Mendelian randomisation analysis provides no evidence for a relationship between adult height and testicular cancer risk

Max Levy[1*], Dominic Hall[1*], Amit Sud[1], Philip Law[1], Kevin Litchfield[1], Darshna Dudakia[1], Trine B. Haugen[2], Robert Karlsson[3], Alison Reid[4], Robert A Huddart[4,5], Tom Grotmol[6], Fredrik Wiklund[3], Richard S Houlston[1], Clare Turnbull[1,7,8]

1. Division of Genetics & Epidemiology, The Institute of Cancer Research, London, SM2 5NG, UK
2. Faculty of Health Sciences, Oslo and Akershus University College of Applied Sciences, N-0130, Oslo, Norway
3. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, 171 77, Sweden
4. Academic Radiotherapy Unit, Institute of Cancer Research, Sutton, Surrey, SM2 5NG, UK
5. Academic Uro-oncology Unit, The Royal Marsden NHS Foundation Trust, Sutton, Surrey, SM2 5PT, UK
6. Department of Research, Cancer Registry of Norway, Oslo, 0369, Norway
7. William Harvey Research Institute, Queen Mary University, London, EC1M 6BQ, UK
8. Department of Clinical Genetics, Guy's and St Thomas' NHS Trust, London, United Kingdom, SE1 9RS

*Denotes equal author contribution.

Correspondence to: Clare Turnbull, Division of Genetics and Epidemiology, The Institute of Cancer Research, London, SM2 5NG, UK; Tel: ++44 (0) 208 722 4485; E-mail: clare.turnbull@icr.ac.uk

Key words: Testicular Cancer, Germ Cell Tumour, GWAS, Growth, Height.

ABSTRACT

Observational studies have suggested anthropometric traits, particularly increased height are associated with an elevated risk of testicular cancer (testicular germ cell tumour, TGCT). However, there is an inconsistency between study findings, suggesting the possibility of the influence of confounding factors. To examine the association between anthropometric traits and TGCT using an unbiased approach, we performed a Mendelian randomisation (MR) study. We used genotype data from genome wide association studies (GWAS) of TGCT totalling 5,518 cases and 19,055 controls. Externally weighted polygenic risk scores were created and used to evaluate associations with TGCT risk per one standard deviation (s.d) increase in genetically-defined adult height, adult BMI, adult waist hip ratio adjusted for BMI (WHRadjBMI), adult hip circumference adjusted for BMI (HIPadjBMI), adult waist circumference adjusted for BMI (WCadjBMI), birth weight (BW) and childhood obesity. MR analysis did not demonstrate an association between any anthropometric trait and TGCT risk. In particular, despite good power, there was no global evidence for association between height and TGCT. However, three SNPs for adult height individually showed association with TGCT (rs4624820: OR=1.47, 95% CI: 1.41-1.55, $P=2.7\times10^{-57}$; rs12228415: OR=1.17, 95% CI: 1.11-1.22, $P=3.1\times10^{-10}$; rs7568069: OR=1.13, 95% CI: 1.07-1.18, $P=1.1\times10^{-6}$). This MR analysis, based on the largest TGCT GWAS dataset to date, does not support a causal etiological association between anthropometric traits and TGCT aetiology. Our findings are more compatible with confounding by shared environmental factors, possibly related to prenatal growth with exposure to these risk factors occurring *in utero*.

