Published in final edited form as:

Gut. 2008 August ; 57(8): 1090–1096. doi:10.1136/gut.2007.140855.

***Insulin-like growth factor-1* promoter polymorphisms and colorectal cancer: a functional genomics approach**

H-L Wong¹, W-P Koh², N M Probst-Hensch³, D Van den Berg⁴, M C Yu⁵, and S A Ingles⁴

¹ Infections and Immunoepidemiology Branch, Division of Cancer Epidemiology and Genetics, NCI, NIH, DDHS, Rockville, Maryland, USA ² Department of Community, Occupational and Family Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore ³ Molecular Epidemiology/Cancer Registry, Institute of Social and Preventive Medicine/Department of Pathology, University of Zurich, Switzerland ⁴ USC/Norris Comprehensive Cancer Center, University of Southern California Keck School of Medicine, Los Angeles, California, USA ⁵ The Cancer Center, University of Minnesota, Minneapolis, Minnesota, USA

Abstract

Rationale—Insulin-like growth factor-1 (IGF1) has been proposed to mediate the obesity-related carcinogenic effects of “Western lifestyle”. While genetic factors explain at least half of inter-individual IGF1 variation, the *IGF1* polymorphisms hypothesised to underlie the variation in cancer incidence rates remain ill-defined.

Methods—We used a comparative genomics approach to identify putative regulatory polymorphisms in the *IGF1* promoter region within a rapidly westernising population, the Singapore Chinese. Association of IGF1 genotype with colorectal cancer risk was assessed among 298 colorectal cancer cases and 1142 controls nested within the Singapore Chinese Health Study.

Results—We identified a common (minor allele frequency = 0.36) single-nucleotide polymorphism (SNP), *IGF1*-2995 C/A, within a consensus domain for an octamer binding factor (Oct1/Oct2) transcription factor binding site. Possession of one or two copies of the minor allele (genotypes AA and CA) conferred an approximate 40% decrease in risk in comparison to genotype CC (odds ratio, 0.59; 95% confidence interval, 0.45 to 0.77). This association was stronger for colon cancer than for rectal cancer ($p_{\text{heterogeneity}} < 0.001$) and for those who were physically active versus inactive ($p_{\text{interaction}} = 0.05$). Models including other previously identified promoter polymorphisms did not provide a better prediction of colorectal cancer risk.

Conclusions—Our results support the hypotheses that IGF1 plays a role in colonic carcinogenesis and that genetically inherited variation in IGF1 expression influences risk of colorectal cancer.

Although a “Western lifestyle” is thought to predispose to colorectal cancer,¹ the mechanistic basis remains unclear. Rapidly westernising populations, ie, Japan, Hong Kong and the southeast Asian island nation of Singapore, have been faced with the doubling of colorectal cancer rates in the past two decades. Insulin-like growth factor-1 (IGF1) is thought to mediate the insulin resistance-related carcinogenic effects of “Western lifestyle”.^{2–4} Adoption of a Western diet is associated with significant increases in serum IGF1 levels.⁵ Extensive evidence,

Correspondence to: Dr Sue A Ingles, USC Norris Cancer Center, 1441 Eastlake Ave, Room 6419, Los Angeles CA, 90033, USA; ingles@usc.edu.

Competing interests: None.

Ethics approval: The Institutional Review Boards at the National University of Singapore, the University of Minnesota and the University of Southern California approved this study.

both in vitro and in vivo, suggests that IGF1 promotes colorectal cancer growth,^{6–15} prevents apoptosis^{16–18} and increases metastasis.¹⁹ IGF1 effects its multiple roles at the endocrine, paracrine and autocrine levels.²⁰

Significant heterogeneity in IGF1 levels, both circulating (endocrine) and local (paracrine and autocrine), exists among normal healthy individuals, both within and between ethnic groups.^{21–25} This population variation in IGF1 levels is hypothesised to underlie the variation in cancer incidence rates. To examine this hypothesis, six prospective population-based studies have examined circulating IGF1 (serum or plasma levels measured at one time point during middle-age) as a proxy for IGF1 mitogenic actions in the colorectum.^{26–31} Individuals with higher levels of circulating IGF1 (uppermost quartile or quintile) were found to be at increased risk of developing cancers of the colorectum (meta-analyses of the first five studies³²).

IGF1 variability is, in part, genetically determined, with genetic factors contributing an estimated 50% of the variability in circulating IGF1 levels.^{33, 34} The genetic basis that underlies IGF1 variability, however, remains poorly defined. Considerable allelic imbalance in *IGF1* gene expression argues for the contribution of cis-acting regulatory polymorphisms.²⁵ Although *IGF1* tag single-nucleotide polymorphisms (SNPs) have been associated with circulating IGF1 levels,³⁵ the functional polymorphisms that might be mediating these associations have not been identified. A cytosine–adenosine dinucleotide repeat sequence (CA_{15–22}) that resides in the promoter region has been inconsistently associated with serum levels and with colorectal cancer risk. However, we have previously argued that this microsatellite is not responsible for variation in IGF1 levels, but rather is in linkage disequilibrium with other, functional, genetic variants.³⁶

The CA_{15–22} polymorphism is localised to a haplotype block that spans 20 kb pairs upstream from the *IGF1* translation start site. We have previously observed a signal in this haplotype block in a case–control study nested within a prospective cohort of Han Chinese in Singapore,³⁶ with recent findings of a history of physician-diagnosed diabetes mellitus as a risk factor for colorectal cancer.³⁷ Hence, we propose that this haplotype block spanning the *IGF1* promoter harbours functional *IGF1* polymorphisms that may impact cancer risk in the population. To localise candidate regulatory sequences within this haplotype block, in the current study we employed a comparative genomics strategy. The assumption is that nucleotide sequences involved in regulating gene expression maintain a higher similarity than do non-critical sequences under the neutral evolution model. Within this 20 kbp haplotype block, we identified four regions that are evolutionarily conserved across vertebrate species (human, mouse, rat and dog). These four regions lie within 10 kbp from the transcription start site (fig 1). No conserved regions were identified in the upstream 10 kbp; the majority of the upstream genomic region harboured repetitive elements specific to humans, eg, Alu sequences. We then re-sequenced these evolutionarily conserved regions (ECRs) in a sub-sample of our study population (60 Han Chinese, Singapore Chinese Health Study controls). We identified one *IGF1* promoter SNP that lies in a potentially functional element. To interrogate the role of insulin resistance and colorectal cancer in a relatively lean population, we examined this newly identified SNP, along with another ECR SNP that we previously studied, for association with colorectal cancer risk.³⁶

