Endocrine CONNECTIONS

A polymorphic repeat in the *IGF1* promoter influences the risk of endometrial cancer

Katherine A Bolton[1,2], Kelly A Avery-Kiejda[1,2], Elizabeth G Holliday[3,4], John Attia[3,4], Nikola A Bowden[1,2] and **Rodney J Scott[1,2,5,6]**

[1]Centre for Bioinformatics, Biomarker Discovery and Information-Based Medicine, Hunter Medical Research Institute, Newcastle, New South Wales, Australia
[2]Priority Research Centre for Cancer, School of Biomedical Sciences and Pharmacy, Faculty of Health and Medicine, University of Newcastle, Newcastle, New South Wales, Australia
[3]Centre for Clinical Epidemiology and Biostatistics, School of Medicine and Public Health, Faculty of Health and Medicine, University of Newcastle, Newcastle, New South Wales, Australia
[4]Clinical Research Design, IT and Statistical Support Unit, Hunter Medical Research Institute, Newcastle, New South Wales, Australia
[5]Molecular Medicine, Pathology North, John Hunter Hospital, Newcastle, New South Wales, Australia
[6]Discipline of Medical Genetics, School of Biomedical Sciences and Pharmacy, Faculty of Health and Medicine, University of Newcastle, University Drive, Newcastle, New South Wales, Australia

Correspondence should be addressed to R J Scott
Email
rodney.scott@newcastle.edu.au

Abstract

Due to the lack of high-throughput genetic assays for tandem repeats, there is a paucity of knowledge about the role they may play in disease. A polymorphic CA repeat in the promoter region of the insulin-like growth factor 1 gene (*IGF1)* has been studied extensively over the past 10 years for association with the risk of developing breast cancer, among other cancers, with variable results. The aim of this study was to determine if this CA repeat is associated with the risk of developing breast cancer and endometrial cancer. Using a case–control design, we analysed the length of this CA repeat in a series of breast cancer and endometrial cancer cases and compared this with a control population. Our results showed an association when both alleles were considered in breast and endometrial cancers ($P=0.029$ and 0.011, respectively), but this did not pass our corrected threshold for significance due to multiple testing. When the allele lengths were analysed categorically against the most common allele length of 19 CA repeats, an association was observed with the risk of endometrial cancer due to a reduction in the number of long alleles ($P=0.013$). This was confirmed in an analysis of the long alleles separately for endometrial cancer risk ($P=0.0012$). Our study found no association between the length of this polymorphic CA repeat and breast cancer risk. The significant association observed between the CA repeat length and the risk of developing endometrial cancer has not been previously reported.

Key Words
- *IGF1* gene
- short tandem repeat
- promoter region
- endometrial cancer
- breast cancer

Endocrine Connections (2016) **5**, 115–122

Introduction

Endometrial cancer is the most common gynaecological malignancy in women from developed countries (1), and breast cancer is the most common cancer that develops in women worldwide and is responsible for the most female cancer deaths (2). The incidence of both of these hormone-related cancers has been steadily increasing in Australia and other Western and developing countries over the past 20 years (3, 4, 5).

http://www.endocrineconnections.org
DOI: 10.1530/EC-16-0003

Oestrogen exposure is a major risk factor for these cancers (6), with other risk factors including obesity and lack of physical activity (7, 8, 9). Body mass index (BMI) is generally used as an indirect measure of an individual's weight or adiposity (10, 11, 12). An increased BMI is proving to be a major risk factor for cancer (13, 14), and increased rates of hormone-related cancer types in particular appear to be associated with increases in obesity (5). An increase in weight and hence adipose tissue leads to altered cytokine and hormone levels, including increased circulating oestrogen (8, 11, 12). The increase in oestrogen is thought to alter the risk of malignancy via its effects on cellular proliferation, differentiation and apoptosis (7).

Insulin-like growth factor I (IGF1) is a peptide hormone with important regulatory roles in cell growth, differentiation and apoptosis (15, 16), and is also required for normal breast development (17). Epidemiological studies have shown that high circulating levels of IGF1 are associated with an increased risk of developing breast cancer, as well as prostate, colorectal and lung cancers (15, 18, 19, 20). Also, mammographic density, which is a strong predictor of breast cancer risk, has been correlated with plasma IGF1 levels in premenopausal women (21). Hence, polymorphisms in the *IGF1* gene have been extensively analysed in relation to breast cancer risk, although the results have been variable and sometimes contradictory (19, 22, 23, 24, 25), most likely due to differences in ethnic composition of the populations studied (26).

The human *IGF1* gene (located at 12q23.2) is under the control of two promoters, one of which (P1) contains a polymorphic short tandem repeat (STR), being a dinucleotide cytosine–adenine (CA) repeat (27, 28, 29). The number of CA repeats ranges from 10 to 25, with the most common allele in Caucasian populations being 19 repeats (30). It was found that the CA repeat length was related to serum IGF1 levels (31). Hence, this polymorphism has been predicted to affect transcription rates of *IGF1* (31) because it is within 1 kb of the transcription start site and is known to contain regulatory elements (32), in much the same way as the polymorphic CA repeat in the regulatory region of the epidermal growth factor receptor gene (*EGFR*) has been shown to influence transcription of *EGFR* (33). This latter polymorphism has also been associated with the risk of developing breast cancer (34) and lung cancer (35). It has hence been postulated that the polymorphic CA repeat in the *IGF1* promoter may influence circulating IGF1 levels and thus be associated with cancer risk (18).

