMe
We would like to thank the research participants and employees of 23andMe for making this work possible.

Nevus GWAS meta-analysis dataset
For full funding support and acknowledgements see [9]. These are briefly summarised here. All studies gratefully acknowledge the support of all research participants and twins, research nurses, research assistants and support staff, without whom this work would not be possible. In addition to specific named thanks below, we also acknowledge the support of G. Clement, Bernet Keto, Pirro Hysi and Emad Qweitin.

Acknowledgements for GWAS contributing to the nevus meta-analysis are as follows:

ALSPAC: The UK Medical Research Council, Wellcome Trust (Grant ref: 102215/2/13/2), the University of Bristol, Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America). 23andMe, ARC Future Fellowship (FT130101709), and a Medical Research Council program grant (MC_UU_12013/4).

Harvard: The NHS and HPFS cohorts, US State Cancer Registries (for the full list of states see [9]), the Odyssey cluster (FAS Division of Science, Research Computing Group at Harvard University.). Funding support by NIH R01 CA49449, P01 CA87969, UM1
CA186107, and UM1 CA167552.

Leeds: Funding from Cancer Research UK (project grant C8216/A6129 and programme award C588/A4994) and by the NIH (R01 CA83115). The UK National Cancer Research supported participant recruitment. We specifically acknowledge the support of Patricia Mack, Kate Gamble, Paul King, and Dr Amy Downing.

QIMR: We specifically thank Dixie Statham, Ann Eldridge, Marlene Grace, Kerrie McAloney, Natalie Garden, Reshika Chand, Lisa Bowdler, Leanne Wallace, David Smyth, Harry Beeby, and Daniel Park. **The Brisbane Twin Nevus Study** was supported by Australian National Health and Medical Research Council (NHMRC) grants (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498, 1031119), Adult twin data collection by NIH grant (AA011998_5978), over-50 twins by a donation from Mr George Landers of Chania, Crete. Genotyping was funded by the NHMRC (552498 and 1049894). We also acknowledge the U.S. National Institutes of Health's support through grants AA07535, AA10248, AA13320, AA13321, AA13326, AA14041, and MH66206. This work was also funded by the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254). Genotyping for individuals using the Illumina 370K was funded by an access award to Dr. Richard Todd and performed at the Center for Inherited Disease Research, Baltimore. Nick Hayward, David C. Whiteman, David. L. Duffy, Grant W. Montgomery were supported by the NHMRC Fellowships scheme during the collection of these samples.

Raine: We acknowledge the support for cohort coordination and data collection from theRaine Study and Lions Eye Institute. The Raine Study is supported and funded by The University of Western Australia (UWA), The Telethon Institute for Child Health Research, Raine Medical Research Foundation, UWA Faculty of Medicine, Dentistry and Health Sciences, Women's and Infant's Research Foundation and Curtin University. Genotyping was funded by the NHMRC (1021105). Support for the REHS was provided by LEI, the Australian Foundation for the Prevention of Blindness and the Ophthalmic Research Institute of Australia.

The Rotterdam Study: We specifically acknowledge the assistance and support of Ada Hooghart, Corina Brussee, Riet Bernaerts-Biskop, Patricia van Hilten, Pascal Arp, Jeanette Vergeer, Maarten Kooijman and Lennart Karssen. The Rotterdam study was supported by the Netherlands Organisation of Scientific Research (NWO); Erasmus Medical Center and Erasmus University, Rotterdam, The Netherlands; Netherlands Organization for Health Research and Development (ZonMw); UitZicht; the Research Institute for Diseases in the Elderly; the Ministry of Education, Culture and Science; the Ministry for Health, Welfare and Sports; the European Commission (DG XII); the Municipality of Rotterdam; the Netherlands Genomics Initiative/NWO; Center for Medical Systems Biology of NGI; Stichting Lijf en Leven; Stichting Oogfonds Nederland; Landelijke Stichting voor Blinden en Slechtzienden; Algemene Nederlandse Vereniging ter Voorkoming van Blindheid; Medical Workshop; Heidelberg Engineering; Topcon Europe BV. Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810.