MATERIALS AND METHODS

Study population

The study design and subject recruitment of the Singapore Chinese Health Study have been described.⁴⁰ Briefly, 63 257 Chinese women and men aged 45–74 years and belonging to the Hokkien or Cantonese dialect group were enrolled in the study between April 1993 and

December 1998. At recruitment, information on lifestyle factors and usual diet over the last year was obtained through in-person interviews.

Incident cancer cases among cohort members were identified through record linkage with the nationwide Singapore Cancer Registry database.³⁹ As of 30 April 2002, 592 colorectal cancer cases had occurred among cohort participants. All cases were histologically confirmed except three who were ascertained by death records and clinical evidence. Between April 1994 and July 1999, we attempted to collect blood (or buccal cells if subjects refused blood donation) and single-void urine specimens from a random 3% sample of cohort enrollees. Details of this biospecimen collection have been described previously.⁴¹ Blood or buccal cell samples were collected from 1194 subjects (63% of total eligible subjects). After excluding five subjects with prevalent history of colorectal cancer and 13 subjects who developed colorectal cancer by 30 April 2002, the remaining 1176 subjects constituted the comparison group for this case-control analysis. In addition, we also attempted to collect blood/buccal cells and urine samples from all incident colorectal cancer cases. Of the 592 colorectal cancer cases, 312 (52%) donated blood/buccal cell samples for genetic analyses. Two cases were later excluded due to misclassification of cancer type. These 310 cases with available DNA samples constituted the case group of our study. Of the 310 available biospecimens, 50 were collected pre-diagnostically. Of the 260 cases that were collected post-diagnostically, the median time from diagnosis to blood draw was 9.5 months. Compared with colorectal cancer cases who did not donate blood or buccal samples, those who donated had a similar mean of age at cancer diagnosis (65 vs 66 years). They were also more likely to be male (57% vs 49%), of the Cantonese dialect group (45% vs 37%) and more educated (69% vs 60% had formal education). There was no difference between these two groups by level of body mass index, cigarette smoking, alcohol drinking or history of diabetes mellitus.

Sample size

Of the 310 cases and 1176 controls with biospecimens available for this study, genotype was missing for 12 cases and 34 controls, leaving a final sample size of 298 cases and 1142 controls.

In silico methods

To identify ECRs in the region upstream of the first *IGF1* translation start site (nucleotide 84199 in NCBI accession # AC010202), we compared the human sequence (NCBI Build 35 in the UCSC human May 2004 assembly, hg17) with mouse, rat and dog (UCSC assembly: mm6, rn3, canFam1 respectively) sequence. Pairwise alignments were constructed and the nucleotide level match-mismatch similarity profiles were compared with the blastz algorithm⁴⁰ incorporated in the ECR Browser portal.⁴¹ Regions with conserved elements were defined as intervals that exceed the threshold of 200 bp with >80% nucleotide identity. Using the MatInspector software (<http://www.genomatrix.de/>, accessed 26 March 2008), putative transcription factor binding sites within the ECRs were predicted, and the potential impact on transcription factor binding was calculated for any SNPs discovered in the re-sequencing process.

Laboratory methods

Re-sequencing—With the exception of ECR1, which we previously re-sequenced,³⁶ the ECRs were re-sequenced in 60 individuals from the Singapore Chinese Health Study. Primers (forward and reverse) were 5'-TGATGTGTCAGTCCCCTG-3' and 5'-GGAGTCTGTGTGCCAGAGTG-3' for ECR2; 5'-TGTTGGCATGTTTATTGCTC-3' and 5'-GCTCGGTGCACAGATATAACC3' for ECR3; and 5'-AAGACTGGGAACATGGCTTG-3' and 5'-AGCCCAAGAGGAGTTCAGGT-3' for ECR4. Sequencing was carried out on an ABI Prism 3700 DNA Sequencer (PE Biosystems, Foster City, CA, USA).

Genotyping—Genomic DNA was isolated from peripheral blood lymphocytes and buccal cell samples using a QIAamp 96 DNA Blood Kit (Qiagen, Valencia, CA, USA). The TaqMan genotyping assay for the *IGF1*-533 SNP (rs5742612, dbSNP build 124) was previously described.³⁶ The genotypes for the *IGF1*-2995 C/A SNP (rs12579108, dbSNP build 124) were determined using the fluorogenic 5'-nuclease assay.⁴² The oligonucleotide primers for amplification of the region surrounding the SNP were 5'-AGG AGT GGA TGT TCT TAT GAT AAG CA-3' and 5'-CTG CAA AAA TCT GAC AGT GAG ATA TG-3'. The fluorogenic allele-specific oligonucleotide probes (TaqMan MGB Probes; Applied Biosystems, Foster City, CA) to detect each of the alleles were: (1) G allele: VIC-labelled 5'-TCC TTG CGG TTA GCT A-3'; (2) 6-carboxy fluorescein (6FAM)-labelled 5'-ATC CTT GCT GTT AGC TAT-3'. (VIC is a proprietary dye, available from Applied Biosystems: $\lambda_{\text{ex}} = 488 \text{ nm}$, $\lambda_{\text{em}} = 552 \text{ nm}$.) PCR amplification using ~10 ng of genomic DNA was performed in a thermal cycler (MWG Biotech, High Point, North Carolina, USA) with an initial step of 95°C for 10 min followed by 95°C for 25 s and 50 cycles of 62°C for 60 s. The fluorescence profile of each well was measured in an ABI 7900HT Detection System and the results analysed with Sequence Detection Software (ABI). All analyses were performed blinded to case or control status. Twenty-four samples were analysed in duplicate, with a concordance rate of 100%.

