Citation: Clinical and Translational Gastroenterology (2014) 5, e50; doi:10.1038/ctg.2014.1

www.nature.com/ctg

Significant Evidence of Linkage for a Gene Predisposing to Colorectal Cancer and Multiple Primary Cancers on 22q11

Craig Teerlink, PhD[1], Quentin Nelson, MS[1], Randall Burt, MD[2,3] and Lisa Cannon-Albright, PhD[1,4]

OBJECTIVES: The genetic basis of colorectal cancer (CRC) is not completely specified. Part of the difficulty in mapping predisposition genes for CRC may be because of phenotypic heterogeneity. Using data from a population genealogy of Utah record linked to a statewide cancer registry, we identified a subset of CRC cases that exhibited familial clustering in excess of that expected for all CRC cases in general, which may represent a genetically homogeneous subset of CRC.
METHODS: Using a new familial aggregation method referred to as the subset genealogic index of familiality (subsetGIF), combined with detailed information from a statewide tumor registry, we identified a subset of CRC cases that exhibited excess familial clustering above that expected for CRC: CRC cases who had at least one other primary tumor at a different site. A genome-wide linkage analysis was performed on a set of high-risk CRC pedigrees that included multiple CRC cases with additional primaries to identify evidence for predisposition loci.
RESULTS: A total of 13 high-risk CRC pedigrees with multiple CRC cases with other primary cancers were identified. Linkage analysis identified one pedigree with a significant linkage signal at 22q11 (LOD (logarithm (base 10) of odds) = 3.39).
CONCLUSIONS: A predisposition gene or variant for CRC that also predisposes to other primary cancers likely resides on chromosome 22q11. The ability to use statewide population genealogy and tumor registry data was critical to identify an informative subset of CRC cases that is possibly more genetically homogeneous than CRC in general, and may have improved statistical power for predisposition locus identification in this study.
Clinical and Translational Gastroenterology (2014) **5,** e50; doi:10.1038/ctg.2014.1; published online 27 February 2014
Subject Category: Colon/Small Bowel

INTRODUCTION

The etiology of colorectal cancer (CRC) includes several well-established genetic factors,[1–4] yet it is likely that additional predisposition variants remain to be discovered. Although the influence of genetic susceptibility to CRC is well documented,[5–7] a central difficulty in the identification of the genetic factors influencing CRC is the inability to adequately cope with the phenotypic heterogeneity present in all complex diseases. To overcome this difficulty, a strategy to identify genetically homogenous subsets of CRC based on data stored in the Utah Cancer Registry (UCR) and linked to population genealogy records from the Utah Population Database (UPDB) was devised to identify subsets of CRC cases showing significantly more relatedness than expected for all CRC cases. It is hypothesized that genetic analysis of these homogeneous pedigrees can be informative for predisposition gene identification.

The genealogical index of familiality (GIF) method that tests for a significant excess of relatedness of a set of cases compared with sets of matched population controls[8] was modified. For the modification, the relatedness of the subset of CRC cases of interest was compared with matched controls selected from *all CRC cases*, rather than from the population. This subsetGIF method allows for prioritization of potential endotypes for prioritization of pedigrees and cases for genetic mapping studies. The endotypes explored were based on information about cancer characteristics at the time of diagnosis, such as stage and grade, as documented in UCR records.

Our approach identified one particular subset of CRC cases that exhibited a significant excess of familial clustering above that observed for CRC in general. The subset with the strongest evidence of increased familial clustering is CRC cases who also have at least one additional primary tumor at another cancer site. This subset of CRC cases may represent a more genetically homogeneous endotype of CRC; a study focus on these cases and pedigrees may be more statistically powerful for genetic mapping because of enhanced phenotype refinement.

To map predisposition loci contributing to CRC that present with multiple primary cancers, we identified informative high-risk CRC pedigrees from a previous study of over 270 Utah high-risk common CRC pedigrees who did not show patterns associated with hereditary nonpolyposis colorectal cancer. Each pedigree included at least three sampled CRC cases who each had at least one additional independent primary.