The polymorphic STR in the *IGF1* promoter has been shown, from analysis in our laboratory and by others, to be associated with earlier age at onset of hereditary non-polyposis colorectal cancer (36, 37, 38). Results of early studies also suggested an association between the absence of the common *IGF1* 19 CA repeat allele and increased breast cancer risk (26, 39). However, recent reports, including several meta-analyses, have found no association between the length of this CA repeat and the risk of developing breast cancer (18, 22, 40). In cervical cancer, a recent study found significant differences in the repeat length between some types of cervical cancer and control groups (41), whereas no association was found with cervical cancer risk in an earlier study (42). These discrepancies may result from ethnic differences between cohorts and some studies having limited power due to small sample sizes.

To our knowledge, the CA repeat polymorphism in the *IGF1* promoter has not been analysed in relation to the risk of developing endometrial cancer. This study was undertaken, using a case–control design, to determine whether changes in the CA repeat length were associated with the risk of developing the oestrogen-driven malignancies, breast cancer or endometrial cancer.

Materials and methods

Patient and control samples

This study included 223 breast cancer cases, 204 endometrial cancer cases and 220 healthy controls from whom blood samples were taken and genomic DNA extracted using the salt extraction method (43). The breast cancer patients were sourced from the Hunter Area Pathology Service as described previously (44). They were all female, with early-onset (<40 years) or bilateral breast cancer, with no known breast cancer predisposition genetic mutations. The DNA from the endometrial cancer cases was collected for previous studies from patients who had presented at the Hunter Centre for Gynaecological Cancer, John Hunter Hospital, Newcastle, NSW, Australia between the years 1992 and 2005 (45). Healthy controls were selected from the Hunter Community Study (HCS) cohort (46), with the requirement that individuals were female and had no history of cancer. The HCS participants were 55–85 years old, randomly recruited from the electoral roll between the years 2004 and 2005 from the Hunter region, Newcastle, NSW, Australia. All participants provided written informed consent for the DNA samples to be used for research. Ethics approval was obtained from the Human Research Ethics Committee, University of Newcastle, and the Hunter New England Human Research

http://www.endocrineconnections.org
DOI: 10.1530/EC-16-0003

Ethics Committee, Hunter New England Health Service, Newcastle, NSW, Australia.

Genotyping the CA repeat

The CA repeat, situated 788 base pairs (bp) upstream of the transcription start site for *IGF1*, was genotyped by polymerase chain reaction (PCR) and fragment analysis using forward (5′-gctagccagctggtgttatt-3′) and reverse (5′-accactctgggagaagggta-3′) primers designed to amplify a 194-bp length fragment as described previously (29). PCR was performed in 10 µL reactions using 40 ng genomic DNA, 1× Platinum *Taq* High Fidelity buffer (Invitrogen), 0.2 mM dNTPs, 1.0 mM $MgSO_4$, 0.4 mM forward primer labelled with 6-carboxyfluorescein (6-FAM), 0.4 mM reverse primer and 0.5 U of Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen). PCRs were performed in a GeneAmp PCR System 9700 (Applied Biosystems) using the following cycles: 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 1 min, and then 72°C for 5 min. Amplified DNA products were used neat or diluted 1:10 or 1:100 in Hi-Di Formamide (Applied Biosystems) dependent upon optimisation results, and denatured for 3 min at 95°C, with the addition of Hi-Di Formamide and 0.5 µL GeneScan 600 LIZ Size Standard (Applied Biosystems). The denatured product was then size separated by capillary electrophoresis on an ABI 3730 DNA Analyzer (Applied Biosystems). The results were analysed using Peak Scanner v1.0 software (Applied Biosystems).

Sanger sequencing was used to confirm the STR length in at least 10% of the samples and to demonstrate that differences in amplified product sizes were due to the length of the STR and not to other sequence changes. PCRs were performed as described previously, but with standard forward and reverse primers (no FAM label).

Table 1 Demographic and clinical characteristics of the participants in this study.

