

Title:

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

Authors:

Matteo Di Giovannantonio^{1*}, Benjamin H.L Harris^{2*}, Ping Zhang¹, Isaac Kitchen-Smith¹, Lingyun Xiong¹, Natasha Sahgal¹, Giovanni Stracquadanio¹, Marsha Wallace¹, Sarah Blagden³, Simon Lord³, David A. Harris⁴, Adrian L. Harris⁵, Francesca M. Buffa^{2†} and Gareth Bond^{1†}

***joint first authors, †corresponding authors**

Email addresses of corresponding authors: francesca.buffa@oncology.ox.ac.uk and gareth.bond@ludwig.ox.ac.uk

Affiliations:

1. Ludwig Institute for Cancer Research, University of Oxford, Nuffield Department of Clinical Medicine, Old Road Campus Research Building, Oxford OX3 7DQ, UK.
2. Computational Biology and Integrative Genomics, Department of Oncology, University of Oxford, Oxford, UK.
3. Early Phase Clinical Trials Unit, Department of Oncology, Churchill Hospital, Headington, Oxford OX3 7LE, UK.
4. St Anne's College, 56 Woodstock Rd, Oxford OX2 6HS, UK.
5. Molecular Oncology Laboratories, Department of Oncology, The Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK.

Acknowledgments: Dawn O'Reilly for proofreading the article

Funding: Ludwig Institute for Cancer Research and Cancer Research UK

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

Abstract:

Inherited genetic variants in tumour suppressors and oncogenes can increase the cancer risk, but little is known about their influence on anthropometric traits. Through the integration of inherited and somatic cancer genetic data, we define functional single nucleotide polymorphisms (SNPs) associated with cancer risk and explore potential pleiotropic associations with anthropometric traits in a cohort of 500,000 individuals. We identify three regulatory SNPs for three important cancer genes that associate with both anthropometric traits and cancer risk. We describe a novel association of a SNP in TP53 (rs78378222) with height, lean body mass measures and basal metabolic rate, as well as validating its known associations with brain and non-melanomatous skin cancer susceptibility. Our results clearly demonstrate that heritable variants in key cancer genes can associate with both differential cancer risk and anthropometric traits in the general population, thereby lending support for a role of genetics in linking these human phenotypes.

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

Introduction

Height and other anthropometric measures have been consistently found to associate with differential cancer risk^{1, 2, 3}. However, both genetic and mechanistic insights into these epidemiological associations are notably lacking^{2, 4}. Increased height has been seen to associate with a number of malignancies including skin, breast colon, rectum, endometrium, ovary, kidney cancers as well as Hodgkin's lymphoma and leukaemia^{2, 5, 6}. Current theories underpinning links between increased height and increased cancer risk centre upon the role of growth hormone, insulin-like growth factor and/or insulin being pro-tumorigenic by up-regulating the Ras-MAPK and PI3K pathways leading to increased cellular proliferation^{7, 8}. It has also been proposed that increased height might correspond to increased cellular number and therefore increased probability of malignancy, simply by chance⁹; however, such allometric scaling of body mass only explains part of the observed effect and it is not observed among mammal species of different size (Peto's paradox)¹⁰. Like height, increased body mass index (BMI), a surrogate measure of obesity based on height and weight, has been associated with an increased risk of number of cancers including post-menopausal breast cancer, colorectal cancer and renal cancer^{3, 11, 12}. Increased BMI within individuals has been mainly attributed to increasing accumulation of fat mass. It is hypothesised that adiposity causes a state of systemic inflammation, a shift of metabolite and adipokine release, and an increase of circulating insulin, secondary to insulin resistance^{13, 14}. This combination may increase cancer risk. However, the exact and predominant mechanisms of how high BMI is linked with increased cancer risk are not understood. Thus, links between height, anthropometric measures and cancer remain an intense area of research.

Anthropometric traits are largely determined by genes that control cellular proliferation, metabolism and apoptosis: these attributes are also required for immortalisation of cancer cells and development of tumours^{15, 16}. It is well established that germline mutations within tumour suppressor genes and oncogenes, affect cancer risk^{17, 18}; however, their influence on anthropometric traits is not well-known. From studying mouse models and rare diseases, there

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

are indications that mutations in such genes may influence body mass composition^{19, 20, 21, 22, 23}. For example, the tumour suppressor gene ARH1 inhibits cell growth. A deletion in ARH1 is commonly associated with breast and ovarian carcinoma while mice that are engineered to overexpress ARH1 are significantly smaller than wild type counterparts²⁴. Mice possessing knockouts of genes within the tumorigenic hypoxia inducible pathway vary in size and respond to high-fat diets differently from those without the mutations^{25, 26}. TP53 is the most frequently mutated gene in human cancers and is a key regulator of a number of cellular activities which prevent tumorigenesis, including maintaining genomic stability, controlling cell growth and metabolism^{19, 27, 28}. Mouse models of TP53 mutations have demonstrated that reduction of p53 activity can increase cancer risk, alter metabolism and influence obesity in a complex and signal dependent manner^{20, 21, 22, 23}.

