IJC
International Journal of Cancer

Molecular analysis of *Ceruloplasmin* in a South African cohort presenting with oesophageal cancer

Natalie J. Strickland[1], Tandi Matsha[2], Rajief T. Erasmus[3] and Monique G. Zaahl[1]

[1] Department of Genetics, University of Stellenbosch, Stellenbosch, South Africa
[2] Obesity and Chronic Disease of Lifestyle Research Unit, Department Biomedical Sciences, Faculty of Health and Wellness, Cape Peninsula University of Technology, Belville, South Africa
[3] Department of Chemical Pathology, University of Stellenbosch, Tygerberg, South Africa

Oesophageal cancer (OC) is a disease characterized by the development of malignant tumors in the epithelial cells lining the oesophagus. It demonstrates marked ethnic variation, with squamous cell carcinoma (SCC) being more prevalent in the Black population and adenocarcinoma (ADC) occurring more often in Caucasians. The etiology of this complex disease has been attributed to a variety of factors, including an excess of iron (resulting in increased tumourigenesis), oesophageal injury and inflammation (due in part to Barrett's oesophagus and smoking among others). The aim of this study was to determine if genetic variations identified in the *ceruloplasmin* (*CP*) gene (implicated in iron homeostasis) contribute to OC pathogenesis or susceptibility. The study cohort consisted of 96 unrelated OC patients from the Black Xhosa-speaking South African population and 88 population-matched control individuals. The promoter and coding regions of the *CP* gene were analyzed for DNA sequence variation using heteroduplex single-strand conformation polymorphism (HEX-SSCP) analysis, restriction fragment length polymorphism (RFLP) analysis and semi-automated bidirectional DNA sequencing analysis. Fourteen previously described and four novel variants were identified. Statistically significant associations were revealed for two of the novel variants with OC in this study and could, therefore, potentially contribute to disease susceptibility. *In silico* analysis of the region of the promoter spanning the identified variants sought to identify putative transcription factor binding sites (TFBSs) that could possibly regulate the expression of *CP*. To our knowledge, this is the first study to examine CP with respect to OC in the Black South African population.

OC is one of the leading causes of cancer-related deaths worldwide; it is the 15th most common cancer in developed nations and the 4th most common in developing countries such as South Africa.[1] OC is one of the most common forms of cancer found in the South African population, particularly in the Transkei region.[2,3] The Transkei is situated on the southeastern coast of South Africa and was previously an independent Xhosa homeland. In 1994, the Transkei was reintegrated into South Africa as part of the Eastern Cape Province. The majority of the population is Xhosa-speaking with less than 10,000 of the estimated 2.3 million inhabitants being from European descent.[4] The Transkei region of South Africa is thought to be the centre of the disease in Africa,[2] with an age standardized incidence rate (ASIR) of 46.7/100,000 for males and 19.2/100,000 for females previously being reported.[3] OC shows clear geographic variation and occurs at a high incidence in certain areas of the world, which are termed "oesophageal cancer belts." These regions are separated into the Asian belt, which is made up of countries such as Iran, Iraq, China, Japan and Turkey.[5,6] Two subtypes of OC exist, SCC and ADC, and both demonstrate marked racial variation which is characteristic of the disease.

Key words: oesophageal cancer, *ceruloplasmin*, iron metabolism, variants, transcription factors

Abbreviations: χ^2: Chi-squared; °C: Degrees Celcius; %TS: percentage transferrin saturation; ADC: adenocarcinoma; AP1: activator protein 1; ASIR: age standardized incidence rate; C/EBPα: CCAAT/enhancer binding protein alpha; CP: ceruloplasmin; D': coefficient of association; Fe^{2+}: ferrous iron; Fe^{3+}: ferric iron; GATA: GATA-binding protein; HEX-SSCP: heteroduplex single-strand conformation polymorphism; HNF: hepatocyte nuclear factor; hr: hour; HWE: Hardy Weinberg equilibrium; LD: linkage disequilibrium; OC: oesophageal cancer; p53: tumour suppressor protein 53; PAA: polyacrylamide; PCR: polymerase chain reaction; r^2: correlation coefficient; RFLP: restriction length fragment polymorphism; SCC: squamous cell carcinoma; SF: serum ferritin; SPI1: spleen focus forming virus proviral integration oncogene 1; TF: transcription factor; TFBS(s): transcription factor binding site(s); TS: transferrin saturation; USF1: upstream regulatory factor 1; UTR: untranslated region; V: volt; YY1: Ying Yang 1

Grant sponsors: The National Research Foundation (Thuthuka); The University of Stellenbosch
DOI: 10.1002/ijc.26418
History: Received 16 Mar 2011; Accepted 23 Aug 2011; Online 7 Sep 2011

Correspondence to: Monique G. Zaahl, Department of Genetics, Private Bag X1, Matieland, 7602, South Africa, Tel.: +27-21-808-5834; Fax +27-21-808-5833, E-mail: mjulies@sun.ac.za

ADC is more prevalent in Caucasians and SCC occurs more frequently in Blacks. In South Africa, SCC is thought to be the leading cause of death among Black males and the 4th most common cause of death in Colored males.[5] In addition, OC incidence rates also increase markedly with age, with the average age of disease diagnosis in South Africa being 60 years of age.[7]

The combined use of tobacco and alcohol is one of the most important risk factors in the etiology of OC in Western countries.[8] Other risk factors include oesophageal injury and inflammation, infectious agents, nutrition, nitrosamines, mycotoxins and exposure to environmental toxins.[5,9–12] The marked differences in the geographic and ethnic incidences of OC have been hypothesized to be due to variations in environmental factors in different populations originating from different regions.[13]

Particular diets from certain regions in the world may result in deficiencies of vitamins and micronutrients. Especially implicated are diets lacking vitamins such as B1 (riboflavin) and the mineral selenium. Individuals with lower levels of selenium have been shown to have an increased risk of developing OC.[14] In a global study, low intakes of fruit and vegetables were found to account for 20% of all cases of OC and 19% of cases of gastric cancers worldwide.[15] This is thought to be due to the fact that the antioxidants, minerals and micronutrients present in fruits and vegetables, suppress the action of carcinogens and prevent oxidative DNA damage.[16] Areas of the Tanskei region of South Africa have diets based primarily on cereal grains, which are low in nutrients such as zinc and iron.