F.L. was supported by the Erasmus University Rotterdam (EUR) fellowship and the Chinese recruiting program "The 1000 Talents Plan" for young scholars. GWAS genotyping was funded by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012).

Twins eye study in Tasmania: This work was supported by an NHMRC Enabling Grant (2004-2009, 350415, 2005-2007); Clifford Craig Medical Research Trust; Ophthalmic Research Institute of Australia; American Health Assistance Foundation; Peggy and Leslie Cranbourne Foundation; Foundation for Children; Jack Brockhoff Foundation, and the National Eye Institute of the National Institute of Health grants RO1EY01824601 from 2007 to 2010. D. Mackey was supported by a Pfizer Australia Senior Research Fellowship and S. MacGregor was a recipient of an Australian NHMRC Career Development Award. Genotyping was funded by an NHMRC Medical Genomics Grant; US NIH/National Eye Institute (1RO1EY018246), Australian sample imputation analyses were carried out on the Genetic Cluster Computer, which is financially supported by the Netherlands Scientific Organization (NWO48005003).

TwinsUK: TwinsUK specifically thanks U. Perks, and acknowledges the funding support from the Wellcome Trust, Medical Research Council, European Union, the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London, and . T.D.S. is an NIHR senior investigator and D.G is an MRC clinical research fellow. Genotyping was supported by the National Eye Institute via an NIH/CIDR genotyping project (PI: Terri Young), the Genotyping Facilities at the Wellcome Trust Sanger Institute.

ENGAGE
We thank the members of the Engage consortium for providing the telomere GWAS data reported in [16]; please see this publication for acknowledgements relevant to this dataset.

MDACC
We thank the individuals who volunteered to participate in this project. This work was supported by the National Cancer Institute of the National Institutes of Health through SPORE grant P50 CA093459 and Cancer Center Support Grant P30 CA016672 (Clinical Trials Support Resource), as well as by philanthropic contributions to The University of Texas MD Anderson Cancer Center Moon Shots Program, The University of Texas MD Anderson Cancer Center Various Donors Melanoma and Skin Cancers Priority Program Fund, the Miriam and Jim Mulva Research Fund, the McCarthy Skin Cancer Research Fund, and the Marit Peterson Fund for Melanoma Research.

Genoa
The Genoa study was supported by The Italian Ministry of Health Grant RF-2016-02362288 and 5x1000 per la ricerca corrente. The authors acknowledge Dr William Bruno for his continuous work on melanoma families.

Barcelona
The research at the Melanoma Unit in Barcelona is partially funded by Spanish Fondo de Investigaciones Sanitarias grants PI15/00716, PI15/00956, PI18/00419 and PI18/01077 ; CIBER de Enfermedades Raras of the Instituto de Salud Carlos III, Spain, co-financed by European Development Regional Fund "A way to achieve Europe" ERDF; AGAUR 2017_SGR_1134 of the Catalan Government, Spain; European Commission under the 6th Framework Programme, Contract No. LSHC-CT-2006-018702 (GenoMEL) and by the European Commission under the 7th Framework Programme, Diagnoptics; The National Cancer Institute (NCI) of the US National Institute of Health (NIH) (CA83115); a grant from "Fundació La Marató de TV3" 201331-30, Catalonia, Spain; a grant from "Fundación Científica de la Asociación Española Contra el Cáncer"

GCB15152978SOEN, Spain, and CERCA Programme / Generalitat de Catalunya. Part of the work was carried out at the Esther Koplowitz Center, Barcelona.

ICR
We thank Breast Cancer Now and the Institute of Cancer Research for funding, and acknowledge National Health Service funding to the Royal Marsden NHS Foundation Trust and Institute of Cancer Research NIHR Biomedical Research Centre.