Statistical methods

To assess strength of association between genotypes and cancer risk, unconditional multiple logistic regression models⁴³ were fit and odds ratio (ORs) and their corresponding 95% confidence intervals (CIs) were reported. All logistic regression models included age at recruitment (continuous), year of recruitment, gender and dialect group (Cantonese, Hokkien) as covariates. Colorectal cancer risk factors considered as potential confounders were body mass index, height, education levels, alcohol intake, smoking history, physical activity, calcium intake, and dietary fat intake. None were included in the final model because inclusion did not substantially alter (>5%) the parameter estimates for the exposures (genotypes).

Colorectal cancer was coded by anatomical subsite according to the *International Classification of Disease Oncology* (second edition) as colon (C18.0–C18.9) or rectal (C19.0–C20.0) cancer. To test for heterogeneity of odds ratios across anatomical subsite, polychotomous logistic regression models were fit and likelihood ratio tests were conducted.

To investigate possible effect modification by sex, measures of energy balance (body mass index and physical inactivity), and factors associated with serum IGF1 in this population (calcium intake, saturated fats, physical inactivity and vitamin intake),⁴⁴ the respective multiplicative interaction terms were included in the regression models and likelihood ratio tests were conducted.

Allele frequencies were determined by gene counting.⁴⁵ Departures from Hardy–Weinberg equilibrium (HWE) were assessed by testing the difference between the observed (sampled) and expected (under HWE) genotype frequencies in controls using a χ^2 test.⁴⁶ Linkage disequilibrium between *IGF1* polymorphisms was assessed by Lewontin's D' and R^2 .⁴⁷

Individuals were assigned a diplotype (pair of haplotypes) based on their combination of genotypes at the –2995 and –533 loci. Since two haplotypes accounted for nearly 100% of haplotype diversity, the phase was known with near certainty and double heterozygotes were assigned as heterozygotes for the two common haplotypes. Diplotypes were categorised into four groups: homozygotes for the common at-risk haplotype, heterozygotes for the two common haplotypes, homozygotes for the common protective haplotype, and all others (rare combinations). Odds ratios were estimated to compare each group to the baseline group (homozygotes for the protective haplotype).

We selected among competing statistical models using the smallest values for the Akaike Information Criteria (AIC)^{48, 49} as an indicator of global fit. AIC was calculated to compare model fit among these four models: (1) baseline covariate model (no genetic effects); (2) baseline model plus the -2995 locus only; (3) baseline model plus the -533 locus only, and (4) baseline model plus diplotype.

Serum IGF1 and IGFBP3 measurements

Of the 1176 controls in this study, 628 had serum IGF1 and 595 had serum IGF binding protein 3 (IGFBP3) measurements available from a previous sero-epidemiological study in which 628 controls were randomly sampled from the entire cohort of nearly 60 000 subjects (described previously by Probst-Hensch and colleagues⁴⁴). Of these 628, twelve had no genotype data available. Analysis of covariance was used to compare distributions of natural log-transformed IGF1 and IGFBP3-adjusted IGF1 by genotype categories. Age, sex, body mass index, dialect group and year of recruitment were also considered as potential confounders but were not reported as inclusion did not substantially alter (>5%) the parameter estimates.

All p-values are two-sided and statistical analyses were performed using STATA 10.0.

RESULTS

Study characteristics

The baseline characteristics of the Singapore Chinese Health Study cohort have been previously described.³⁶ For the case-control study nested within the cohort, table 1 describes the distributions of demographic and risk factors by case-control status. Cases were more often male, were heavier, slightly but significantly taller (2 cm), less educated, and more likely to be smokers.

Identification of evolutionary conserved regions and polymorphisms

We identified four ECRs in the region upstream of the *IGF1* translational start site (fig 1). Locations are: nt: 1587–2053 in NCBI#S85346 for ECR1; nt 82 871–83 140, 81 301–81 600, and 80 831–81 100 in NCBI#AC010202 for ECR2, ECR3 and ECR4.

We previously identified two SNPs (IGF1 -533C/T and -484G/A) in ECR1 in complete linkage disequilibrium with each other.³⁶ In this study, we identified one additional common SNP (minor allele frequency, MAF = 0.35), IGF1 -2995G/A (rs12579108, dbSNP build 127), which has been reported in the Perlegen polymorphism database to be common in Han Chinese (MAF = 0.42) and Japanese but rare in European-Americans and African-Americans.

Association between IGF1 genotype and colorectal cancer

Table 2 presents genotype distributions for the -2995C/A SNP in Singapore Chinese cases and controls. Genotype distributions did not deviate from Hardy-Weinberg expectations among cohort controls. Possession of one or two copies of the A allele (genotypes AA and CA) conferred an approximate 40% decrease in risk in comparison to genotype CC. This association was stronger for colon than for rectal cancer (heterogeneity $p < 0.001$).

Effect modification by energy balance

We examined two measures of energy balance as potential modifiers of the *IGF1* genotype-cancer risk association: (1) physical inactivity (no weekly vigorous or strenuous sports and ≥ 9 h of sitting/day), and (2) body mass index. The effect of *IGF1* -2995C/A genotype was approximately 2-fold stronger in individuals who were physically active (table 3; $p_{\text{interaction}} = 0.05$). No effect modification was detected by body mass index. In addition, no effect

modification was observed for the dietary factors that were previously reported to predict serum IGF1 levels in this population:⁴⁴ calcium intake, saturated fats or vitamin intake (data not shown).