The relationship between iron and OC was previously described during a study of Black South African OC patients who were shown to have iron overload as a consequence of routinely drinking traditional beer brewed in nongalvanized steel drums.[17] It was demonstrated that excess dietary iron can act as a potential risk factor for the development of OC, and further studies have shown that iron is involved in the development of other cancers such as liver cancer.[18] In addition, studies using rat models showed that iron supplementation could be indicated as a risk factor for developing OC.[19,20] Studies have also shown that an increase in dietary iron plays a role in the development of OC in groups other than the Black South African population.[21,22] It has also been demonstrated that loss of iron from a target cell results in antitumor activity, which is reduced when there is excess iron present. For this reason, iron can be said to have an indirect carcinogenic effect on cells. Increased levels of iron can also inhibit the growth of tumouricidal-activated macrophages.[23]

To further investigate iron as a risk factor for OC, this study focused on the analysis of the regulatory and coding regions of the *CP* gene, which is involved in iron homeostasis, in patients diagnosed with OC. CP is a ferroxidase enzyme synthesized in the liver which catalyses the oxidation of ferrous iron (Fe^{2+}) to ferric iron (Fe^{3+}), thereby creating an ion gradient favoring iron export from the cells. Therefore, CP which is located primarily in the plasma is responsible for regulating iron transport from stores in various tissues.[24]

In this study, we aim to demonstrate the relationship between *CP* and the development of OC, by identifying gene variations that could potentially contribute toward iron dysregulation and subsequent disease pathogenesis. It is anticipated that the results obtained from this study will lead to a greater understanding of the role that iron homeostasis plays in the etiology of OC.

Material and Methods

Ethical approval for this study was obtained from the Ethics Review Committee of the University of Stellenbosch (reference number: N07/06/147). Prior to sampling written informed consent was obtained from all study participants.

Subjects

The study cohort for this study included 96 patients and 88 control individuals from the Black Xhosa-speaking population of South Africa. All individuals are of Central African descent and observe cultural traditions originating from Xhosa tribes of South Africa.

Patient samples ($n = 96$) were recruited from collaborating clinicians, who were responsible for clinical assessment. These patients were referred for a barium swallow followed by biopsies in the theatre to confirm the presence of squamous cell carcinoma (SCC) or adenocarcinoma (ADC) based on the histology. All of the patients in this study presented with SCC.

The control individuals in this study consist of 88 unrelated, healthy individuals from the same population. These individuals were interviewed for a potential family history of relevant diseases and early screening for OC was performed using a Nabeya capsule for brush biopsy.

Biochemical assessment of iron status

In the current study, the transferrin saturation (TS) and the serum ferritin (SF) levels were measured in the patient and control cohorts according to standard methodology. Iron overload parameters were defined as the percentage transferrin saturation (%TS) > 45 and/or serum ferritin exceeding 200 μg l^{-1} in females and 300 μg l^{-1} in males.[25,26] The following iron parameters were obtained for all participating patient and control individuals: serum iron, transferrin, transferrin saturation (percentage), ferritin and C-reactive protein (CRP) levels. The normal ranges for healthy adults for these parameters were as follows: serum iron—males 65–176 μg dl^{-1} and females 50–170 μg dl^{-1}, transferrin—0.21–0.36 g l^{-1} and CRP—10–40 mg l^{-1} (mild inflammation and viral infection), 40–200 mg l^{-1} (active bacterial infection) and >200 mg l^{-1} (severe bacterial infection).

Methods

DNA was extracted from whole blood using a standard method.[27] The promoter and coding region of the *CP* gene were amplified by polymerase chain reaction (PCR) amplification using intronic oligonucleotide primers. Primer sequences are available upon request. The PCR products were subjected to a combined heteroduplex single-strand conformation polymorphism (HEX-SSCP) analysis[28] method and resolved on a 12% polyacrylamide (PAA) gel supplemented with 7.5% urea at 4°C (350 V) for 18 hr. Following electrophoresis, the gels were stained with ethidium bromide and the DNA visualized by ultraviolet light transillumination. Semiautomated bidirectional DNA sequencing was performed on PCR products demonstrating mobility or conformation shifts in the PAA gels using an ABI 3100 PRISM automated sequencer (Applied Biosystems).

The PCR products of Exon 2, incorporating the T83 variant, were subjected to restriction fragment length polymorphism (RFLP) analysis, using the *SfcI* enzyme (recognition site: 5′ C↓TRYAG 3′; New England Biolabs, Beverly, USA), to distinguish between the various genotypes. This variant abolishes the *SfcI* restriction site.

Statistical analysis

Allele and genotype frequencies were estimated by allele counting and statistical differences between patient and control groups were tested for significance by the Fisher exact test and/or chi-squared (χ^2) analysis. A probability value smaller than 0.05 was regarded as statistically significant. The Hardy–Weinberg equilibrium test was performed to test equilibrium for the genetic traits investigated in the respective populations.

Haplotype analysis and linkage disequilibrium (LD) analysis was performed on each of the variants identified (Haploview 4.0).[29] Default parameters were applied to test for LD [the coefficient of association (D'): where $D' = 1$ for perfect linkage disequilibrium; logarithm of the likelihood odds ratio, measure of confidence (LOD): LOD > 3; correlation coefficient between two loci (r^2): $r^2 > 0.8$, where $r^2 = 1$ for prefect linkage disequilibrium]. The default block definition was applied[30] upon haplotype analysis.

Box plots were used to identify outliers in the iron parameter data of patients and controls, which were then subsequently removed from further analysis. The *F* test for variance was used to assess the type of *t* test to be used. For *F* tests with a *p* value of greater than 0.05 the student's *t* test for equal variances was carried out; for *F* tests with a *p* value of less than 0.05 the student's *t* test for unequal variances was carried out to determine statistically significant associations between patients and controls. All statistical analyses with regard to the iron parameter data were conducted on males and females separately, followed by males and females combined. These analyses were conducted using Microsoft Office 2007 software with the Analysis Toolpak Add-In.

