Mutations linked to the pro $\alpha 2(I)$ collagen gene are responsible for several cases of osteogenesis imperfecta type I

G WALLIS*, P BEIGHTON*, C BOYD†, AND C G MATHEW‡

From *the MRC Research Unit for Inherited Skeletal Disorders, Department of Human Genetics, Medical School, University of Cape Town, Observatory 7925, South Africa; †the Department of Surgery, UMDNJ-Rutgers Medical School, CN 19 New Brunswick, New Jersey 08903, USA; and ‡the MRC Molecular and Cellular Cardiology Research Unit, University of Stellenbosch Medical School, Tygerberg 7505, South Africa.

SUMMARY We have analysed six South African families with osteogenesis imperfecta type I using three DNA polymorphisms associated with the pro $\alpha 2(I)$ collagen gene. In four of these families linkage of the pro $\alpha 2(I)$ gene and the osteogenesis imperfecta phenotype was suggested, whereas in the remaining two families there was a lack of linkage. No distinct correlation could be made between the phenotypic features of the families studied and linkage or lack of linkage to the pro $\alpha 2(I)$ gene. Two different haplotypes were found to be associated with the mutant pro $\alpha 2(I)$ alleles. These findings suggest that molecular heterogeneity exists within osteogenesis imperfecta type I and that in a significant proportion of cases the defect is linked to the pro $\alpha 2(I)$ gene.

Osteogenesis imperfecta (OI) is a heterogeneous disorder characterised by fragility of the bones and variable abnormalities of other collagen containing extraskeletal tissues.¹ The autosomal dominant forms of OI have been divided into two broad types, OI type I and OI type IV, on the basis of their clinical manifestations.² OI type I is the mildest form of the disorder and is characterised by normal stature, bone fragility without marked deformity, blue sclerae, early hearing loss in some patients,³ and the presence of dentinogenesis imperfecta⁴ in some families. OI type I is relatively common and constitutes the majority of all known cases of OI.^{2,5} There is considerable phenotypic variability within OI type I and it is very probable that there is further heterogeneity. The features of OI type IV are white or greyish sclerae, a moderate tendency to fracture, and variable bone deformity.⁶ This disorder has been reported in only a few families and its phenotypic differentiation from OI type I is far from absolute. Indeed, the phenotypic features of scleral colour and fracture severity are poor discriminants and their value in accurate subclassification is questionable. The problem is compounded by difficulties in subjective assessment and the probable influence of multiple epistatic factors.

Defects in the biosynthesis or structure of type I collagen have been found in several variants of OI.⁷⁻⁹ Biochemical studies have suggested that the OI type I phenotype is due to alterations in the synthesis of one of the pro $\alpha 1(I)$ chains of type I collagen.¹⁰ This has been substantiated by Barsh *et al*¹¹ who showed that the decreased production of type I procollagen by cells from three subjects with OI type I was the consequence of a non-functional allele for pro $\alpha 1(I)$.

A recent report has implicated a mutation in a pro $\alpha 2(I)$ chain in a patient with phenotypic features compatible with OI type IV.¹² Linkage studies using an *EcoRI* restriction fragment length polymorphism (RFLP) in the pro $\alpha 2(I)$ gene in families with mild forms of OI have also suggested that mutations in the pro $\alpha 2(I)$ gene are involved in OI type IV but not in OI type I.¹³

Recently, we have described linkage between an *RsaI* RFLP in the pro $\alpha 2(I)$ collagen gene and the OI phenotype in a family with OI type I.¹⁴ We have now used three pro $\alpha 2(I)$ collagen gene RFLPs¹³⁻¹⁶ in the analysis of this family and a further five families with OI type I. Our data suggest that defects in the pro $\alpha 2(I)$ gene are responsible for this phenotype in four of the families and that further molecular heterogeneity exists within this category of OI.