Further evidence for the relation between body size, genetics and cancer comes human genetic disease, such as Turner syndrome. Turner Syndrome is defined by complete or partial chromosome X monosomy and is associated with a distinct clinical phenotype, including gonadal dysgenesis, cubitus valgus and short stature²⁹. Short stature within this condition has been attributed to haplo-insufficiency of the SHOX tumour suppressor gene³⁰. Deletions and mutations in SHOX have also been identified as an explanation for short stature within the general population³¹. Another example comes from patients with Simson–Golabi–Behmel syndrome (SGB), caused by mutations in the GPC3 tumour suppressor gene. Patients with SGB tend to have pre- and post-natal overgrowth and thus ultimately taller stature as well as an increased risk of embryonic tumours. In line with this, GPC3-null mice also display overgrowth³². Effects in both humans and mice are thought to be due to unshackling of the hedgehog pathways, and increasing cellular proliferation^{33, 34, 35}.

These rarer diseases provide evidence that major differences in the expression of cancer genes can influence both anthropometric traits and cancer risk in specific cohorts of patients. However, it is less clear if such associations occur in the broader population. The availability of large phenotypic and genetically linked datasets, has allowed unsupervised

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

approaches, such as genome-wide association studies (GWAS) of common single nucleotide polymorphisms (SNPs), to be deployed to further understand the link between genetics and anthropometric traits. Pleiotropy is the phenomenon whereby a single SNP or genetic mutation can influence multiple traits. Until now, it has been difficult to look for pleiotropy between functional SNPs in genes associated with cancer risk and anthropometric traits, as very few cohorts possess both comprehensive genetic data and detailed anthropometric data in the same population. In fact, the UK Biobank provides a unique opportunity to investigate such pleiotropic associations in a large prospective cohort of over 500,000 participants.

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

Results:

Through integration and curation of the GWAS catalog, eQTL databases and the Cancer Gene Census we identified 100 SNPs, which have been shown to associate with a differential risk of developing a total of 21 different cancer types and differential gene expression in at least one tissue type of (i) proto-oncogenes (8 genes) (ii) oncogenic fusion proteins (8), (iii) tumour suppressors (16) and (iv) 15 genes that span two or more of these groups (supplementary tables 1 and 2). We define these SNPs as *cancer eSNPs*. In our analyses of the UK Biobank cohort, we found 13 cancer eSNPs associated with differences in anthropometric traits and 31 with differential cancer risk. Interestingly, 7 of these SNPs overlapped and thus displayed some level of pleiotropy between cancer risk and anthropometric traits (Fig. 1, supplementary tables 3&4). These 7 cancer eSNPs are found on three different chromosomes and in linkage disequilibrium (see methods) with at least one other identified cancer eSNP (Fig. 2). They have been found to be associated with differential expression levels of (i) Fanconi Anemia, complementation group A, FANCA [rs1805007, rs258322], (ii) tumor suppressor p53, TP53 [rs78378222, rs35850753, rs8753], and Mitogen-Activated Protein Kinase 3, Kinase 1 [MAP3K1: rs889312, rs1862626]). Both cancer eSNPs associated with FANCA expression reside in neighbouring genes; rs258322 is a non-coding variant within CDK10 and rs1805007 is a missense variant within MC1R. eSNPs associated with MAP3K1 expression are found in an intergenic region close to ZNF296 and MAP3K1 (5q11.2). In contrast, the two cancer eSNPs associated with differential TP53 expression are found in untranslated regions of the TP53 gene itself. rs78378222 is found in the 3'-UTR of TP53 and rs35850753 is found in the 5'-UTR of the d133 isoform of TP53. The third TP53 eSNP is in the 3'-UTR of the neighbouring POLR2A gene. The cancer eSNPs that showed the strongest association with anthropometric measures and cancer risk within their loci (lead cancer eSNP) were rs1805007 (FANCA, C>T, minor allele frequency in UK Biobank ³⁶= 0.102), rs78378222 (TP53, T>G, MAF= 0.012) and rs889312 (MAP3K1, C>A, MAF= 0.284).

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

In the UK Biobank cohort, the minor allele of rs1805007 (FANCA) was associated with an increased risk of melanoma (OR=1.63 [1.52-1.75], adjusted-p=2.57E⁻⁴¹) and non-melanomatous skin cancer (OR=1.36 [1.31-1.41], adjusted-p=7.62E⁻⁶³). The minor allele of rs78378222 (TP53) was associated with an increased risk of brain malignancy (OR= 3.12 [2.22-4.37], adjusted p= 1.43E⁻¹²) and non-melanomatous skin cancer (OR= 1.46 [1.34-1.60], adjusted p= 5.20E⁻¹⁸). The minor-allele of rs889312 (MAP3K1) was associated with an increased risk of breast cancer (OR= 1.1 [1.07-1.13], adjusted p= 2.82E⁻¹¹). Of note, as we selected these SNPs due to their noted association with differential cancer susceptibility in GWAS studies, our results provide an independent validation of these associations (Fig. 3a, supplementary table 5).