Table 1. Patient demographics

Demographic	OC patients	
	Males ($n = 48$)	Females ($n = 48$)
Average age	59 ± 13	66 ± 12
Smoking		
Yes	9 (19%)	4 (8%)
Stopped	27 (56%)	18 (36%)
Never	4 (8%)	13 (27%)
Unknown	8 (17%)	13 (37%)
Alcohol consumption		
Yes	31 (65%)	25 (52%)
Stopped	2 (4%)	0
Never	6 (13%)	10 (21%)
Unknown	9 (19%)	13 (27%)
Normal iron levels (Male: SF < 300 μg l^{-1}; Female: SF < 200 μg l^{-1})	31 (64.5%)	25 (52.08%)
Iron deficient (SF < 20 μg l^{-1})	0	3 (6.25%)
Increased iron levels (Male: SF > 300 μg l^{-1}; Female: SF > 200 μg l^{-1})		
%TS<45	17 (35.4%)	19 (39.5%)
%TS>45	0	1 (2.08%)

Computational analysis

In silico analysis was performed on each of the variants identified in the promoter region in this study. Several bioinformatic databases are available for *in silico* analysis of the promoter region, of which the following two were used in this study: JASPAR CORE[31] and TRANSFAC®7.[32] From TRANSFAC®7, two programs were used, namely PATCH and MATCH™ (v1.0).[33] The default parameters were employed in the use of both of these databases.

Results

The patient cohort included 48 (50%) males and 48 (50%) females, with a mean age of 59 years [standard deviation (SD) ± 13 years] and 66 years [standard deviation (SD) ± 12 years], respectively (Table 1). Demographic information on alcohol consumption and the smoking status of the patients is summarized in Table 1. Smoking status was restricted to cigarette smoking (including shop-bought and home-grown cigarettes). Alcohol consumption was restricted to beer (including shop-bought and home-made beer). The iron status of each individual was classified into one of four groups (see Table 1): iron deficiency (SF < 20 μg l^{-1}), normal serum ferritin levels (females: SF between 20 and 200 μg l^{-1} and males: SF between 20 and 300 μg l^{-1}), raised serum ferritin levels (females: SF > 200 μg l^{-1} and males: SF

Table 2. Genotypic and polymorphic allele frequencies of variants identified in the *CP* promoter region in the Black South African population

Variant	Study cohort	*n*	Genotype			*p*	*2n*	Variant allele[1]	*p*
			CC	CG	GG			G	
5′UTR-567C→G[2,5]	Patients	84	1.00	0.00	0.00	0.157	168	0.00	0.159
rs34053109	Controls	85	0.98	0.02	0.00		170	0.01	
			TT	TC	CC			C	
5′UTR-563T→C[3]	Patients	90	0.41	0.51	0.08	0.229	180	0.33	0.127
rs17838834	Controls	85	0.52	0.45	0.04		170	0.23	
			CC	CT	TT			T	
5′UTR-439C→T[2]	Patients	79	0.99	0.01	0.00	0.305	158	0.01	0.308
rs701749	Controls	78	0.96	0.04	0.00		156	0.02	
			TT	T-	–			–	
5′UTR-364delT[2,4]	Patients	79	0.98	0.03	0.00	0.157	158	0.01	0.159
rs17838833	Controls	78	1.00	0.00	0.00		156	0.00	
			TT	TC	CC			C	
5′UTR-354T→C[3]	Patients	80	0.43	0.50	0.08	0.321	160	0.33	0.481
rs17838832	Controls	78	0.50	0.42	0.08		156	0.30	
			CC	CT	TT			T	
5′UTR-350C→T[3]	Patients	80	0.76	0.20	0.04	0.284	160	0.14	0.125
rs34334174	Controls	78	0.85	0.14	0.01		156	0.08	
			GG	GA	AA			A	
5′UTR-308G→A[2,4]	Patients	52	0.90	0.10	0.00	**0.011**	104	0.05	**0.012**
This study	Controls	64	1.00	0.00	0.00		128	0.00	
			AA	AG	GG			G	
5′UTR-282A→G[3]	Patients	53	0.60	0.38	0.02	0.539	106	0.21	0.835
rs17838831	Controls	64	0.63	0.31	0.06		128	0.22	

Statistically significant associations are indicated in bold.
[1]Allele frequencies of only the polymorphic allele denoted. [2]Variants identified only in the heterozygous state. [3]Variants identified in both the heterozygous and homozygous states. [4]Variants identified only in the patient group. [5]Variants identified only in the control group.
Abbreviations: 5′: 5 prime end; A: adenine; C: cytosine; del: deletion; G: guanine; *n*: number of individuals; *2n*: number of alleles; *p*: probability value; T: thymine; UTR: untranslated region.

$> 300\ \mu g\ l^{-1}$) with %TS < 45 and raised serum ferritin levels with %TS > 45.

The *CP* gene implicated in iron metabolism was subjected to mutation analysis. Several variants (14 previously described and four novel) were identified following HEX-SSCP analysis of the promoter and coding region of this gene (Tables 2 and 3). RFLP analysis was performed for specific variants where HEX-SSCP analysis proved inconclusive. All of the variants detected were verified by semiautomated bidirectional DNA sequencing analysis to confirm their presence. The patient and control cohorts were tested for Hardy Weinberg equilibrium (HWE) at the variant loci. All groups were found to conform to HWE in this study.

Mutation analysis of the *CP* promoter region resulted in the detection of six previously described single nucleotide polymorphisms (5′UTR-567C→G, 5′UTR-563T→C, 5′UTR-439C→T, 5′UTR-354T→C, 5′UTR-350C→T and 5′UTR-282A→G), one previously described single base pair deletion (5′UTR-364delT) and one novel single nucleotide substitution (5′UTR-308G→A). Four previously described (V223, R367C, Y425 and D544E) and three novel (T83, V246A and G633) variants were identified in the coding region of the *CP* gene following HEX-SSCP analysis. The T83 variant was subjected to RFLP analysis following HEX-SSCP analysis due to difficulties distinguishing between banding patterns on the HEX-SSCP gels. Three previously described intronic variants (IVS4-14C→T, IVS7+9T→C and IVS15-12T→C) were also identified in this study.

Two of the novel variants were demonstrated to be significantly associated with the OC phenotype in this study; the promoter variant 5′UTR-308G→A ($p = 0.012$) and the exonic variant G633 ($p = 0.0003$) (indicated bold in Tables 2 and 3).