Characteristic	Breast cancer (*n* = 223)	Endometrial cancer (*n* = 204)	Healthy controls (*n* = 220)
Sex	All female	All female	All female
Age (at ascertainment; in years)			
Range	NA	40–92	67–86
Median		68	73
Mean (S.D.)		67.9 (9.5)	73.4 (4.6)
Age (at diagnosis; in years)			
Range	22–57	37–86	NA
Median	41	63.5	
Mean (S.D.)	39.8 (7.3)	63.2 (9.0)	
BMI (in kg/m^2)			
Range	NA	16.9–66.6	17.4–47.1
Median		30.0	27.9
Mean (S.D.)		31.3 (7.8)	28.5 (5.3)
Underweight (BMI < 18.5)		*n* = 1	*n* = 1
Normal (18.5 ≤ BMI < 25)		*n* = 37	*n* = 58
Overweight (25 ≤ BMI < 30)		*n* = 56	*n* = 91
Obese (BMI ≥ 30)		*n* = 94	*n* = 70
Not specified		*n* = 16	*n* = 0
Type of endometrial cancer			
Type I (endometrioid adenocarcinoma, squamous)	NA	144 (70.6%)	NA
Type II (clear cell, UPSC, sarcomas, mucinous, MMMT)		23 (11.3%)	
Other (leiomyomata)		2 (0.9%)	
Unknown		35 (17.2%)	
Ethnicity			
Caucasian	NA	197 (96.6%)	208 (94.5%)
Asian		4 (1.9%)	2 (0.9%)
Middle Eastern		1 (0.5%)	0 (0.0%)
South American		1 (0.5%)	0 (0.0%)
Indigenous Australian		0 (0.0%)	4 (1.8%)
Black African		0 (0.0%)	1 (0.5%)
Pacific Islander		0 (0.0%)	1 (0.5%)
Missing		1 (0.5%)	4 (1.8%)

NA, not available; BMI, body mass index; UPSC, uterine papillary serous carcinoma; MMMT, malignant mixed Mullerian tumours.

http://www.endocrineconnections.org
DOI: 10.1530/EC-16-0003

Table 2 Distribution of allele lengths for the *IGF1* STR in the healthy controls, breast cancer cohorts and endometrial cancer cohorts.

	CA < 19	CA = 19	CA > 19	*n*
Healthy controls	58 (13.2%)	299 (67.9%)	83 (18.9%)	440
Breast cancer cohorts	67 (15.2%)	299 (67.6%)	76 (17.2%)	442
Endometrial cancer cohorts	61 (15.6%)	284 (72.8%)	45 (11.5%)	390

The actual number of allele lengths observed is shown with the percentage of total number of alleles enclosed by parentheses.

They were treated with ExoSAP mix (exonuclease I and shrimp alkaline phosphatase; Thermo Fisher Scientific) before being bidirectionally sequenced with the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). The sequencing products were cleaned up using Agencourt CleanSEQ (Beckman Coulter) and then size separated by capillary electrophoresis on an ABI 3730 DNA Analyzer (Applied Biosystems). The sequencing traces were analysed using Mutation Surveyor v3.97 software (SoftGenetics, State College, PA, USA). A line of best fit was formed from the direct sequencing results to correct the lengths obtained from fragment analysis as described by Pasqualotto and coworkers (47).

Statistical analysis

Statistical analyses were performed using the Stata 11.1 software package (StataCorp LP, College Station, TX, USA). The *t*-tests were used to test case–control differences in continuous STR length when STR length was normally distributed. The non-parametric equivalent (Mann–Whitney rank-sum/*U* test) was used for analysing STR lengths when deviation from a normal distribution was observed. Cox proportional hazard regression analysis was used to assess association between STR length and age at diagnosis. A Pearson χ^2 test was used to analyse STR length when it was treated as a categorical variable, being divided into three groups: STR length equal to 19 (being 19 CA repeats, the most common allele length; CA = 19), CA length less than 19 (CA < 19) and CA length greater than 19 (CA > 19), as is most commonly done when analysing this particular STR (18, 41). The significance levels of all tests were set at a P value <0.05 (two-sided) and then corrected for multiple comparisons using the Bonferroni method.

The hypotheses we set out to test were that an altered length of the *IGF1* STR would be associated with the risk of breast or endometrial cancer and/or the age at diagnosis of breast or endometrial cancer. We considered STR length as a continuous variable (for both alleles and separately by long and short allele) as well as a categorical variable (CA < 19, CA = 19 and CA > 19). According to the Bonferroni method, when we consider our original two outcomes (being occurrence of cancer and age at diagnosis) and three predictors (both short and long alleles or categories; 2 × 3 = 6 tests), our adjusted significance threshold is $\alpha = 0.05/6 = 0.008$. We have adjusted for three predictors as the categories are not independent predictors.

Results

The demographic and clinical characteristics of the participants used in this study are shown in Table 1. The majority of individuals in these cohorts had Caucasian ethnicity; ethnicity details were not available for the breast cancer cohort. The non-Caucasian samples and those with missing ethnicity (contributing 3.4% of the endometrial cancer cohort and 5.5% of the healthy controls) are described in Table 1.

The length of the CA repeat in the IGF1 promoter

The length of the *IGF1* repeat (number of copies of the CA motif) ranged from 12 to 22 across all three cohorts. The mean numbers of copies of the CA motif were 19.06, 19.02 and 18.98 for the healthy control, breast cancer and endometrial cancer cohorts, respectively. The most common allele length seen was 19 CA repeats, which was observed in 67.9, 67.6 and 72.8% of alleles in the healthy controls, breast cancer patients and endometrial cancer patients, respectively (Table 2). The distribution of allele lengths when grouped against the most common allele length of 19 CA repeats is shown in Table 2. The most common genotype was the homozygote CA19/19, which was observed with a frequency of 59.5, 62.0 and 68.7% in the healthy control, breast cancer and endometrial cancer cohorts, respectively.