INTRODUCTION

Testicular cancer (testicular germ cell tumour, TGCT) has doubled in incidence over the last 30 years in Western Countries and has an unusual incidence rate in comparison to many other neoplasms with rates peaking at ~30 years and declining rapidly thereafter (Bray *et al.* 2006; Le Cornet *et al.* 2014; Ruf *et al.* 2014). Tumours with this kind of age-incidence profile are likely to have onset early in life, and previous studies have therefore focussed on perinatal and adolescent stages of development as well as adult anthropometric traits (Akre *et al.* 2000; Bjørge *et al.* 2006; Dieckmann & Pichlmeier, 2002; Dieckmann *et al.* 2009; Garner *et al.* 2003; McGlynn *et al.* 2007; Rasmussen *et al.* 2003; Richiardi *et al.* 2003). Although there is disagreement as to the effect of adult body mass index (BMI) on TGCT risk many, but not all, studies have reported increased

height as a risk factor for the disease (Akre *et al.* 2000; Bjørge *et al.* 2006; Dieckmann & Pichlmeier, 2002; McGlynn *et al.* 2007; Rasmussen *et al.* 2003; Richiardi *et al.* 2003). However, none of these studies have established causal pathways linking traits such as height with TGCT, with some studies suggesting that confounding environmental factors such as childhood nutrition could play a greater role.

An alternative approach to a traditional epidemiology observational study is the Mendelian randomisation (MR) approach. The strategy of MR uses as instrumental variables (IV) genetic markers known to be associated with a factor in the assessment of its effect on another trait (Palmer *et al.* 2012). This methodology permits the nature of the relation between the trait and the factor to be assessed and, importantly, establishes whether an association is causal. A further attribute of MR is the avoidance of the influence of factors whose effect may be time sensitive, for example temporal trends in calorific intake. Therefore, a genetically defined IV has the potential to more accurately assess lifetime exposure when compared to measurements of potential risk factors recorded in an observational study.

Recent genome-wide association studies (GWAS) have identified single nucleotide polymorphisms (SNPs) at multiple loci that are robustly associated with height and other anthropometric traits (Locke *et al.* 2015; Wood *et al.* 2014). Given the unexplained recent rapid increase in TGCT incidence, the paucity of established risk factors for TGCT and the mixed observational epidemiology on the association of anthropometric traits with TCGT risk, we performed a MR study by analysing genotype data from three GWAS conducted on TGCT (Kristiansen *et al.* 2015; Turnbull *et al.* 2010). The traits we considered were adult height, adult BMI, adult waist-hip ratio adjusted for BMI (WHRadjBMI), adult waist circumference adjusted for BMI (WCadjBMI), adult hip circumference adjusted for BMI (HIPadjBMI), childhood obesity and birth weight (BW).

MATERIALS AND METHODS

Testicular cancer GWAS data sources

Our MR analysis was based on data from three previously reported GWAS of TGCT totalling 5,518 cases and 19,055 controls. Briefly, these GWAS were based on individuals with European ancestry and comprised: UK GWAS1 (985 cases, 4,946 controls), UK Oncoarray (3,206 cases, 7,422 controls), and Scandinavian GWAS1 (1,327 cases, 6,687 controls). Details of the genotyping, quality control and imputation of untyped SNPs have been previously published (Kristiansen *et al.* 2015; Turnbull *et al.* 2010). Briefly, we excluded SNPs with a minor allele frequency of <1%, a call rate <95%, those SNPs violating Hardy-Weinberg equilibrium, and individuals with non-European ancestry as assessed using data from HapMap v2 (International HapMap *et al.* 2007). IMPUTEv2 software (Howie *et al.* 2012) was used to recover untyped SNP genotypes using a merged reference panel consisting of UK10K and 1000 Genomes Project data. Poorly imputed SNPs, defined by an INFO score of <0.9, were excluded. Summary statistics from these GWAS were used to calculate the ratio estimates for the anthropometric-related SNPs.

Instrumental variables

For each anthropometric trait we used data from recent GWAS of individuals of European-descent. Specifically for height, we used data from the Genetic Investigation of Anthropometric Traits (GIANT) consortium, which was based on an analysis of up to 253,288 individuals (Wood *et al.* 2014). For adult BMI, data was obtained from a consortium meta-analysis comprising 339,224 individuals from 125 studies (Locke *et al.* 2015). Data for WHRadjBMI, WCadjBMI and HIPadjBMI were obtained from related consortium meta-analyses comprising up to 224,459 individuals (Shungin *et al.* 2015). Data reported by the Early Growth Genetics (EGG) consortium of a meta-analysis of 5,530 cases and 8,318 controls was used for childhood obesity (Bradfield *et al.* 2012). Birth weight data was obtained from a meta-analysis of up to 69,308 individuals from 43 studies (Horikoshi *et al.* 2013).