[1]Division of Genetic Epidemiology, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah, USA; [2]Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah, USA; [3]Division of Gastroenterology, Department of Medicine, University of Utah School of Medicine, Salt Lake City, Utah, USA and [4]George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, Utah, USA
Correspondence: Craig Teerlink, PhD, Division of Genetic Epidemiology, Department of Internal Medicine, University of Utah School of Medicine, 391 Chipeta Way, Suite D, Salt Lake City, Utah 84108, USA. E-mail: craig.teerlink@hsc.utah.edu
Received 13 November 2013; revised 15 November 2013; accepted 23 Decemeber 2013

In total, 96 cases in 13 such high-risk CRC pedigrees were selected. Genome-wide genotyping with dense single-nucleotide polymorphisms (SNPs) was performed in the 96 CRC already sampled cases in these pedigrees. Parametric linkage analysis identified statistically significant evidence for linkage at cytogenetic band 22q11.1.

METHODS

The Utah population database. The UPDB contains genealogical and demographic data representing the Utah population from the mid-1800s. The genealogy data have been record-linked to several statewide data resources including the UCR and vital statistics records. The genealogy records in the UPDB span up to 15 generations. The original genealogy was completed in the 1970s and included 1.6 million genealogy records for the Utah pioneers and their descendants.[8] The UPDB genealogy data have since been expanded to include current generations through the inclusion of vital statistics records. It now contains 9 million individual records, not all of which represent genealogy data. Analysis was restricted to the 1.3 million individuals in the UPDB who have at least 12 of their 14 immediate ancestors in order to mitigate biases that may be incurred from including people who have few relationships represented. The Utah population is genetically representative of Northern Europeans[9] and has the same (low) inbreeding levels as other areas of the United States.[10] The UCR is a statewide cancer registry that became a National Cancer Institute Surveillance, Epidemiology, and End-Results (SEER) registry in 1973 and has tracked the occurrence of all cancer cases occurring in Utah by law since 1966. All cases have histopathologic confirmation and only independent primary cancers are reported.

Identifying genetically homogeneous subsets of CRC. The GIF is a familial aggregation technique that can be used to measure the extent of familial clustering among a cohort of cases relative to expected level of relatedness as estimated in the UPDB population. The GIF statistic for a set of individuals is the average of the kinship coefficient estimated for each pair of cases in the set.[11] To perform the GIF analysis, the GIF average relatedness statistic is calculated for the set of cases of interest. Then, an empirical distribution of average relatedness is estimated from 1,000 sets of matched controls, matched on age (5-year birth cohort), sex, and birthplace (in or out of Utah). The distribution of average relatedness from the 1,000 matched controls sets represents the expected relatedness in the UPDB population and can be compared with the case GIF for an empirical test of the hypothesis of no excess relatedness in the set of cases of interest. Diseases with significant excess relatedness are more likely to have predisposing genetic factors that contribute to the observed familial clustering.

We hypothesized that it may be possible to identify a subset of CRC cases based on some clinically relevant characteristics that demonstrates a higher level of relatedness than all CRC cases. The SubsetGIF method was used to perform the analysis (Nelson *et al.*[12]). The SubsetGIF analysis is a modification of the GIF analysis that has an additional matching requirement that the controls are themselves CRC cases. The additional matching requirement removes confounding that may exist between the subset in question and familial excess that is due to the heritability of CRC more generally. For instance, in order to show that a subset of CRC cases with some characteristic has a heritable component, the subsetGIF statistic must exceed the GIF for all CRC cancer cases taken together, otherwise the observed clustering may simply be an artifact of the heritable nature of CRC itself. The SubsetGIF method was employed in order to identify subsets of CRC that are potentially more genetically homogeneous than CRC in general.

We analyzed several subsets of CRC cases defined by data available from tumor records in the UCR. Characteristics considered included age at diagnosis, presence of multiple primaries (either CRC or other primaries), stage at diagnosis, grade at diagnosis, survival months after diagnosis, and body mass index.