The association of *IGF1* STR length with breast and endometrial cancer risk

As the distribution of STR lengths is slightly skewed, the Mann–Whitney rank-sum test was used to

http://www.endocrineconnections.org
DOI: 10.1530/EC-16-0003

Table 3 HR, 95% CI and *P* values for breast and endometrial cancer case–control analysis in relation to *IGF1* STR lengths.

		Breast cancer ($n=221$)		**Endometrial cancer** ($n=195$)	
Category	**Statistical test**	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Both allele lengths with cancer risk	Mann–Whitney *U* test	NA	**0.029**	NA	**0.011**
Both allele lengths with age at diagnosis	Cox proportional hazard regression	0.989 (0.901–1.086)	0.824	0.952 (0.863–1.052)	0.335
Both allele lengths with age at diagnosis; BMI considered	Cox proportional hazard regression	NA	NA	0.951 (0.854–1.060)	0.366
Short allele length with cancer risk	Mann–Whitney *U* test	NA	0.184	NA	0.780
Short allele length with age at diagnosis	Cox proportional hazard regression	1.009 (0.884–1.150)	0.897	1.065 (0.908–1.249)	0.442
Short allele length with age at diagnosis; BMI considered	Cox proportional hazard regression	NA	NA	1.051 (0.886–1.248)	0.567
Long allele length with cancer risk	Mann–Whitney *U* test	NA	0.105	NA	**0.0012**
Long allele length with age at diagnosis	Cox proportional hazard regression	0.960 (0.821–1.122)	0.608	**0.794** (0.659–0.956)	**0.015**
Long allele length with age at diagnosis; BMI considered	Cox proportional hazard regression	NA	NA	**0.806** (0.659–0.986)	**0.036**
Categorical analysis (three groups; Table 2)	Pearson's χ^2 test	NA	0.621	NA	**0.013**

Significant *P* values and associated HRs are highlighted in bold

test for association between *IGF1* STR length and breast or endometrial cancer occurrence. The results tended towards an association when both allele lengths were analysed in relation to breast cancer risk ($P=0.029$) and for endometrial cancer risk ($P=0.011$; Table 3), although they did not pass our stringent Bonferroni-adjusted threshold for multiple testing ($\alpha=0.008$). There was no association between the length of both alleles and cancer risk when considering age at diagnosis for breast cancer (hazard ratio (HR) = 0.989 (95% confidence interval (CI) = 0.901–1.086), $P=0.824$) or endometrial cancer (HR = 0.952 (95% CI = 0.863–1.052), $P=0.335$), nor when BMI was adjusted for in the endometrial cancer cohort (HR = 0.951 (95% CI = 0.854–1.060), $P=0.366$; Table 3). BMI data were not available for the breast cancer cohort.

When the allele lengths were categorised according to the most common allele length of 19 CA repeats, a sizeable decrease in the number of long alleles (CA > 19) in the endometrial cancer cohort (Table 2) led to an association between this polymorphic CA repeat and the risk of developing endometrial cancer that tends towards significance ($P=0.013$; Table 3). This association was due to less long alleles observed (CA >19; 11.5%; Table 2) than expected (15.4%) in the endometrial cancer cases compared with more long alleles (18.9%) than expected (15.4%) in the healthy controls ($P=0.013$; Table 3).

This association was verified when the STR allele lengths were considered separately. No association was seen when the short allele lengths were considered separately for both cancer types (for breast cancer, $P=0.184$; for endometrial cancer, $P=0.780$), nor when the long allele was considered for breast cancer risk ($P=0.105$; Table 3). However, a significant association was observed between the long alleles and the risk of developing endometrial cancer ($P=0.0012$; Table 3). In order to check that varied ethnicity between cohorts is not confounding this analysis, the 7 non-Caucasian endometrial cancer cases and 12 non-Caucasian healthy controls were removed from the case–control analysis (data not shown). This resulted in HRs that were unchanged from those presented in Table 3, and the above-mentioned *P* value ($P=0.0012$) for association between the long alleles and the risk of developing endometrial cancer remained significant ($P=0.0015$). The size of this association was moderated when age at diagnosis ($P=0.015$) and BMI ($P=0.036$) were considered, suggesting that these may be confounding factors in the association. For this reason, larger cohorts need to be analysed to confirm these associations.

Discussion

When present in regulatory regions, polymorphic tandem repeats (TRs) such as the CA repeat in the promoter of the

http://www.endocrineconnections.org
DOI: 10.1530/EC-16-0003

IGF1 gene can have profound effects on gene expression and hence protein abundance and, therefore, impact on the risk and severity of disease (48). However, TRs remain an under-explored source of genomic variation, largely due to their inability to be analysed on a large scale, with recent emphasis being on genome-wide association studies, which are unable to detect polymorphisms produced by variable TRs (49). High-throughput techniques, such as next-generation sequencing, are not readily amenable to analysing TRs (50) mainly due to problems with alignment and assembly of the regions containing stretches of TRs from short reads (51, 52). Other problems that occur when analysing TRs are due to inconsistencies in their definition, their heterogeneity and the fast evolution of some TR-containing regions (53). Also, selecting appropriate statistical methods is problematic due to the multitude of allele lengths generated by polymorphic TRs, which can be treated as continuous numerical variables or grouped to form alternate variables to fit with standard bi-allelic genetic analysis (54). Hence, these technical limitations make the analysis of TRs tedious and difficult to obtain large sample numbers.