To calculate the genetic scores we considered only SNPs that were genome-wide significant ($P < 5\times10^{-8}$) for each anthropometric trait and were present in all three GWAS TCGT datasets. To avoid co-linearity between SNPs we imposed a threshold R^2 value of ≥0.8 for linkage disequilibrium (LD) including only the SNPs with the strongest effect on the trait in the genetic risk score. After imposing these metrics we analysed 379 SNPs for adult height, 69 SNPs for adult BMI, 44 SNPs for

WHRadjBMI, 25 SNPs for WCadjBMI and 10 SNPs for HIPadjBMI, seven SNPs for birth weight and seven SNPs for childhood obesity.

We also examined male-specific risk-associated SNPs. However, because of a paucity of data, we relaxed the significance threshold and considered all SNPs at $P < 1.0\times10^{-7}$. Imposing this criteria provided seven male-specific SNPs for WHRadjBMI, four for HIPadjBMI and nine for WCadjBMI (Randall *et al.* 2013). Fifteen SNPs displaying significant male dimorphism for adult height were also considered (Randall *et al.* 2013). For all SNPs used, we recovered the chromosome position, risk alleles, association estimates and standard errors. None of the GWAS that comprised our dataset had reported non-additive effects of the SNPs, hence per allele effects were considered additive. For each SNP, we extracted the respective effect estimates and *P*-values from the TGCT GWAS.

Statistical Analysis

A central assumption of MR analysis is that there is no pleiotropism between the genetic instruments influencing TGCT and the trait being tested. Prior to performing MR, we used linkage disequilibrium regression (LDR) to test for global pleiotropy and used MR-Egger and inverse-variance weighted (IVW) regression to examine standard IV assumptions (Bowden *et al.* 2015). LDR was performed for TGCT against each trait to estimate genetic covariance with TGCT (Bulik-Sullivan *et al.* 2015).

We performed MR analysis to assess the association between each anthropometric trait and TGCT using summary statistics (Burgess *et al.* 2015). Using this method, the combined ratio estimate $(\hat{\beta})$ of all SNPs associated with the particular risk factor on TGCT was calculated under a fixed effects model:

$$\hat{\beta} = \frac{\sum_k X_k Y_k \sigma_{Y_k}^{-2}}{\sum_k X_k^2 \sigma_{Y_k}^{-2}}$$

where X_k represents the anthropometric trait effect size for SNP *k*, Y_k represents the TGCT risk effect size for SNP *k* and σ_{Y_k} denotes the standard error in the value of Y_k. The standard error of the TGCT risk effect size, denoted by $se(\hat{\beta})$, is given by:

$$se(\hat{\beta}) = \sqrt{\frac{1}{\sum_k X_k \sigma_{Y_k}^{-2}}}$$

With the statistics generated by following these calculations on the three different cohorts in the TGCT data, we performed a meta-analysis under a fixed-effects model to derive the final odds ratios and confidence intervals.

For each statistical test we considered a global nominal significance level of $P \leq 0.05$ as satisfactory to derive conclusions. A Bonferroni correction was applied to significance levels to compensate for multiple testing. All analyses were undertaken using R software (Version 2.14.1).

To gain biological insight into individual SNPs associated with both TGCT and anthropometric traits, we investigated the effects of significantly associated individual SNPs on gene expression using the GTEx resource, examining association with expression of all genes within 1MB of the associated SNP (Lonsdale *et al.* 2013).

Finally, we estimated the power of our MR analysis to demonstrate an association between TGCT and anthropometric traits tested over a range of effect sizes (Brion *et al.* 2013).