Iron parameters were obtained for all the patients and control participants in this study (see Materials and Methods). To further assess the effect of iron on disease expression, statistical analyses were performed on this data by comparing the OC patients and control individuals. In addition,

Table 3. Genotypic and polymorphic allele frequencies of variants identified in the *CP* exonic and intronic regions in the Black South African population

Exon/Intron	Variant	Study cohort	*n*	Genotype			*p*	*2n*	Variant allele[1]	*p*
2	**T83**[2]			**TT**	**TC**	**CC**			**C**	
	This study	Patients	92	0.97	0.03	0.00	0.357	184	0.02	0.360
		Controls	84	0.99	0.01	0.00		168	0.01	
4	**V223**[2]			**GG**	**GC**	**CC**			**C**	
	rs35438054	Patients	95	0.99	0.01	0.00	0.937	190	0.01	0.937
		Controls	85	0.99	0.01	0.00		170	0.01	
4	**V246A**[2,4]			**TT**	**TC**	**CC**			**C**	
	This study	Patients	95	0.98	0.02	0.00	0.179	190	0.01	0.180
		Controls	85	1.00	0.00	0.00		170	0.00	
4	**IVS4-14C→T**[2]			**CC**	**CT**	**TT**			**T**	
	rs34067682	Patients	84	0.69	0.31	0.00	0.959	168	0.12	0.963
		Controls	83	0.69	0.31	0.00		166	0.12	
6	**R367C**[2]			**CC**	**CT**	**TT**			**T**	
	rs34624984	Patients	94	0.97	0.03	0.00	0.356	188	0.02	0.359
		Controls	86	0.99	0.01	0.00		172	0.01	
7	**Y425**[2,4]			**TT**	**TC**	**CC**			**C**	
	rs34237139	Patients	91	0.99	0.01	0.00	0.327	182	0.01	0.328
		Controls	87	1.00	0.00	0.00		174	0.00	
7	**IVS7+9T→C**[2,4]			**TT**	**TC**	**CC**			**C**	
	rs35272481	Patients	91	0.99	0.01	0.00	0.327	182	0.01	0.328
		Controls	87	1.00	0.00	0.00		174	0.00	
9	**D544E**[3]			**TT**	**TA**	**AA**			**A**	
	rs701753	Patients	95	0.55	0.44	0.01	0.380	190	0.23	0.463
		Controls	81	0.56	0.36	0.09		162	0.27	
11	**G633**[2,4]			**TT**	**TC**	**CC**			**C**	
	This study	Patients	88	0.85	0.15	0.00	**0.0003**	176	0.07	**0.0004**
		Controls	84	1.00	0.00	0.00		168	0.00	
15	**IVS15-12T→C**[3]			**TT**	**TC**	**CC**			**C**	
	rs16861582	Patients	93	0.28	0.56	0.16	0.453	186	0.44	0.616
		Controls	88	0.33	0.51	0.16		176	0.41	

Statistically significant associations indicated in bold.
[1]Allele frequencies of only the polymorphic allele denoted. [2]Variants identified only in the heterozygous state. [3]Variants identified in both the heterozygous and homozygous states. [4]Variants identified only in the patient group.
Abbreviations: C: cytosine; IVS: intervening sequence; *n*: number of individuals; *2n*: number of alleles; *p*: probability value; T: thymine.

males and females within each group were compared separately due to the fact that the normal ranges of the various iron parameters may differ markedly between sexes. The data was further analyzed by subdividing all individuals in terms of presence or absence of the variants and then analyzing iron parameters based on these groupings. Several statistically significant observations were made and are summarized in Table 4.

LD and haplotype analysis, using the Haploview 4.0 software, was performed on all of the variants identified in this study using association and case-control studies.[29] The LD test was applied, and a haplotype block consisting of the variant alleles of the promoter variants 5′UTR-563T→C, 5′UTR-439C→T, 5′UTR-354T→C, 5′UTR-350C→T, 5′UTR-308G→A and 5′UTR-282A→G was predicted ($D' = 0.947$, LOD = 30.81, $r^2 = 0.875$; Fig. 1). This haplotype spans a genomic region of 281 bp. Haplotype association frequencies demonstrated that this haplotype is significantly associated ($p = 0.01$) with OC in the Black South African population.

All the variants identified in the 5′UTR region of the *CP* gene were subjected to *in silico* analysis to determine if they disrupted or created any putative transcription factor binding sites (TFBS) predicted to occur within this regulatory region. Identification of the same TFBS by more than one of the

Table 4. Iron parameter statistics

	Patients *vs.* controls			All samples	
				Variant	
	Males	Females	Combined	5′UTR-308G→A	G633
Iron parameter	*t*-test, *p* value				
Serum iron	**0.0000**[3]	**0.0000**[3]	**0.0000**[3]	0.0663[1,5]	**0.0010**[1,5]
Transferrin	**0.0000**[3]	**0.0000**[1,3]	**0.0000**[3]	**0.0202**[1,5]	**0.0090**[1,5]
% Transferrin saturation	0.2432[3]	0.3619[1,3]	0.4322[1,3]	0.1918[4]	**0.0492**[1,5]
Ferritin	**0.0000**[2]	**0.0000**[2]	**0.0000**[2]	0.4733[5]	0.0911[4]
CRP	**0.0000**[2]	**0.0000**[2]	**0.0000**[2]	0.4631[1,4]	0.0612[1,4]

Statistically significant associations are indicated in bold.
[1]Use of *t*-test for assuming equal variances, all other instances use *t*-test assuming unequal variances. [2]Measured level is higher in patients relative to controls. [3]Measured level is lower in patients relative to controls. [4]Measured level is higher in individuals with the variant relative to wild-type individuals. [5]Measured level is lower in individuals with the variant relative to wild-type individuals.