Pedigree identification. Since 1980, over 4,000 individuals were recruited and sampled in 272 Utah high-risk CRC pedigrees identified in the UPDB. A high-risk CRC pedigree is defined as a set of descendants of a founder in which there is a statistically significant excess of CRC cases observed compared with the expected number of CRC cases based on age-, sex-, and birth state-corrected rates of CRC estimated from the UPDB. Previously studied high-risk CRC pedigrees with at least three CRC cases with additional primary cancers were selected for analysis.

Genotyping. Study subjects were genotyped at the University of Utah Health Sciences Center DNA Sequencing and Genomics Core Facility on the Illumina 720K Omni-Express SNP platform (San Diego, CA, USA). Typical quality control measures were applied to all genotype data before analysis (removal of individuals with low call rates ($<98\%$ of genotypes called); exclusion of SNPs with low call rates ($<98\%$), with low minor allele frequencies ($<1\%$), without all 3 genotypes observed, or with significant deviation from Hardy–Weinberg equilibrium in controls ($P<0.01$)) identified using Plink software.[13]

The presence of linkage disequilibrium between markers in a linkage analysis can artificially inflate LOD (logarithm (base 10) of odds) scores because certain alleles are more frequently encountered than expected by chance that (falsely) increases the likelihood that they were inherited from a common ancestor. To remedy this, the set of 720 k SNPs was reduced to a non-linkage disequilibrium set before linkage analysis by removing SNPs that exceed an R^2 threshold of 0.1 and heterozygosity of >0.3. Previous analyses have demonstrated that this strategy results in an ideal set of $\sim$27,000 genome-wide SNPs for linkage without loss of information.[14]

Linkage analysis. Linkage analysis was conducted using a robust parametric multipoint LOD score, referred to as the TLOD.[15,16] The LOD score is essentially a likelihood ratio test comparing the probability that a trait is linked with a genetic marker or not. The LOD score has a known distribution and provides a signal showing the strength of

cosegregation of a trait and an observed genetic marker through a pedigree. The multipoint LOD score uses information from several markers at once to inform the likelihood estimation that linkage between the trait and the marker is more likely than independent segregation of both. However, the multipoint LOD score is typically underpowered in the presence of model parameter misspecification concerning the mode of inheritance or the sporadic rate.[16] The TLOD statistic is comparable to the multipoint LOD score in that it uses multipoint information to estimate the likelihood function at genetic marker loci, but is superior in that it also optimizes the likelihood function over the recombination fraction (theta), similar to conventional two-point linkage score. The inclusion of the additional parameter, theta, allows for the statistic to absorb model misspecification, particularly with regard to the sporadic rate. McLink software was used for linkage analyses, a package specifically designed to analyze extended pedigrees and incorporating the TLOD statistic.[17] A statistically powerful general parametric modeling strategy including a dominant and a recessive model was pursued.[18] The conventional thresholds for interpreting LOD scores were used (3.3 for significant and 1.89 for suggestive).[19] In an analysis of all pedigrees simultaneously (assuming a common predisposition locus exists among multiple pedigrees), evidence was combined across pedigrees using the heterogeneity LOD score function[20] applied to the TLOD statistic (het-TLOD). High het-TLOD scores are an indication that multiple pedigrees are contributing evidence for linkage at a given genetic locus. Each pedigree was assumed to be singly informative for linkage and was also analyzed separately.

This study was approved by the University of Utah Institutional Review Board as of 1996.

RESULTS

Identifying genetically homogeneous subsets of CRC. There were 8,277 CRC cases with genealogy data in the UPDB. Traditional GIF analysis of all CRC cases showed evidence of significant excess relatedness compared with the Utah populations ($P<0.001$, Table 1). Traditional GIF analysis of the various CRC subsets considered also concluded that most of the subsets of cases exhibited relatedness in excess of expected when they were compared with randomly matched controls from the UPDB (Table 1, traditional GIF *P* value). These traditional GIF results do not allow discrimination of whether any subset is more valuable for predisposition gene identification.