Analysis of the dinucleotide repeat in the promoter of the *IGF1* gene in our healthy population and in breast and endometrial cancer patient samples confirmed its variability. There was no significant association seen between the length of this CA repeat and the risk of developing breast cancer, which is in agreement with the results of recent meta-analyses in mainly Caucasian populations, including one that involved nine studies with a total of 5641 cases and 10,471 controls (40) and another that included seven studies with 3533 breast cancer cases and 7771 controls (18). A meta-analysis conducted recently on the effect of this repeat on the risk of developing breast cancer (which included 11 studies of various ethnicity groups with a total of 7047 cases and 12,096 controls) reported a decreased risk of breast cancer in the Caucasian population associated with the CA 19/19 repeat genotype; however, the significance of the *P* values was borderline in this analysis, and the threshold had not been corrected for multiple testing (55).

A significant association was observed between the long alleles for the CA repeat in the *IGF1* promoter and the risk of developing endometrial cancer in our study ($P=0.0012$), which was because the length of the long alleles is significantly shorter in the endometrial cancer cohort compared with the healthy controls. The direction of this association is in agreement with that seen for the same STR in the *IGF1* promoter in a previous breast cancer risk study, which found shorter alleles to be a risk factor (26), as well as with an association between shorter alleles and an earlier age of onset for colorectal cancer (36, 37, 38). There remains the possibility of ethnically distinct allele sizes confounding this type of analysis (56). As it was recognised that these cohorts were not homogeneous, the analysis was repeated after removal of the results for the small number of non-Caucasian endometrial cancer and control samples, which resulted in unchanged hazard ratios and equivalent *P* values.

It may be that a particular length of this STR is required for optimal transcriptional activity and hence optimal expression of this gene. The presence of variable TRs in regulatory regions can affect gene expression (57) due to alterations in DNA structure away from the usual double-helix B-DNA formation (58). An unusual DNA structure, known as Z-DNA, is predicted to form when negative torsional strain occurs as a result of STRs with motif sequences composed of alternating purine-pyrimidine bases, such as the CA repeat which is commonly found in the human genome, especially in promoter regions (58). This could affect transcription by blocking RNA polymerase activity, altering accessibility of transcription factor binding sites or altering binding of regulatory elements due to changed chromatin structure (58). Hence, it is important to consider all allele lengths (18, 26), and not only whether the most common allele length (of 19 CA repeats) is present or absent as done in some previous studies (19, 39, 55).

In summary, our results support the notion that there is no association between this polymorphic dinucleotide CA repeat in the *IGF1* promoter and the risk of developing breast cancer. However, the association of this STR with the risk of developing endometrial cancer is a new and interesting finding. Although further analyses of larger cohorts and those with varied ethnicity are required to confirm this association, it raises the possibility of this polymorphic STR being used as a biomarker for screening individuals at an increased risk of developing endometrial cancer in the future.

Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding
This work was supported by an Australian Postgraduate Award (KAB), National Health and Medical Research Council Training (post-doctoral) Fellowship (grant number 569317; NAB) and Cancer Institute NSW Career Development Fellowship (grant number 14/CDF/1-17; NAB).

http://www.endocrineconnections.org
DOI: 10.1530/EC-16-0003

Author contribution statement

R J Scott conceived this study. K A Bolton, K A Avery-Kiejda, N A Bowden and R J Scott performed the study design. K A Bolton performed the experimental work and statistical analysis and wrote the manuscript. E G Holliday assisted with the statistical analysis. K A Avery-Kiejda, E G Holliday, N A Bowden and R J Scott assisted with interpretation of results. E G Holliday, K A Avery-Kiejda, N A Bowden and R J Scott reviewed the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank the breast cancer patients (Hunter Area Pathology Service, HAPS), endometrial cancer patients (John Hunter Hospital, JHH) and healthy controls (Hunter Community Study, HCS) who kindly consented to the use of their DNA for research. We also thank Dr Michelle Wong-Brown for helping to source the breast cancer samples from HAPS; Drs K Ashton, A Proietto and G Otton for sourcing and collecting the endometrial cancer samples from JHH; and Mr M McEvoy and Mr S Hancock for sourcing samples and clinical data for the HCS cohort.