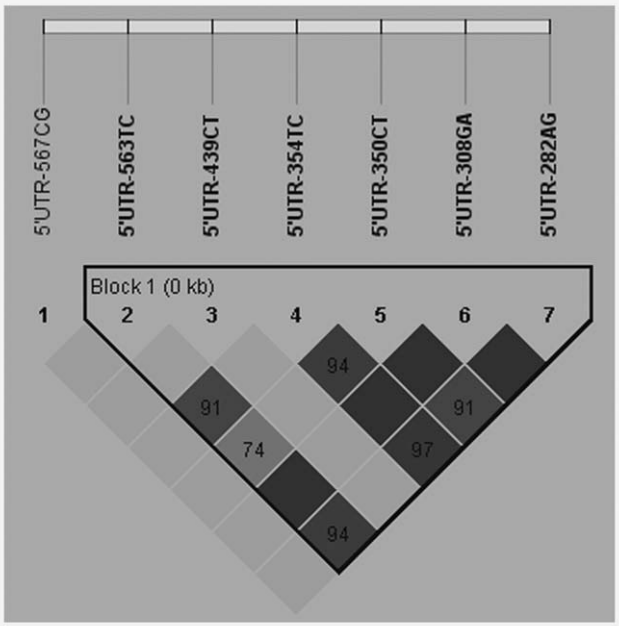

Figure 1. LD plot of the statistically significant haplotype predicted using Haploview 4.0.

programs utilized increases the likelihood that the motif exists. JASPAR contains annotated, matrix-based transcription factor binding site profiles for eukaryotic organisms in an open-access database. Experimentally validated sets of nucleotide sequences that bind transcription factors form the basis of these profiles.[31] The TRANSFAC®7 database contains data on transcription factors, their experimentally proven binding sites, and regulated genes. The TRANSFAC®7[32] programs (such as PATCH and MATCH™) use the integrated matrices and site sequences in TRANSFAC®7 to perform matrix-or pattern-based searches of factor binding sites in regulatory DNA sequences. The results obtained through *in silico* analysis of the promoter variants identified in the *CP* gene are shown in Table 5.

Discussion

OC, like the majority of cancers, demonstrates an extremely complicated etiology, with both genetic and environmental factors involved in disease development and progression.[4] Previous studies have attempted to identify possible risk factors for OC as well as candidate genes involved in disease development.[19,20] However, contributing factors leading to OC are still poorly understood on a molecular level and many genetic aberrations that could infer disease risk or resistance remain to be identified.

Mutation analysis of the *CP* gene revealed several sequence changes. Only four of these variants were identified as novel, whereas the majority have previously been documented. Most of these variants are classified as single nucleotide polymorphisms.

A novel promoter variant (5′UTR-308G→A) occurs as a result of a G to A substitution 308 nucleotides upstream of the translation initiation site (ATG). This novel polymorphism presented only in the heterozygous state and was not identified in the control individuals. It was detected in five of 52 (9.6%) of the OC patient individuals (Table 2). The novel 5′UTR-308G→A variant was found to be statistically associated ($p = 0.01$) with OC in this study (Table 2). Statistical analyses were performed on this variant in conjunction with the iron parameter data by grouping all subjects together, whether patients or controls, with or without the variant (Table 4). Statistical significance was identified for transferrin ($p = 0.02$) and borderline significance was achieved for serum iron ($p = 0.06$) with the heterozygous 5′UTR-308G→A variant (Table 4). The levels of both serum iron and transferrin were lower in the individuals with the heterozygous variant, which indicates less circulating iron in the plasma and a decrease in the mobilization of iron stores. The CP enzyme is synthesized in the liver and it catalyzes the oxidation of ferrous iron (Fe^{2+}) to ferric iron (Fe^{3+}), thereby creating an ion gradient favoring iron export from the cells. Transferrin (TF) is the main iron transport molecule in the plasma, but it can

Table 5. Predicted TFBS in the promoter region of the *CP* gene

	Putative transcription factor binding sites					
	PATCH		MATCH™		JASPAR CORE	
Variant	Abolished	Created	Abolished	Created	Abolished	Created
5′UTR-567C→G	YY1	GATA1	–	HNF-1	FOXL1, YY1	GATA3
5′UTR-563T→C	YY1, HiNF-D, HiNF-M, HiNF-P, TFIID, TMF, SRF	RXRα, RXRγ, T3Rα	–	–	FOXL1	FOXC1, SP1
5′UTR-439C→T	–	–	–	AP-1	SPI1	SRY, NKX3–1
5′UTR-364delT	FOXM1a, FOXM1b, HNF-3α, HNF-3β	GR	HNF-3β	–	SRY, FOXD1	–
5′UTR-354T→C	SP1	c-myc	–	–	–	–
5′UTR-350C→T	GR, SPI1, NF-Y	–	–	GR, C/EBPα	-	–
5′UTR-308G→A	p53	GR	–	–	MAX, USF1, SPI1	FOXL1,
5′UTR-282A→G	C/EBPα, HNF-1, HNF-3α	–	C/EBPα	–	HLF, FOXL1, SRY, SOX9	GATA2, GATA3, SPIB

Abbreviations: 5′, 5 prime end; AP-1, Activator protein 1; C/EBPα, CCAAT/enhancer binding protein alpha; c-myc, Myelocytomatosis viral oncogene c; FOXC1, Forkhead box protein C1; FOXD1, Forkhead box protein D1; FOXL1, Forkhead box protein L1; FOXM1a, Forkhead box M1a; FOXM1b, Forkhead box M1b; GATA1, GATA-binding protein 1; GATA2, GATA-binding protein 2; GATA3, GATA-binding protein 3; GR, Glucocorticoid receptor; HiNF-D, Histone nuclear factor D; HiNF-M, Histone nuclear factor M; HiNF-P, Histone nuclear factor P; HLF, Hepatic leukaemia factor; HNF-1, Hepatocyte nuclear factor 1; HNF-3α, Hepatocyte nuclear factor alpha; HNF-3β, Hepatocyte nuclear factor 3 beta; MAX, Myc-associated factor X; NF-Y, Nuclear transcription factor Y; NKX3–1, NK 3 homeobox 1; p53, Tumour suppressor protein 53; RXRα, Retinoid X receptor alpha; RXRγ, Retinoid X receptor gamma; SOX9, SRY (sex determining region Y)-box-9; SP1, Specificity protein 1; SPI1, spleen focus forming virus (SFFV) proviral integration oncogene 1; SPIB, SPI-B transcription factor; SRF, Serum response factor; SRY, Sex-determining region Y; T3Rα, Thyroid hormone receptor alpha; TFIID, Transcription factor IID; TMF, TATA element modulatory factor; USF1, Upstream transcription factor 1; UTR, untranslated region; YY1, Ying Yang 1.

only carry iron in the ferric state, therefore CP acts as an aid in iron transport in the plasma. Enzymatic oxidation of ferrous iron by CP has been hypothesized as an important step in the formation of TF,[34] *via* the incorporation of iron into apo-transferrin. A decrease in CP expression (perhaps due to the 5′UTR-308G→A variant) could potentially explain the decreased levels of serum iron and transferrin that were identified in the individuals with the variant.