Subsets of CRC cases selected were based on available data concerning various characteristics available from tumor records in the UCR. The subsets of CRC cases that were analyzed and the available sample sizes are shown in Table 1.

The SubsetGif analysis identified several CRC subsets with significant evidence for excess relatedness, including early diagnosis ($P=0.001$), CRC and other cancer primaries ($P=0.002$), and multiple independent CRC primaries ($P=0.002$). Both early diagnosed CRCs and multiple independent CRC primaries have previously been suggested as characteristics associated with CRCs more likely to be due to an inherited predisposition.[21–23] However, the subset of CRC cases with at least one primary cancer of another site has not been previously identified to be of interest. These findings indicate that this subset of CRC cases are observed to cluster more than expected in relatives.

Table 1 GIF results for subsets of CRC

Category of CRC cases	*n*	Traditional GIF *P* value	SubsetGIF *P* value
All CRC[a]	8,277	<0.001	NA
Early diagnosis (<50 years)	682	<0.001	0.001
Distant stage at diagnosis	1,527	<0.001	0.35
Grade at diagnosis (3 or 4)	1,260	0.121	0.939
CRC and ≥1 primary cancer of another site	1,549	<0.001	0.002
Multiple independent primary CRCs	270	<0.001	0.003
Long survival (>240 months)	641	0.004	0.107
Short survival (<10 months)	1,918	0.009	0.964
Under and normal weight (BMI <25 kg/m^2)	1,328	0.002	0.287
Obese (BMI >30 kg/m^2)	858	<0.001	0.073

BMI, body mass index; CRC, colorectal cancer; GIF, genealogic index of familiality; NA, not applicable.
[a]For the analysis of all CRC cases, control sets were selected from the population. For all other subsets, controls were selected from the set of all CRC cases (SubsetGIF method).

Table 2 Characteristics of 13 pedigrees containing a significant excess of colorectal cancer (CRC) and including multiple CRC cases with at least one other primary tumor at another site

No. of CRC cases with at least one other primary[a]	Other cancer sites observed in CRC cases diagnosed with at least one other primary	Total no. of CRC cases	Total no. of genotyped cases
3	Breast, lip, stomach	6	5
3	Prostate	5	5
3	Breast, prostate, melanoma	6	5
3	Prostate, stomach	5	5
3	Prostate, thyroid, lip	8	8
3	Breast, prostate	13	11
4	Lip, prostate, stomach	8	8
3	Breast, prostate	5	5
3	Breast, prostate, lymphoma	5	3
3	Prostate, lymphoma, stomach	3	3
3	Breast, lymphoma, thyroid	3	3
3	Breast, gallbladder, bladder	3	3
4	Breast, prostate, lymphoma	4	2

[a]Not all CRC cases with at least one other primary tumor had samples available for genotyping.

Selection of study participants for genotyping. A total of 13 pedigrees met the inclusion criteria of exhibiting a statistical excess of CRC and having at least 3 previously sampled CRC cases with other primary tumors. Details about

these pedigrees are given in Table 2. A total of 9 different primary tumors were observed, with the 2 most common cancers in the UPDB, prostate cancer and breast cancer, occurring in 10 and 8 of the pedigrees, respectively.

Linkage results. The results of the genome-wide linkage analysis for all pedigrees using the het-TLOD statistic and dominant and recessive general models are graphically depicted in Figure 1. No het-TLOD scores exceeded the threshold for suggestive evidence for linkage (het-TLOD >1.9), showing no significant evidence for multiple pedigrees linked to a specific region. Because many Utah pedigrees are singly informative for linkage, we also considered evidence for linkage by pedigree.

In the analysis of individual pedigrees, one pedigree achieved genome-wide significance at 22q11.1 (TLOD = 3.39). A 1-LOD unit support interval defines a region of interest from 4.1 to 15.3 cM (17.7 to 21.4 Mbp). Figure 2 shows this pedigree, including the affected pedigree members who share a haplotype at the linked locus for the pedigree, and indicates which CRC cases have been diagnosed with additional primaries of a different site. There were no other linked pedigrees (TLOD >0.588) in the region.

