

A CLINICAL AND MOLECULAR GENETIC STUDY
OF THE SKELETAL DYSPLASIA
DYGGVE MELCHIOR CLAUSEN SYNDROME

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Esther Kinning MbChB (Leicester) MRCPCH
Department of Genetics
University of Leicester

April 2008

UMI Number: U526032

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U526032

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

I would like to thank the Medical Research Council who funded the Clinical Research Training Fellowship which allowed me to undertake this study.

I am grateful to my supervisors, Professor Richard Trembath and Dr J Howard Pringle and to the many staff within the Departments of Genetics, Biochemistry, Cancer Studies and Molecular Medicine at the University of Leicester who provided help and advice and without whom it would not have been possible to complete this work.

A Clinical and Molecular Genetic Study of the Skeletal Dysplasia Dyggve Melchior Clausen Syndrome

Abstract

Esther Kinning

Dyggve Melchior Clausen (DMC) syndrome is an autosomal recessive skeletal dysplasia caused by mutations in the *Dymeclin* (*DYM*) gene on chromosome 18q12-21. Affected individuals have multiple bony abnormalities and mental retardation. Histological examination of affected cartilage reveals failure of chondrocyte columnarisation, dilated rough endoplasmic reticulum and multiple vacuoles. The molecular basis of these pathological findings remains unclear.

The aim of this work was to elucidate the function of the *DYM* gene product and determine the mechanisms by which mutation of the disease gene lead to the cellular and clinical phenotype. Ten affected individuals from eight families were recruited to the study and five *DYM* mutations identified. These included two novel and complex genomic duplication/repetition events each predicted to result in a truncated transcript.

In-silico analyses of *DYM* suggest it encodes a transmembrane protein involved in protein sorting and targeting within the cell. The *DYM* transcript was shown by *in-situ* hybridisation to be expressed at high levels in cartilage and brain, particularly in resting and hypertrophic chondrocytes. Sub-cellular localisation demonstrated the *DYM* gene product to be located within the endoplasmic reticulum.

Yeast two-hybrid analysis performed to detect Dymeclin interacting proteins identified EGF-containing fibulin-like extracellular matrix protein and vacuolar protein sorting protein 25 (human homologue known as EAP20), respectively an extracellular matrix (ECM) protein and a component of the endosomal sorting complex required for transport which acts in the transport of transmembrane proteins for export and recycling.

Taken together, these findings indicate that Dymeclin is an endoplasmic reticulum transmembrane protein required for cargo transport through the endosomal compartment. Abnormal chondrocyte differentiation and brain function occur in the absence of adequate functional Dymeclin. Given that cartilage and brain both have substantial requirements for extracellular matrix, it is suggested that Dymeclin contributes to the transport of components of the ECM.

Table of Contents

List of abbreviations	4
Chapter 1- Introduction	7
1.1 General introduction	7
1.2 Overview of the skeletal dysplasias	7
1.3 Features of Dyggve Melchior Clausen syndrome	9
1.3.1 Clinical features	9
1.3.2 Radiological features	10
1.3.3 Histological features	11
1.4 Mapping of DMC gene by autozygosity mapping	12
1.5 Identification of the DMC gene	13
1.6 From disease gene identification to elucidation of the underlying molecular mechanism	15
1.7 Development and structure of the skeletal system	16
1.8 Protein synthesis, transport and degradation	20
1.9 Yeast two-hybrid system	21
1.10 Summary	22
Chapter 2 – Materials and Methods	24
2.1 Materials	24
2.1.1 General reagents	24
2.1.2 Enzymes	24
2.1.3 Cell culture media and reagents	24
2.1.4 Oligonucleotides	24
2.1.5 Molecular Biology kits	25
2.1.6 Patient samples	25
2.1.7 Common solutions	25
2.2 Methods	30
2.2.1 Multi-centre Research Ethical Committee (MREC) approval	30
2.2.2 Family recruitment	30
2.2.3 DNA extraction	30
2.2.3.1 Blood	30
2.2.3.2 Cell lines	30
2.2.4 RNA extraction	31
2.2.4.1 Blood	31
2.2.4.2 Cell lines	31
2.2.5 Oligonucleotide design	31
2.2.5.1 Genomic DNA	31
2.2.5.2 cDNA	31
2.2.5.3 Cloning oligonucleotides	31
2.2.6 PCR	32
2.2.6.1 PCR of microsatellite markers	32
2.2.6.2 PCR for genomic sequencing	32
2.2.6.3 PCR for cloning	32
2.2.6.4 Fluorescent dosage PCR	33
2.2.7 Agarose gel electrophoresis	33
2.2.8 Genotyping of microsatellites	33
2.2.9 Fluorescent DNA sequencing	34
2.2.10 RT-PCR	34
2.2.11 Fluorescent dosage PCR- calculation of dosage quotients	35
2.2.12 Southern blot hybridisation	35
2.2.13 Cloning methods	36

2.2.13.1 Preparation of competent <i>E.coli</i>	36
2.2.13.2 Restriction enzyme digestion	36
2.2.13.3 DNA blunt ending	36
2.2.13.4 DNA phosphorylation	37
2.2.13.5 DNA ligation	37
2.2.13.6 Bacterial transformation	37
2.2.13.7 Storage and growth of bacteria.....	38
2.2.13.8 Preparation of plasmid DNA	38
2.2.13.8.1 Small scale.....	38
2.2.13.8.2 Large scale.....	38
2.2.14 Cell culture methods.....	38
2.2.14.1 Cell culture	38
2.2.14.2 Preparation of cell extracts	39
2.2.14.3 Transient transfection	39
2.2.15 <i>In-vitro</i> translation of plasmids	39
2.2.16 SDS-PAGE analysis of proteins.....	40
2.2.17 Western blot analysis.....	40
2.2.18 <i>In-situ</i> hybridisation	41
2.2.18.1 Probe generation.....	41
2.2.18.2 Hybridisation.....	41
2.2.19 Yeast two-hybrid analysis	42
2.2.19.1 Growth and maintenance of yeast	42
2.2.19.2 Small scale yeast transformation.....	42
2.2.19.3 Large scale yeast transformation.....	43
2.2.19.4 Plasmid rescue.....	44
2.2.20 MBP pulldown	44
2.2.20.1 MBP protein expression and purification.....	44
2.2.20.2 MBP pulldown assay	44
2.2.21 Database analyses.....	45
Chapter 3 - Identification and characterisation of mutant <i>DYM</i> alleles	46
3.1 Introduction	46
3.2 Family resource	47
3.3 Microsatellite marker analysis.....	47
3.4 Identification of <i>DYM</i> mutations by direct sequencing.....	48
3.5 RT-PCR of patient cDNA	48
3.6 Fluorescent dosage PCR.....	50
3.7 Southern blot analysis	50
3.8 Discussion	51
3.8.1.1 <i>DYM</i> mutations.....	51
3.8.1.2 Missense mutations	52
3.8.1.3 Duplication mutations	52
3.8.1.4 Limitations of techniques used for mutation identification.....	53
3.8.2 Exon duplication/repetition as a disease mechanism	54
3.8.3 Premature termination codons and nonsense mediated mRNA decay.....	56
3.8.4 Implications for other recessive disorders.....	57
3.8.5 Summary	57
Chapter 4 – Characterisation of the <i>DYM</i> gene product.....	58
4.1 Introduction	58
4.2 <i>In-silico</i> bioinformatic analysis	58
4.3 RT-PCR.....	59
4.4 <i>In-vitro</i> translation.....	60
4.5 Transient transfection studies.....	60

4.6 Western blot analysis.....	61
4.7 Discussion	62
Chapter 5 - <i>In-situ</i> hybridisation studies	64
5.1 Introduction	64
5.2 Preparation of mouse tissue.....	66
5.3 Probe design	66
5.4 Embryology and histology of the skeletal system.....	67
5.5 Embryology and histology of the nervous system	68
5.6 <i>DYM</i> expression in fetal mouse tissues	69
5.7 <i>DYM</i> expression in adult mouse tissues	69
5.8 Discussion	70
Chapter 6 - Identification of Dymeclin interactors by yeast two-hybrid analysis.....	74
6.1 Introduction	74
6.2 Characterisation of a prey cDNA library.....	74
6.3 Construction of <i>DYM</i> yeast two-hybrid bait.....	75
6.4 Verification of DYM1-98 and 449-669 as suitable bait.....	76
6.5 Optimisation of transformation efficiency	77
6.6 Library transformation and analysis of positive clones.....	78
6.7 Sequence identification of interacting clones.....	78
6.8 Confirmation of the interaction of DYM with EAP20 and EFEMP1	79
6.9 Discussion	81
6.9.1 Dymeclin interactors	81
6.9.2 Other skeletal disorders attributed to defects in ER transport.....	84
6.9.3 Experimental limitations	85
6.9.4 Summary	86
Chapter 7 - General discussion.....	87
Appendix A	91
Appendix B.....	92
Appendix C.....	93
References	94

List of abbreviations

ABP	actin binding protein
ACAA	acetyl-coenzyme A acyltransferase
AD	activation domain
AER	apical ectodermal ridge
Amp	ampere
APS	ammonium persulphate
ARSE	arylsulphatase E
ASPM	abnormal spindle-like microcephaly associated protein
AZM	autozygosity mapping
BAC	bacterial artificial chromosome
BCIP	bromo-chloro-indolyl phosphate
BLAST	basic local alignment search tool
βME	beta-mercaptoethanol
BMP	bone morphogenetic protein
BMPR	bone morphogenetic protein receptor
bp	base pairs
BSA	bovine serum albumin
cDNA	complementary DNA
CGH	comparative genomic hybridisation
CLSD	cranio-lenticulo-sutural dysplasia
cM	centimorgan
CNV	copy number variation
COL	collagen
COMP	cartilage oligomeric matrix protein
dATP	deoxy-adenosine triphosphate
DBD	DNA binding domain
dCTP	deoxy-cytidine triphosphate
DEPC	diethylpyrocarbonate
dGTP	deoxy-guanosine triphosphate
DHPLC	denaturing high performance liquid chromatography
DHRD	Doyne honeycomb retinal dystrophy
Dig	Digoxigenin
DMC	Dyggve Melchior Clausen syndrome
DMEM	Dulbecco's modified eagle media
DMSO	di-methyl sulphoxide
DNA	deoxy-ribonucleic acid
dNTP	deoxy-ribonucleotide triphosphate
DQ	dosage quotient
dTTP	deoxy-thymidine triphosphate
dUTP	deoxy-uridine triphosphate
DYM	dymeclin
EAP	ELL associated protein
ECACC	European collection of cell cultures
ECM	extracellular matrix
EDTA	disodium ethylenediaminetetraacetate
EFEMP	EGF-containing fibulin-like extracellular matrix protein
EGF	epidermal growth factor
EJC	exon-exon junction complex
EMBL	European molecular biology laboratory
ENC	ectodermal neural cortex
ER	endoplasmic reticulum
ERAD	endoplasmic reticulum associated degradation