RESULTS

Using LDR we found no evidence for global pleiotropism between TGCT and any of the anthropometric traits after correction for multiple testing (Supplementary Table 1). Of the seven tested anthropometric traits and four male-specific components of traits, the strongest correlation was found between TGCT and childhood obesity (Genetic Correlation (GC) = 0.14, 95% confidence interval (CI) -0.26, -0.012, *P* = 0.032). MR-Egger and IVW regression tests provided no evidence of infringement of the standard IV assumptions (Supplementary Fig. 1 - 11, Supplementary Tables 2 and 3). All genetic variants used were therefore considered valid IVs and the following MR analysis valid.

MR analysis found no evidence for a genetic relationship between any of the seven anthropometric traits and TGCT risk (Table 1). To explore the effect of male dimorphism we additionally performed MR using IVs generated from male-specific SNPs for four of the anthropometric traits. This restricted analysis found no relationship between any IV and TGCT risk (Table 1).

Although in the LDR analysis we found no evidence for a global pleiotropism between height and TGCT, three of the IV SNPs for height showed a strong association with TGCT risk, even after applying a Bonferroni correction for testing 379 SNPs (rs4624820 (5q31.3): OR 1.47, 95% CI 1.41-1.55, $P=2.7\times10^{-57}$; rs12228415 (12p13.1): OR 1.17, 95% CI 1.11-1.22, $P=3.1\times10^{-10}$; and rs7568069 (2p13.3): OR 1.13, 95% CI 1.07-1.18, $P=1.1\times10^{-6}$). Only rs7568069 showed a significant eQTL effect in testicular tissue (for *DYSF, P* = 5.8×10^{-6}). No other trait tested had any SNP associated with TGCT risk.

Analyses of the power afforded by these MR analyses for each anthropometric trait analysed are shown in Supplementary Figs. 12-15. Since the IV SNPs for adult height explain ~20% of variance of the trait (Wood *et al.* 2014), our study has 90% power to detect an OR≥1.1. In contrast, the SNPs used as IVs for adult BMI only account for 3% of the variation (Locke *et al.* 2015), and hence we had <30% power to detect an OR≥1.1. Similarly, we had <20% power to demonstrate associations for WHRadjBMI and birth weight for effect sizes of OR≥1.1.

DISCUSSION

Following up upon associations reported in observational epidemiological studies between anthropometric traits and TGCT, our MR analysis using a TGCT GWAS dataset did not find an association between any anthropometric trait tested and risk. Although based on large datasets, aside from adult height, the SNP sets used as IVs only account for a modest proportion of the genetic variance underlying the anthropometric traits analysed. Hence we cannot exclude the possibility that null findings were simply a consequence of limited study power.

The primary motivation for this study was to further explore the epidemiological evidence for a positive relationship between adult height and TGCT, which has been reproduced across a number of studies (Akre *et al.* 2000; Bjørge *et al.* 2006; Dieckmann & Pichlmeier, 2002; McGlynn *et al.* 2007; Rasmussen *et al.* 2003; Richiardi *et al.* 2003). Despite 90% power to detect an OR>1.1, in our MR analysis we did not find global genetic evidence supporting a role of height in TGCT aetiology. These results suggest that the observed association between height and TGCT, if true, may be due to the influence of confounding due to shared environmental factors (Dieckmann *et al.* 2008).

Although we found no evidence of global pleiotropism between height and TGCT through LDR, our results, however, do not discount pleiotropic influences of individual SNPs on both adult height and TGCT development. Interestingly, three SNPs associated with increased height also showed a strong association with TGCT risk. The strongest of these SNPs, rs4624820 (5q31.3), is within the previously well-established TGCT-associated locus containing *SPRY4* (Kanetsky *et al.* 2009; Rapley *et al.* 2009; Turnbull *et al.* 2010). Another of these SNPs, rs7568069 (2p13.3), showed a significant eQTL effect in testicular tissue of *DYSF* (Dystrophy-Associated Fer-1-Like 1), which has been implicated in both muscular growth and developmental processes including sperm activation during fertilization and early embryogenesis (Smith & Wakimoto, 2007). While these results are consistent with overlapping pleiotropic pathways for these variants, these three SNPs are among 379 instrumental variables for height and contribute very modestly individually or in combination to each trait (Cole, 2000; Le Cornet *et al.* 2014).