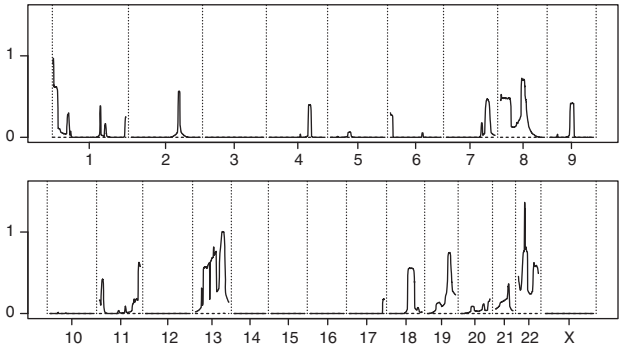

Figure 1 Genome-wide heterogeneity-TLOD scores for all pedigrees combined for general dominant (solid line) and recessive (dashed line) models.

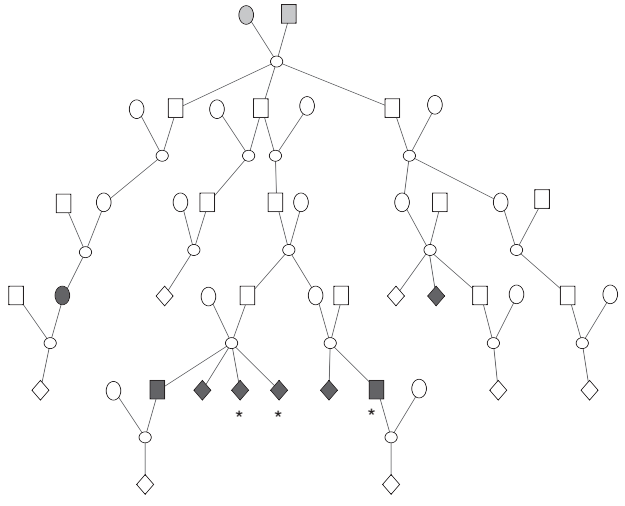

Figure 2 Representation of the pedigree with significant evidence for linkage at chromosome 22q11. Solid filled nodes indicate haplotype carriers diagnosed with colorectal cancer (CRC), and those with an asterisk indicate the CRC cases with primary tumors at other cancer sites. Sex has been intentionally obscured to prevent identification.

To determine whether the CRC cases in the 13 high-risk subset pedigrees with an excess of multiple primary cancers at other sites were similar to all CRC cases in the UPDB ($n=8{,}277$), we compared the two groups for several characteristics including body mass index, stage, grade, age at diagnosis, and survival. The CRC cases in the subset pedigrees had a statistically shorter survival (mean subset survival = 78 months; mean CRC survival = 103 months; *t*-test *P* value = 4E − 6), and a statistically younger age at diagnosis (mean subset age at diagnosis = 66.3 years; mean CRC age at diagnosis = 69.3 years; *t*-test *P* value = 4E − 5), but were not different with respect to stage, grade, or body mass index. The 8 CRC cases in the linked pedigree had an average age at diagnosis of 61 years ($P=0.06$ compared with all cases); average survival of 191 months ($P=0.24$ compared with all cases); and 2/8 cases classified as overweight, obese, or morbidly obese ($P=0.26$ compared with the distribution for all cases). Survival time after CRC diagnosis and age at diagnosis of CRC might not be independent of the presence of multiple independent primaries.

DISCUSSION

The genetic contribution to CRC is well recognized and many high-risk CRC pedigree studies have been performed. We hypothesize that genetic heterogeneity, even within high-risk pedigrees, could have added considerable noise to any signal of linkage. We present a method identifying subsets of CRC cases with the most evidence for excess familial clustering. This new approach has identified homogeneous informative high-risk CRC pedigrees on which to focus gene identification studies. Analysis of these pedigrees provides significant evidence for a CRC predisposition gene on chromosome 22 that also predisposes to primary cancers of other sites.