ESCRT	endosomal complex required for transport
EST	expressed sequence tag
EXT	exostosin
FISH	fluorescent <i>in-situ</i> hybridisation
FGD	facio-genital dysplasia
FGF	fibroblast growth factor
FGFR	fibroblast growth factor receptor
GAG	glycosaminoglycan
GFP	green fluorescent protein
GH	growth hormone
GRP	glycine rich RNA binding protein
hCGBP	human CpG binding protein
HEPES	hydroxyethyl piperazine N-2- ethane sulphonic acid
HNPP	hereditary neuropathy and liability to pressure palsy
HOX	homeobox
hr	hour
HSPG	heparan sulphate proteoglycan
IGF	insulin like growth factor
IHH	indian hedgehog
IMAGE	integrated molecular analysis of genomes and their expression
INCDB	International nomenclature of constitutional diseases of bone
IPTG	isopropothiolactose
INCDB	International nomenclature for constitutional disorders of bone
kb	kilobase
kD	kilodalton
l	litre
LB	luria broth
LCR	low copy repeat
LMAN	lectin mannose binding
LOD	logarithm of the odds
M	molar
MADH	mothers-against-decapentaplegic homologue
MAPH	multiplex amplifiable probe hybridisation
MATN	matrilin
MBD	methyl-CpG binding domain
MBP	maltose binding protein
MED	multiple epiphyseal dysplasia
MEPS	minimal efficient processing segment
mg	milligram
µg	microgram
min	minute
ml	millilitre
µl	microlitre
µM	micromolar
MMP	matrix metalloproteinase
MOPD	microcephalic osteodysplastic primordial dwarfism
MREC	multi-centre research ethics committee
mRNA	messenger RNA
MYO	myosin
NAHR	non-allelic homologous recombination
NBT	nitro-blue tetrazolium
NCBI	national centre for biotechnology information
NHEJ	non-homologous end joining
ng	nanogram

NMD	nonsense mediated mRNA decay
OD	optical density
OMIM	online Mendelian inheritance in man
PAGE	poly-acrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEX	peroxisome biogenesis factor
PFGE	pulsed field gel electrophoresis
PMSF	phenylmethanesulphonyl-fluoride
PSACH	pseudoachondroplasia
PTH	parathyroid hormone
PTHrP	parathyroid hormone related peptide
Rcf	relative centrifugal force
RER	rough endoplasmic reticulum
RFLP	restriction fragment length polymorphism
RMRP	mitochondrial RNA-processing endoribonuclease
RNA	ribonucleic acid
RPMI	Roswell Park memorial institute
RS	replication slippage
RT-PCR	reverse transcriptase polymerase chain reaction
RUNX	runt related transcription factor
sec	second
SDS	sodium dodecyl sulphate
SMC	Smith McCort dysplasia
SNP	single nucleotide polymorphism
SOX	SRX related homeobox gene
SRP	signal recognition particle
SRX	sex determining region of Y chromosome
TEMED	tetramethylethylenediamine
TRIS	tris(hydroxymethane) aminomethane
U	units
UTR	untranslated region
VEGF	vascular endothelial growth factor
VPS	vacuolar sorting protein
WNT	wingless type MMTV integration
YNB	yeast nutrient broth
YPD	yeast extract, peptone and dextrose
ZPA	zone of polarising activity

Chapter 1- Introduction

1.1 General introduction

This study focuses on the skeletal dysplasia Dyggve Melchior Clausen syndrome (DMC), a developmental bone disorder caused by recessive mutations in the gene *Dymeclin* (*DYM*) and named after the clinicians who first described the condition in 1962 (El Ghouzzi *et al.*, 2003, Dyggve *et al.*, 1962, Ehtesham *et al.*, 2002, Cohn *et al.*, 2003).

The broad aim of the work is to elucidate the function of the *DYM* gene product and determine the functional mechanisms relating disease gene mutation to cellular and clinical phenotype. This broad aim can be subdivided into the following components;

- To identify and characterise mutant *DYM* alleles.
- To perform detailed subcellular and histological localisation of the *DYM* transcript and gene product.
- To identify proteins interacting with DYM using the yeast two-hybrid system.

The introduction will include an overview of the skeletal dysplasias and the specific features of Dyggve Melchior Clausen syndrome will be discussed. In order to understand the aetiology of these disorders of bone, it is necessary to have some knowledge of the processes involved in skeletal development and the control and homeostasis of bone and cartilage. The embryological development of the skeleton, the structure and function of bone and cartilage and the control of bone growth will each be reviewed.

Mutations in *DYM* have been identified as part of the project hence the relationship between disease causing mutation and clinical phenotype is approached. In addition to studying *DYM* mutations to gain further insight as to disease causing mechanisms, functional studies of the *DYM* gene product were undertaken and DYM was found to lie within the endoplasmic reticulum (ER). A yeast two-hybrid analysis was performed in order to identify proteins which interact with DYM which identified a component of the extracellular matrix and a protein required for endosomal transport. Hence protein production, transport and degradation in relation to these cellular compartments will be discussed.

1.2 Overview of the skeletal dysplasias

The skeletal dysplasias are a heterogeneous group of at least 370 different disorders, affecting approximately 1 in 4000 of European populations (Superti-Furga and Unger, 2007, Stoll *et al.*, 1989). Skeletal dysplasias are characterised as abnormalities of bone

structure and function, typically resulting in altered size and shape of the skull, limbs and/or trunk (Beighton *et al.*, 1992). They have traditionally been classified clinically according to whether the limbs and/or trunk are affected and subdivided by which segments of the long bones are most severely affected. Radiological classification is based upon the region of long bone involved and whether the spine is affected. Long bone can be divided into the diaphysis or shaft, the epiphysis which is located at the end of the bone and the metaphysis which separates the two and during childhood contains the growth plate (see Figure 1.1). The International Nomenclature of Constitutional Diseases of Bone (INCDB) was set up in 1970 in order to establish a worldwide classification system for skeletal dysplasias and has subsequently been updated regularly. More recently, following the rapid increase in our knowledge of the molecular basis of these conditions, a molecular pathogenetic classification has been proposed and has been incorporated to some extent in the most recent INCDB. Of the 372 conditions included in this classification, 140 genes have been implicated in 215 different phenotypes. The skeletal dysplasias offer good examples of conditions which show allelic heterogeneity. For example, mutations in *Fibroblast growth factor receptor 3* (*FGFR3*) cause the lethal condition thanatophoric dysplasia in addition to achondroplasia characterised by severe short stature and the milder hypochondroplasia phenotype. The Nosology does not offer a purely molecular classification as there are many conditions which are genetically heterogeneous and others for which the underlying genetic cause has not yet been identified (Superti-Furga *et al.*, 2001, Superti-Furga and Unger, 2007). The 2006 classification is summarised in Table 1.1, the first six diagnostic categories having a molecular basis and the remainder based on clinical or radiological findings. There are a total of 37 classes, reflecting the broad range of phenotypes within the group.

Genes causative of conditions within the skeletal dysplasia group encode a wide range of molecules with different functions and can be broadly classified into extracellular structural proteins (eg. *Collagen I* $\alpha 1/2$ (*COL1A1/2*) in osteogenesis imperfecta), enzymes (eg. *arylsulphatase E* (*ARSE*) in X-linked chondrodysplasia punctata), defects in macromolecular folding/degradation (eg. *lysosomal acid hydrolases* in lysosomal storage disorders), hormones (eg. *parathyroid hormone* (*PTH*) in metaphyseal dysplasia Jansen type), transcription factors (*sex determining region of Y chromosome* (*SRY*) *related homeobox gene 9* (*SOX9*) in campomelic dysplasia), oncogenes (eg. *exostosin* (*EXT1*) in multiple exostoses syndrome) and deoxy-ribonucleic acid (DNA) or ribonucleic acid (RNA) processors (eg. *RNAse mitochondrial RNA processing component* in cartilage-hair hypoplasia) (Superti-Furga *et al.*, 2001). This molecular

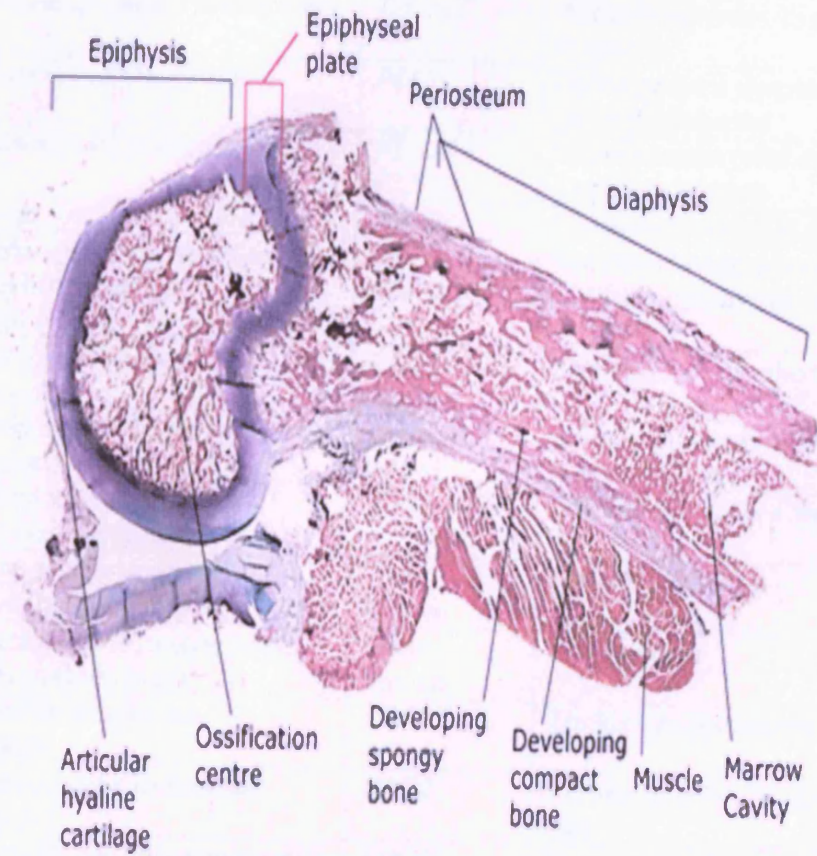


Figure 1.1. Structure of long bone. The long bone can be divided into the diaphysis (shaft), and epiphysis separated by the epiphyseal (growth) plate. From www.nlm.nih.gov/medlineplus/imagepages.