Patients, materials, and methods

PATIENTS

The patients from families who participated in this study had been examined and investigated by PB and were derived from a previously reported series of 153 affected South Africans.⁵ Following evaluation of their clinical and radiographical manifestations and their pedigrees, the patients were grouped according to the classification of OI proposed by Sillence *et al.*² The families in this study were categorised as having type I OI on the basis of these accepted criteria.² All families had autosomal dominant inheritance, a mild to moderate tendency to fracture, blue sclerae, and normal stature. Wormian bones in skull radiographs were a consistent feature, while dentinogenesis imperfecta and deafness were variable components.

MATERIALS

Restriction enzymes were obtained from Bethesda Research Laboratories, Boehringer Mannheim, and New England Biolabs, and were used as directed by the suppliers. ³²P labelled dCTP was obtained from New England Nuclear. Nitrocellulose was obtained from Schleicher and Schuell.

GENOMIC DNA HYBRIDISATION ANALYSIS

High molecular weight DNA was prepared from 10 ml of EDTA anticoagulated blood.¹⁷ Samples of 10

µg DNA were digested with the appropriate restriction enzymes. Digested DNA was separated by electrophoresis in 0.6% (w/v) agarose gels. The DNA fragments were transferred to nitrocellulose filters and hybridised with the human pro α2(I) probes for 48 hours at 65°C as previously described.¹⁷ Stringency washes for the cDNA probes were in 0.5×SSC at 65°C and genomic probes in 0.1×SSC at 65°C. The probes were labelled to a specific activity of 2 to 5×10⁸ cpm/µg by nick translation.

DNA PROBES FOR THE HUMAN PRO α2(I) GENE

The human pro α2(I) gene has been cloned,¹⁸ characterised,^{19,20} and localised.²¹ Three RFLPs of high frequency have been detected in this gene with the restriction enzymes *Eco*RI, *Rsa*I, and *Msp*I.¹³⁻¹⁶ The positions of the genomic and cDNA probes used in the detection of these RFLPs are shown in fig 1. Two contiguous *Eco*RI genomic clones were used to detect the *Eco*RI RFLP in the 5' region of the gene. The cDNA Hf32¹⁸ was used for the detection of the *Rsa*I RFLP. A 4 kb *Eco*RI genomic clone^{20,22} was used to detect the *Msp*I RFLP.

LINKAGE ANALYSIS

Subjects from 12 families with autosomal dominant OI type I were screened for the presence (+) or absence (-) of the *Eco*RI, *Msp*I, and *Rsa*I RFLPs

FIG 1 Restriction map indicating polymorphic sites in the human pro α2(I) collagen gene. (a) *Eco*RI RFLP detected with genomic clones NJ-3' and NJ-3''. (b) *Rsa*I and *Msp*I RFLPs in region of gene spanned by the cDNA, Hf32. E=*Eco*RI, R=*Rsa*I, M=*Msp*I, *=polymorphic site (only the relevant *Rsa*I and *Msp*I sites have been indicated).

(fig 2). Ten of the 12 families were informative at one or more sites. Four of these families were excluded owing to the lack of cooperative or accessible family members. The genotypes of the available subjects from the remaining six families were determined for the *EcoRI*, *RsaI*, and *MspI* RFLPs. Lod scores were calculated at various recombination fractions (θ) using the computer programme LIPED.²³

Results

HETEROZYGOTE FREQUENCIES FOR THE RFLPS IN THE PRO $\alpha 2(I)$ GENE

In order for an RFLP to be useful in a linkage analysis of an autosomal dominant disorder, affected subjects should be heterozygous. We have therefore determined the frequencies of the *EcoRI* (this paper), *MspI*,¹⁶ and *RsaI*¹⁴ RFLPs in three South African population groups. The results are given in table 1 for the Europeans, the so-called 'Cape Coloureds'²⁴ (a population of mixed ancestry), and the South African Blacks. The *EcoRI* frequencies were calculated from the genotypes of 18 unrelated healthy subjects from each population group. It is significant that the minor allele in the European population (+) is the major allele in the

Black population. This probably contributes to the high degree of heterozygosity for the *EcoRI* RFLP in the Cape Coloured population.