The ability to identify such cohorts requires the availability of several key data types in one resource: population genealogy records (such as that available from the UPDB), rich phenotype data (such as that available from the UCR), and a large cohort of sampled cancer cases and their relatives such as is available in the Utah family study. The selection of high-risk CRC pedigrees with a statistical excess of a potentially more genetically homogeneous subset of cases can be extended to other phenotype settings in this and other similar resources.

The significantly linked region of chromosome 22q11 has been previously implicated in contributing to metastasis of colorectal cancer, usually observed as chromosomal rearrangements.[24–29] Furthermore, one previous study identified nominal evidence for CRC linkage to chromosome 22q11 in the Swedish population.[30]

The 1-LOD support interval for the linkage evidence on chromosome 22q11 spans a gene-rich 3.7 Mb region of the genome. The region contains several genes that are known to be involved in the regulation of the cell division cycle, including

CDC45 (*cell division cycle homolog 45*, which is required for DNA replication[31]), *SEPT5* (*septin 5*, the disruption of which is shown to produce polyploid cells[32]), and *GNB1L* (*guanine nucleotide binding protein*, which has been shown to be involved in many cellular regulatory activities including cell cycle progression, gene regulation, and apoptosis[33]). The region also contains several tumor suppressors including BID (a cell death activator that has been implicated in colon cancer[34]), BCL2L13 (an apoptosis facilitator implicated in the progression of leukemia[35]), and AIFM3 (apoptosis-inducing factor associated with mitochondrial function[36]). The region also contains a potential oncogene, CRKL,[37] and a transcription coactivator of RNA polymerase II, MED15,[38] both of which have been proposed as contributing to the progression of various cancers.

Several genes spread across the region of interest are related to DiGeorge Syndrome, the features of which include arrested cardiac, craniofacial, and mental development. DiGeorge Syndrome typically results from constitutional rearrangements at the 22q11 locus. There are several reported associations of DiGeorge syndrome with various cancers.[39–42]

This identification of linkage evidence for a new CRC predisposition locus that includes predisposition for other cancers, to date observed in 1 of 13 pedigrees (8%), could lead to identification of new genes or variants responsible for CRC and other cancers. Although it is always likely that environmental effects have contributed to the familial clustering observed, the evidence for the contribution of a genetic contribution to the subset of CRC observed in this study is significant. These results warrant further investigation of this locus to identify the responsible causal variants. The results also support the general approach used here of identifying homogeneous subsets and restricting analysis to a limited set of informative pedigrees.

CONFLICT OF INTEREST

Guarantor of the article: Lisa Cannon-Albright, PhD.
Specific author contributions: Craig Teerlink and Quentin Nelson: data analysis and manuscript preparation; Randall Burt; interpretation of study outcomes and study supervision; Lisa Cannon-Albright: sample ascertainment, study design, interpreting data, and supervision of data analysis.
Financial support: This study was funded by the Huntsman Cancer Institute Cancer Control and Population Sciences pilot award program (CCT). This research was supported by the Utah Cancer Registry that is funded by Contract Number HHSN261201000026C from the National Cancer Institute's SEER Program with additional support from the Utah State Department of Health and the University of Utah. Partial support for all data sets within the Utah Population Database (UPDB) was provided by Huntsman Cancer Institute, University of Utah, and the Huntsman Cancer Institute's Cancer Center Support grant, P30 CA42014, from National Cancer Institute. L.C.-A. acknowledges support from the Huntsman Cancer Foundation. Study sponsors had no role in the study design, collection, analysis, and interpretation of the data, and in the writing of the report.
Potential competing interests: None.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- The genetic factors influencing colon cancer are not completely known.

WHAT IS NEW HERE

- Significant statistical evidence suggests that a genetic predisposition locus exists on chromosome 22q11 that predisposes to colorectal cancer as well as other primary cancers.