A number of studies have suggested that BMI plays a role in TGCT risk (Akre *et al.* 2000; Dieckmann *et al.* 2009; Garner *et al.* 2003), although others have found no significant positive or negative association (Bjørge *et al.* 2006; Dieckmann & Pichlmeier, 2002; Rasmussen *et al.* 2003). Physical activity was found in one study to have a preventative effect on TGCT risk (Gallagher *et al.* 1995). There is some evidence for a negative association of BMI from a dataset of 500,000 Norwegian men (Akre *et al.* 2000); examining our Scandinavian cohort alone, waist circumference

adjusted for BMI displayed a nominally significant (P = 0.04) preventative association with TGCT risk but there were no other associations of note. Globally we did not find genetic evidence for a role in TGCT of BMI, waist circumference, or any other related traits. However, we acknowledge our power to detect an association of BMI and related anthropometric traits with TGCT using MR was limited. Accordingly, our null results therefore do not further clarify whether the reported association between BMI and related anthropometric traits with TGCT are true or not, and if true, whether causal or due to the confounding influence of common environmental factors. Similarly, MR analyses for association of both birth weight and childhood obesity on TGCT risk were subject to low power. To be more informative, further MR analysis of these traits would require significantly larger SNP sets capturing a higher proportion of phenotypic variation.

In summary, in our MR analysis we found no evidence for influence of adult height, BMI or other anthropometric traits in TGCT aetiology. Our MR analysis of height with TGCT was well-powered, and our null result suggests that the previously reported positive association between height and TGCT, if a true association, is unlikely to be driven by a common aetiology but instead is more likely to reflect confounding. Several potential mechanisms underlying the association between height and TGCT have previously been postulated, such as individual variation in the insulin-like growth factor I system (Zavos *et al.* 2004), age at puberty (Akre *et al.* 2000), and childhood nutrition (Dieckmann *et al.* 2008). Molecular and epidemiological data support a model whereby initial neoplastic transformation into the TGCT precursor lesion arises during the foetal period with the pre-invasive TGCT lesion lying dormant through childhood into adolescence and malignant growth finally occurring in early adulthood (Jacobsen & Henriques, 1992; Kristensen *et al.* 2008; Oosterhuis & Looijenga, 2005; Rajpert-De Meyts, 2006; Skakkebaek *et al.* 1987). This model of prenatal determination of TGCT would therefore be consistent with proposed exposure to an environmental factor common to adult height and TGCT risk occurring *in utero*. It is therefore plausible that foetal nutrition or another maternal or *in utero* exposure is positively associated with each of adult height and TGCT, thus potentially driving the observed association between these traits. Although much epidemiological study has examined association of maternal and *in utero* exposures with TGCT, to date no compelling evidence that implicates any specific pre-natal risk factors in the aetiology of TGCT has emerged from these studies (McGlynn & Cook, 2009).

In conclusion, although the genomic basis of TGCTs is becoming increasingly well understood (Litchfield *et al.* 2016), the environmental factors underlying the rapid increase in incidence of TGCT over the last four decades remain unclear. From the MR analyses we have presented, it is likely that if any of the associations of TGCT with anthropometric variables are true, they are a

product of confounding influence from early-life environmental exposures which are yet to be elucidated.

Table 1 Meta-analysis results of Mendelian Randomisation (MR) for anthropometric phenotypes against TGCT risk. MR was performed for each phenotype against three cohorts.