Unphased multilocus genotype analysis (IGF1 –533C/T and IGF1 –2995C/A)

We previously reported an association between colorectal cancer risk and a SNP downstream from the –2995 locus, *IGF1* –533C/T.³⁶ The two SNPs are in tight linkage disequilibrium ($D' = 0.97$ and $R^2 = 0.90$). Thus, in this population, only two common –2995/–533 haplotypes exist, C_T and A_C, having estimated frequencies of 0.66 and 0.33, respectively. Odds ratios for the single locus and diplotype models were similar (table 4). The best fitting model, as indicated by the lowest AIC (table 5), was the model containing baseline covariates plus the –2995 genotype.

IGF1 genotype–phenotype (serum IGF1) association

IGF1 genotype did not predict serum IGF1 levels among controls ($p = 0.15$). The geometric mean values for the –2995CC, CA and AA genotypes were 122.6, 131.1 and 129.7 ng/ml respectively. Adjustment for IGFBP3 did not alter the results.

DISCUSSION

This is the first report of the characterisation of the allelic architecture surrounding the *IGF1* promoter region in relation to cancer risk. A putative regulatory *IGF1* SNP in the promoter region is associated with reduced colorectal cancer risk of approximately 40%. The effect is stronger in colon cancer versus rectum cancer and is more apparent in persons who are physically active.

The human *IGF1* gene spans approximately 90 kb and encodes five exons, two alternative first exons (exon 1 and exon 2) and four coding exons. Two common IGF1 splice mRNA variants arise from two *IGF1* promoters. No common sequence variation (frequency >0.02) was observed in the coding regions of the *IGF1* gene. Thus, cis-acting genetic variants that affect IGF1 levels are probably in the regulatory regions, ie, upstream enhancers, promoter, introns or 3' untranslated regions.

Little is known about regulatory *IGF1* polymorphisms and their impact on cancer risk. The only promoter region variant that has been studied is a cytosine–adenosine dinucleotide [*IGF1* (CA)_n] repeated 15–22 times, approximately 1 kb upstream of the first transcription start site of the more common splice variant. This *IGF1*(CA)_n microsatellite is associated with colorectal cancer in the Singapore Chinese (this population) as well as in some^{50, 51} but not all Caucasian populations.⁵² In those studies reporting an association, different alleles were associated with risk. The hypermutable nature of the microsatellite may result in an allele being linked to various haplotypic backgrounds. Thus we considered the possibility that these inconsistent associations are due to linkage disequilibrium with another functional polymorphic variant.

To dissect the haplotype block for these functional SNPs, we identified and re-sequenced regions with a high probability to harbour functional elements. The two SNPs identified in our previous report³⁶ are in a conserved region that covers the 5' untranslated region (ECR1, fig 1). The functionality of these SNPs is unknown. Upon re-sequencing the remaining conserved DNA stretches (ECR2–4, fig 1), we identified one additional SNP, *IGF1* –2995C/A. Based on transcription factor binding site prediction algorithms, this *IGF1* –2995C/A resides in a consensus domain for a transcription factor: octamer binding factor (Oct1/Oct2). The A allele (protective allele) destroys the predicted binding site for Oct1/Oct2. Thus, *in silico* prediction suggests that the –2995C (risk) allele is more likely than the –2995A (protective) allele to bind

octamer binding factor proteins. Oct2 proteins are found mainly in cells of the lymphoid system, and are reported to be over-expressed in colon tumours (curated database repository of gene expression profiles, NCBI Gene Expression Omnibus⁵³).

IGF1 is thought to mediate the effect of energy balance on cancer. We previously reported that high physical activity was associated with decreased serum IGF1 in this population.³⁰ The *IGF1* -2995C/A effect on colorectal cancer is more apparent in physically active as compared to physically inactive persons. In agreement with our observations, among highly active individuals in a large (1346 colon cancer cases and 1544 controls) predominantly Caucasian study, a 40% reduction in colon cancer risk was observed with the *IGF1*(CA)_n 192/192 genotype (interaction $p = 0.01$),⁵¹ which is in partial linkage disequilibrium with SNPs in the *IGF1* promoter. Neither study detected effect modification by BMI, another measure of energy balance that may not be a sensitive indicator of insulin sensitivity.⁵⁴

The association between *IGF1* promoter genotype and colon cancer observed in our study might be expected to act through plasma/serum level changes. Further, if the circulating IGF1 levels are considered as an intermediate phenotype, the proportion of variance due to the *IGF1* SNP is expected to be larger for serum IGF1 levels as compared to cancer incidence rates. However, in our study, the -2995C/A SNP did not predict serum IGF1 levels in 628 controls. Serum IGF1 concentration parallels hepatic IGF1 expression but does not correlate with non-hepatic expression.⁵⁵ Serum levels may not be a good marker of local tissue exposure. Rather, *IGF1* genotypes may be a better marker for colorectal cancer risk.

Using association methods, we could not definitively distinguish between several causal models (table 5). Differences between the models, as assessed by information criteria, are small. The best fitting model contains a single SNP at position -2995, which is located in an evolutionarily conserved sequence and is predicted to alter the binding site of a transcription factor that appears to be upregulated in colorectal cancer. Further functional studies of this polymorphism are warranted.

In any case-control study, the potential for population stratification exists. The Singapore Chinese Health Study is a population-based cohort comprised of subjects drawn from an ethnically homogeneous southern Chinese population (Han Chinese). Specifically, this population originates from the contiguous coastal provinces, Fujian and Guangdong, and was shown to cluster tightly into a genetically homogeneous subcluster within the relatively genetically similar southern Chinese population.^{56, 57} Therefore, we believe that population stratification is unlikely to explain our observed genetic association between IGF1 and colorectal cancer.⁵⁸ One limitation of our study is that it is based on a case-control that was sampled from an ongoing cohort with differential clinical stage between cases who provided biospecimens (46.8% with stage 3 and 4 cancer) and those who did not give biospecimens (62.0%). However, we do not have any strong evidence that cancer stage or mortality differs by *IGF1* genotypes. Another limitation is the relatively modest sample size. However, the ongoing cohort will continue to accrue cases and we will eventually be able to confirm or refute our finding. In the meantime, we believe that the finding of a new, potentially functional SNP in Asian populations warrants independent replication in other Han Chinese population.