References

1 Siegel R, Ma J, Zou Z & Jemal A. Cancer statistics, 2014. *CA: A Cancer Journal for Clinicians* 2014 **64** 9–29. (doi:10.3322/caac.21208)
2 Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D & Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer* 2015 **136** E359–E386. (doi:10.1002/ijc.29210)
3 Australian Institute of Health and Welfare. *Australian Cancer Incidence and Mortality Books: Breast Cancer*. Canberra, ACT, Australia: Australian Institute of Health and Welfare, 2015*a*.
4 Australian Institute of Health and Welfare. *Australian Cancer Incidence and Mortality Books: Uterine Cancer*. Canberra, ACT, Australia: Australian Institute of Health and Welfare, 2015*b*.
5 World Cancer Research Fund. *Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective*. Washington, DC, USA: American Institute for Cancer Research, 2007.
6 Henderson BE & Feigelson HS. Hormonal carcinogenesis. *Carcinogenesis* 2000 **21** 427–433. (doi:10.1093/carcin/21.3.427)
7 Kaaks R, Lukanova A & Kurzer MS. Obesity, endogenous hormones, and endometrial cancer risk: a synthetic review. *Cancer Epidemiology, Biomarkers & Prevention* 2002 **11** 1531–1543.
8 Park SL, Goodman MT, Zhang ZF, Kolonel LN, Henderson BE & Setiawan VW. Body size, adult BMI gain and endometrial cancer risk: the multiethnic cohort. *International Journal of Cancer* 2010 **126** 490–499. (doi:10.1002/ijc.24718)
9 Travis RC & Key TJ. Oestrogen exposure and breast cancer risk. *Breast Cancer Research* 2003 **5** 239–247. (doi:10.1186/bcr628)
10 Bianchini F, Kaaks R & Vainio H. Overweight, obesity, and cancer risk. *Lancet Oncology* 2002 **3** 565–574. (doi:10.1016/S1470-2045(02)00849-5)
11 Bray GA. The underlying basis for obesity: relationship to cancer. *Journal of Nutrition* 2002 **132** 3451S–3455S.
12 Calle EE & Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. *Nature Reviews Cancer* 2004 **4** 579–591. (doi:10.1038/nrc1408)
13 Crosbie EJ, Zwahlen M, Kitchener HC, Egger M & Renehan AG. Body mass index, hormone replacement therapy, and endometrial cancer risk: a meta-analysis. *Cancer Epidemiology, Biomarkers & Prevention* 2010 **19** 3119–3130. (doi:10.1158/1055-9965.EPI-10-0832)
14 Renehan AG, Tyson M, Egger M, Heller RF & Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet* 2008 **371** 569–578. (doi:10.1016/S0140-6736(08)60269-X)
15 Furstenberger G & Senn HJ. Insulin-like growth factors and cancer. *Lancet Oncology* 2002 **3** 298–302. (doi:10.1016/S1470-2045(02)00731-3)
16 LeRoith D, Werner H, Beitner-Johnson D & Roberts CT Jr. Molecular and cellular aspects of the insulin-like growth factor I receptor. *Endocrine Reviews* 1995 **16** 143–163. (doi:10.1210/edrv-16-2-143)
17 Laban C, Bustin SA & Jenkins PJ. The GH-IGF-I axis and breast cancer. *Trends in Endocrinology & Metabolism* 2003 **14** 28–34. (doi:10.1016/S1043-2760(02)00003-6)
18 Chen X, Guan J, Song Y, Chen P, Zheng H, Tang C & Wu Q. IGF-I (CA) repeat polymorphisms and risk of cancer: a meta-analysis. *Journal of Human Genetics* 2008 **53** 227–238. (doi:10.1007/s10038-007-0241-3)
19 DeLellis K, Ingles S, Kolonel L, McKean-Cowdin R, Henderson B, Stanczyk F & Probst-Hensch NM. IGF1 genotype, mean plasma level and breast cancer risk in the Hawaii/Los Angeles multiethnic cohort. *British Journal of Cancer* 2003 **88** 277–282. (doi:10.1038/sj.bjc.6600728)
20 Hankinson SE, Willett WC, Colditz GA, Hunter DJ, Michaud DS, Deroo B, Rosner B, Speizer FE & Pollak M. Circulating concentrations of insulin-like growth factor-I and risk of breast cancer. *Lancet* 1998 **351** 1393–1396. (doi:10.1016/S0140-6736(97)10384-1)
21 Byrne C, Colditz GA, Willett WC, Speizer FE, Pollak M & Hankinson SE. Plasma insulin-like growth factor (IGF) I, IGF-binding protein 3, and mammographic density. *Cancer Research* 2000 **60** 3744–3748.
22 Fehringer G, Boyd NF, Knight JA, Paterson AD, Dite GS, Giles GG, Southey MC, Andrulis IL, Hopper JL & Ozcelik H. Family-based genetic association study of insulin-like growth factor I microsatellite markers and premenopausal breast cancer risk. *Breast Cancer Research and Treatment* 2009 **118** 415–424. (doi:10.1007/s10549-009-0336-y)
23 Figer A, Karasik YP, Baruch RG, Chetrit A, Papa MZ, Sade RB, Rizel S & Friedman E. Insulin-like growth factor I polymorphism and breast cancer risk in Jewish women. *Israel Medical Association Journal* 2002 **4** 759–762.
24 Missmer SA, Haiman CA, Hunter DJ, Willett WC, Colditz GA, Speizer FE, Pollak MN & Hankinson SE. A sequence repeat in the insulin-like growth factor-1 gene and risk of breast cancer. *International Journal of Cancer* 2002 **100** 332–336. (doi:10.1002/ijc.10473)
25 Yu H, Li BD, Smith M, Shi R, Berkel HJ & Kato I. Polymorphic CA repeats in the IGF-I gene and breast cancer. *Breast Cancer Research and Treatment* 2001 **70** 117–122. (doi:10.1023/A:1012947027213)
26 Cleveland RJ, Gammon MD, Edmiston SN, Teitelbaum SL, Britton JA, Terry MB, Eng SM, Neugut AI, Santella RM & Conway K. IGF1 CA repeat polymorphisms, lifestyle factors and breast cancer risk in the Long Island Breast Cancer Study Project. *Carcinogenesis* 2006 **27** 758–765. (doi:10.1093/carcin/bgi294)
27 LeRoith D & Roberts CT Jr. Insulin-like growth factors. *Annals of the New York Academy of Sciences* 1993 **692** 1–9. (doi:10.1111/j.1749-6632.1993.tb26200.x)
28 Rotwein P, Pollock KM, Didier DK & Krivi GG. Organization and sequence of the human insulin-like growth factor I gene. Alternative RNA processing produces two insulin-like growth factor I precursor peptides. *Journal of Biological Chemistry* 1986 **261** 4828–4832.
29 Weber JL & May PE. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* 1989 **44** 388–396.
30 Rietveld I, Janssen JA, Van Duijn CM & Lamberts SW. A polymorphic CA repeat in the promoter region of the insulin-like growth factor-I (IGF-I) gene. *European Journal of Epidemiology* 2003 **18** 191–193. (doi:10.1023/A:1023309321705)