Phenotype	OR (95% CI)	*P-value*	I^2	P_{het}
Adult Height	1.066 (0.968 – 1.173)	0.197	25.70	0.260
Adult Height (Male only)	0.922 (0.600 – 1.416)	0.710	62.89	0.068
Adult BMI	0.848 (0.658 – 1.093)	0.203	9.79	0.330
WHRadjBMI	0.956 (0.712 – 1.284)	0.767	0	0.649
WHRadjBMI (Male only)	1.388 (0.755 – 2.551)	0.292	0	0.579
WCadjBMI	1.163 (0.787 – 1.719)	0.449	58.83	0.088
WCadjBMI (Male only)	1.245 (0.720 – 2.150)	0.433	27.46	0.252
HIPadjBMI	1.664 (0.936 – 2.957)	0.083	53.83	0.115
HIPadjBMI (Male only)	0.996 (0.498 – 1.994)	0.991	0	0.507
Childhood Obesity	0.863 (0.673 – 1.107)	0.246	0	0.517
Birth Weight	1.149 (0.783 – 1.687)	0.478	0	0.599

ACKNOWLEDGEMENTS

We thank the patients and all clinicians forming part of the UK Testicular Cancer Collaboration (UKTCC) for their participation in the studies from which the TGCT GWAS datasets were generated. Dominic Hall was supported by Wellcome Trust. K. Litchfield and A. Sud are supported by a PhD fellowship from Cancer Research UK. R.S.H. is supported by Cancer Research UK (C1298/A8362 Bobby Moore Fund for Cancer Research UK). We acknowledge National Health Service funding to the National Institute for Health Research Biomedical Research Centre. This study was supported by the Movember foundation and the Institute of Cancer Research.

DISCLOSURE

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

C.T. and R.S.H. designed the study. C.T and R.H lead patient recruitment for UK GWAS1 and UK Oncoarray, for which D.D. coordinated assembly and tracking of samples. T.B.H., R.K., T.G. and F.W. supplied GWAS data. D.H. conducted bioinformatic and statistical analyses with assistance from A.S., P.L. and K.L.. M.L. and D.H. drafted the manuscript with assistance from C.T. and R.S.H. All authors reviewed and contributed to the manuscript.

REFERENCES

Akre O, Ekbom A, Sparén P & Tretli S. (2000) Body Size and Testicular Cancer. *Journal of the National Cancer Institute* **92**, 1093-1096.

Bjørge T, Tretli S, Lie A K & Engeland A. (2006) The impact of height and body mass index on the risk of testicular cancer in 600,000 Norwegian men. *Cancer Causes & Control* **17**, 983-987.

Bowden J, Davey Smith G & Burgess S. (2015) Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *International journal of epidemiology* **44**, 512-525.

Bradfield J P, Taal H R, Timpson N J, Scherag A, Lecoeur C, Warrington N M, *et al.* (2012) A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat Genet* **44**, 526-531.

Bray F, Ferlay J, Devesa S S, McGlynn K A & Moller H. (2006) Interpreting the international trends in testicular seminoma and nonseminoma incidence. *Nature clinical practice. Urology* **3**, 532-543.

Brion M-J A, Shakhbazov K & Visscher P M. (2013) Calculating statistical power in Mendelian randomization studies. *International Journal of Epidemiology* **42**, 1497-1501.

Bulik-Sullivan B, Finucane H K, Anttila V, Gusev A, Day F R, Loh P-R, ReproGen Consortium, Psychiatric Genomics Consortium, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3, Duncan L, Perry J R B, Patterson N, Robinson E B, Daly M J, Price A L & Neale B M. (2015) An atlas of genetic correlations across human diseases and traits. *Nature genetics* **47**, 1236-1241.

Burgess S, Scott R A, Timpson N J, Davey Smith G, Thompson S G & Consortium E-I. (2015) Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *European journal of epidemiology* **30**, 543-552.

Cole T J. (2000) Secular trends in growth. *The Proceedings of the Nutrition Society* **59**, 317-324.