Group (type of classification)	Examples of Gene(s)	Examples of condition(s)
1. FGFR3 (molecular)	<i>FGFR3</i>	Thanatophoric dysplasia. Achondroplasia
2. Type II Collagen (molecular)	<i>COL2A1</i>	Achondrogenesis Type 2, Kniest dysplasia
3. Type I collagen (molecular)	<i>COL1A1</i>	Stickler syndrome
4. Sulphation disorders (molecular)	<i>DTDST</i>	Achondrogenesis Type 1B. Diastrophic dysplasia
5. Perlecan group (molecular)	<i>PLC</i>	Dyssegmental dysplasia, Schwartz-Jampel syndrome
6. Filamin group (molecular)	<i>FLNA</i> <i>FLNB</i>	Frontometaphyseal dysplasia Larsen syndrome
7. Short rib dysplasia (radiological)	<i>EVC1/2</i>	Chondroectodermal dysplasia
8. Multiple epiphyseal dysplasia Pseudoachondroplasia (radiological)	<i>COMP</i> <i>MATN3</i>	Pseudoachondroplasia Multiple epiphyseal dysplasia
9. Metaphyseal dysplasia (radiological)	<i>COL10A1</i>	Metaphyseal dysplasia Schmid type
10. Spondylometaphyseal dysplasia (radiological)	None known	Spondylometaphyseal dysplasia Kozlowski type
11. Spondylometaphyseal dysplasia (radiological)	<i>DYM</i>	Dyggve Melchior Clausen syndrome
12. Severe spondylodysplastic dysplasias (radiological)	None known	Opsismodysplasia
13. Moderate spondylodysplastic dysplasias (radiological)	None known	Brachyolmia
14. Acromelic dysplasias (radiological)	<i>TRPS</i>	Trichorhinophalangeal dysplasia
15. Acromesomelic dysplasias (radiological)	<i>NPR2</i>	Acromesomelic dysplasia Maroteaux type
16. Mesomelic and rhizomesomelic dysplasias (radiological)	<i>SHOX</i>	Dyschondrosteosis
17. Bent bone dysplasias (radiological)	<i>SOX9</i>	Campomelic dysplasia
18. Slender bone dysplasias (radiological)	<i>CUL7</i>	3M syndrome
19. Dysplasias with multiple dislocations (radiological)	None known	Desbuquois syndrome
20. Chondrodysplasia punctata (radiological)	<i>NSDHL</i>	CHILD syndrome
21. Neonatal osteosclerotic dysplasias (radiological)	<i>PTHR1</i>	Blomstrand dysplasia
22/23. Increased bone density group (radiological)	<i>CLCN7</i> <i>CAI</i>	Osteopetrosis
24. Decreased bone density group (radiological)	<i>COL1A1/2</i>	Osteogenesis imperfecta
25. Defective mineralization (radiological)	<i>ALPL</i>	Hypophosphatasia

26. Lysosomal storage disorders (radiological)	<i>HSS</i>	Mucopolysaccharidoses
27. Osteolysis group (radiological)	<i>LMNA</i>	Mandibuloacral dysplasia type A
28. Disorganised development of skeletal elements (radiological)	<i>GNAS1</i>	Fibrous dysplasia polyostotoc form
29. Cleidocranial dysplasias (radiological)	<i>RUNX2</i>	Cleidocranial dysplasia
30. Craniosynostosis syndromes (radiological)	<i>FGFR2</i>	Apert syndrome
31. Dysostoses with craniofacial involvement (radiological)	<i>TCOF1</i>	Mandibulofacial dysostosis
32. Dysostoses with vertebral/costal involvement (radiological)	<i>HLXB9</i>	Currarino syndrome
33. Patellar dysostoses (radiological)	<i>LMX1B</i>	Nail-patella syndrome
34. Brachydactylies (radiological)	<i>HHH</i>	Brachydactyly type A1
35. Limb hypoplasia/reduction defects (radiological)	<i>NIPBL</i>	De Lange syndrome
36. Polydactyly-syndactyly-triphalangism group (radiological)	<i>GLI3</i>	Grieg cephalopolysyndactyly
37. Defects in joint formation and synostoses (radiological)	<i>NOG</i>	Proximal symphalangism type 1

Table 1.1. Summary of ISDN Nosology of skeletal dysplasias. Superti-Furga, A. and Unger, S., 2007.

classification is summarised in Table 1.2 and simplifies the categorisation but does not account for those conditions for which a molecular basis has not yet been found. Clear genotype-phenotype correlation is not seen although the broad categorisation of underlying molecular defects results largely in groups of phenotypes which would be expected. Hence transcription factor defects result in phenotypes with multiple developmental abnormalities, for example campomelic dysplasia caused by mutations in *SOX9* is a skeletal dysplasia also characterised by other developmental abnormalities including ambiguous genitalia, heart defects and learning difficulty. Conditions where degradation of molecules is not able to proceed normally such as the lysosomal storage disorders due to lysosomal acid hydrolase deficiency result in substrate accumulation in many organ systems with consequent effect upon multiple organ systems and deterioration over time.

The severity and clinical impact of the skeletal dysplasias varies greatly from neonatal lethality in disorders such as thanatophoric dysplasia caused by mutations in *FGFR3*, to minor physical anomalies such as short digits in brachydactyly type E due to mutations in *homeobox D13* (*HOXD13*). The inclusion of 140 different genes causative of a variety of skeletal dysplasias in the revised Nosology of 2006 indicates the complexity of the processes required for normal bone growth and development (Superti-Furga and Unger, 2007).

1.3 Features of Dyggve Melchior Clausen syndrome

1.3.1 Clinical features

The first description of DMC (online Mendelian inheritance in man, OMIM 223800) was as 'Morquio-Ullrich's disease' in the offspring of an uncle-niece marriage in Greenland (Dyggve *et al.*, 1962). Subsequently, the name Dyggve Melchior Clausen syndrome was adopted using the names of the clinicians who produced the initial report in order to differentiate the condition from the mucopolysaccharidosis, Morquio's disease. DMC has most frequently been reported in largely inbred populations, particularly in Lebanon suggesting a founder effect (Naffah, 1976, Bonafede and Beighton, 1978). The characteristic clinical phenotype of short stature with short limbs, a short trunk and a barrel shaped chest, is not apparent at birth but typically evident by 18 months of age. Affected individuals have multiple bony abnormalities visible clinically and radiologically in addition to microcephaly, mental retardation and a coarse facial appearance (Beighton, 1990). Affected individuals are shown in Figure 1.2. Features seen on clinical examination include short stature with a short trunk and

Group	Examples of Gene(s)	Examples of condition(s)
Extracellular structural proteins	<i>COL1A1/2</i> <i>COMP</i> <i>MATN3</i>	Osteogenesis imperfecta Pseudoachondroplasia Multiple epiphyseal dysplasia
Metabolic pathways	<i>DTDST</i> <i>ARSE</i>	Diastrophic dysplasia X-linked chondrodysplasia punctata
Folding/degradation macromolecules	<i>Lysosomal acid hydrolases</i> <i>Sedlin</i>	Lysosomal storage disorders X-linked spondyloepiphyseal dysplasia
Hormones/signal transduction	<i>IHH</i> <i>GNAS1</i>	Brachydactyly Type A1 Allbright hereditary osteodystrophy
Nuclear proteins/transcription factors	<i>SOX9</i> <i>TBX5</i>	Campomelic dysplasia Holt Oram syndrome
Oncogenes/tumoursuppresoor genes	<i>EXT1</i>	Multiple exostosis syndrome Type 1
RNA/DNA processing/metabolism	<i>ADA</i>	Severe Combined immunodeficiency with metaphyseal changes

Table 1.2. Summary of Molecular Genetic Classification of skeletal dysplasias proposed in Superti-Furga, A. and Unger, S., 2001.



Figure 1.2. Photographs of individuals affected with DMC.
From Thauvin-Robinet *et al*, 2002.

short limbs, short hands and feet, a barrel shaped chest with protrusion of the sternum, scoliosis, exaggerated lumbar lordosis and microcephaly. Patients with DMC display a variable degree of mental retardation, yet a number of individuals with identical clinical and radiological skeletal features have been reported to have normal development, a condition given the name Smith-McCort dysplasia (SMC OMIM 607326). This condition has subsequently been shown to be allelic with DMC (Cohn *et al.*, 2003, Smith and McCort, 1958, Spranger *et al.*, 1976). The underlying cause of the learning difficulty is not known but the microcephaly seen in DMC reflects an underlying small brain with no structural abnormalities seen on computerised tomography (Girisha *et al.*, 2007). Mental retardation is reported in association with a number of other skeletal dysplasias, often also in association with microcephaly. Examples of such conditions include the spondylometaphyseal dysplasia anauxetic type caused by mutations in the *mitochondrial RNA-processing endoribonuclease (RMRP)* gene, an RNA processor. Mutations in this gene lead to impaired cell growth by altering cell cycle regulation with subsequent abnormal brain growth and function (Thiel *et al.*, 2005). Mutations in the centrosomal protein *Pericentrin* have been identified in the overlapping conditions microcephalic osteodysplastic primordial dwarfism Type II (MOPDII) and Seckel syndrome, characterised by extreme growth retardation and microcephaly, multiple bony abnormalities in MOPDII and sometimes learning difficulty. *pericentrin* is a centrosomal protein required for mitosis hence it has been postulated that the microcephaly is caused by decreased cell number in the brain (Rauch *et al.*, 2008, Griffith *et al.*, 2008). Other centrosomal proteins such as microcephalin and abnormal spindle-like microcephaly associated protein (ASPM) have been implicated in autosomal recessive primary microcephaly suggesting that centrosomal function is essential for normal brain development (Bond and Woods, 2006). The cause of microcephaly and mental retardation in DMC is not yet clear.