As regards the anthropometric measures, we identified an unexpected high level of pleiotropy and very strong association between the TP53 cancer eSNPs and three measures of height and six lean body mass, as well as basal metabolic rate with p-values as low as 7.54E⁻³⁶ (Fig. 3b). After quality control (QC) procedures, we identified 59 participants homozygous for the minor allele (increased cancer risk) of the lead TP53 eSNP (rs78378222), 9,253 heterozygous for the minor allele and 370,046 homozygous for the major allele. The minor allele carriers were on average taller, leaner and have a higher basal metabolic rate. The associations of these anthropometric traits with the minor allele of the rs78378222 SNP were markedly more significant (up to almost 3 fold) and with greater effect sizes (up to almost 5 fold) than for the other pleiotropic eSNPs (Fig 3c, supplementary table 3).

The cancer eSNPs for TP53 demonstrated strikingly strong associations with 10 different anthropometric traits in the UK Biobank cohort. In contrast, the cancer eSNPs related to FANCA [rs1805007] and MAP3K1 [rs889312] were associated with standing height but not lean body mass measures. After QC, we identified 4271 participants homozygous for the minor allele of the lead FANCA eSNP, 68,956 heterozygous for the minor allele and 306,131 homozygous for the major allele. The minor allele for the lead MAP3K1 eSNP was found in 185,184 participants (30,394 homozygous, 154,790

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

heterozygous). The FANCA eSNP [rs1805007] only associated with standing height (beta= -0.02 ± 0.002, adjusted-p=9.20E⁻¹⁵), while the MAP3K1 eSNP [rs889312] associated with broader range of height measures (standing height beta= -0.02 ± 0.004, adjusted p= 6.11E⁻¹⁴ and sitting height beta= -0.01 ± 0.001, adjusted p= 3.34E⁻⁹). We sought validation for these associations with standing height in the independent GIANT consortium dataset. To do so, we examined at all lead and linked SNPs for each pleiotropic eSNP in the results of the meta-analysis conducted by Wood et al. (2014)³⁷. We were able to identify eSNPs for MAP3K1 (rs889312, rs1862626) and FANCA (rs1805007, rs258322), but none for TP53, due to the low minor allele frequency. Three of these eSNPs significantly associated with height in the Wood et al. meta-analysis: (i) rs258322 FANCA eQTL (p=1.5E-09, Beta=-0.029), (ii) rs889312 MAP3K1 eQTL (p=2.3E-08, Beta=-0.018) and (iii) rs1862626 MAP3K1 eQTL (p=3.4E-08, Beta=-0.018). Interestingly, both linked MAP3K1 eSNPs and a FANCA eSNP passed the significance threshold and, reassuringly, the directions of the allelic associations are consistent with our findings. The other FANCA eSNP only just fell short of the GWAS significance threshold (rs1805007, 4.3E-04, Beta = -0.024). Together, these data clearly link eSNPs for MAP3K1 and FANCA with height.

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

Discussion:

This is the first comprehensive study providing evidence that functional common genetic variants in oncogenes and tumour suppressor genes can associate with both anthropometric traits and cancer risk in the general European population and in the same cohort. SNPs exhibiting these pleiotropic associations in our study are found in three different loci: (i) MAP3K1 (two SNPs in linkage disequilibrium), (ii) FANCA (two SNPs in linkage disequilibrium) and (iii) TP53 (three SNPs in linkage disequilibrium). Observations gained from mouse models designed to alter signalling pathways involving *FANCA* and *TP53*, suggested that such associations with anthropometric traits in humans could be possible. For example, targeted disruption of exons of the Fanconi Anemia group A (*FANCA*) gene in mice results in altered anthropometric traits, including growth retardation as well as differential cancer risk³⁸.³⁹ Furthermore, mouse models of TP53 mutations have clearly demonstrated that reduction of p53 activity can result in increased cancer risk, altered metabolism and influence obesity in a complex and signal dependent manner^{20, 21, 22, 23}.