Dieckmann K-P & Pichlmeier U. (2002) Is Risk of Testicular Cancer Related to Body Size? *European Urology* **42**, 564-569.

Dieckmann K P, Hartmann J T, Classen J, Diederichs M & Pichlmeier U. (2009) Is increased body mass index associated with the incidence of testicular germ cell cancer? *Journal of cancer research and clinical oncology* **135**, 731-738.

Dieckmann K P, Hartmann J T, Classen J, Ludde R, Diederichs M & Pichlmeier U. (2008) Tallness is associated with risk of testicular cancer: evidence for the nutrition hypothesis. *British journal of cancer* **99**, 1517-1521.

Gallagher R P, Huchcroft S, Phillips N, Hill G B, Coldman A J, Coppin C & Lee T. (1995) Physical activity, medical history, and risk of testicular cancer (Alberta and British Columbia, Canada). *Cancer Causes & Control* **6**, 398-406.

Garner M J, Birkett N J, Johnson K C, Shatenstein B, Ghadirian P, Krewski D & Canadian Cancer Registries Epidemiology Research Group (2003) Dietary risk factors for testicular carcinoma. *International journal of cancer* **106**, 934-941.

Horikoshi M, Yaghootkar H, Mook-Kanamori D O, Sovio U, Taal H R, Hennig B J, *et al.* (2013) New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nat Genet* **45**, 76-82.

Howie B, Fuchsberger C, Stephens M, Marchini J & Abecasis G R. (2012) Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature genetics* **44**, 955-959.

International HapMap Consortium, Frazer K A, Ballinger D G, Cox D R, Hinds D A, Stuve L L, *et al.* (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851-861.

Jacobsen G K & Henriques U V. (1992) A fetal testis with intratubular germ cell neoplasia (ITGCN). *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **5**, 547-549.

Kanetsky P A, Mitra N, Vardhanabhuti S, Li M, Vaughn D J, Letrero R, Ciosek S L, Doody D R, Smith L M, Weaver J, Albano A, Chen C, Starr J R, Rader D J, Godwin A K, Reilly M P, Hakonarson H, Schwartz S M & Nathanson K L. (2009) Common variation in KITLG and at 5q31.3 predisposes to testicular germ cell cancer. *Nature genetics* **41**, 811-815.

Kristensen D M, Sonne S B, Ottesen A M, Perrett R M, Nielsen J E, Almstrup K, Skakkebaek N E, Leffers H & Rajpert-De Meyts E. (2008) Origin of pluripotent germ cell tumours: the role of microenvironment during embryonic development. *Molecular and cellular endocrinology* **288**, 111-118.

Kristiansen W, Karlsson R, Rounge T B, Whitington T, Andreassen B K, Magnusson P K, Fossa S D, Adami H O, Turnbull C, Haugen T B, Grotmol T & Wiklund F. (2015) Two new loci and gene sets related to sex determination and cancer progression are associated with susceptibility to testicular germ cell tumor. *Human molecular genetics* **24**, 4138-4146.

Le Cornet C, Lortet-Tieulent J, Forman D, Beranger R, Flechon A, Fervers B, Schuz J & Bray F. (2014) Testicular cancer incidence to rise by 25% by 2025 in Europe? Model-based predictions in 40 countries using population-based registry data. *European journal of cancer* **50**, 831-839.

Litchfield K, Levy M, Huddart R A, Shipley J & Turnbull C. (2016) The genomic landscape of testicular germ cell tumours: from susceptibility to treatment. *Nature reviews. Urology* **13**, 409-419.

Locke A E, Kahali B, Berndt S I, Justice A E, Pers T H, Day F R, *et al.* (2015) Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197-206.

Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, *et al.* (2013) The Genotype-Tissue Expression (GTEx) project. *Nature genetics* **45**, 580-585.

McGlynn K A & Cook M B. (2009) Etiologic factors in testicular germ-cell tumors. *Future oncology* **5**, 1389-1402.