1.3.2 Radiological features

Radiographs of DMC affected individuals show multiple abnormalities; vertebral bodies are flattened (platyspondyly) with anterior pointing and notch-like ossification of the superior and inferior end plates leading to a characteristic double humped appearance. The odontoid is often hypoplastic and the pelvis small with a lacy appearance of the iliac crests due to irregular ossification (Nakamura *et al.*, 1997). The ilia are short and broad with hypoplasia of the basilar portions. The femora show a prominent medial, spur-like portion of the femoral neck and late ossification of the epiphyses. The long

bones are shortened with irregular metaphyses and epiphyses. The carpal bones show delayed and irregular ossification and accessory ossification centres are present in the metacarpals and phalanges (Spranger *et al.*, 1975). Radiologically, DMC is classified as a spondylometaphyseal dysplasia as the spine, epiphyses and metaphyses are all affected. Typical radiographs are shown in Figure 1.3, 1.4 and 1.5. Spinal cord compression from atlantoaxial instability due to odontoid hypoplasia is a potentially preventable complication and affected individuals are advised to avoid potential precipitating factors such as contact sport. An experienced anaesthetist should be present if surgery is required as the neck should not be hyperextended during intubation to reduce the risk of spinal cord compression (Kandziora *et al.*, 2002, Girisha *et al.*, 2007). Skeletal changes tend to worsen with increasing age, with patients developing dorsal kyphosis, lumbar lordosis, scoliosis and limitation of joint mobility as well as early onset osteoarthritis (Naffah, 1976).

1.3.3 Histological features

Histological findings in DMC and SMC offer some clues as to disease pathogenesis. Irregular deposition of bone in a wavy pattern occurs at the osteochondral junction, the likely explanation for the lacy margins. Abnormal endochondral ossification takes place at growth plates, with poor chondrocyte columnarisation and clustering of chondrocytes, many showing signs of degeneration (see Figure 1.6). This is in contrast to normal endochondral ossification and chondrocyte differentiation whereby chondrocytes form columns and progress through stages of differentiation from resting through hypertrophic and proliferating before apoptosing (see Figure 1.6). Electron microscopy, which provides ultra-structural analysis of chondrocytes and fibroblasts demonstrates multiple vacuoles and dilated cisternae of rough endoplasmic reticulum (RER), which appear to contain an amorphous granular material (see Figure 1.7) (El Ghouzzi *et al.*, 2003, Engfeldt *et al.*, 1983, Horton and Scott, 1982, Nakamura *et al.*, 1997). RER dilatation has also been observed in other skeletal dysplasias including pseudoachondroplasia (PSACH) and multiple epiphyseal dysplasia (MED), both caused by mutations in the gene encoding *cartilage oligomeric matrix protein (COMP)*. *In-vivo* studies have found COMP and type IX collagen to localise to the dilated RER in these conditions with the accumulation of mutant COMP and other extracellular matrix proteins resulting in compromised chondrocyte function and cell death. It has been suggested that mutant COMP is disrupting the processing of other proteins leading to their accumulation (Maddox *et al.*, 1997, Delot *et al.*, 1998, Hecht *et al.*, 1995). Irregular ossification in association with decreased chondrocyte number, abnormal

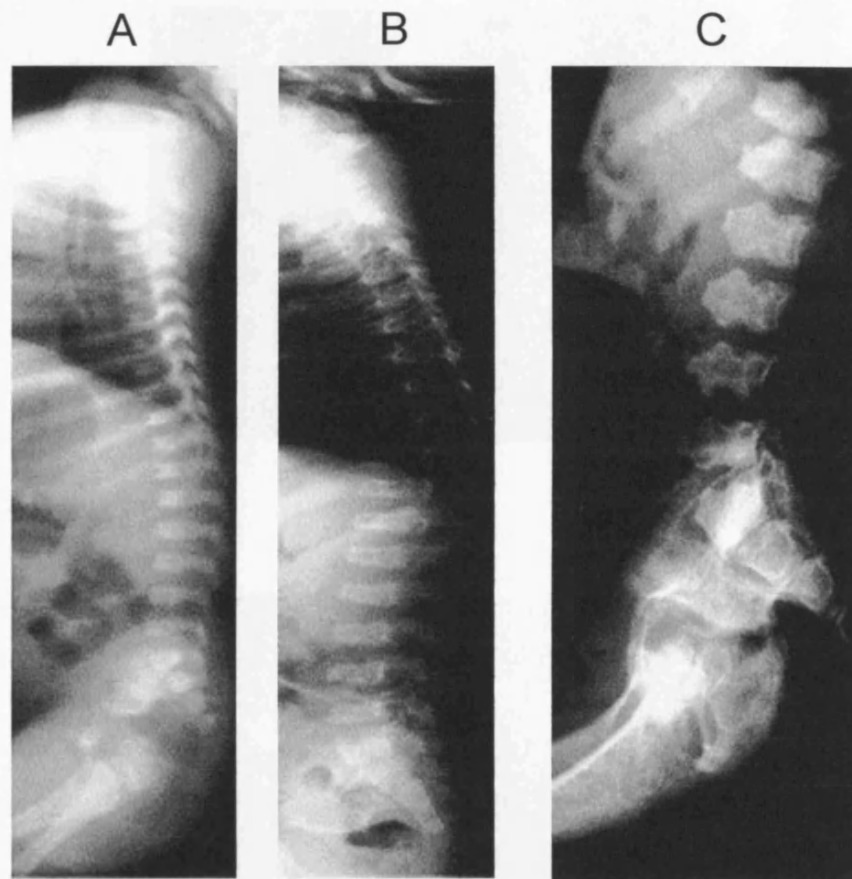


Figure 1.3. Spinal radiographs of DMC affected individuals showing the progression of X-ray findings over time. Flattened vertebral bodies are seen at birth (A) which have become more pronounced by age 14 months (B). By age 4 years (C) vertebral bodies have developed the characteristic double humped appearance. From Paupe *et al*, 2004.

A



B



Figure 1.4. Pelvic radiographs of DMC affected individuals showing the progression of X-ray findings over time. Lacy iliac crests are seen at age 4 years, in addition to hypoplastic acetabulae and short femoral necks (A). By age 18 years, the iliac crests broaden but remain lacy (B). From Paupe *et al*, 2004.

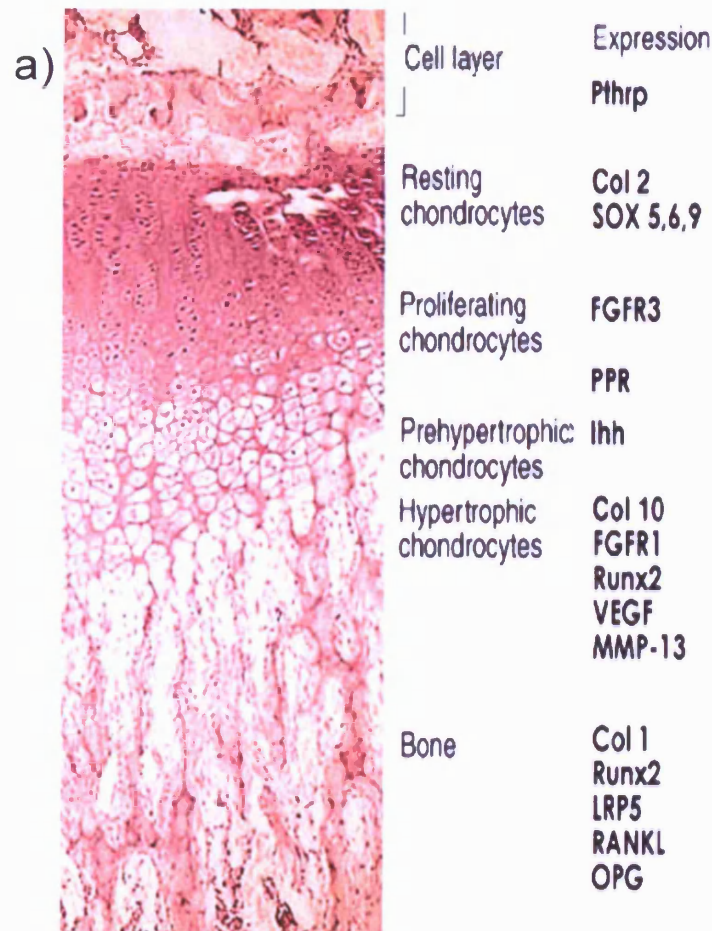
A



B



Figure 1.5. Hand radiographs of DMC affected individual seen in A. Hands are short and broad with irregular metacarpals and phalanges. Skull X-ray in B shows a hypoplastic odontoid. From Paupe *et al*, 2004.



b)

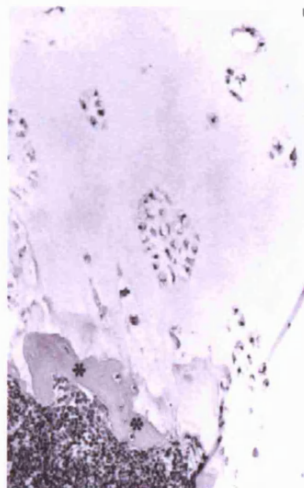


Figure 1.6. a). Light microscope image of 2 week old mouse growth plate Chondrocytes form columns and differentiate before ossification occurs. Genes expressed at different stages are shown in the right hand panel. Adapted from Kornak and Mundlos 2003.

b). Chondrocytes fail to form columns and are arranged irregularly in DMC cartilage in an 8 year old child. Stars indicate degenerating cells. From Nakamura *et al* 1997.