1. De la Chapelle A. Genetic predisposition to colorectal cancer. *Cancer* 2004; **4**: 769–780.
2. Lichtenstein P, Holm N, Verkasalo P. Environmental and heritable factors in the causation of cancer- analyses of cohorts of twins from Sweden, Denmark and Finland. *N Engl J Med* 2000; **343**: 78–85.
3. Grady WM. Genetic testing for high-risk colon cancer patients. *Gastroenterology* 2003; **124**: 1574–1594.
4. Jasperson KW, Tuohy TM, Neklason DW *et al.* Hereditary and familial colon cancer. *Gastroenterology* 2010; **138**: 2044–2058.
5. Teerlink CC, Albright FS, Lins L *et al.* A comprehensive survey of cancer risks in extended families. *Genet Med* 2012; **14**: 107–114.
6. Taylor DP, Burt RW, Williams MS *et al.* Population-based family history-specific risks for colorectal cancer: a constellation approach. *Gastroenterology* 2010; **138**: 877–885.
7. Cannon-Albright LA, Thomas TC, Bishop DT *et al.* Characteristics of familial colon cancer in a large population data base. *Cancer* 1989; **64**: 1971–1975.
8. Skolnick MH. The Utah genealogical database: a resource for genetic epidemiology. In: Cairns J, Lyon JL, Skolnick M (eds). *Banbury Report No. 4: Cancer Incidence in Defined Populations*. Cold Spring Harbor Laboratories: New York, 1980, pp 285–297.
9. McLellan T, Jorde LB, Skolnick MH. Genetic distances between the Utah Mormons and related populations. *Am J Hum Genet* 1984; **36**: 836–857.
10. Jorde LB. Inbreeding in the Utah Mormons: an evaluation of estimates based on pedigrees, isonomy, and migration matrices. *Ann Hum Genet* 1989; **53**: 339–355.
11. Skolnick MH, Bishop DT, Carmelli D *et al.* A population-based assessment of familial cancer risk in Utah Mormon genealogies. In: Arrighi FE, Rao PN, Stubblefield E (eds). *Genes, Chromosomes and Neoplasia*. Raven Press: New York, 1981, pp 477–500.
12. Nelson Q, Agarwal N, Stephenson R *et al.* A population-based analysis of clustering identifies a strong genetic contribution to lethal prostate cancer. *Front Genet* 2013; **4**: 152.
13. Purcell S, Neale B, Todd-Brown K *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
14. Allen-Brady K, Norton PA, Farnham JM *et al.* Significant linkage evidence for a predisposition gene for pelvic floor disorders on chromosome 9q21. *Am J Hum Genet* 2009; **84**: 678–682.
15. Abkevich V, Camp NJ, Gutin A *et al.* A robust multipoint linkage statistic (tlod) for mapping complex trait loci. *Genet Epidemiol* 2001; **21**(Suppl 1): S492–S497.
16. Göring HH, Terwilliger JD. Linkage analysis in the presence of errors II: marker-locus genotyping errors modeled with hypercomplex recombination fractions. *Am J Hum Genet* 2000; **66**: 1107–1118.
17. Thomas A, Gutin A, Abkevich V *et al.* Multilocus linkage analysis by blocked Gibbs sampling. *Stat Comput* 2000; **10**: 259–269.
18. Risch N, Claus E, Giuffra L. Linkage and mode of inheritance in complex traits. *Prog Clin Biol Res* 1989; **329**: 183–188.
19. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; **11**: 241–247.
20. Ott J. Linkage probability and its approximate confidence interval under possible heterogeneity. *Genet Epidemiol* 1986; **3**: S251–S257.
21. Groden J, Thliveris A, Samowitz W *et al.* Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* 1991; **66**: 589–600.
22. Lynch HT, Guirgis HA, Lynch PM *et al.* Familial cancer syndromes: a survey. *Cancer* 1977; **39**: 1867–1881.
23. Lynch HT, Krush AJ. Cancer family 'G' revisited: 1895–1970. *Cancer* 1971; **27**: 1505–1511.
24. González-González M, Muñoz-Bellvis L, Mackintosh C *et al.* Prognostic impact of del(17p) and del(22q) as assessed by interphase FISH in sporadic colorectal carcinomas. *PLoS One* 2012; **7**: e42683.
25. Yamamoto S, Midorikawa Y, Morikawa T *et al.* Identification of chromosomal aberrations of metastatic potential in colorectal carcinoma. *Genes Chromosomes Cancer* 2010; **49**: 487–496.
26. Bartos JD, Gaile DP, McQuaid DE *et al.* aCGH local copy number aberrations associated with overall copy number genomic instability in colorectal cancer: coordinate involvement of the regions including BCR and ABL. *Mutat Res* 2007; **615**: 1–11.