The strongest associations we observed with both anthropometric traits and cancer risk are loci related to TP53. p53 is key regulator of a number of cellular activities which prevent tumorigenesis including maintaining genomic stability, controlling cell growth and metabolism^{27, 28}. TP53 is the most frequently mutated gene in human cancers¹⁹. Moreover, in all families with similar TP53 mutations in their heritable genomes, a dramatic increase in cancer risk is observed (Li-Fraumeni Syndrome, LFS)^{17, 40}. In recent studies, it has been shown that LFS patients not only have an increased risk of developing cancer but also an increased capacity for oxidative phosphorylation⁴¹, providing a potential link with anthropometric traits and basal metabolic rate. Furthermore, the well-tolerated anti-diabetic drug metformin, which is thought to inhibit mitochondrial complex 1, increases cancer-free survival in a mouse model of LFS and reduces proliferation in cancer cell lines^{42, 43}. Metformin is now being trialled in LFS patients to hopefully provide a preventative option for these high-risk patients (ClinicalTrials.gov number: NCT01981525). Based on the potential link with

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

oxidative phosphorylation, this intervention might also be trialled in those carrying the minor allele of the TP53 cancer eSNP (rs78378222).

The TP53 mutations found in LFS are rare in the general population⁴⁴. However, here we show relatively frequent SNPs related to p53 affect cancer risk and anthropometric traits. Notably, the allele of a SNP in the polyadenylation signal of p53 (rs78378222[C]) which is found in approximately 1% of populations of European descent, has been shown to impair 3'-end processing of p53 mRNA, resulting in a reduction of p53 protein and an increased risk for glioma and basal cell carcinoma as well as affecting head circumference and intracranial volume^{45, 46}. Here we not only validate these cancer associations in a separate cohort (non-melanomatous skin cancer, OR= 1.46 [1.34-1.60], adjusted p= 5.20E⁻¹⁸, brain malignancy, OR= 3.12 [2.23-4.37], adjusted p= 1.43E⁻¹²) but also show that carriers of this allele tend to be taller, leaner and have a higher basal metabolic rate (standing height, adjusted-p= 2.18E⁻²⁴, beta= 0.073 ± 0.007, whole body fat free mass, adjusted-p= 8.34E⁻³⁷, beta= 0.073 ± 0.005, basal metabolic rate, adjusted-p= 1.13E⁻³¹, beta= 0.076 ± 0.006).

Prior to this study, the association between increased height and non-melanomatous skin cancer/brain malignancy were well established, however, mechanistic explanations were lacking^{2, 47, 48}. Our study proves the concept that functional loci in well-characterised tumour suppressors and oncogenes alter both cancer risk and anthropometric traits and, of particular interest, identifies a strong, new, association between the rs78378222[C] SNP in the polyadenylation site of p53 with both increased risk for developing non-melanomatous skin cancer/brain malignancy and increased height, lean body mass and basal metabolic rate, thereby offering a novel genetic link between these anthropometric traits and cancer risk.

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

Methods

The UK Biobank:

The UK Biobank is a cohort of ~500,000 UK residents who volunteered to have their clinical, lifestyle, anthropometric and genetic data collected for research. Data was collected in a number of ways, including face-to-face interview, touchscreen assessment and from centralised clinical registers (e.g. the Cancer register and Death register). Participants were aged between 40-69 years at recruitment⁴⁹. UK Biobank obtained informed consent from all participants and all protocols were approved by the National Research Ethics Service Committee. All baseline data used in this study was collected in 22 UK centres, between 2006-2010. This study was conducted under the UK Biobank approved application (#24456, PI Gareth Bond).

Anthropometric data:

During the baseline assessment, participants had various anthropometric traits measured directly or by bioelectrical impedance. Bioimpedance data was measured using a Tanita BC418MA body composition analyser. Participants stood barefoot on the analyser and held the metal handles. This device produced measurements of fat mass, fat-free mass and basal metabolic rate. Further information on anthropometric data collected is found here⁵⁰.

Cancer data:

Cancer occurrences were defined by presence of a cancer international classification of diseases (ICD) code in the UK Cancer register or the UK Death register. To maximise the number of individual cancers cases, we combined ICD9 and ICD10 codes of identical cancers. This was done by a clinician to ensure the matching was accurate, for instance malignant neoplasm of brain was defined as C71 (ICD10) and 191 (ICD9). With the goal of increasing

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

power, when appropriate we merged ICD codes into clinically relevant groups. We ran individual cancers (combined ICD9 and ICD10 codes) and clinically relevant groups in our analysis.

Genetic data:

Blood samples were collected when participants were recruited, and DNA extracted⁵¹. DNA was then genotyped on either the Affymetrix UK BiLEVE Axiom array or the Affymetrix UK Biobank Axiom array (Santa Clara, CA, USA). Imputation was based upon a merged reference panel of ~90 million biallelic variants, from the 1000 Genomes Phase 3⁵² and the UK10K⁵³ haplotype panels. Imputation was performed using IMPUTE2 as described⁵⁴, producing 488,295 genotyped participants.