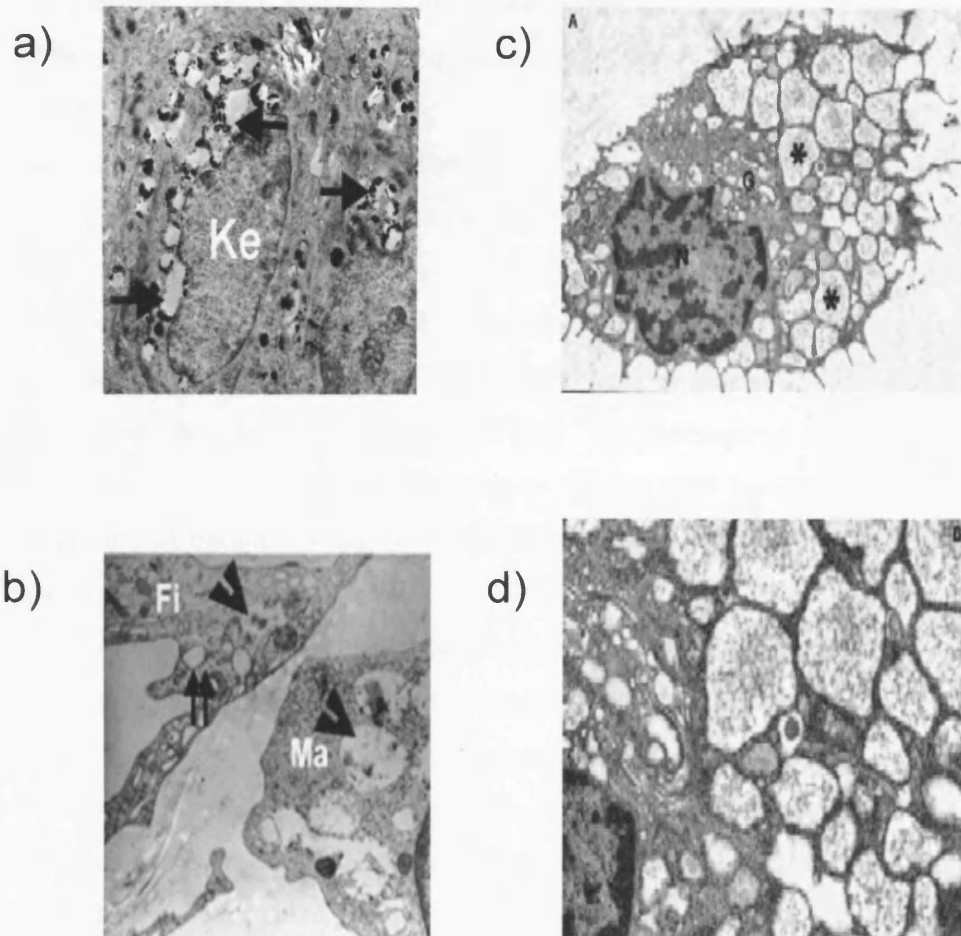


Figure 1.7. Electron microscopy of skin from a DMC affected individual. a) shows a keratinocyte containing a large number of vacuoles. b) demonstrates a fibroblast and macrophage again containing multiple vacuoles. From El Ghouzzi *et al.* c) and d) show fibroblasts with dilated endoplasmic reticulum containing amorphous material. From Nakamura *et al.*

columnarisation and dilated RER are also seen in spondylometaepiphyseal dysplasia anauxetic type (Thiel *et al.*, 2005). Given that this condition has been shown to be caused by mutations in a gene required for RNA processing and cell cycle regulation whereas PSACH and MED are caused by mutations in a structural protein, this suggests that the same histological end point can result from a broad range of underlying functional abnormalities. The nature of the accumulated substance in the RER of DMC affected cells is not known, but if elucidated may give insight as to the function of the *DYM* gene product.

The differential diagnosis of DMC includes the storage disorder Morquio syndrome (Mucopolysaccharidosis Type IV) characterised by the accumulation of keratan sulphate and its excretion in the urine due to the absence of galactosamine-6-sulphatase or beta-galactosidase. However the two disorders can be distinguished by the absence of lens opacity, aortic valve disease and keratosulphaturia in patients with DMC and normal intellectual function in Morquio syndrome. The overlapping clinical features taken together with the microscopic appearance of dilated RER and multiple vacuoles have led a number of authors to question a possible metabolic abnormality underlying DMC (Rastogi *et al.*, 1977, Rastogi *et al.*, 1978, Rastogi *et al.*, 1980, Roesel *et al.*, 1991). Findings have included increased serum pipecolic acid suggesting a peroxisomal disorder although other markers for such a condition were normal (Roesel *et al.*, 1991). The skeletal dysplasia rhizomelic chondrodysplasia punctata is caused by a defect in uptake of matrix proteins into peroxisomes due to mutations in the gene *peroxisome biogenesis factor 7 (PEX7)* and also has microcephaly and mental retardation as clinical features thus disorders of peroxisomal function can result in a skeletal phenotype (Shimozawa *et al.*, 1999). The reports of metabolic abnormality in DMC have been based on findings in a very small number of individuals, and in some cases on a single subject, hence must be interpreted with caution. Taken together, no specific underlying metabolic defect has been identified.

1.4 Mapping of DMC gene by autozygosity mapping

Preliminary work leading up to this project mapped the disease gene for DMC/SMC to a 1.8 centimorgan (cM) region on chromosome 18q12-21 using an autozygosity mapping (AZM) approach (Ehtesham *et al.*, 2002, Thauvin-Robinet *et al.*, 2002). AZM is a powerful method of mapping disease genes in autosomal recessive disorders. It relies upon the assumption that affected offspring from consanguineous partnerships will have inherited two copies of the same mutant allele from a common ancestor (identity by descent). Homozygous regions of the genome which may harbour the

disease gene are identified using, for example, microsatellite markers (Lander and Botstein, 1987, Mueller and Bishop, 1993). The technique was first suggested by Lander and Botstein in 1987 using restriction fragment length polymorphism (RFLP) linkage maps, even though at the time of writing, a dense enough map was not yet available. They showed that offspring of a first cousin marriage are expected to be homozygous for one-sixteenth of their genome and calculated that only ten such families would be required in order to obtain a significant logarithm of the odds (LOD) score of over three (Lander and Botstein, 1987). The main advantage of autozygosity mapping is that it allows rare recessive diseases to be mapped using a small number of affected children from consanguineous partnerships. Given that such families will be overrepresented among those affected with rare recessive diseases it should be easier to collect enough families for autozygosity mapping than to gather enough non-inbred families for traditional linkage studies. Although the area of autozygosity may be large in any given family, by identifying a shared region of autozygosity between families, the region of linkage can be narrowed. AZM has since been used to map the causative genes in a huge number of conditions, initially using RFLPs but progressing as the technology advanced to use microsatellite markers and more recently single nucleotide polymorphisms (SNPs) in microarray format. For example, three loci (Bardet-Biedl syndrome (BBS) 9, 10 and 11) for the multi-system disorder Bardet-Biedl syndrome have been identified using SNP microarray analysis (Chiang *et al.*, 2006, White *et al.*, 2007, Laurier *et al.*, 2006).

The DMC/SMC locus was mapped simultaneously to chromosome 18q12-21 in nine families affected by DMC and two families affected by SMC with respective LOD scores of 9.65 and 3.04 (Ehtesham *et al.*, 2002, Thauvin-Robinet *et al.*, 2002). Linkage to the region in three DMC kindreds was also demonstrated by Ehtesham *et al* who identified a 10.7 cM region of homozygosity whereas Thauvin-Robinet *et al* found homozygosity for a 1.8 cM region within the larger mapped locus as illustrated in Figure 1.8.

1.5 Identification of the DMC gene

Following identification of the mapped region, a positional cloning approach was used. Within the smaller 1.8 cM linked region were eight known genes and one predicted transcript, namely *mothers-against-decapentaplegic-homologue 7* (*MADH7*), *signal recognition particle 72 kD-like* (*SRP72-like*), *acetyl-coenzyme A acyltransferase 2* (*ACAA2*), *myosin Vb* (*MYO5B*), *MAP-kinase 4* (*MAPK4*), *methyl-CpG binding domain protein 1* (*MBD1*) genes, and *human CpG binding protein gene* (*hCGBP*). Genes which