In silico analysis of this variant using the PATCH program from the TRANSFAC®7 database[32] predicted that the transcription factor binding site (TFBS) for the tumor suppressor protein 53 (p53) would be abolished in the presence of the 5′UTR-308G→A variant (Table 5). The gene that encodes the p53 protein, *TP53,* possesses sequence-specific DNA binding properties. It is thought that when cells are exposed to damaging agents, p53 is able to activate target genes involved in cell repair and apoptosis. This therefore prevents cells from replicating after damage or induces apoptosis of cells containing damaged DNA.[34] The p53 transcription factor (TF) therefore functions indirectly as a tumor suppressor. Mutations in the DNA binding domain of p53 are common in a variety of cancers and may prevent the activation of genes involved in DNA repair. This could enhance the genomic instability of tumor cells.[35] It is feasible therefore to consider that variants in p53 TFBSs could result in decreased levels of gene expression of target genes and possibly result in subsequent tumor development. However, functional studies are required to determine if p53 does interact with this region of the *CP* promoter *in vivo*.

A haplotype, consisting of six of the promoter variants detected in this study, was predicted to exist following analysis using the Haploview 4.0 program (Fig. 1). This haplotype was shown to be significantly associated with OC ($p = 0.01$) in this study. The variants comprising the haplotype span a genomic region of 281 bp that was identified as an important promoter region in the rat (*Rattus norvegicus*).[36] An area of ~300 bp was determined to be critical for gene expression to occur and has been shown to be highly conserved in humans. The region from −393 to −348 bp upstream of the initiating ATG was found to exert a positive effect on gene expression and also showed sequence homology to the rat albumin D site, known to bind to CCAAT/enhancer binding protein alpha (C/EBPα). In this study, the 5′UTR-350C→T variant, which lies within the recognition site, was predicted to create a binding site for this TF (Table 5). *CP* levels have been shown to be upregulated in response to certain aggressive cancers which may be mediated by C/EBP binding.[37] 5′UTR-282A→G was predicted by two of the databases to abolish a C/EBPα binding domain (Table 5). C/EBP TFs are part of the bZIP family of TFs which have an important role in genes involved in the inflammatory response pathway. CP is an acute-phase protein that is involved in this inflammatory response pathway.[38] Therefore, disruption to C/EBPα binding domains may result in the down-regulation of *CP* expression leading to impaired inflammatory response.

In addition to C/EBPα, many TFBSs for the GATA-binding protein (GATA) family of TFs were predicted for the *CP* promoter region (Table 5). The 5′UTR-567C→G variant was

predicted by the *in silico* software programs to create a GATA1 and a GATA3 TFBS. The 5′UTR-282A→G variant was also predicted to create GATA2 and GATA3 TFBSs.

GATA1 has been shown to play a role in the regulation of genes involved in the haem biosynthesis pathway. GATA2 is involved in the control of haematopoietic progenitor cells. The function of GATA3 is less understood but it has been shown to play a role in erythroid development.[39] Genes expressed in the liver, such as *CP*, are known to be regulated by the GATA TFs, although the exact role of these TFs in *CP* expression is unknown. It is therefore important not to disregard the effect that these TFs may exert on *CP* regulation. Additional research will aid in determining the role (if any) of the GATA TFs in *CP* expression.

Hepatocyte nuclear factors (HNFs) regulate genes that are expressed in the cells of the liver. Correct expression of these genes is determined by the binding of a wide variety of HNFs to the regulatory regions of these genes (reviewed by Costa *et al.*[40]). HNFs are members of the steroid/thyroid nuclear receptor family which are expressed predominately in the liver and acts as an essential regulator of liver metabolism and development. The 5′UTR-567C→G variant creates a putative HNF-1 site in the *CP* promoter (Table 5). Creation of this TFBS may result in increased expression of the *CP* gene. The 5′UTR-282A→G variant abolishes a putative HNF-1 motif (Table 5). The deletion of the T nucleotide at position −364 is predicted to abolish a HNF-3α and HNF-3β TFBS (Table 5).

CP is an acute-phase protein that plays an important antioxidant role protecting cells against damage that may be caused by oxidative stress.[41] It has been shown that members of the HNF family show an increase in expression in the presence of oxidative stress as a result of hepatitis C virus infection.[40] Excess iron catalyses the conversion of hydrogen peroxide to free radicals and leads to oxidative stress in various tissues.[42] Disruptions to HNF TFBSs may perturb *CP* expression and therefore result in incorrect functioning of the antioxidant mechanism. It is possible that oxidative damage from the presence of excess iron could lead to cell damage and the subsequent formation of tumor cells.

The presence of the 5′UTR-567C→G and the 5′UTR-563T→C variants are predicted to abolish a putative TFBS for Ying Yang 1 (YY1), (Table 5). YY1 is a zinc finger protein that has been shown to play a role in iron metabolism, particularly by interacting with transferring.[43] It has also been implicated in gene silencing in the liver.[43] CP is known to interact with transferrin during the transfer of iron molecules across the basolateral membrane of the enterocyte. This TF could be involved in simultaneously mediating expression of these two proteins.

A number of other putative TFBSs were predicted to be altered following *in silico* analysis of the variants identified in the *CP* promoter (Table 5). These include the binding motif for activator protein 1 (AP1) which has shown to be involved in cell proliferation, differentiation and tumourigenesis,[44] upstream regulatory factor 1 (USF1) which has defined roles in regulation of genes expressed in the liver[45] and spleen focus forming virus proviral integration oncogene 1 (SPI1) which plays a role in the development of myeloid and β-lymphoid cells which are important during an immune response.[46] The functional importance of these TFs in regulation of *CP* expression should not be overlooked.