Several factors may explain why our study is the only one to date to report an association between risk of colorectal cancer and an *IGF1* SNP. The SNPs examined in this study were chosen for high prior probability of functionality. They also happen to be rare in Caucasians but common in Chinese. Furthermore, IGF1 has been reported to be differentially expressed by ethnicity, with a 2.2-fold higher expression in Asians than in Caucasians.⁵⁹ Thus the association between *IGF1* genetic variation and colorectal cancer risk may have been more readily detectable in this cohort of Han Chinese than in some other studies. However, the

implication of this study, that variation in IGF1 expression is related to risk of colorectal cancer, may be more broadly generalisable. What remain to be determined are the genetic and non-genetic determinants of IGF expression that are most relevant in other populations.

Acknowledgments

We thank Siew-Hong Low of the National University of Singapore for supervising the field work of the Singapore Chinese Health Study, and Kazuko Arakawa of the University of Minnesota for the development and management of the cohort study database. We also thank the Ministry of Health in Singapore for assistance with the identification of cancer cases via database linkages.

Funding: This work was supported by grants R01 CA55069, R35 CA53890, R01 CA80205 and R01 CA98497 from the United States National Cancer Institute, Bethesda, Maryland. H-LW was supported by the National Institute of Health-Oak Ridge Institute for Science and Education fellowship.

References

1. Potter JD. Colorectal cancer: molecules and populations. *J Natl Cancer Inst* 1999;91:916–32. [PubMed: 10359544]
2. Komninou D, Ayonote A, Richie JP Jr, et al. Insulin resistance and its contribution to colon carcinogenesis. *Exp Biol Med* (Maywood) 2003;228:396–405. [PubMed: 12671184]
3. McKeown-Eyssen G. Epidemiology of colorectal cancer revisited: are serum triglycerides and/or plasma glucose associated with risk? *Cancer Epidemiol Biomarkers Prev* 1994;3:687–95. [PubMed: 7881343]
4. Giovannucci E. Insulin and colon cancer. *Cancer Causes Control* 1995;6:164–79. [PubMed: 7749056]
5. Heald A, Sharma R, Anderson S, et al. Dietary intake and the insulin-like growth factor system: effects of migration in two related populations in India and Britain with markedly different dietary intake. *Public Health Nutr* 2005;8:620–7. [PubMed: 16236191]
6. Lahm H, Suardet L, Laurent PL, et al. Growth regulation and co-stimulation of human colorectal cancer cell lines by insulin-like growth factor I, II and transforming growth factor alpha. *Br J Cancer* 1992;65:341–6. [PubMed: 1558785]
7. Whithead RH, Nice EC, Lloyd CJ, et al. Detection of colonic growth factors using a human colonic carcinoma cell line (LIM1215). *Int J Cancer* 1990;46:858–63. [PubMed: 2228314]
8. Tricoli JV, Rall LB, Karakousis CP, et al. Enhanced levels of insulin-like growth factor messenger RNA in human colon carcinomas and liposarcomas. *Cancer Res* 1986;46:6169–73. [PubMed: 3779638]
9. Hague A, Hicks DJ, Bracey TS, et al. Cell–cell contact and specific cytokines inhibit apoptosis of colonic epithelial cells: growth factors protect against c-myc-independent apoptosis. *Br J Cancer* 1997;75:960–8. [PubMed: 9083330]
10. Remacle-Bonnet MM, Garrouste FL, Heller S, et al. Insulin-like growth factor-I protects colon cancer cells from death factor-induced apoptosis by potentiating tumor necrosis factor alpha-induced mitogen-activated protein kinase and nuclear factor kappaB signaling pathways. *Cancer Res* 2000;60:2007–17. [PubMed: 10766192]
11. Guo YS, Jin GF, Houston CW, et al. Insulin-like growth factor-I promotes multidrug resistance in MCLM colon cancer cells. *J Cell Physiol* 1998;175:141–8. [PubMed: 9525472]
12. Sekharam M, Zhao H, Sun M, et al. Insulin-like growth factor 1 receptor enhances invasion and induces resistance to apoptosis of colon cancer cells through the Akt/Bcl-x(L) pathway. *Cancer Res* 2003;63:7708–16. [PubMed: 14633695]
13. Fukuda R, Hirota K, Fan F, et al. Insulin-like growth factor 1 induces hypoxia-inducible factor 1-mediated vascular endothelial growth factor expression, which is dependent on MAP kinase and phosphatidylinositol 3-kinase signaling in colon cancer cells. *J Biol Chem* 2002;277:38205–11. [PubMed: 12149254]
14. Koenuma M, Yamori T, Tsuruo T. Insulin and insulin-like growth factor 1 stimulate proliferation of metastatic variants of colon carcinoma 26. *Jpn J Cancer Res* 1989;80:51–8. [PubMed: 2540132]