The combined use of the three RFLPs results in a predicted level of heterozygosity which exceeds 50% in all three population groups (table 1). This predicted level of heterozygosity assumes that no linkage disequilibrium exists between the marker loci. Although there is as yet no conclusive evidence of linkage disequilibrium, the close proximity of the polymorphic loci raises the possibility that the various alleles may associate preferentially with one another in the general population to form a frequently occurring haplotype.

TABLE 1 Frequencies and combined percentage heterozygosity for the *EcoRI*, *MspI*, and *RsaI* RFLPs in randomly selected, unaffected subjects in three South African populations. The minor allele is indicated in parentheses.

	<i>EcoRI</i>	<i>MspI</i> *	<i>RsaI</i> †	% heterozygosity
European	0.42 (+)	0.17 (-)	0.38 (-)	80.5
'Cape Coloured'	0.44 (+)	0.13 (-)	0.11 (-)	60.8
SA Black	0.37 (-)	0.05 (-)	0.00	51.7

*Previously reported in reference 16.

†Previously reported in reference 14.

FIG 2 Southern blots showing RFLPs associated with the pro $\alpha 2(I)$ collagen gene. DNA samples digested with (a) *EcoRI* and hybridised to NJ-3' and NJ-3", (b) *RsaI* and hybridised to Hf32, and (c) *MspI* and hybridised to a 4 kb *EcoRI* genomic subclone. The presence and absence of the polymorphic sites are indicated by + and - respectively.

LINKAGE ANALYSIS

The pedigrees of the six dominant OI type I families where linkage analysis was feasible are given in fig 3. The polymorphic markers used in establishing linkage or non-linkage are indicated on the pedigrees. The lod scores obtained in families a to f are given in table 2.

In families a to d positive lod scores were obtained, providing presumptive evidence that OI in these families is caused by mutations within or close to the pro $\alpha 2(I)$ gene. The lod score for family d is low as the family is small and is not the result of the

TABLE 2 Lod scores and haplotypes in families with OI type I.

Family	θ	Lod scores	Haplotypes of the OI-pro $\alpha 2(I)$ allele*		
			EcoRI	RsaI	MspI
a	0	3.91	-	-	+
b	0	2.7	-	-	+
c	0	1.5	-	-	+
d	0	0.602	-	+	+
e	0.05	-1.442			
f	0.05	-1.442			

*, + and - indicate presence and absence of the cutting site respectively.
† θ =recombination fraction.

FIG 3 Pedigrees of families with OI type I. Informative restriction sites are designated by E=EcoRI, M=MspI, and R=RsaI. Subjects heterozygous (+/-) and homozygous for the presence (+/+) and absence (-/-) of the polymorphic sites are indicated. The shaded symbols indicate affected subjects. The probands are arrowed. Family a is as previously reported in reference 14.

detection of recombinants. In families e and f the negative lod scores indicate a lack of linkage between the pro $\alpha 2(I)$ gene and the OI phenotype. In families a to c the same haplotype (—+) for the *EcoRI*, *RsaI*, and *MspI* RFLPs respectively was coinherited with the OI phenotype, whereas in family d the haplotype of the pro $\alpha 2(I)$ allele is different (—++).

CLINICAL PHENOTYPES

The four families in whom linkage was demonstrated all had normal stature, a mild to moderate tendency to fracture, and definite but variable blueness of the sclerae. Dentinogenesis imperfecta was present in three of these kindreds, but none had deafness.

The two families in whom linkage was not demonstrated had a similar phenotype, but differed in that deafness had developed in early adulthood in some family members. The significance of this observation, if any, is uncertain, although it may be indicative of heterogeneity.