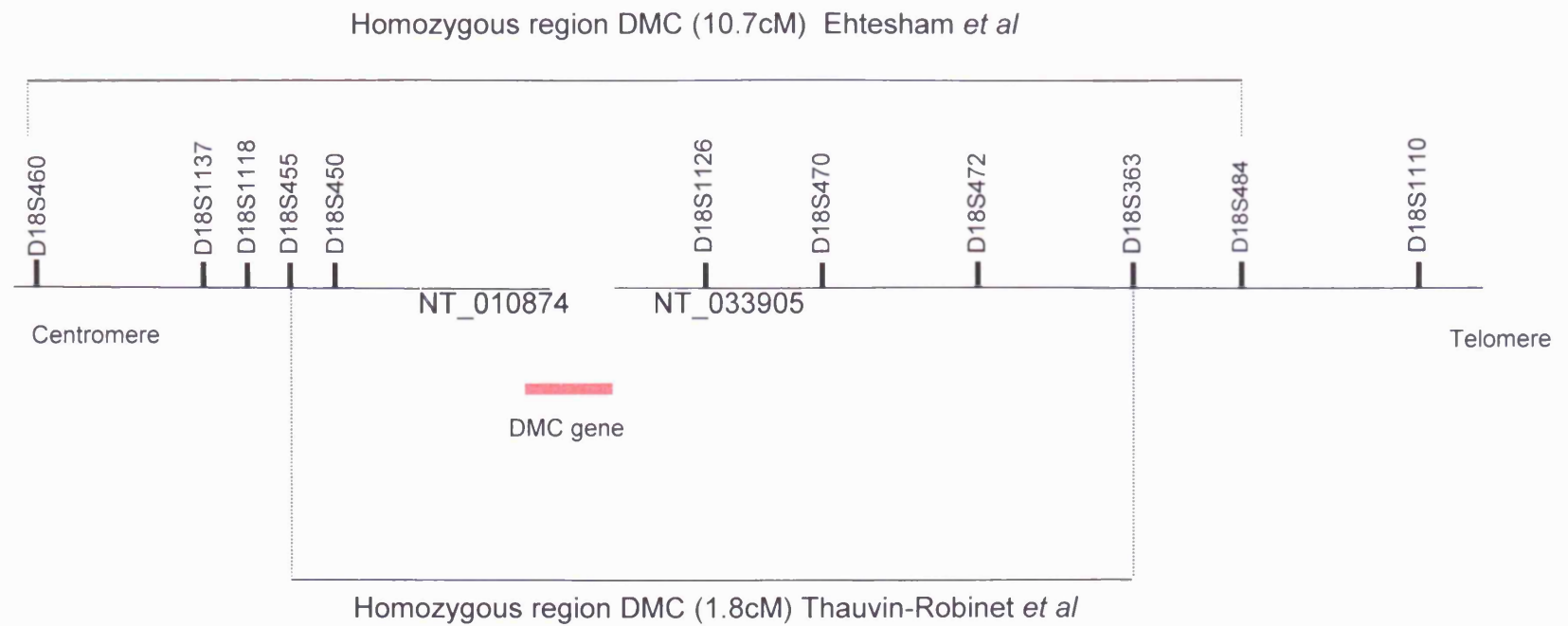


Figure 1.8. Schematic illustrating chromosomal position of microsatellite markers and DMC gene. Overlapping regions of homozygosity identified by autozygosity mapping studies by Ehtesham *et al* and Thauvin Robinet *et al* are also shown.

were at the time considered to be good candidates given that DMC had been suggested to be a storage disorder included *ACAA2* which encodes an enzyme catalysing a step in mitochondrial fatty acid beta oxidation and *MYO5B* which belongs to the class V myosin family that function as motors for actin dependent organelle trafficking. Additionally, *MBD1* and *hCGBP* were shown to be highly expressed in fetal brain, osteoblasts, and chondrocytes also making them good candidates (Thauvin-Robinet *et al.*, 2002).

The protein product SMAD2 of the *MADH7* gene has been shown to have a role in chondrocyte development and was sequenced in DMC affected individuals from five families but no mutations identified (Cohn *et al.*, 2003). Direct sequencing of the coding regions and exon-intron boundaries failed to detect any deleterious sequence variants in the seven other known genes. However, mutations were identified in affected individuals within the predicted transcript known as *FLJ20071* (Thauvin-Robinet *et al.*, 2002). Initial predictions using computational exon prediction software suggested this gene consisted of 9 exons, however, by alignment of expressed sequence tags (ESTs) against genomic sequence using *in-silico* databases, another 8 exons were identified resulting in a gene comprising 17 exons over 400 kb of genomic sequence. The ATG initiator codon is located in exon 2, flanked by a Kozak consensus sequence and the polyA tail found 250 base pairs (bp) downstream of the termination codon in exon 17 (Thauvin-Robinet *et al.*, 2002). This gene has been given the name *Dymeclin*, or *DYM*. For the purposes of this report, the gene causative of DMC will subsequently be referred to as *DYM* and the protein product as Dymeclin.

To date, 23 different *DYM* mutations have been identified, five as part of this study (Cohn *et al.*, 2003, El Ghouzzi *et al.*, 2003, Neumann *et al.*, 2006, Paupe *et al.*, 2004, Kinning *et al.*, 2005, Girisha *et al.*, 2007). Of these 23 mutations, eight result in a premature stop codon (two identified in this study). Five splice site mutations have been shown and two missense mutations demonstrated as compound heterozygotes in association with mutations predicting a truncated protein product. Additionally, a single mutation 3' to the untranslated exon 1 has been demonstrated, but the consequence upon the protein is unknown. The remaining deletions/insertions predict a frameshift and protein truncation or extension. Two novel, complex exon duplication mutations have been identified as part of this project and will be discussed later. The sequence changes and predicted consequences on the protein are summarised in Table 1.3. The location of mutations throughout *DYM* are illustrated in Figure 1.9.

Nucleotide change	<i>DYM</i> exon	Amino acid change	Consequence on protein	Phenotype	Ethnic origin	Reference
G>C 34 bp 3' to exon 1	Intron 1 (untranslated)		Unknown	DMC	Portuguese	Paupé 2004
c.48C>G	2	Y16X	Truncated	DMC	Dominican	Cohn 2003
c.IVS 3-1G>A	4 (5' end) splice acceptor		Skipping exon 3	DMC	Lebanese	Paupé 2004, Neumann 2006
c.259G>A	4	E87K	Missense	SMC	Guamese	Cohn 2003
c.IVS 4 -2A>G	5 (5' end) splice acceptor		Skipping exon 4	DMC	Spanish	Paupé 2004
c.369T>A	5	Y132X	Truncated	DMC	Not stated	Cohn 2003
c.IVS 5 -2A>G	6		Skipping exon 6	DMC	Spanish	Pogue 2005
c.580C>T	7	R194X	Truncated	DMC	Tunisian	El Ghouzzi 2003
c.610C>T	7	R204X	Truncated	DMC	Moroccan	El Ghouzzi 2003
c.656T>G	8	R219X	Truncated	DMC	Moroccan	El Ghouzzi 2003
c.763delA	8	FS	Frameshift and truncation	DMC	Not stated	Cohn 2003
c.IVS 7-2 A>G	8 (5'end) splice acceptor		Skipping exon 8 and frameshift	SMC	Guamese	Cohn 2003, Pogue 2005
c.IVS 10 +1G>T	10(3'end) splice donor		Skipping exon 11 (in frame)	DMC	Moroccan	El Ghouzzi 2003
c.1172_1173insC	11	FS	Frameshift and truncation	DMC	Indian	Girisha 2008
c.IVS 11 -1G>A	12(5'end) splice acceptor		Skipping exon 11 (in frame)	DMC	Lebanese	El Ghouzzi 2003
c.1405A>T	13	N469Y	Missense	DMC	Not stated	Cohn 2003
c.1447C>T	13	Q483X	Truncated	DMC	Moroccan	El Ghouzzi 2003
c.1877delA	17	FS	Frameshift and extension	DMC	Moroccan	El Ghouzzi 2003
c.1937_1942 delTGTCT	17	FS	Frameshift and extension	DMC	Georgian	Neumann 2006

Table 1.3. *DYM* mutations reported in the literature prior to and during the course of this study. Mutations identified as part of this project are not included in the table.

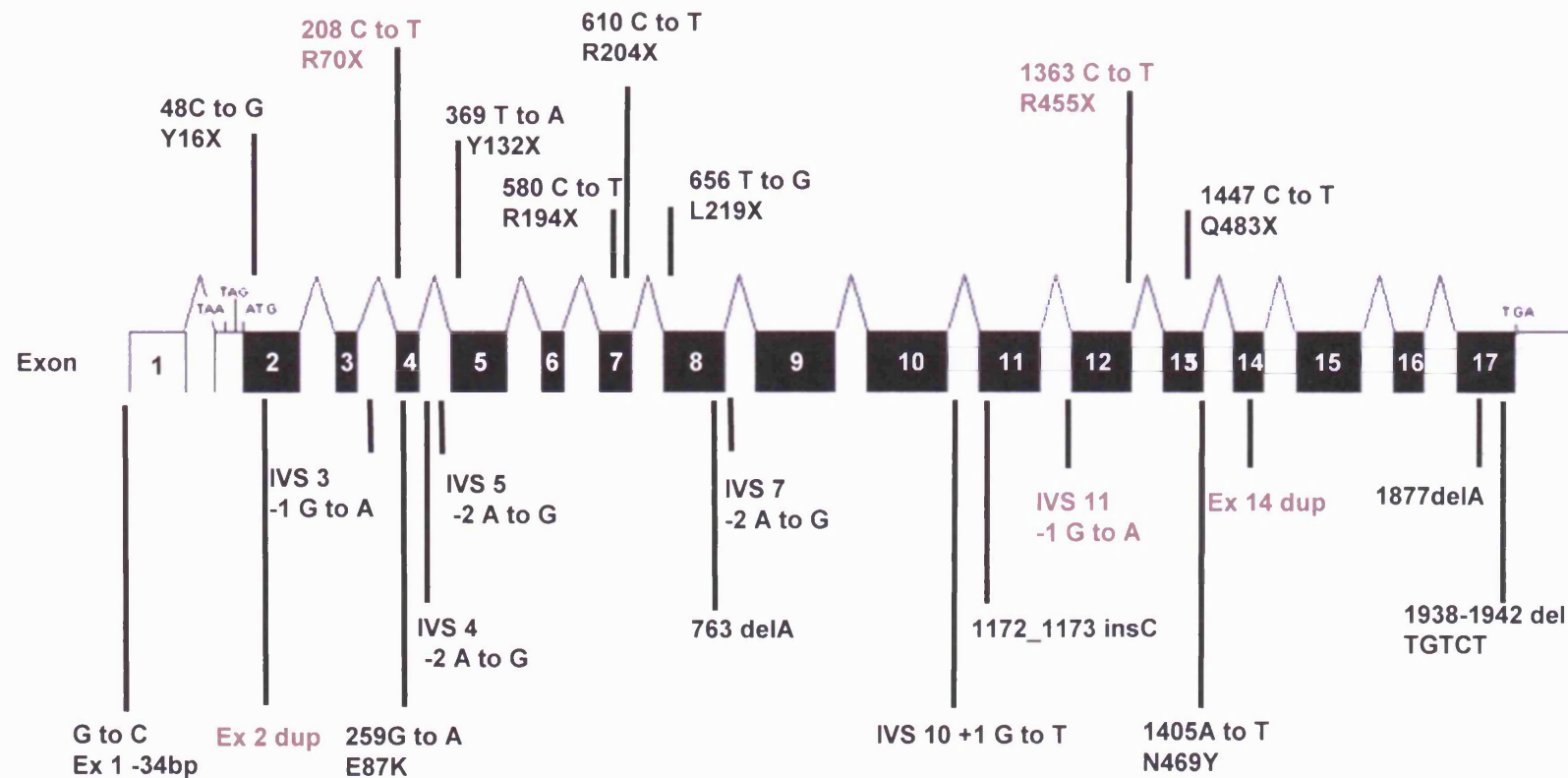


Figure 1.9. Distribution of mutations in the DMC gene reported prior to and during the course of this study. The DMC gene structure is illustrated with exons shown as black boxes drawn approximately to scale. The ATG initiator is located in exon 2 and the termination codon in exon 17. Nonsense mutations are shown above the gene structure and all other types of mutation below. Mutations identified as part of this project are shown in red. El Ghouzzi *et al*, Cohn *et al*, Paupe *et al*, Kinning *et al*, Neumann *et al*, Girisha *et al*.