15. Reinmuth N, Fan F, Liu W, et al. Impact of insulin-like growth factor receptor-I function on angiogenesis, growth, and metastasis of colon cancer. *Lab Invest* 2002;82:1377–89. [PubMed: 12379772]
16. Wu Y, Yakar S, Zhao L, et al. Circulating insulin-like growth factor-I levels regulate colon cancer growth and metastasis. *Cancer Res* 2002;62:1030–5. [PubMed: 11861378]
17. Hirose Y, Hata K, Kuno T, et al. Enhancement of development of azoxymethane-induced colonic premalignant lesions in C57BL/KsJ-db/db mice. *Carcinogenesis* 2004;25:821–5. [PubMed: 14729596]
18. Hassan AB, Howell JA. Insulin-like growth factor II supply modifies growth of intestinal adenoma in Apc(Min/+) mice. *Cancer Res* 2000;60:1070–6. [PubMed: 10706126]
19. Bjorndahl M, Cao R, Nissen LJ, et al. Insulin-like growth factors 1 and 2 induce lymphangiogenesis in vivo. *Proc Natl Acad Sci U S A* 2005;102:15593–8. [PubMed: 16230630]
20. D'Ercole AJ, Stiles AD, Underwood LE. Tissue concentrations of somatomedin C: further evidence for multiple sites of synthesis and paracrine or autocrine mechanisms of action. *Proc Natl Acad Sci U S A* 1984;81:935–9. [PubMed: 6583688]
21. Platz EA, Pollak MN, Rimm EB, et al. Racial variation in insulin-like growth factor-1 and binding protein-3 concentrations in middle-aged men. *Cancer Epidemiol Biomarkers Prev* 1999;8:1107–10. [PubMed: 10613344]
22. Girgis R, Abrams SA, Castracane VD, et al. Ethnic differences in androgens, IGF-I and body fat in healthy prepubertal girls. *J Pediatr Endocrinol Metab* 2000;13:497–503. [PubMed: 10803867]
23. DeLellis K, Ingles S, Kolonel L, et al. IGF1 genotype, mean plasma level and breast cancer risk in the Hawaii/Los Angeles multiethnic cohort. *Br J Cancer* 2003;88:277–82. [PubMed: 12610514]
24. Cruickshank JK, Heald AH, Anderson S, et al. Epidemiology of the insulin-like growth factor system in three ethnic groups. *Am J Epidemiol* 2001;154:504–13. [PubMed: 11549555]
25. Pastinen T, Hudson TJ. Cis-acting regulatory variation in the human genome. *Science* 2004;306:647–50. [PubMed: 15499010]
26. Ma J, Pollak MN, Giovannucci E, et al. Prospective study of colorectal cancer risk in men and plasma levels of insulin-like growth factor (IGF)-I and IGF-binding protein-3. *J Natl Cancer Inst* 1999;91:620–5. [PubMed: 10203281]
27. Nomura AMSG, Lee J, Pollak MN. Serum insulin-like growth factor I and subsequent risk of colorectal cancer among Japanese–American men. *Am J Epidemiol* 2003;158:423–31.
28. Giovannucci E, Pollak MN, Platz EA, et al. A prospective study of plasma insulin-like growth factor-1 and binding protein-3 and risk of colorectal neoplasia in women. *Cancer Epidemiol Biomarkers Prev* 2000;9:345–9. [PubMed: 10794477]
29. Palmqvist R, Hallmans G, Rinaldi S, et al. Plasma insulin-like growth factor 1, insulin-like growth factor binding protein 3, and risk of colorectal cancer: a prospective study in northern Sweden. *Gut* 2002;50:642–6. [PubMed: 11950809]
30. Probst-Hensch NM, Yuan JM, Stanczyk FZ, et al. IGF-1, IGF-2 and IGFBP-3 in prediagnostic serum: association with colorectal cancer in a cohort of Chinese men in Shanghai. *Br J Cancer* 2001;85:1695–9. [PubMed: 11742490]
31. Kaaks R, Toniolo P, Akhmedkhanov A, et al. Serum C-peptide, insulin-like growth factor (IGF)-I, IGF-binding proteins, and colorectal cancer risk in women. *J Natl Cancer Inst* 2000;92:1592–600. [PubMed: 11018095]
32. Renehan AG, Zwahlen M, Minder C, et al. Insulin-like growth factor (IGF)-I, IGF binding protein-3, and cancer risk: systematic review and meta-regression analysis. *Lancet* 2004;363:1346–53. [PubMed: 15110491]
33. Harrela M, Koistinen H, Kaprio J, et al. Genetic and environmental components of interindividual variation in circulating levels of IGF-I, IGF-II, IGFBP-1, and IGFBP-3. *J Clin Invest* 1996;98:2612–5. [PubMed: 8958225]
34. Kao PC, Matheny AP Jr, Lang CA. Insulin-like growth factor-I comparisons in healthy twin children. *J Clin Endocrinol Metab* 1994;78:310–2. [PubMed: 8106617]
35. Al-Zahrani A, Sandhu MS, Luben RN, et al. IGF1 and IGFBP3 tagging polymorphisms are associated with circulating levels of IGF1, IGFBP3 and risk of breast cancer. *Hum Mol Genet* 2006;15:1–10. [PubMed: 16306136]