Within all the families studied there was considerable intrafamilial variation including frequency of fracturing, depth of scleral blueness, and presence and severity of dentinogenesis imperfecta.

Discussion

By establishing linkage and lack of linkage between the pro $\alpha 2(I)$ gene and the OI type I phenotype, we have demonstrated that molecular heterogeneity exists within this form of OI and that in a significant proportion of cases the defect is in or near the pro $\alpha 2(I)$ gene. No distinct correlation could be made between the phenotypic features of the families studied and linkage or lack of linkage to the pro $\alpha 2(I)$ gene. Thus, the non-allelic mutations causing the OI type I phenotype possibly contribute to the observed heterogeneity, but do not subdivide this form of OI at the clinical level.

Mutations in pro $\alpha 2(I)$ alleles producing the OI type I phenotype are likely to be structural, as a decreased rate of synthesis of pro $\alpha 2(I)$ chains has been demonstrated in subjects without OI.²⁵ To date, structural mutations in the pro $\alpha 2(I)$ chains have been shown to produce mild variant forms of OI,^{26 27} as well as the OI type II,²⁸ OI type III,²⁵ and OI type IV phenotypes.^{11 29} Only mutations which reduce the expression of one of the pro $\alpha 1(I)$ alleles resulting in a decreased production of normal collagen have been shown to produce the OI type I phenotype. Structural mutations in pro $\alpha 2(I)$ could conceivably result in a similar decreased production of normal collagen if the type I collagen trimers

incorporating the abnormal pro $\alpha 2(I)$ chain were not secreted.²⁹ Although mutations of this nature are possible, it is likely that many other structural mutations in pro $\alpha 2(I)$ exist which contribute to the spectrum of phenotypic variation which encompasses both OI type I and OI type IV.

Two different haplotypes were recognised in the mutant pro $\alpha 2(I)$ alleles in families where linkage of OI to the pro $\alpha 2(I)$ gene was established. Preliminary results have shown that the mutations in family a and family d map to different regions of the pro $\alpha 2(I)$ collagen chain (unpublished data). Families a to c have the same linked haplotype and it is possible that an association may exist between this haplotype and a specific mutation within the gene. A close association between DNA haplotypes and specific thalassaemia mutations has been described.³⁰ This association, although not invariant, was used to develop a strategy for the characterisation of β thalassaemia mutations. With the detection of an increasing number of RFLPs in the pro $\alpha 2(I)$ gene, such a strategy may well be applicable in the efficient characterisation of the mutations causing the OI type I phenotype. The lack of linkage to the pro $\alpha 2(I)$ gene in families e and f suggests the involvement of the pro $\alpha 1(I)$ gene.^{10 11} This has been confirmed in the case of family e where a decreased production of pro $\alpha 1(I)$ has been observed (unpublished results).

Apart from interfamilial variation, considerable intrafamilial variation was evident in the families studied. The phenotypic variation observed in these families is important as it emphasises the range of phenotypes that can result from the same molecular defect. This intrafamilial variation must result from the expression of other genes that modify the phenotype. A precedent for the presence of modifying genes exists in the β thalassaemias and sickle cell anaemia. The severity of these disorders has been shown to be decreased where, for example, there is an increased expression of fetal haemoglobin.³¹ Possibly, the intrafamilial variation observed in OI could similarly result from an increased expression of other collagen types in some subjects. Thus, the clinical heterogeneity in OI type I can be the result of mutations in different alleles, different mutations in the same allele, and the variable expression of modifying genes.

We are grateful to Drs J Myers, F Ramirez, P Tsipouras, and D Prockop for the pro $\alpha 2(I)$ cDNA, Hf32, and the pro $\alpha 2(I)$ *EcoRI* genomic subclones NJ-3' and NJ-3'', and to Drs R Dalgleish and R Kaufman for the 4 kb pro $\alpha 2(I)$ genomic subclone. We would like to thank the genetic nursing sisters, G Christy and L Merckel, for obtaining the blood

samples, J Schram and L van Herwerden for technical assistance, and E Lavin and D Cannell for helping to prepare the manuscript. This work was supported by grants from the Medical Research Council of South Africa, the University of Cape Town Staff Research Fund, the Mauerberger Foundation, and the Harry Crossley Foundation.