Group Authorships

AMFS investigators

Anne E. Cust[1,2], Helen Schmid[2], Elizabeth A. Holland[2], Joanne F. Aitken[3], Bruce K. Armstrong[1], Graham G. Giles[3,4], Richard F. Kefford[2], John L. Hopper[5] Mark A. Jenkins[5], Graham J. Mann[2]

1) Cancer Epidemiology and Prevention Research, Sydney School of Public Health, The University of Sydney, Sydney, Australia 2) Melanoma Institute Australia, The University of Sydney, Sydney, Australia 3) Centre for Cancer Research, Westmead Institute for Medical Research, The University of Sydney, Sydney, Australia 4) Viertel Centre for Research in Cancer Control, Cancer Council Queensland, Spring Hill, Brisbane, Australia 5) Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia 6) Centre for Epidemiology & Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia

Queensland study of Melanoma: Environmental and Genetic Associations (Q-MEGA) and QTWIN investigators

Q- MEGA PIs: Nicholas G. Martin[1], Grant W. Montgomery[1], David Duffy[1], David C. Whiteman[1], Matthew H. Law[1], Stuart MacGregor[1], Nicholas K. Hayward[1]. The Australian Cancer Study (ACS) Principal Investigators are: David C. Whiteman[1], Penny Webb[1], Adele Green[1], Peter Parsons[1], David Purdie[1], Nicholas K.Hayward[1]. Additional QTWIN PIs: Zhen Zhen Zhao[1], Joanne F Aitken[2], Anjali K.Henders[1], Mitchell Stark[1], David L. Duffy[1], Jodie N. Painter[1]

1) QIMR Berghofer, Brisbane, QLD 4029, Australia. 2) Viertel Centre for Research in Cancer Control, Cancer Council Queensland, Spring Hill, Brisbane, Australia

The SDH Study Group

Chief Investigators: David Whiteman[1], Adele Green[1], Nicholas Hayward[1], Peter Parsons[1], Sandra Pavey[1], David Purdie[1], Penny Webb[1], David Gotley[2], Mark Smithers[2], Paul Drew[3], Glyn Jamieson[3], Paul Drew[4], David Watson[4], Andrew Clouston[5]

1) QIMR Berghofer Medical Research Institute. 2) University of Queensland / Princess Alexandra Hospital 3) University of Adelaide 4) Flinders University of South Australia 5) Mayne Pathology

Clinical Collaborators: A Clouston[1], I Brown[2], N Walker[3], Justin Bessell[4], William Tam[5], Andrew Ruskowicz[6]

1) Envoi Pathology, QLD 2) S&N Pathology, QLD 3) QML Pathology, QLD 4) Flinders Medical Centre, SA, 5) Royal Adelaide Hospital, SA, 6) Institute of Medical and Veterinary Science, SA

Research Staff: D Nancarrow, D Hussey, E Smith, G Mayne. QLD Research Nurses: A McMurtrie, L Terry, M Connard, L Jackman, S Perry, M Davis. SA Research Nurses: D Roffe, M Martin, L Smith. QIMR Berghofer Project Manager: S O'Brien. QIMR Berghofer Data Manager: T Sadkowsky.

IBD investigators

Lisa Simms[1], Grant W. Montgomery[2], Peter Visscher[3]

1) Inflammatory Bowel Diseases Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, Australia 2) Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia 3) The Queensland Brain Institute, The University of Queensland, QBI Building, St Lucia, Queensland 4071, Australia.

Essen-Heidelberg investigators

Rajiv Kumar, Dirk Schadendorf[1], Hans-Joachim Schulze[2], Kari Hemminki[3], Antje Sucker[1], Thomas Vogt[4], Johan Hansson[5], Ralf Gutzmer[6], Helen Gogas[7], Dave Hoon[8], Eduardo Nagore[9], John Kirkwood[10], Benjamin Weide[11], Piotr Rutkowski[12], Selma Ugurel[1]