36. Wong HL, Delellis K, Probst-Hensch N, et al. A new single nucleotide polymorphism in the insulin-like growth factor I regulatory region associates with colorectal cancer risk in Singapore Chinese. *Cancer Epidemiol Biomarkers Prev* 2005;14:144–51. [PubMed: 15668488]
37. Seow A, Yuan JM, Koh WP, et al. Diabetes mellitus and risk of colorectal cancer in the Singapore Chinese Health Study. *J Natl Cancer Inst* 2006;98:135–8. [PubMed: 16418516]
38. Hankin JH, Stram DO, Arakawa K, et al. Singapore Chinese Health Study: development, validation, and calibration of the quantitative food frequency questionnaire. *Nutr Cancer* 2001;39:187–95. [PubMed: 11759279]
39. Chia, KS.; Lee, HP.; Shanmugaratnam, K. Cancer incidence in Singapore 1993–1997. Singapore: Singapore Cancer Registry; 2000. Report No.5
40. Schwartz S, Kent WJ, Smit A, et al. Human-mouse alignments with BLASTZ. *Genome Res* 2003;13:103–7. [PubMed: 12529312]
41. Ovcharenko I, Nobrega MA, Loots GG, et al. ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. *Nucl Acids Res* 2004;32:W280–6. [PubMed: 15215395]
42. Lee LG, Connell CR, Bloch W. Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nucl Acids Res* 1993;21:3761–6. [PubMed: 8367293]
43. Breslow, NE.; Day, NE. Statistical methods in cancer research. Lyon: IARC; 1980.
44. Probst-Hensch NM, Wang H, Goh VH, et al. Determinants of circulating insulin-like growth factor I and insulin-like growth factor binding protein 3 concentrations in a cohort of Singapore men and women. *Cancer Epidemiol Biomarkers Prev* 2003;12:739–46. [PubMed: 12917205]
45. Li, CC. First course in population genetics. CA: The Boxwood Press, Pacific Grove; 1976.
46. Chiano MN, Clayton DG. Genotypic relative risks under ordered restriction. *Genetic Epidemiol* 1998;15:135–46.
47. Lewontin RC. On measures of gametic disequilibrium. *Genetics* 1988;120:849–52. [PubMed: 3224810]
48. Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, AC 1974;19:716–23.
49. Akaike, H. Canonical correlation analysis of time series and the use of an information criterion. In: Mehra, RKaLD., editor. *System identification: advances and case studies*. New York: Academic Press; 1976. p. 52-107.
50. Slattery ML, Murtaugh M, Caan B, et al. Energy balance, insulin-related genes and risk of colon and rectal cancer. *Int J Cancer* 2005;115:148–54. [PubMed: 15688407]
51. Morimoto LM, Newcomb PA, White E, et al. Insulin-like growth factor polymorphisms and colorectal cancer risk. *Cancer Epidemiol Biomarkers Prev* 2005;14:1204–11. [PubMed: 15894673]
52. Slattery ML, Samowitz W, Curtin K, et al. Associations among IRS1, IRS2, IGF1, and IGFBP3 genetic polymorphisms and colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 2004;13:1206–14. [PubMed: 15247132]
53. Barrett T, Troup DB, Wilhite SE, et al. NCBI GEO: mining tens of millions of expression profiles – database and tools update. *Nucl Acids Res* 2007;35L:D760–5. [PubMed: 17099226]
54. Alberti KG, Zimmet P, Shaw J. The metabolic syndrome – a new worldwide definition. *Lancet* 2005;366:1059–62. [PubMed: 16182882]
55. Iida K, Rosen CJ, Ackert-Bicknell C, et al. Genetic differences in the IGF-I gene among inbred strains of mice with different serum IGF-I levels. *J Endocrinol* 2005;186:481–9. [PubMed: 16135668]
56. Cavalli-Sforza LL. The Chinese human genome diversity project. *Proc Natl Acad Sci U S A* 1998;95:11501–3. [PubMed: 9751692]
57. Chu JY, Huang W, Kuang SQ, et al. Genetic relationship of populations in China. *Proc Natl Acad Sci U S A* 1998;95:11763–8. [PubMed: 9751739]
58. Deng HW, Chen WM, Recker RR. Population admixture: detection by Hardy–Weinberg test and its quantitative effects on linkage–disequilibrium methods for localizing genes underlying complex traits. *Genetics* 2001;157:885–97. [PubMed: 11157005]
59. Spielman RS, Bastone LA, Burdick JT, et al. Common genetic variants account for differences in gene expression among ethnic groups. *Nat Genet* 2007;39:226–31. [PubMed: 17206142]

Figure 1.

Conservation profiles of the human insulin-like growth factor-1 (IGF1) upstream regulatory region in reference to the mouse, rat and dog. The vertical axis for each panel indicates the percent identity. The human IGF1 DNA sequence lies along the horizontal axis. Coloured peaks depict regions of conservation with blue indicating exon 1, yellow indicating the 5'UTR and red indicating the upstream regulatory region. Green bars in the last panel indicate regions of repeat sequences. Evolutionarily conserved regions (exceed at 200 bp/80% identity threshold) are indicated by red bars. Black boxes delimit the regions that were re-sequenced for polymorphisms in 60 Han Chinese. ECR, evolutionary conserved region.

Table 1
Selected characteristics of colorectal cancer cases and controls nested within the Singapore Chinese Health Study

Characteristic	Controls (n = 1142) [*]	Cases (n = 298) [*]	p Value [†]
Sex			
Male	489 (42.8)	169 (56.7)	<0.001
Female	653 (57.2)	129 (43.3)	
Dialect group			
Cantonese	561 (49.1)	133 (44.6)	0.17
Hokkien	581 (51.9)	165 (55.4)	
Body mass index (kg/m ²)			
≤20	180 (15.8)	46 (15.4)	0.03
20–<24	638 (55.9)	141 (47.3)	
24–<28	261 (22.8)	90 (30.2)	
28+	63 (5.5)	21 (7.1)	
Height [‡]	159.3 (7.5)	160.9 (8.0)	0.02
Physical inactivity [§]			
Yes	942 (82.5)	247 (82.9)	0.87
No	200 (17.5)	51 (17.1)	
Education level			
None (formal)	305 (26.7)	93 (31.5)	0.003
High school	770 (67.5)	195 (65.2)	
Post-high school	44 (3.8)	6 (2.0)	
University	23 (2.0)	4 (1.4)	
Cigarette smoking (cigarette and/or water pipe)			
Never	829 (72.6)	178 (59.7)	<0.001
Former	127 (11.1)	49 (16.5)	
Current	186 (16.3)	72 (23.8)	
Frequency of alcohol consumption			
Non-drinkers	935 (81.9)	236 (79.2)	0.20
Monthly	85 (7.4)	19 (6.4)	
Weekly	89 (7.8)	26 (8.7)	
Daily	33 (2.9)	17 (5.7)	
Total calories (kcal/day)	1484.6 (831.4, 2483.5)	1491.0 (808.3, 2592.5)	0.72
Total calcium intake (mg/day)	373.5 (158.7, 858.1)	360.7 (153.7, 793.3)	0.53
Total fibre intake (g/day)	12.3 (5.1, 23.2)	11.9 (4.2, 22.9)	0.66
Total fat (g/day)	40.6 (19.8, 82.2)	39.8 (18.3, 77.5)	0.58

^{*} Number of subjects (%), median (5th percentile, 95th percentile) for skewed distribution (dietary factors) or mean (standard deviation) for normal distribution (height).