In addition to the eight promoter variants identified, a statistically significant ($p = 0.0004$) synonymous variant (Table 3), G633, was detected after HEX-SSCP analysis of exon 11 of the *CP* gene. It occurs from a T to C transition at nucleotide position 1891 (g.1891 T→C). This polymorphism does not result in the replacement of glycine (G) at amino acid position 633 (G633). The heterozygous state of this synonymous variant was observed in 13 of 88 (14.8%) of the patient group (Table 3). It was not detected in the homozygous state and was absent from the control group. The serum ferritin (SF) levels of the patients displaying the G633 variant differed markedly from each other. Statistical significance was identified for serum iron ($p = 0.001$), transferrin ($p = 0.009$) and %TS ($p = 0.049$) with the heterozygous G633 variant (Table 4). The levels of serum iron, transferrin and %TS were lower in the individuals with the heterozygous variant relative to wildtype individuals. It is possible that the G633 variant could be acting in conjunction with other variants identified in the *CP* gene, or other genes in the iron metabolism pathway, to produce the differences in iron levels observed in each of the patients. This data, along with the fact the G633 variant was not detected in the control individuals, could indicate that this variant is associated with OC in the Black South African population. Further studies using the minigene system are warranted to determine what effect this variant has on protein expression in patients presenting with OC.[47]

With regards to the iron parameter statistics, the serum iron, transferrin, ferritin and C-reactive protein levels (CRP) were found to be statistically significantly different between patient and control subjects regardless of gender (Table 5). Mean serum iron and mean transferrin levels were found to be higher in the control subjects relative to the OC patient group, whereas the mean ferritin and CRP levels were significantly higher in the patient group relative to the control group. This represents a higher level of intracellular iron storage (bound to ferritin) in OC patients, as well as a higher level of inflammation as indicated by the increased levels of CRP protein.[48] The amount of transferrin present is dependent on the amount of serum iron present as a direct relationship exists between the two. This is demonstrated in this study as both the serum iron levels and the transferrin levels are lower in the OC patient cohort relative to the control group. The high level of iron stored as ferritin seen in cancer patients in general correlates with the reduced serum levels of iron available, characteristic of an inability to mobilize iron stores.

It is important to note that at the time of OC diagnosis, the majority of patients may be suffering from

malnutrition due to the location of the tumors in the upper region of the digestive tract and the subsequent inability to ingest sufficient nutrients. Iron is absorbed predominately from the diet and the body iron levels of OC patients may be compromised as a result of insufficient iron uptake. The relationship between iron parameters, measured at this stage of the disease, and the effects of variations in the iron metabolizing genes are therefore difficult to correlate.

This study constitutes an integral part of a much larger long-term analysis of the correlation between aberrations in iron regulatory genes and disease. The data generated from this study, in conjunction with that of the larger study, may potentially lead to the identification of molecular markers for OC. These markers could ultimately serve as an effective tool in the presymptomatic diagnosis and subsequent treatment of this disease.

Acknowledgements

The authors thank the University of Stellenbosch and the Department of Genetics for providing the infrastructure and facilities utilized in the completion of this study.