1) Department of Dermatology, University Hospital Essen, 45122 Essen, Germany and German Consortium Translational Cancer Research (DKTK), 69120 Heidelberg, Germany
2) Department of Dermatology, Fachklinik Hornheide, Institute for Tumors of the Skin at the University of Münster, Germany
3) Division of Molecular Genetic Epidemiology, German Cancer Research Center, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany
4) University Hospital Saarland, Department of Dermatology, Venerology and Allergology, Building 18, Kirrberger Straße, D - 66424 Homburg/Saar, Germany
5) Department of Oncology Pathology, Karolinska Institutet, Karolinska University Hospital, Solna S 171 76, Stockholm, Sweden.
6) Department of Dermatology and Allergy, Skin Cancer Center Hannover, Hannover Medical School, Carl-Neuberg-Str. 1, 30625 Hannover, Germany
7) 1st Department of Medicine, University of Athens Medical School, Laiko Hospital, PO 14120, 11510, Athens, Greece
8) John Wayne Cancer Institute, 2200 Santa Monica Blvd, Santa Monica, CA 90404, USA
9) Department of Dermatology,Instituto Valenciano de Oncologia, Valencia 46009, Spain
10) Dermatology & Translational Science, Melanoma and Skin Cancer Program, 5117 Centre Avenue, Suite 1.32, Pittsburgh, PA 15213, USA
11) Department of Dermatology, University Medical Center, Liebermeisterstr. 25, 72076 Tübingen, Germany
12) Department of Soft Tissue/Bone Sarcoma and Melanoma, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Roentgena 5, 02-781 Warsaw, Poland

The 23andMe Research Team

Michelle Agee, Babak Alipanahi, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah L. Elson, Pierre Fontanillas, Nicholas A. Furlotte, David A. Hinds, Karen E. Huber, Aaron Kleinman, Nadia K. Litterman, Jennifer C. McCreight, Matthew H. McIntyre, Joanna L. Mountain, Elizabeth S. Noblin, Carrie A.M. Northover, Steven J. Pitts, J. Fah Sathirapongsasuti, Olga V. Sazonova, Janie F. Shelton, Suyash Shringarpure, Chao Tian, Joyce Y. Tung, Vladimir Vacic, and Catherine H. Wilson.

MelaNostrum Consortium

Paula Aguilera, Llucia Alos, Celia Badenas, Alicia Barreiro, Neus Calbet-Llopart, Cristina Carrera, Marc Combalia, Natalia Espinosa, Daniel Gabriel, Adriana Garcia, Melinda Gonzalez, Pablo Iglesias, Josep Malvehy, Judith Mateu, Sebastian Podlipnik, Miriam Potrony, Susana Puig, Joan Anton Puig-Butille, Gemma Tell-Marti, Pol Gimenez-Xavier