1.6 From disease gene identification to elucidation of the underlying molecular mechanism

A variety of approaches can be used in order to investigate the function of a novel gene. These can be divided into *in-silico* and experimental methods. The amino acid sequence of the protein of interest can be compared *in-silico* to those available in online databases. Homologous amino acid sequences likely to represent functional motifs may be detected or the gene product identified as part of a known protein family. Insight as to the function of the protein may be obtained in this way. However, online databases will only provide predictions and are subject to several limiting factors including the quality and quantity of data available and the extent of homology between a novel protein and those in the databases. Similarly, the localisation of disease causing mutations within the gene can be used to identify critical functional domains, although this may be difficult in a rare disorder as the number of reported mutations is likely to be limited. Experimental methods which can be employed to help determine function include localisation of the protein at a sub-cellular level as well as determining the expression pattern of the gene within different cell and tissue types.

The identification of disease gene mutation leads the investigator to question how such a mutation may assist in understanding the disease process and clinical phenotype. Mutations will usually result in either reduced levels or loss of function of the protein product, or the production of an altered or novel protein which may itself have a deleterious effect or interfere with the function of other proteins. DMC is an autosomal recessive disorder hence loss of protein function is considered to be the likely underlying molecular mechanism with nonsense-mediated messenger ribo-nucleic acid (mRNA) decay acting to destroy the transcript and prevent production of a potentially deleterious protein product.

DYM is a novel protein which has no homology to any known protein and contains motifs which only give limited insight as to its potential action. In order to determine DYM function, experiments were undertaken to elucidate the gene and protein expression patterns at a cellular and tissue level by RT-PCR, *in-situ* hybridisation and Western blot analysis. Additionally, transient transfection studies were performed to reveal DYM localisation within the cell. Unfortunately, despite the generation of antibodies against different DYM epitopes in two species, their experimental use was unsuccessful but will be discussed briefly in Chapter 5.

The identification of proteins which interact with DYM may give further clues as to its function, given that they may be better characterised and with assigned functions. There are many techniques which can be employed to detect protein–protein interactions such

as co-immunoprecipitation and *in-vitro* protein pulldown assays. These techniques rely on both interacting proteins being known whereas the yeast two-hybrid technique can be used to detect an unknown interacting protein partner of a single protein of interest. Yeast two-hybrid analysis has proven to be essential in the elucidation of disease mechanisms not only in identifying novel interactions but also allowing study of the effects of known disease causing mutations, for example in determining the changes in interaction ability of mutant proteins. The yeast two-hybrid technique has been used successfully to detect proteins interacting with those implicated in recessive disorders including juvenile onset Parkinson's disease and Fanconi anaemia (Hussain *et al.*, 2003, Huynh *et al.*, 2003). Successful yeast two-hybrid studies in skeletal dysplasias include the identification of cortactin and mouse actin binding protein 1 (mABP1) as faciogenital dysplasia 1 interactors in faciogenital dysplasia. The interaction of these proteins has been shown to be essential in the maintenance of cell shape (Hou *et al.*, 2003). Yeast two-hybrid analysis can also be used to determine whether protein products of genes causing similar disorders interact. When this technique was employed in the hereditary ataxias, 770 novel interactions of 54 proteins causing 23 different ataxias were identified (Lim *et al.*, 2006). However, the yeast two-hybrid technique is not a mammalian cell based system and hence it is important to remember that proteins may not be available for, or capable of interaction in the same way within both yeast and human cells.

This study has adopted a broad range of approaches in order to try and gain maximum information about the function of *DYM*.

1.7 Development and structure of the skeletal system

In order to understand the pathogenesis of the skeletal dysplasias including DMC, it is necessary to understand the mechanisms underlying the development of cartilage and bone, and the genes which control these processes. It may then be possible to link clinical phenotype to the cellular processes of proliferation, differentiation and apoptosis essential in skeletal formation.

The skeleton originates from three different embryonic tissues; somites, mesoderm and neural crest. The somites, transient developmental structures of the embryo located on either side of the neural tube, give rise to the sclerotome which forms the axial skeleton (vertebrae and ribs). Mesenchymal cells from the lateral plate mesoderm form the appendicular skeleton (limbs) in a complex process which requires the developing limb bud to differentiate in three dimensions, each axis controlled by a separate set of signalling molecules produced by a localised cell population. Proximal to distal limb

growth is controlled by the apicoectodermal ridge (AER), dorsoventral patterning is under control of ectoderm overlying the limb bud and the zone of polarising activity (ZPA) is responsible for anterior to posterior differentiation. Together, these pathways act to produce limbs with the pre-determined number of segments and digits in an appropriate configuration. Craniofacial bones are formed from neural crest cells.

Skeletal development can be divided into three stages; formation of the template which will go on to form the skeletal elements is followed by condensation then differentiation of mesenchymal precursor cells into prechondrocytes. The mesenchymal cells which are committed to become chondrocytes aggregate into chondrogenic nodules and differentiate into chondrocytes. Defects in genes functioning at these different stages will cause alternative patterns of abnormality. Genes active in the initial stage of template formation will result in skeletal dysostoses, that is disorders of the skeletal pattern with abnormality in an embryologically defined skeletal element. Many of these genes are transcription factors, for example the *homeobox (HOX)* group of genes which are essential in limb patterning, with mutations in *HOXD13* causing a form of synpolydactyly inherited in a dominant manner (Muragaki *et al.*, 1996, Shiang *et al.*, 1994). In contrast, genes functioning in the regulation of organogenesis will cause skeletal dysplasias, disorders with a generalised skeletal defect such as achondroplasia, caused by mutations in *fibroblast growth factor receptor type 3 (FGFR3)* (Shiang *et al.*, 1994). The distinction between these types of disorder is illustrated schematically in Figure 1.10.

Bone can be formed by two different processes, endochondral ossification involves the formation of a cartilage template which is then ossified to become bone as illustrated in Figure 1.11 (Cohen, 2006). Mesenchyme condenses to form a cartilage model, chondrocytes in the centre of the model hypertrophy and ossify to form a collar of bone and then proliferate and form secondary ossification centres. This process forms all bones except the skull and part of the clavicles in which bone is laid down directly during intramembranous ossification. Endochondral ossification will be considered in more detail as it known to be disrupted in DMC with chondrocytes failing to form columns, being randomly oriented and seen to degenerate (Engfeldt *et al.*, 1983, Horton and Scott, 1982, Nakamura *et al.*, 1997).

The initial event in chondrogenesis is the aggregation of mesenchymal chondroprogenitor cells into pre-cartilage condensations, a process dependent upon signalling initiated by both cell-cell and cell-matrix interactions. Before condensation begins, an extracellular matrix rich in collagens type I and IIA and hyaluronan is excreted. Increased hyaluronidase activity is associated with condensation of cells and

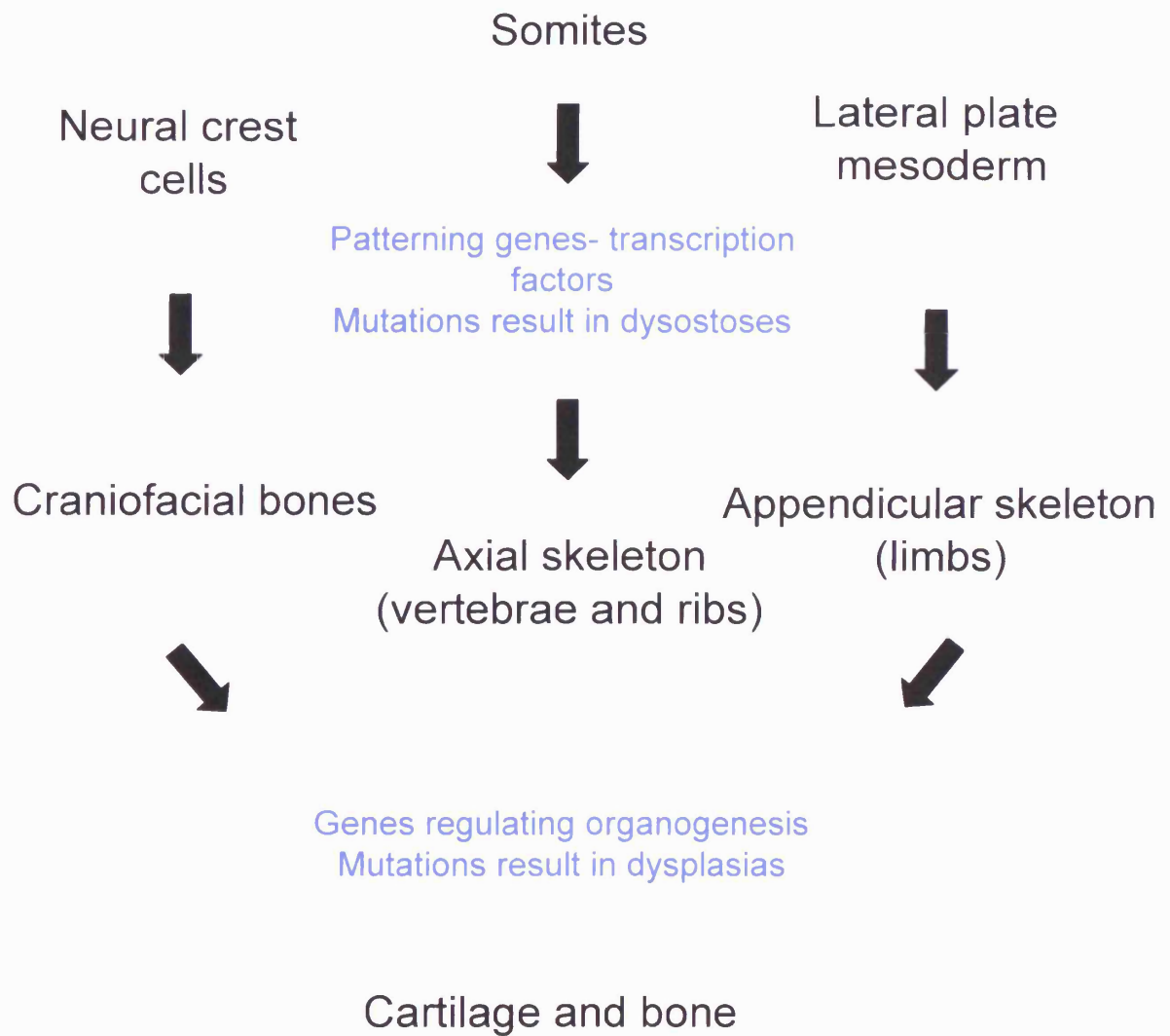


Figure 1.10. Embryological development of the skeleton.