References

1. Crespi M, Bogomoletz VW, Munoz N. Cancer of the esophagus. *Gastroenterol Int* 1994;7: 24–35.
2. Sammon AM. Carcinogens and endemic squamous cancer of the oesophagus in Transkei, South Africa. Environmental initiation is the dominant factor; tobacco and other carcinogens of low potency or concentration are sufficient for carcinogenesis in the predisposed mucosa. *Med Hypot* 2007;69:125–31.
3. Makaula AN, Marasas WF, Venter FS, Badenhorst CJ, Bradshaw D, Swanevelder S. Oesophageal and other cancer patterns in four selected districts of the Transkei, Southern Africa: 1985–1990. *Afr J Health Sci* 1996;3:11–15.
4. Barber J. South Africa in the Twentieth Century. Oxford: Blackwell Publishers, 1999.
5. Blot WJ. Esophageal cancer trends and risk factors. *Semin Oncol* 1994;21:403–10.
6. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005;55: 74–108.
7. Lagergren J, Bergström R, Lindgren A, Nyrén O. Symptomatic gastoesophageal reflux as a risk factor for esophageal adenocarcinoma. *N Engl J Med* 1999; 340:825–31.
8. Parkin DM, Bray FI, Devessa SS. Cancer burden in the year 2000. The global picture. *Eur J Cancer* 2001;37:4–66.
9. Marasas WFO, van Rensburg SJ, Mirocha CJ. Incidence of Fusarium species and the mycotoxins, deoxynivalenol and zearalenone, in corn produced in esophageal cancer areas in Transkei. *J Agric Food Chem* 1979;27:1108.
10. Syrjanen KJ. Histological changes identical to those of condylomatous lesions found in esophageal squamous cell carcinomas. *Arch Geschwulstforsch* 1982;52:283–92.
11. Parent M-E, Siemiatycki J, Fritschi L. Workplace exposures and oesophageal cancer. *Occup Environ Med* 2000;57:325–34.
12. Matsha T, Erasmus R, Kafuko AB, Mugwanya D, Stepien A, Parker MI. Human papillomavirus associated with oesophageal cancer. *J Clin Pathol* 2002;55:587–90.
13. Marasas WF, Jaskiewicz K, Venter FS, Van Schalkwyk DJ. Fusarium moniliforme contamination of maize in oesophageal cancer areas in Transkei. *S Afr Med J* 1988;74:110–14.
14. Cai L, Mu L-N, Lu H, Lu Q-Y, Yuko You N-C, Yu S-Z, Le AD, Zhao J, Zhou X-F, Marshall J, Heber D, Zhang Z-F. Dietary selenium intake and genetic polymorphisms of the *GSTP1* and *p53* genes on the risk of esophageal squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev* 2006;15:294–300.
15. Lock K, Pomerleau, J, Causer L, McKee M. Low fruit and vegetable consumption. In: Ezzati M, Lopez AD, Rodgers A, Murray CJ, eds. Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors. Geneva: World Health Organization, 2004. 597–728.
16. Farrow DC, Vaughan TL, Sweeney C. Gastroesophageal reflux disease, use of H2 receptor antagonists, and risk of esophageal and gastric cancer. *Cancer Causes Control* 2000;11: 231–8.
17. Isaacson C, Bothwell TH, MacPhail AP, Simon M. The iron status of urban Black subjects with carcinoma of the oesophagus. *S Afr Med J* 1985; 67:591–3.
18. Mandishona E, MacPhail AP, Gordeuk VR, Kedda MA, Paterson AC, Rouault TA, Kew MC. Dietary iron overload as a risk factor for hepatocellular carcinoma in Black Africans. *Hepatology* 1998;27:1563–6.
19. Goldstein SR, Yang G-Y, Chen X, Curtis SK, Yang CS. Studies of iron deposits, inducible nitric oxide synthase and nitrotyrosine in a rat model for esophageal adenocarcinoma. *Carcinogenesis* 1998;19:145–1449.
20. Chen X, Ding YW, Yang G-Y, Bondoc F, Lee M-J, Yang CS. Oxidative damage in an esophageal adenocarcinoma model with rats. *Carcinogenesis* 2000;21:257–63.
21. Amer MH, El-Yazigi A, Hannan MA, Mohamed ME. Water contamination and esophageal cancer at Gassim Region, Saudi Arabia. *Gastroenterology* 1990;98:1141–7.
22. Rogers AE, Zeisel SH, Groopman J. Diet and carcinogenesis. *Carcinogenesis* 1993;14: 2205–17.
23. Weiss G, Fuchs D, Hausen A, Reibnegger G, Werner ER, Werner-Felmayer G, Wachter H. Iron modulates interferon-gamma effects in the human myelomonocytic cell line THP-1. *Exp Hematol* 1992;20:605–10.
24. Sargent PJ, Farnaud S, Evans RW. Structure/ function overview of proteins involved in iron storage and transport. *Curr Med Chem* 2005;12: 2683–93.
25. Looker AC, Johnson CL. Prevalence of elevated serum transferrin saturation in adults in the United States. *Ann Int Med* 1988;129:940–5.
26. Adams PC, Chakrabarti S. Genotypic/phenotypic correlation in genetic haemochromatosis: evolution of diagnostic criteria. *Gastroenterology* 1998;114:319.
27. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl Acids Res* 1988;16: 1215.
28. Kotze MJ, Theart L, Callis M, Peeters A, Thiart R, Langenhoven E. Nonradioactive multiplex PCR screening strategy for the simultaneous detection of multiple low-density lipoprotein receptor gene mutations. *PCR Methods Appl* 1995;4:352–6.
29. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–5.
30. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M. The structure of haplotype blocks in the human genome. *Science* 2002;296:2225–9.
31. Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucl Acids Res* 2004;32:91–4.
32. Wingender E, Chen X, Fricke E, Geffers R, Hehl R, Liebich I, Krull M, Matys V, Michael H, Ohnauser R, Pruss M, Schacherer F, et al. The TRANSFAC system on gene expression regulation. *Nucl Acids Res* 2001;29:281–3.
33. Kel A, Kel-Margoulis O, Borlak J, Tchekmenev D, Wingender E. Databases and tools for in silico analysis of regulation of gene expression. In: Borlak J, ed. Handbook of toxicogenomics: strategies and applications. Germany; Wiley-VCH, 2005. 253–85
34. Osaki S, Johnson DA, Frieden E. The possible significance of ferrous oxidase activity of ceruloplasmin in normal human serum. *J Biol Chem* 1966;241:2746–51.
35. Kastan MB, Onyekwere O, Sidransky D, Vogelstein B, Craig RW. Participation of p53

protein in the cellular response to DNA damage. *Cancer Res* 1991;51:6304–11.
36. Cho Y, Gorina S, Jeffrey PD, Pavletich NP. Crystal structure of a p53 tumour suppressor-DNA complex: understanding tumorigenic mutations. *Science* 1994;265:346–55.
37. Fleming RE, Gitlin JD. Structural and functional analysis of the 5′-flanking region of the rat ceruloplasmin gene. *Biol Chem* 1992;267: 479–86.
38. Linder MC, Moor JR, Wright K. Ceruloplasmin assays in diagnosis and treatment of human lung, breast, and gastrointestinal cancers. *J Natl Cancer Inst* 1981;67:263–75.
39. Rice EW. Evaluation of the role of ceruloplasmin as an acute-phase reactant. *Clin Chim Acta* 1961; 6:652–5.
40. Costa RH, Kalinichenko VV, Holterman A-XI, Wang X. Transcription factors in liver development, differentiation and regeneration. *Hepatology* 2003;38:1331–47.
41. Mukhopadhyay CK, Attieh ZK, Fox PL. Role of ceruloplasmin in cellular iron uptake. *Science* 1998;279:714–17.
42. Andrews NC. Disorders of iron metabolism. *N Engl J Med* 1999;341:1986–95.
43. Adrian GS, Seto E, Fischbach KS, Rivera EV, Adrian EK, Herbert DC, Walter FJ, Weaker FJ, Bowman BH. YY1 and SP1 transcription factors bind the human transferrin gene in an age-related manner. *J Gerontol A Biol Sci Med Sci* 1996;51: B66–B75.
44. Hilberg F, Aguzzi A, Howells N, Wagner EF. c-Jun is essential for normal mouse development and hepatogenesis. *Nature* 1993;365:179–81.
45. Vallet VS, Casado M, Henrion AA, Bucchini D, Raymondjean M, Kahn AS, Vaulont S. Differential roles of upstream stimulatory factors 1 and 2 in the transcriptional response to liver genes. *J Biol Chem* 1998;273:20175–9.
46. Tondravi MM, McKercher SR, Anderson K, Erdmann JM, Quiroz M, Maki R, Teitelbaum SL. Osteoporosis in mice lacking haematopoietic transcription factor PU.1. *Nature* 1997;386: 81–4.
47. Baralle M, Baralle D, De Conti L, Mattocks C, Whittaker J, Knezevich A, French-Constant C, Baralle FE. Identification of a mutation that perturbs *NF1* gene splicing using genomic DNA samples and a minigene assay. *J Med Genet* 2003;40:220–2.
48. Kalantar-Zadeh K, Rodriguez RA, Humphreys MH. Association between serum ferritin and measures of inflammation, nutrition and iron in haemodialysis patients. *Nephrol Dial Transplant* 2004;19:141–9.