References

1. Brown, K. M. et al. Common sequence variants on 20q11.22 confer melanoma susceptibility. Nat. Genet. **40**, 838–840 (2008).
2. Bishop, D. T. et al. Genome-wide association study identifies three loci associated with melanoma risk. Nat. Genet. **41**, 920–925 (2009).
3. Amos, C. I. et al. Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. Hum. Mol. Genet. **20**, 5012–5023 (2011).
4. Barrett, J. H. et al. Genome-wide association study identifies three new melanoma susceptibility loci. Nat. Genet. **43**, 1108–1113 (2011).
5. Macgregor, S. et al. Genome-wide association study identifies a new melanoma susceptibility locus at 1q21.3. Nat. Genet. **43**, 1114–1118 (2011).
6. Iles, M. M. et al. A variant in FTO shows association with melanoma risk not due to BMI. Nat. Genet. (2013). doi:10.1038/ng.2571
7. Law, M. H. et al. Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. Nat. Genet. **47**, 987–995 (2015).
8. Ransohoff, K. J. et al. Two-stage genome-wide association study identifies a novel susceptibility locus associated with melanoma. Oncotarget **8**, 17586–17592 (2017).
9. Duffy, D. L. et al. Novel pleiotropic risk loci for melanoma and nevus density implicate multiple biological pathways. Nat. Commun. **9**, 4774 (2018).
10. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. Nat. Genet. (2015). doi:10.1038/ng.3406
11. Carvalho-Silva, D. et al. Open Targets Platform: new developments and updates two years on. Nucleic Acids Res. **47**, D1056–D1065 (2019).
12. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience **4**, 7 (2015).
13. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. **88**, 76–82 (2011).
14. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat. Genet. **44**, 369–75, S1–3 (2012).
15. Higgins, J. P. T. & Thompson, S. G. Quantifying heterogeneity in a meta-analysis. Stat. Med. (2002). doi:10.1002/sim.1186
16. Codd, V. et al. Identification of seven loci affecting mean telomere length and their association with disease. Nat. Genet. **45**, 422–7, 427e1–2 (2013).
17. Delgado, D. A. et al. Genome-wide association study of telomere length among South Asians identifies a second RTEL1 association signal. J. Med. Genet. **55**, 64–71 (2018).
18. Zhang, T. et al. Cell-type-specific eQTL of primary melanocytes facilitates identification of melanoma susceptibility genes. Genome Res. **28**, 1621–1635 (2018).
19. Hormozdiari, F. et al. Colocalization of GWAS and eQTL Signals Detects Target Genes. Am. J. Hum. Genet. **99**, 1245–1260 (2016).
20. Baxter, A. J. et al. The Queensland Study of Melanoma: Environmental and Genetic Associations (Q-MEGA); study design, baseline characteristics, and repeatability of phenotype and sun exposure measures. Twin Res. Hum. Genet. (2008). doi:10.1375/twin.11.2.183
21. Zhu, G. et al. A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: a maximum-likelihood combined linkage and association analysis in twins and their sibs. Am. J. Hum. Genet. **65**, 483–492 (1999).
22. McGregor, B. et al. Genetic and environmental contributions to size, color, shape, and other characteristics of melanocytic naevi in a sample of adolescent twins. Genet. Epidemiol. **16**, 40–53 (1999).
23. Painter, J. N. et al. Genome-wide association study identifies a locus at 7p15.2 associated with endometriosis. Nat. Genet. **43**, 51–54 (2011).
24. Kendall, B. J. et al. Leptin and the risk of Barrett's oesophagus. Gut (2008). doi:10.1136/gut.2007.131243
25. Cust, A. E. et al. Population-based, case-control-family design to investigate genetic and environmental influences on melanoma risk. Am. J. Epidemiol. (2009). doi:10.1093/aje/kwp307
26. Ward, S. V. et al. The Western Australian Melanoma Health Study: study design and participant characteristics. Cancer Epidemiol. **35**, 423–431 (2011).