27. Knösel T, Schlüns K, Stein U *et al.* Genetic imbalances with impact on survival in colorectal cancer patients. *Histopathology* 2003; **43**: 323–331.
28. Alcock HE, Stephenson TJ, Royds JA *et al.* Analysis of colorectal tumor progression by microdissection and comparative genomic hybridization. *Genes Chromosomes Cancer* 2003; **37**: 369–380.
29. Knösel T, Petersen S, Schwabe H *et al.* Incidence of chromosomal imbalances in advanced colorectal carcinomas and their metastases. *Virchows Arch* 2002; **440**: 187–194.
30. Djureinovic T, Skoglund J, Vandrovcova J *et al.* A genome wide linkage analysis in Swedish families with hereditary non-familial adenomatous polyposis/non-hereditary non-polyposis colorectal cancer. *Gut* 2006; **55**: 362–366.
31. Kang YH, Galal WC, Farina A *et al.* Properties of the human Cdc45/Mcm2-7/GINS helicase complex and its action with DNA polymerase epsilon in rolling circle DNA synthesis. *Proc Natl Acad Sci USA* 2012; **109**: 6042–6047.
32. Cerveira N, Santos J, Teixeira MR. Structural and expression changes of septins in myeloid neoplasia. *Crit Rev Oncog* 2009; **15**: 91–115.
33. Kim W, Bennett EJ, Huttlin EL *et al.* Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol Cell* 2011; **44**: 325–340.
34. Jalmar O, François-Moutal L, García-Sáez AJ *et al.* Caspase-8 binding to cardiolipin in giant unilamellar vesicles provides a functional docking platform for BID. *PLoS One* 2013; **8**: e55250.
35. Kim JY, So KJ, Lee S *et al.* Bcl-rambo induces apoptosis via interaction with the adenine nucleotide translocator. *FEBS Lett* 2012; **586**: 3142–3149.
36. Urbano A, Lakshmanan U, Choo PH *et al.* AIF suppresses chemical stress-induced apoptosis and maintains the transformed state of tumor cells. *EMBO J* 2005; **24**: 2815–2826.
37. Bell ES, Park M. Models of crk adaptor proteins in cancer. *Genes Cancer* 2012; **3**: 341–352.
38. Jedidi I, Zhang F, Qiu H *et al.* Activator Gcn4 employs multiple segments of Med15/Gal11, including the KIX domain, to recruit mediator to target genes in vivo. *J Biol Chem* 2010; **285**: 2438–2455.
39. Murray JC, Donahue DJ, Malik SI *et al.* Temporal lobe pleomorphic xanthoastrocytoma and acquired BRAF mutation in an adolescent with the constitutional 22q11.2 deletion syndrome. *J Neurooncol* 2011; **102**: 509–514.
40. Beddow RA, Smith M, Kidd A *et al.* Diagnosis of distal 22q11.2 deletion syndrome in a patient with a teratoid/rhabdoid tumour. *Eur J Med Genet* 2011; **54**: 295–298.
41. Chang VY, Quintero-Rivera F, Baldwin EE *et al.* B-acute lymphoblastic leukemia and cystinuria in a patient with duplication 22q11.21 detected by chromosomal microarray analysis. *Pediatr Blood Cancer* 2011; **56**: 470–473.
42. Toth G, Zraly CB, Thomson TL *et al.* Congenital anomalies and rhabdoid tumor associated with 22q11 germline deletion and somatic inactivation of the SMARCB1 tumor suppressor. *Genes Chromosomes Cancer* 2011; **50**: 379–388.