http://www.endocrineconnections.org
DOI: 10.1530/EC-16-0003

31 Rosen CJ, Kurland ES, Vereault D, Adler RA, Rackoff PJ, Craig WY, Witte S, Rogers J & Bilezikian JP. Association between serum insulin growth factor-I (IGF-I) and a simple sequence repeat in IGF-I gene: implications for genetic studies of bone mineral density. *Journal of Clinical Endocrinology and Metabolism* 1998 **83** 2286–2290. (doi:10.1210/jcem.83.7.4964)

32 McCarthy TL, Ji C, Shu H, Casinghino S, Crothers K, Rotwein P & Centrella M. 17beta-estradiol potently suppresses cAMP-induced insulin-like growth factor-I gene activation in primary rat osteoblast cultures. *Journal of Biological Chemistry* 1997 **272** 18132–18139. (doi:10.1074/jbc.272.29.18132)

33 Gebhardt F, Zanker KS & Brandt B. Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1. *Journal of Biological Chemistry* 1999 **274** 13176–13180. (doi:10.1074/jbc.274.19.13176)

34 Brandt B, Hermann S, Straif K, Tidow N, Buerger H & Chang-Claude J. Modification of breast cancer risk in young women by a polymorphic sequence in the egfr gene. *Cancer Research* 2004 **64** 7–12. (doi:10.1158/0008-5472.CAN-03-2623)

35 Yoshida K, Nakachi K, Imai K, Cologne JB, Niwa Y, Kusunoki Y & Hayashi T. Lung cancer susceptibility among atomic bomb survivors in relation to CA repeat number polymorphism of epidermal growth factor receptor gene and radiation dose. *Carcinogenesis* 2009 **30** 2037–2041. (doi:10.1093/carcin/bgp247)

36 Reeves S, Meldrum C & Scott RJ. Re: IGF-1 gene polymorphism and risk for hereditary nonpolyposis colorectal cancer. *Journal of the National Cancer Institute* 2006 **98** 1664–1665. (doi:10.1093/jnci/djj452)

37 Reeves SG, Rich D, Meldrum CJ, Colyvas K, Kurzawski G, Suchy J, Lubinski J & Scott R. IGF1 is a modifier of disease risk in hereditary non-polyposis colorectal cancer. *International Journal of Cancer* 2008 **123** 1339–1343. (doi:10.1002/ijc.23668)

38 Zecevic M, Amos CI, Gu X, Campos IM, Jones JS, Lynch PM, Rodriguez-Bigas MA & Frazier ML. IGF1 gene polymorphism and risk for hereditary nonpolyposis colorectal cancer. *Journal of the National Cancer Institute* 2006 **98** 139–143. (doi:10.1093/jnci/djj016)

39 Jernstrom H, Sandberg T, Bageman E, Borg A & Olsson H. Insulin-like growth factor-1 genotype predicts breast volume after pregnancy and hormonal contraception and is associated with circulating insulin-like growth factor-1 levels: implications for risk of early-onset breast cancer in young women from hereditary breast cancer families. *International Journal of Gynecological Cancer* 2006 **16** (Supplement 2) 497.