McGlynn K A, Sakoda L C, Rubertone M V, Sesterhenn I A, Lyu C, Graubard B I & Erickson R L. (2007) Body Size, Dairy Consumption, Puberty, and Risk of Testicular Germ Cell Tumors. *American Journal of Epidemiology* **165**, 355-363.

Oosterhuis J W & Looijenga L H. (2005) Testicular germ-cell tumours in a broader perspective. *Nature reviews. Cancer* **5**, 210-222.

Palmer T M, Lawlor D A, Harbord R M, Sheehan N A, Tobias J H, Timpson N J, Davey Smith G & Sterne J A. (2012) Using multiple genetic variants as instrumental variables for modifiable risk factors. *Statistical methods in medical research* **21**, 223-242.

Rajpert-De Meyts E. (2006) Developmental model for the pathogenesis of testicular carcinoma in situ: genetic and environmental aspects. *Human reproduction update* **12**, 303-323.

Randall J C, Winkler T W, Kutalik Z, Berndt S I, Jackson A U, Monda K L, *et al.* (2013) Sex-stratified Genome-wide Association Studies Including 270,000 Individuals Show Sexual Dimorphism in Genetic Loci for Anthropometric Traits. *PLoS Genet* **9**, e1003500.

Rapley E A, Turnbull C, Al Olama A A, Dermitzakis E T, Linger R, Huddart R A, Renwick A, Hughes D, Hines S, Seal S, Morrison J, Nsengimana J, Deloukas P, UK Testicular Cancer Collaboration, Rahman N, Bishop D T, Easton D F & Stratton M R. (2009) A genome-wide association study of testicular germ cell tumor. *Nature genetics* **41**, 807-810.

Rasmussen F, Gunnell D, Ekbom A, Hallqvist J & Tynelius P. (2003) Birth weight, adult height, and testicular cancer: cohort study of 337,249 Swedish young men. *Cancer Causes & Control* **14**, 595-598.

Richiardi L, Askling J, Granath F & Akre O. (2003) Body Size at Birth and Adulthood and the Risk for Germ-cell Testicular Cancer. *Cancer Epidemiology Biomarkers & Prevention* **12**, 669-673.

Ruf C G, Isbarn H, Wagner W, Fisch M, Matthies C & Dieckmann K P. (2014) Changes in epidemiologic features of testicular germ cell cancer: age at diagnosis and relative frequency of seminoma are constantly and significantly increasing. *Urologic oncology* **32**, 33 e31-36.

Shungin D, Winkler T W, Croteau-Chonka D C, Ferreira T, Locke A E, Magi R, Strawbridge R J, *et al.* (2015) New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187-196.

Skakkebaek N E, Berthelsen J G, Giwercman A & Muller J. (1987) Carcinoma-in-situ of the testis: possible origin from gonocytes and precursor of all types of germ cell tumours except spermatocytoma. *International journal of andrology* **10**, 19-28.

Smith M K & Wakimoto B T. (2007) Complex regulation and multiple developmental functions of misfire, the Drosophila melanogaster ferlin gene. *BMC developmental biology* **7**, 21.

Turnbull C, Rapley E A, Seal S, Pernet D, Renwick A, Hughes D, Ricketts M, Linger R, Nsengimana J, Deloukas P, Huddart R A, Bishop D T, Easton D F, Stratton M R, Rahman N & UK Testicular Cancer Collaboration (2010) Variants near DMRT1, TERT and ATF7IP are associated with testicular germ cell cancer. *Nature genetics* **42**, 604-607.

Wood A R, Esko T, Yang J, Vedantam S, Pers T H, Gustafsson S, *et al.* (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature genetics* **46**, 1173-1186.

Zavos C, Andreadis C, Diamantopoulos N & Mouratidou D. (2004) A hypothesis on the role of insulin-like growth factor I in testicular germ cell tumours. *Medical hypotheses* **63**, 511-514.