27. Hume, G. E. et al. Novel NOD2 haplotype strengthens the association between TLR4 Asp299gly and Crohn's disease in an Australian population. Inflamm. Bowel Dis. (2008). doi:10.1002/ibd.20362
28. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. Nature **562**, 203–209 (2018).
29. Bellenguez, C. et al. A robust clustering algorithm for identifying problematic samples in genome-wide association studies. Bioinformatics **28**, 134–135 (2012).
30. Fritsche, L. G. et al. Association of Polygenic Risk Scores for Multiple Cancers in a Phenome-wide Study: Results from The Michigan Genomics Initiative. Am. J. Hum. Genet. **102**, 1048–1061 (2018).
31. Li, J. Z. et al. Worldwide human relationships inferred from genome-wide patterns of variation. Science **319**, 1100–1104 (2008).
32. Carroll, R. J., Bastarache, L. & Denny, J. C. R PheWAS: data analysis and plotting tools for phenome-wide association studies in the R environment. Bioinformatics **30**, 2375–2376 (2014).
33. Ma, C., Blackwell, T., Boehnke, M., Scott, L. J. & GoT2D investigators. Recommended joint and meta-analysis strategies for case-control association testing of single low-count variants. Genet. Epidemiol. **37**, 539–550 (2013).
34. Newton-Bishop, J. A. et al. Melanocytic nevi, nevus genes, and melanoma risk in a large case-control study in the United Kingdom. Cancer Epidemiol. Biomarkers Prev. **19**, 2043–2054 (2010).
35. Newton-Bishop, J. A. et al. Relationship between sun exposure and melanoma risk for tumours in different body sites in a large case-control study in a temperate climate. Eur. J. Cancer **47**, 732–741 (2011).
36. Falchi, M. et al. Genome-wide association study identifies variants at 9p21 and 22q13 associated with development of cutaneous nevi. Nat. Genet. **41**, 915–919 (2009).
37. Pooley, K. A. et al. Common single-nucleotide polymorphisms in DNA double-strand break repair genes and breast cancer risk. Cancer Epidemiol. Biomarkers Prev. **17**, 3482–3489 (2008).
38. Pooley, K. A. et al. No association between TERT-CLPTM1L single nucleotide polymorphism rs401681 and mean telomere length or cancer risk. Cancer Epidemiol. Biomarkers Prev. (2010). doi:10.1158/1055-9965.EPI-10-0281
39. Song, F. et al. Exonuclease 1 (EXO1) gene variation and melanoma risk. DNA Repair (2012). doi:10.1016/j.dnarep.2011.12.005
40. Li, Y., Willer, C., Sanna, S. & Abecasis, G. Genotype imputation. Annu. Rev. Genomics Hum. Genet. **10**, 387–406 (2009).
41. Li, Y., Willer, C. J., Ding, J., Scheet, P. & Abecasis, G. R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet. Epidemiol. **34**, 816–834 (2010).
42. Kvaskoff, M. et al. Site-specific determinants of cutaneous melanoma: a case-case comparison of patients with tumors arising on the head or trunk. Cancer Epidemiol. Biomarkers Prev. **22**, 2222–2231 (2013).
43. Amos, C. I. et al. The OncoArray Consortium: A Network for Understanding the Genetic Architecture of Common Cancers. Cancer Epidemiol. Biomarkers Prev. **26**, 126–135 (2017).
44. Kuchenbaecker, K. B. et al. Identification of six new susceptibility loci for invasive epithelial ovarian cancer. Nat. Genet. **47**, 164–171 (2015).
45. Merritt, M. A. et al. Talcum powder, chronic pelvic inflammation and NSAIDs in relation to risk of epithelial ovarian cancer. Int. J. Cancer **122**, 170–176 (2008).
46. Bolton, K. L. et al. Common variants at 19p13 are associated with susceptibility to ovarian cancer. Nat. Genet. **42**, 880–884 (2010).
47. Chaudru, V. et al. Influence of Genes, Nevi, and Sun Sensitivity on Melanoma Risk in a Family Sample Unselected by Family History and in Melanoma-Prone Families. JNCI Journal of the National Cancer Institute **96**, 785–795 (2004).
48. Hercberg, S. et al. The SU.VI.MAX Study: a randomized, placebo-controlled trial of the health effects of antioxidant vitamins and minerals. Arch. Intern. Med. **164**, 2335–2342 (2004).
49. Olsen, C. M. et al. Cohort profile: the QSkin Sun and Health Study. Int. J. Epidemiol. **41**, 929–929i (2012).
50. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. **44**, 821–824 (2012).
51. Pirinen, M., Donnelly, P. & Spencer, C. C. A. EFFICIENT COMPUTATION WITH A LINEAR MIXED MODEL ON LARGE-SCALE DATA SETS WITH APPLICATIONS TO GENETIC STUDIES 1. Ann. Appl. Stat. **7**, 369–390 (2013).
52. Kypreou, K. P. et al. Prediction of Melanoma Risk in a Southern European Population Based on a