References

- ¹ Smith R, Francis M, Houghton G. *The brittle bone syndrome*. London: Butterworths, 1983.
- ² Silience DO, Senn A, Danks DM. Genetic heterogeneity in osteogenesis imperfecta. *J Med Genet* 1979;16:101-16.
- ³ Shapiro JR, Pikus A, Weiss G, Rowe DW. Hearing and middle ear function in osteogenesis imperfecta. *JAMA* 1982;247:2120-6.
- ⁴ Levin LS, Salinas CF, Jorgenson RJ. Classification of osteogenesis imperfecta by dental characteristics. *Lancet* 1978;i:332-3.
- ⁵ Beighton P, Spranger J, Versfeld G. Skeletal complications in osteogenesis imperfecta: a review of 153 South African patients. *S Afr Med J* 1983;64:565-68.
- ⁶ Paterson CR, McAllion S, Miller R. Osteogenesis imperfecta with dominant inheritance and normal sclerae. *J Bone Joint Surg (Br)* 1983;65:35-9.
- ⁷ Prockop DJ. Osteogenesis imperfecta: phenotypic heterogeneity, protein suicide, short and long collagen. *Am J Hum Genet* 1984;36:499-505.
- ⁸ Byers PM, Bonadio JP, Steinmann B. Osteogenesis imperfecta: update and perspective. *Am J Med Genet* 1984;17:429-35.
- ⁹ Prockop DJ, Kivirikko KI. Heritable diseases of collagen. *N Engl J Med* 1984;311:376-86.
- ¹⁰ Francis MJO, Williams KJ, Sykes BC, Smith R. The relative amounts of the collagen chains $\alpha 1(I)$, $\alpha 2(I)$ and $\alpha 1(III)$ in the skin of 31 patients with osteogenesis imperfecta. *Clin Sci* 1981;60:617-23.
- ¹¹ Barsh GS, David KE, Byers PM. Type I osteogenesis imperfecta: a non-functional allele for pro $\alpha 1(I)$ chains of type I procollagen. *Proc Natl Acad Sci USA* 1982;79:3838-42.
- ¹² Wenstrup RJ, Tsiouras P, Byers PH. Osteogenesis imperfecta (OI) type IV: a mutation in pro $\alpha 2(I)$ produces an unstable type I collagen. *Am J Hum Genet* 1984;36:23S.
- ¹³ Tsiouras P, Borrenson AL, Dickson LA, Berg K, Prockop DJ, Ramirez F. Molecular heterogeneity in the mild autosomal dominant forms of osteogenesis imperfecta. *Am J Hum Genet* 1984;36:1172-9.
- ¹⁴ Grobler-Rabie AF, Wallis G, Brebner DK, Beighton P, Bester AJ, Mathew CG. Detection of a high frequency RsaI polymorphism in the human pro $\alpha 2(I)$ collagen gene which is linked to an autosomal dominant form of osteogenesis imperfecta. *EMBO J* 1985;4:1745-8.
- ¹⁵ Tsiouras P, Myers JL, Ramirez F, Prockop DJ. Restriction fragment length polymorphisms associated with the pro $\alpha 2(I)$ gene of human type I collagen. Application to a family with an autosomal dominant form of osteogenesis imperfecta. *J Clin Invest* 1983;72:1261-7.
- ¹⁶ Grobler-Rabie AF, Brebner DK, Vandenplas S, et al. Polymorphism of DNA sequence in the human pro $\alpha 2(I)$ collagen gene. *J Med Genet* 1985;22:182-6.
- ¹⁷ Vandenplas S, Wiid I, Grobler-Rabie A, et al. Blot hybridisation analysis of genomic DNA. *J Med Genet* 1984;21:164-72.
- ¹⁸ Myers JC, Chu ML, Faro SM, Clark WJ, Prockop DJ, Ramirez F. Cloning a cDNA for the pro $\alpha 2$ chain of human type I collagen. *Proc Natl Acad Sci USA* 1981;78:3516-20.