Sample quality control:

In addition to the standard quality control, we carried out further quality control steps to ensure robustness of our analyses. We excluded individuals based on: a mismatched value between self-reported and genetic sex (data-field: 22001 and 31), level of genotype missingness of >0.05 (data-field: 22005), genetic relatedness factor with kinship coefficient of >0.0442, sex chromosome aneuploidy (data-field 22019), outliers for heterozygosity or missing rate (data-field: 22027). We selected the European population based on self-reported ethnicity (data-field: 21000) by excluding non-white ethnic background. This left a study population of 379,358 suitable genotyped individuals.

Identification of functional cancer gene SNPs:

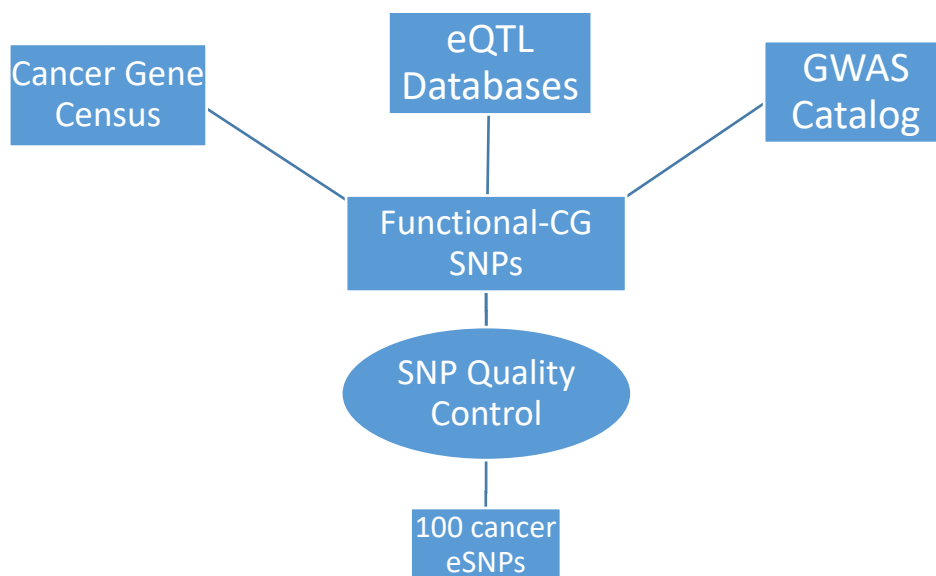
We identified genes confirmed to be involved in carcinogenesis in the COSMIC Cancer Gene Census (release v88, 19th March 2019): the current reference record of genes containing cancer driver mutations. We then selected all SNPs annotated to these genes that have been

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

significantly associated with differential cancer risk in genome-wide association studies in the GWAS catalog⁵⁵. Significant associations with cancer risk were defined with a p-values of $<5E^{-08}$ cut-off. Finally, we required that these SNPs were also associated with differential expression of the cancer gene in at least one expression quantitative trait loci (eQTL) database. cis-eQTL databases utilised were GTEX, NESDA/NTR and PancanQTL^{56, 57, 58}.

SNP quality control:

SNP exclusions were Hardy-Weinberg equilibrium with p-value less than $1E^{-10}$, a minor allele frequency less than 0.0001, level of missingness more than 0.05 or an imputation score less than 0.8 (as per <http://www.nealelab.is/uk-biobank/>).



Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

Analysis:

To carry out the SNP-wise analysis on all functional-CG SNPs we used SNPTTEST (version 2.5.4)⁵⁹, and employed the frequentist approach under dominant, additive and recessive inheritance models, using sex, age and genetic principal components (1-20 PCs) as covariates. Genetic PCs were included as covariates to control for hidden population structure. We controlled the genotype uncertainty by implementing the missing data likelihood score test. Within SNPTTEST, we used the frequentist approach under an additive inheritance model. P-values of SNP-wise association were adjusted by the stringent Bonferroni correction based on number of tested traits (28 anthropometric traits, 50 cancer types) multiplied by the number of eSNPs (100). Significant associations were defined Bonferroni correction p values below the threshold of $1E^{-5}$.

Linkage and lead SNPs

R^2 and D' coefficients were calculated to evaluate the degree of Linkage Disequilibrium (LD) between different loci using LDlink v3.7 ($D' > 0.8$, $R^2 > 0.4$ <https://ldlink.nci.nih.gov/>)⁶⁰. Lead SNPs were defined as the SNPs that were most strongly associated with the traits in question for each locus⁶¹.

Due to the low value of r^2 in some of the potential LD eSNPs we carried out a leave-one-out analysis:

- For each lead SNP:
 - All participants from the cohort carrying the lead SNP were removed.
 - The association for each of the potential LD eSNPs was performed (e.g. lead SNP rs783222 carriers were removed whilst rs35850753 and rs8753 were tested for their associations with cancer and anthropometric traits).
 - Adjusted p-values from the whole cohort association analysis with those obtained in this analysis were compared.