Weighted Genetic Risk Score. J. Invest. Dermatol. (2016). doi:10.1016/j.jid.2015.12.007
53. Pectasides, D. et al. Randomized phase III study of 1 month versus 1 year of adjuvant high-dose interferon alfa-2b in patients with resected high-risk melanoma. J. Clin. Oncol. **27**, 939–944 (2009).
54. Puig, S., Ruiz, A., Castel, T., Volpini, V. & Malvehy, J. Inherited susceptibility to several cancers but absence of linkage between dysplastic nevus syndrome and CDKN2A in a melanoma family with a mutation in the …. Hum. Genet. (1997).
55. Ruiz, A. et al. CDKN2A mutations in Spanish cutaneous malignant melanoma families and patients with multiple melanomas and other neoplasia. J. Med. Genet. **36**, 490–493 (1999).
56. Puig, S. et al. Role of the CDKN2A locus in patients with multiple primary melanomas. J. Clin. Oncol. **23**, 3043–3051 (2005).
57. Goldstein, A. M. et al. Features associated with germline CDKN2A mutations: a GenoMEL study of melanoma-prone families from three continents. J. Med. Genet. **44**, 99–106 (2007).
58. Harland, M. et al. Prevalence and predictors of germline CDKN2A mutations for melanoma cases from Australia, Spain and the United Kingdom. Hered. Cancer Clin. Pract. **12**, 20 (2014).
59. Potrony, M. et al. Increased prevalence of lung, breast, and pancreatic cancers in addition to melanoma risk in families bearing the cyclin-dependent kinase inhibitor 2A mutation: implications for genetic counseling. J. Am. Acad. Dermatol. **71**, 888–895 (2014).
60. Potrony, M. et al. Prevalence of MITF p.E318K in Patients With Melanoma Independent of the Presence of CDKN2A Causative Mutations. JAMA Dermatol. **152**, 405–412 (2016).
61. Puig, S. et al. Characterization of individuals at high risk of developing melanoma in Latin America: bases for genetic counseling in melanoma. Genet. Med. **18**, 727–736 (2016).
62. Potrony, M. et al. Inherited functional variants of the lymphocyte receptor CD5 influence melanoma survival. Int. J. Cancer **139**, 1297–1302 (2016).
63. Potrony, M. et al. IRF4 rs12203592 functional variant and melanoma survival. International Journal of Cancer (2017). doi:10.1002/ijc.30605
64. Hadjisavvas, A. et al. An investigation of breast cancer risk factors in Cyprus: a case control study. BMC Cancer **10**, 447 (2010).
65. Ghiorzo, P. et al. Impact of E27X, a novel CDKN2A germ line mutation, on p16 and p14ARF expression in Italian melanoma families displaying pancreatic cancer and neuroblastoma. Hum. Mol. Genet. **15**, 2682–2689 (2006).
66. Ghiorzo, P. et al. CDKN2A is the main susceptibility gene in Italian pancreatic cancer families. J. Med. Genet. **49**, 164–170 (2012).
67. Ghiorzo, P. et al. Prevalence of the E318K MITF germline mutation in Italian melanoma patients: associations with histological subtypes and family cancer history. Pigment Cell Melanoma Res. **26**, 259–262 (2013).
68. Bruno, W. et al. Multiple primary melanomas (MPMs) and criteria for genetic assessment: MultiMEL, a multicenter study of the Italian Melanoma Intergroup. J. Am. Acad. Dermatol. **74**, 325–332 (2016).
69. Fargnoli, M. C. et al. Constitutional and environmental risk factors for cutaneous melanoma in an Italian population. A case–control study. Melanoma Research **14**, 151–157 (2004).
70. Concetta Fargnoli, M. et al. Contribution of melanocortin-1 receptor gene variants to sporadic cutaneous melanoma risk in a population in central Italy: a case–control study. Melanoma Res. **16**, 175 (2006).
71. Fargnoli, M. C. et al. MC1R variants increase risk of melanomas harboring BRAF mutations. J. Invest. Dermatol. **128**, 2485–2490 (2008).
72. Pellegrini, C. et al. Characterization of melanoma susceptibility genes in high-risk patients from Central Italy. Melanoma Res. **27**, 258–267 (2017).
73. Menin, C. et al. Contribution of susceptibility gene variants to melanoma risk in families from the Veneto region of Italy. Pigment Cell Melanoma Res. **24**, 728–730 (2011).
74. Fitzpatrick, T. B. The validity and practicality of sun-reactive skin types I through VI. Arch. Dermatol. **124**, 869–871 (1988).
75. Nagore, E. et al. TERT promoter mutations in melanoma survival. Int. J. Cancer **139**, 75–84 (2016).