40 Huang Q, Wang C, Qiu LJ, Shao F & Yu JH. The association between IGF1 CA repeat polymorphisms and breast cancer risk: a meta-analysis. *Breast Cancer Research and Treatment* 2011 **129** 9. (doi:10.1007/s10549-011-1434-1)

41 Kwasniewski W, Gozdzicka-Jozefiak A, Kotarska M, Polak G, Barczynski B, Broniarczyk J, Nowak W, Wolun-Cholewa M, Kwasniewska A & Kotarski J. Analysis of cytosine-adenine repeats in P1 promoter region of IGF-1 gene in peripheral blood cells and cervical tissue samples of females with cervical intraepithelial lesions and squamous cervical cancer. *Molecular Medicine Reports* 2015 **11** 766–774. (doi:10.3892/mmr.2014.2916)

42 Pacholska-Bogalska J, Jozefiak A, Nowak W, Kedzia W, Kwasniewska A & Gozdzicka-Jozefiak A. Association of the IGF-I promoter P1 polymorphism with risk of cervical cancer. *European Journal of Gynaecological Oncology* 2011 **32** 393–398.

43 Miller SA, Dykes DD & Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* 1988 **16** 1215. (doi:10.1093/nar/16.3.1215)

44 Wong MW, Nordfors C, Mossman D, Pecenpetelovska G, Avery-Kiejda KA, Talseth-Palmer B, Bowden NA & Scott RJ. BRIP1, PALB2, and RAD51C mutation analysis reveals their relative importance as genetic susceptibility factors for breast cancer. *Breast Cancer Research and Treatment* 2011 **127** 853–859. (doi:10.1007/s10549-011-1443-0)

45 Ashton KA, Proietto A, Otton G, Symonds I, McEvoy M, Attia J, Gilbert M, Hamann U & Scott RJ. The influence of the Cyclin D1 870 G>A polymorphism as an endometrial cancer risk factor. *BMC Cancer* 2008 **8** 272. (doi:10.1186/1471-2407-8-272)

46 McEvoy M, Smith W, D'Este C, Duke J, Peel R, Schofield P, Scott R, Byles J, Henry D, Ewald B, *et al*. Cohort profile: The Hunter Community Study. *International Journal of Epidemiology* 2010 **39** 1452–1463. (doi:10.1093/ije/dyp343)

47 Pasqualotto AC, Denning DW & Anderson MJ. A cautionary tale: Lack of consistency in allele sizes between two laboratories for a published multilocus microsatellite typing system. *Journal of Clinical Microbiology* 2007 **45** 522–528. (doi:10.1128/JCM.02136-06)

48 Gemayel R, Vinces MD, Legendre M & Verstrepen KJ. Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annual Review of Genetics* 2010 **44** 445–477. (doi:10.1146/annurev-genet-072610-155046)

49 Hannan AJ. Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for 'missing heritability'. *Trends in Genetics* 2010 **26** 59–65. (doi:10.1016/j.tig.2009.11.008)

50 Jorgenson E & Witte JS. Microsatellite markers for genome-wide association studies. *Nature Reviews Genetics* 2007 **8** c2. (doi:10.1038/nrg1962-c2)

51 Treangen TJ & Salzberg SL. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nature Reviews Genetics* 2011 **13** 36–46. (doi:10.1038/nrg3117)

52 Zavodna M, Bagshaw A, Brauning R & Gemmell NJ. The accuracy, feasibility and challenges of sequencing short tandem repeats using next-generation sequencing platforms. *PLoS ONE* 2014 **9** e113862. (doi:10.1371/journal.pone.0113862)

53 Anisimova M, Pecerska J & Schaper E. Statistical approaches to detecting and analyzing tandem repeats in genomic sequences. *Frontiers in Bioengineering and Biotechnology* 2015 **3** 31. (doi:10.3389/fbioe.2015.00031)

54 Press MO, Carlson KD & Queitsch C. The overdue promise of short tandem repeat variation for heritability. *Trends in Genetics* 2014 **30** 504–512. (doi:10.1016/j.tig.2014.07.008)

55 He B, Xu Y, Pan Y, Li R, Gao T, Song G, Gu L, Nie Z, Chen L & Wang S. Differential effects of insulin-like growth factor-1 CA repeat polymorphism on breast cancer risk along with race: a meta-analysis. *Gene* 2013 **525** 92–98. (doi:10.1016/j.gene.2013.04.041)

56 Jernstrom H, Chu W, Vesprini D, Tao Y, Majeed N, Deal C, Pollak M & Narod SA. Genetic factors related to racial variation in plasma levels of insulin-like growth factor-1: implications for premenopausal breast cancer risk. *Molecular Genetics and Metabolism* 2001 **72** 144–154. (doi:10.1006/mgme.2000.3130)

57 Lawson MJ & Zhang L. Housekeeping and tissue-specific genes differ in simple sequence repeats in the 5′-UTR region. *Gene* 2008 **407** 54–62. (doi:10.1016/j.gene.2007.09.017)

58 Sawaya SM, Bagshaw AT, Buschiazzo E & Gemmel NJ. Promoter microsatellites as modulators of human gene expression. In *Tandem Repeat Polymorphisms: Genetic Plasticity, Neural Diversity and Disease*. Ed AJ Hannan. Austin, TX, USA: Landes Bioscience, 2012.

Received in final form 13 April 2016
Accepted 18 April 2016

http://www.endocrineconnections.org
DOI: 10.1530/EC-16-0003