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

- If the association of the potential LD eSNPs was not significant after the removal of lead SNP carriers (p -values >0.05), we considered the eSNP in Linkage Disequilibrium.

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

○

References:

1. Albanes D, Jones DY, Schatzkin A, Micozzi MS, Taylor PR. Adult stature and risk of cancer. *Cancer Res* 48, 1658-1662 (1988).
2. Renehan AG. Height and cancer: consistent links, but mechanisms unclear. *Lancet Oncol* 12, 716-717 (2011).
3. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. *Lancet* 384, 755-765 (2014).
4. Park J, Morley TS, Kim M, Clegg DJ, Scherer PE. Obesity and cancer--mechanisms underlying tumour progression and recurrence. *Nat Rev Endocrinol* 10, 455-465 (2014).
5. Green J, Cairns BJ, Casabonne D, Wright FL, Reeves G, Beral V. Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies of height and total cancer risk. *Lancet Oncol* 12, 785-794 (2011).
6. Benyi E, Linder M, Adami J, Kieler H, Palme M, Savendahl L. Adult height is associated with risk of cancer and mortality in 5.5 million Swedish women and men. *Journal of epidemiology and community health*, (2019).
7. Clayton PE, Banerjee I, Murray PG, Renehan AG. Growth hormone, the insulin-like growth factor axis, insulin and cancer risk. *Nat Rev Endocrinol* 7, 11-24 (2011).
8. Magruder JT, Elahi D, Andersen DK. Diabetes and pancreatic cancer: chicken or egg? *Pancreas* 40, 339-351 (2011).
9. Nunney L. Size matters: height, cell number and a person's risk of cancer. *Proceedings Biological sciences* 285, (2018).
10. Peto R. Epidemiology, multistage models, and short-term mutagenicity tests (1977).
11. Arnold M, et al. Obesity and cancer: An update of the global impact. *Cancer epidemiol* 41, 8-15 (2016).
12. Kyrgiou M, et al. Adiposity and cancer at major anatomical sites: umbrella review of the literature. *BMJ* 356, j477 (2017).
13. Renehan AG, Zwahlen M, Egger M. Adiposity and cancer risk: new mechanistic insights from epidemiology. *Nat Rev Cancer* 15, 484-498 (2015).
14. Dalamaga M, Diakopoulos KN, Mantzoros CS. The role of adiponectin in cancer: a review of current evidence. *Endocr Rev* 33, 547-594 (2012).
15. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 100, 57-70 (2000).
16. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 144, 646-674 (2011).
17. Li FP, Fraumeni JF, Jr. Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? *Annals of internal medicine* 71, 747-752 (1969).

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

18. Swift M, Morrell D, Massey RB, Chase CL. Incidence of cancer in 161 families affected by ataxia-telangiectasia. *The New England journal of medicine* 325, 1831-1836 (1991).
19. Soussi T, Dehouche K, Beroud C. p53 website and analysis of p53 gene mutations in human cancer: forging a link between epidemiology and carcinogenesis. *Hum Mutat* 15, 105-113 (2000).
20. Liu S, Kim T-H, Franklin DA, Zhang Y. Protection against High-Fat-Diet-Induced Obesity in MDM2(C305F) Mice Due to Reduced p53 Activity and Enhanced Energy Expenditure. *Cell Rep* 18, 1005-1018 (2017).
21. Vergoni B, et al. DNA Damage and the Activation of the p53 Pathway Mediate Alterations in Metabolic and Secretory Functions of Adipocytes. *Diabetes* 65, 3062-3074 (2016).
22. Kung CP, et al. The P72R Polymorphism of p53 Predisposes to Obesity and Metabolic Dysfunction. *Cell Rep* 14, 2413-2425 (2016).
23. Labuschagne CF, Zani F, Vousden KH. Control of metabolism by p53 - Cancer and beyond. *Biochim Biophys Acta Rev Cancer* 1870, 32-42 (2018).
24. Bignon YJ, et al. Expression of a retinoblastoma transgene results in dwarf mice. *Genes & development* 7, 1654-1662 (1993).
25. Matsuura H, et al. Prolyl hydroxylase domain protein 2 plays a critical role in diet-induced obesity and glucose intolerance. *Circulation* 127, 2078-2087 (2013).
26. Zhang X, et al. Adipose tissue-specific inhibition of hypoxia-inducible factor 1{alpha} induces obesity and glucose intolerance by impeding energy expenditure in mice. *J Biol Chem* 285, 32869-32877 (2010).
27. Levine AJ. p53, the cellular gatekeeper for growth and division. *Cell* 88, 323-331 (1997).
28. Vogelstein B, Lane D, Levine AJ. Surfing the p53 network. *Nature* 408, 307-310 (2000).
29. Al Alwan I, et al. Turner Syndrome Genotype and phenotype and their effect on presenting features and timing of Diagnosis. *Int J Health Sci (Qassim)* 8, 195-202 (2014).
30. Rao E, et al. Pseudoautosomal deletions encompassing a novel homeobox gene cause growth failure in idiopathic short stature and Turner syndrome. *Nat Genet* 16, 54-63 (1997).
31. Stuppia L, et al. SHOX mutations detected by FISH and direct sequencing in patients with short stature. *J Med Genet* 40, E11 (2003).
32. Song HH, Shi W, Xiang YY, Filmus J. The loss of glypican-3 induces alterations in Wnt signaling. *J Biol Chem* 280, 2116-2125 (2005).
33. Filmus J, Capurro M. The role of glypican-3 in the regulation of body size and cancer. *Cell Cycle* 7, 2787-2790 (2008).