[†] Two-sided p-values derived from the Kruskal–Wallis test.

[‡] Physical inactivity: yes = no weekly vigorous or strenuous sports and ≥9 h of sitting/day; no = all others.

Table 2Odds ratios (ORs) and 95% confidence intervals (95% CIs) for the *IGF1* -2995C/A genotype and colorectal cancer

Genotype	Controls (%)	Cases (%)	OR (95% CI) [*]
CC	482 (42.2)	160 (53.7)	1.00 (referent)
CA	532 (46.6)	106 (35.6)	0.56 (0.42 to 0.75)
AA	128 (11.2)	32 (11.7)	0.68 (0.43 to 1.07)
CA + AA vs CC			0.59 (0.45 to 0.77)
Subsite			
Colon [†]			
CC	482 (42.2)	98 (56.3)	1.00 (referent)
CA	532 (46.6)	59 (33.9)	0.51 (0.36 to 0.73)
AA	128 (11.2)	17 (9.8)	0.59 (0.33 to 1.04)
CA + AA vs CC			0.53 (0.38 to 0.74)
Rectum [†]			
CC	482 (42.2)	62 (50.0)	1.00 (referent)
CA	532 (46.6)	47 (37.9)	0.65 (0.43 to 0.98)
AA	128 (11.2)	15 (12.1)	0.83 (0.45 to 1.53)
CA + AA vs CC			0.68 (0.47 to 1.00)
			p<0.001 [‡]

^{*} Odds ratio from unconditional logistic regression; adjusted for age at recruitment, gender, dialect groups (Cantonese, Hokkien), and year of recruitment.

[†] Odds ratios from polychotomous logistic regression; adjusted for age at recruitment, gender, dialect groups (Cantonese or Hokkien), and year of recruitment.

[‡] The p value for heterogeneity.

Table 3
Odds ratios (ORs) and 95% confidence intervals (95% CIs) for the *IGF1* -2995C/A genotype and colorectal cancer by physical activity

Genotype	Controls (%)	Cases (%)	OR (95% CI) [*]
Physically inactive [†]			
CC	400 (42.5)	127 (51.4)	1.00 (referent)
CA + AA	542 (57.5)	120 (48.6)	0.66 (0.49 to 0.89)
Physically active			
CC	82 (41.0)	33 (64.7)	1.00 (referent)
CA + AA	118 (59.0)	18 (35.3)	0.30 (0.15 to 0.60)
			p = 0.05 [‡]

^{*} Odds ratio from unconditional logistic regression; adjusted for age at recruitment, gender, dialect groups (Cantonese, Hokkien), and year of recruitment.

[†] No weekly vigorous or strenuous sports and ≥ 9 h of sitting/day.

[‡] The p-value for interaction.

Table 4

Diplotypes of *IGF1* loci -2995C/A and -533C/T and colorectal cancer risk

Genotype combination	Diplotype	Controls/cases [†]	# risk (CT)haplotypes	OR (95% CI) [‡]
-2995CC, -533TT	C_T/C_T	435/154	2	1.00
-2995CA, -533CT	C_T/A_C	448/98	1	0.61 (0.45 to 0.83)
-2995AA, -533CC	A_C/A_C	106/30	0	0.72 (0.45 to 1.15)
C_T/A_C + A_C/A_C		554/128	0-1	0.63 (0.48 to 0.83)
All others	All others	42/13	Unphased	0.73 (0.37 to 1.44)
				p = 0.01 [§]

* At loci -2995C/A and -533C/T

[†] Total of 1031 controls and 295 cases. One hundred and eleven controls and three cases had missing genotypes for IGF1 -533.

[‡] Odds ratio estimated from logistic regression, adjusted for age at recruitment, gender, dialect groups (Cantonese, Hokkien), and year of recruitment.

[§] Trend-test assessing the linear trend of the A_C haplotype dosage; ie, 0 vs 1 vs 2 copies of the A_C haplotype.

OR, odds ratio; CI, confidence interval.

Table 5

Comparison of the model fit among the single marker and diplotype models (294 cases and 1035 controls having complete genotype data)

Model	OR (95% CI)	Akaike's score [*]
Baseline model: baseline covariates [†]	Reduced model	0.970
Baseline model plus -533T/C [‡]	TC + CC vs TT 0.66 (0.50 to 0.86)	0.966
Baseline model plus -2995C/A [§]	CA + AA vs CC 0.61 (0.47 to 0.81)	0.963
Baseline model plus diplotypes [¶]	C_T/A_C + A_C/A_C vs C_T/C_T 0.63 (0.48 to 0.83)	0.966

CI, confidence interval; OR, odds ratio.

^{*} Akaike's information score.

[†] Age at recruitment, gender, dialect groups (Cantonese, Hokkien), and year of recruitment.

[‡] TC+CC vs TT: individuals with at least one copy of the risk allele -533C (genotypes TC or CC as a referent group) in comparison to individuals with genotype TT.

[§] CA + AA vs CC: individuals with at least one copy of the risk allele -2995A (genotypes CA and AA) versus individuals with genotype CC.

[¶] C_T/A_C + A_C/A_C vs C_T/C_T: individuals with at least one copy of the risk haplotype -533A_-2995C (haplotype pairs of -533C_-2995T/-533A_-2995C or haplotypes pairs of -533A_-2995C/-533A_-2995C) as a referent group in comparison to individuals with diplotype -533C_-2995T.