- ¹⁹ Bernard MP, Myers KC, Chu ML, Ramirez F, Eikenberry EF, Prockop DJ. Structure of a cDNA for the pro $\alpha 2(I)$ chain of human type I procollagen. Comparison with chick cDNA for pro $\alpha 2(I)$ identifies structurally conserved features of the protein and gene. *Biochemistry* 1983;22:1139-45.
- ²⁰ Tajima S, Ting JPY, Pinnell SR, Kaufman RE. Isolation and characterisation of a human pro $\alpha 2(I)$ collagen gene segment. *J Invest Dermatol* 1984;82:265-9.
- ²¹ Henderson AS, Myers JC, Ramirez F. Localization of the human pro $\alpha 2(I)$ collagen gene (COLIA2) to chromosome 7q22. *Cytogenet Cell Genet* 1983;36:586-7.
- ²² Dalglish R, Trapnell BC, Crystal RG, Tolstoshev P. Copy number of human type I $\alpha 2$ collagen gene. *J Biol Chem* 1982;257:13816-22.
- ²³ Ott J. Estimation of the recombination fraction in human pedigrees: efficient computation of the likelihood for human linkage studies. *Am J Hum Genet* 1974;26:588-97.
- ²⁴ Botha MC, Pritchard J. Blood group gene frequencies. An indication of the genetic constitution of population samples in Cape Town. *S Afr Med J* 1972;suppl 1:1-27.
- ²⁵ Deak SB, Nicholls AC, Pope FM, Prockop DJ. The molecular defect in a nonlethal variant of osteogenesis imperfecta: synthesis of pro $\alpha 2(I)$ chains which are not incorporated into trimers of type I procollagen. *J Biol Chem* 1983;258:15192-7.
- ²⁶ Byers PM, Shapiro JR, Rowe DW, Davids KE, Holbrook KA. Abnormal $\alpha 2$ chain in type I collagen from a patient with a form of osteogenesis imperfecta. *J Clin Invest* 1983;71:689-97.
- ²⁷ Sippola M, Kaffe S, Prockop DJ. A heterozygous defect for structurally altered pro $\alpha 2$ chains of type I procollagen in mild variant of osteogenesis imperfecta. *J Biol Chem* 1984;259:14094-100.
- ²⁸ de Wet WJ, Pihlajaniemi T, Myers J, Kelly TE, Prockop DJ. Synthesis of a shortened pro $\alpha 2(I)$ chain and decreased synthesis of pro $\alpha 2(I)$ chains in a proband with osteogenesis imperfecta. *J Biol Chem* 1983;258:7721-8.
- ²⁹ Byers PM, Bonadio JF. The molecular basis of clinical heterogeneity in osteogenesis imperfecta: mutations of type I collagen genes have different effects on collagen processing. In: Scriver CR, Lloyd J, eds. *Genetic and metabolic disease*. International Medical Reviews. Paediatrics 5. London: Butterworths (in press).
- ³⁰ Kazazian HH, Orkin SH, Markham AF, Chapman CR, Yousoufian H, Waber PG. Quantification of the close association between haplotypes and specific β -thalassaemia mutations in Mediterraneans. *Nature* 1984;310:152-4.
- ³¹ Nienhuis AW, Anagnou NP, Ley TJ. Advances in thalassaemia research. *Blood* 1984;63:738-58.

Correspondence and requests for reprints to Dr G Wallis, MRC Research Unit for Inherited Skeletal Disorders, Department of Human Genetics, Medical School, University of Cape Town, Observatory 7925, South Africa.