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

34. Pilia G, et al. Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome. *Nat Genet* 12, 241-247 (1996).
35. Neri G, Gurrieri F, Zanni G, Lin A. Clinical and molecular aspects of the Simpson-Golabi-Behmel syndrome. *American journal of medical genetics* 79, 279-283 (1998).
36. Polverosi R, Maffesanti M, Dalpiaz G. Organizing pneumonia: typical and atypical HRCT patterns. *La Radiologia medica* 111, 202-212 (2006).
37. Wood AR, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* 46, 1173-1186 (2014).
38. Wong JC, Alon N, McKerlie C, Huang JR, Meyn MS, Buchwald M. Targeted disruption of exons 1 to 6 of the Fanconi Anemia group A gene leads to growth retardation, strain-specific microphthalmia, meiotic defects and primordial germ cell hypoplasia. *Hum Mol Genet* 12, 2063-2076 (2003).
39. Parmar K, D'Andrea A, Niedernhofer LJ. Mouse models of Fanconi anemia. *Mutat Res* 668, 133-140 (2009).
40. Malkin D. Li-Fraumeni Syndrome. *Genes & Cancer* 2, 475-484 (2011).
41. Wang PY, et al. Increased oxidative metabolism in the Li-Fraumeni syndrome. *N Engl J Med* 368, 1027-1032 (2013).
42. Wang PY, et al. Inhibiting mitochondrial respiration prevents cancer in a mouse model of Li-Fraumeni syndrome. *J Clin Invest* 127, 132-136 (2017).
43. Wheaton WW, et al. Metformin inhibits mitochondrial complex I of cancer cells to reduce tumorigenesis. *Elife* 3, e02242 (2014).
44. Valdez JM, Nichols KE, Kesserwan C. Li-Fraumeni syndrome: a paradigm for the understanding of hereditary cancer predisposition. *Br J Haematol* 176, 539-552 (2017).
45. Stacey SN, et al. A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nat Genet* 43, 1098-1103 (2011).
46. Haworth S, et al. Low-frequency variation in TP53 has large effects on head circumference and intracranial volume. *Nat Commun* 10, 357 (2019).
47. Li X, et al. Height, height-related SNPs, and risk of non-melanoma skin cancer. *Br J Cancer* 116, 134-140 (2017).
48. Moore SC, et al. Height, body mass index, and physical activity in relation to glioma risk. *Cancer Res* 69, 8349-8355 (2009).
49. Bycroft C, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562, 203-209 (2018).
50. UK Biobank. Anthropometry. (2014).
51. UK Biobank. Genotyping of 500,000 UK Biobank participants: Description of sample processing workflow and preparation of DNA for genotyping. (2015).
52. Auton A, Salcedo, T. *Assessing Rare Variation in Complex Traits*. Springer (2015).

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

53. UK Biobank Consortium, et al. The UK10K project identifies rare variants in health and disease. *Nature* 526, 82-90 (2015).
54. UK Biobank. Genotype imputation and genetic association studies of UK Biobank. (2015).
55. Welter D, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic acids research* 42, D1001-D1006 (2014).
56. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 45, 580-585 (2013).
57. Gong J, et al. PancanQTL: systematic identification of cis-eQTLs and trans-eQTLs in 33 cancer types. *Nucleic Acids Res* 46, D971-D976 (2018).
58. Jansen R, et al. Conditional eQTL analysis reveals allelic heterogeneity of gene expression. *Human molecular genetics* 26, 1444-1451 (2017).
59. Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nature reviews Genetics* 11, 499-511 (2010).
60. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 31, 3555-3557 (2015).
61. Schaid DJ, Chen W, Larson NB. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nature reviews Genetics* 19, 491-504 (2018).

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

Fig. 1:

A circos plot indicating the functional SNPs that significantly associate with anthropometric traits and cancer risk. SNPs are located on the top half of the plot whilst cancer types and anthropometric traits are on the bottom half. Significant associations (Bonferroni corrected p-values $<1E^{-5}$) are shown as solid lines. Blue lines indicate significant associations with anthropometric traits. Red lines depict significant associations with differential cancer risk.

Fig. 2:

A gene map of the chromosomal regions containing the pleiotropic functional SNPs. Lead SNPs are highlighted in red. Tables summarise salient GWAS catalog, eQTL, and Cancer Gene Census information for each variant.

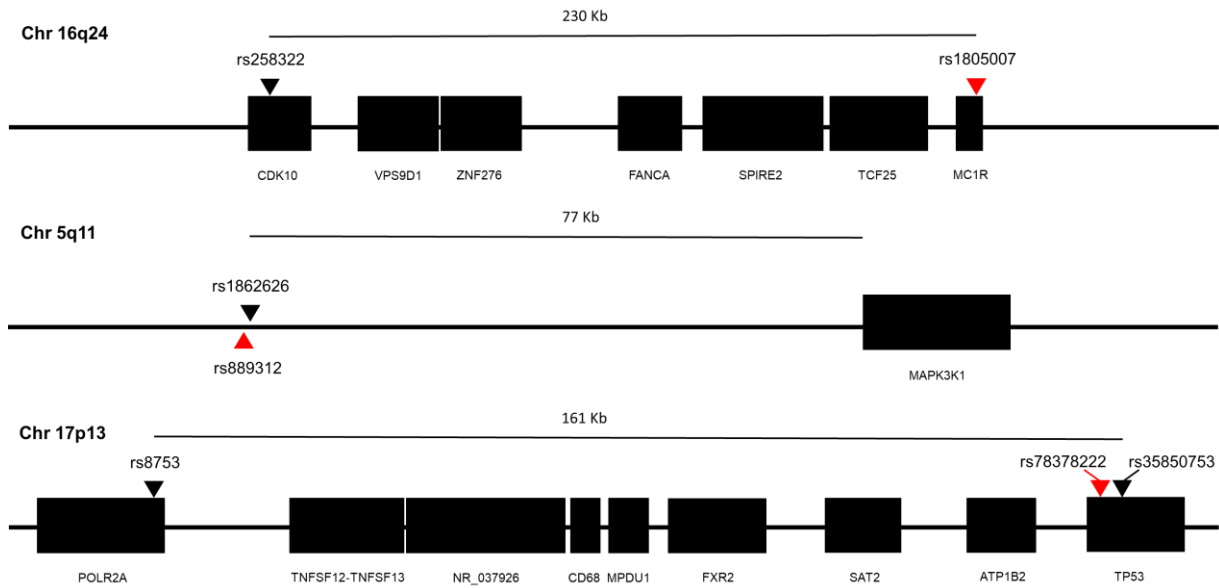
Fig. 3:

Detailed association results for cancer eSNPs with anthropometric traits and cancer risk in the UK Biobank cohort. A heatmap (a) depicting the significance of the associations of the three lead pleiotropic SNPs with differential cancer risk. The colour scale represents the $-\log_{10}$ adjusted-p: with the darkest red on the scale being $7.62E^{-63}$ and the lightest yellow being $1E^{-5}$. A radar plot (b) illustrating the level of significance of the associations of the three lead pleiotropic SNPs with anthropometric traits (radial axis: $-\log_{10}$ adjusted p). The darkest two pink categories are bio-electrical impedance measures (darkest being measures of fat). Traditional bedside anthropometric measures are in blush. SNPs related to FANCA (rs1805007[†]) and MAP3K1 (rs889312[†]) associate with standing height. The lead functional SNP related to TP53, rs78378222[†], markedly associates with multiple measures of height and lean body mass, as well as basal metabolic rate. An error bar plot (c) of the beta coefficients (y-axis) of the significant associations with anthropometric traits (x-axis). Error bars denote the 95% confidence interval of the beta value.

[†] denotes lead cancer eSNP

Heritable genetic variants in key cancer genes link cancer risk with anthropometric traits

Fig. 2



rs1805007[†] / rs258322

GWAS catalog traits	Skin (Melanoma, Non-melanomatous skin cancer)
eQTL Database	PancanQTL
eQTL gene	FANCA
Germline Cancer Mutations	AML, leukaemia
Somatic Cancer Mutations	-

rs889312[†] / rs1862626

GWAS catalog traits	Breast cancer
eQTL Database	NESDA.NTR
eQTL gene	MAP3K1
Germline Cancer Mutations	-
Somatic Cancer Mutations	Luminal A breast

rs78378222[†] / rs35850753 / rs8753

GWAS catalog traits	Brain (Glioma), Skin (Non-melanomatous skin cancer)
eQTL Database	NESDA.NTR
eQTL gene	TP53
Germline Cancer Mutations	Breast, sarcoma, adrenocortical carcinoma, glioma, multiple other tumour types
Somatic Cancer Mutations	Breast, colorectal, lung, sarcoma, adrenocortical, glioma, Spitzoid tumour, multiple other tumour types

Fig. 3