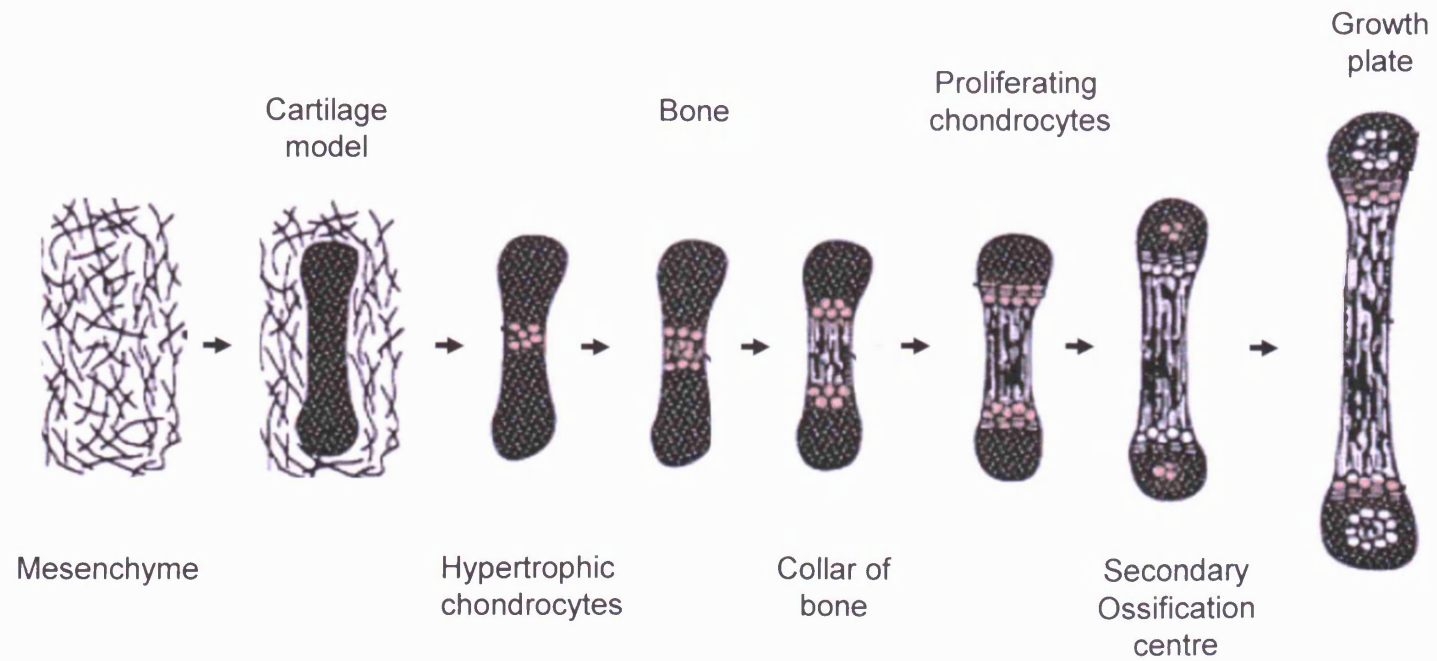


Figure 1.11. Endochondral bone development. Mesenchyme condenses to form a cartilage model. Chondrocytes in the centre of the model hypertrophy and ossify to form a collar of bone, they then proliferate and form secondary ossification centres. Labelled areas are highlighted in red. Adapted from Cohen.

increased levels of cell adhesion molecules. Extracellular matrix molecules such as the tenascins and thrombospondins interact with cell adhesion molecules to induce intracellular signalling pathways which induce pre-chondrocytes to become chondrocytes (DeLise *et al.*, 2000). In the limbs, complex patterning of these condensed regions is needed in order for the correct components of the limb to develop, a process controlled by many signalling pathways including *fibroblast growth factor (FGF)*, *hedgehog*, *wingless type MMTV integration (WNT)* and *bone morphogenetic protein (BMP)* (Tickle, 2002). *BMP* signalling is not only required for this process, but also for condensation, formation and differentiation of pre-chondrocytes into chondrocytes (Yoon and Lyons, 2004). Once chondroprogenitors have differentiated into chondrocytes, the extracellular matrix becomes rich in aggrecan and collagens II, IX and XI. Type II collagen expression is largely under the control of *SOX9* itself controlled by *BMP* signalling (Shinomura *et al.*, 1984, Ng *et al.*, 1997).

Chondrocyte proliferation is controlled by *growth hormone (GH)* and *insulin like growth factor 1 (IGF1)*, with disorders of this pathway resulting in proportionate short stature as chondrocyte proliferation is largely responsible for long bone growth (Ballock and O'Keefe, 2003). This is in contrast to disorders affecting chondrocyte differentiation that cause disproportionate dwarfism. The *indian hedgehog/parathyroid hormone related peptide (IHH/PTHrP)* pathway is involved in chondrocyte proliferation and hypertrophy in the lower proliferative and pre-hypertrophic zones through the action of PTHrP which keeps chondrocytes in a non-hypertrophic proliferative state. *IHH* and *PTHrP* form a feedback loop with *IHH* leading to increased PTHrP secretion and therefore slowing chondrocyte hypertrophy (Vortkamp *et al.*, 1996). Mutations in *IHH* cause the isolated hand abnormality brachydactyly type A1 (OMIM 112500) whereas mutations in *PTHrP* result in a more widespread skeletal dysplasia. Jansen's chondrodysplasia is characterised by generalised poor skeletal growth with abnormal metaphyses (Schipani *et al.*, 1995, Kirkpatrick *et al.*, 2003). FGF and BMP additionally modify chondrocyte differentiation rate in relation to the rate of proliferation (Minina *et al.*, 2002). In order for endochondral ossification to occur, chondrocytes undergo terminal differentiation and through hypertrophy increase their cellular volume by approximately 20 times. Once the chondrocytes have become hypertrophied, type X collagen and alkaline phosphatase secretion predominate (Schmid 1985). *Runt-related transcription factor 2 (RUNX2)* is expressed in prehypertrophic chondrocytes and required for cells to enter the hypertrophic phase (Enomoto *et al.*, 2000). Heterozygous mutation in *RUNX2* result in the condition cleidocranial dysplasia characterised by

failure of closure of the skull sutures and abnormal clavicles and teeth (Mundlos *et al.*, 1997).

The final stage of cartilage differentiation is the process of ossification whereby hypertrophic cartilage is replaced by bone. Vascular endothelial growth factor is produced in the extracellular matrix by matrix metalloproteinases which stimulate vessel invasion subsequently allowing chondroclasts, osteoblasts and osteoclasts to enter and remove the cartilage matrix and form bone (Colnot *et al.*, 2004). Chondrocyte maturation proceeds towards the ends of the developing bone from the newly formed ossification centre and a growth plate is formed (see Figure 1.6). Mineralisation begins between the hypertrophic chondrocytes as osteoblasts deposit bone matrix. Hypertrophic cartilage is invaded by vascular tissue and finally hypertrophic chondrocytes apoptose, thought to be induced by changes in the ECM (Gerber *et al.*, 1999). This process continues towards the end (epiphysis) of the bone, with chondrocytes at the opposite pole of the growth plate proliferating and producing new cells which synthesise cartilage matrix. Longitudinal growth is accounted for mostly by chondrocyte hypertrophy. Gene expression patterns in chondrocyte differentiation are summarised in Figure 1.12 (Ballock and O'Keefe, 2003).

Cartilage is a highly specialised connective tissue, which allows compressive and shearing forces to be resisted. It acts prenatally to provide support to the developing embryo and forms a template for the development of endochondral bones. Postnatally it allows the skeleton to grow rapidly, protects the joints and is essential for repair of fractured bones. Cartilage can be classified based on its morphologic appearance into hyaline, elastic and fibrocartilage. Hyaline cartilage is the most common type and forms the temporary skeleton in the embryo as well as the epiphyseal plate. Chondrocytes secrete a specialised extracellular matrix (ECM), with a complex make-up which varies through development. ECM constituents include fibrous collagen to provide tensile strength and proteoglycans (a core protein with glycosaminoglycans side chains), which provide elasticity and resistance to compression. Collagens are the most abundant extracellular matrix protein and provide a scaffold structure whereas glycoproteins bind large amounts of water and give cartilage its gelatinous, compression resistant texture. There are many different types of collagen providing structural support in bone and leading to a broad range of conditions with very variable severity if the encoding genes are mutated. Collagen type I is the main constituent of bone matrix, hence bone fragility results if *collagen type I* genes are mutated, leading to the phenotype of osteogenesis imperfecta (Pope *et al.*, 1985). Aggrecan is the most abundant proteoglycan found in

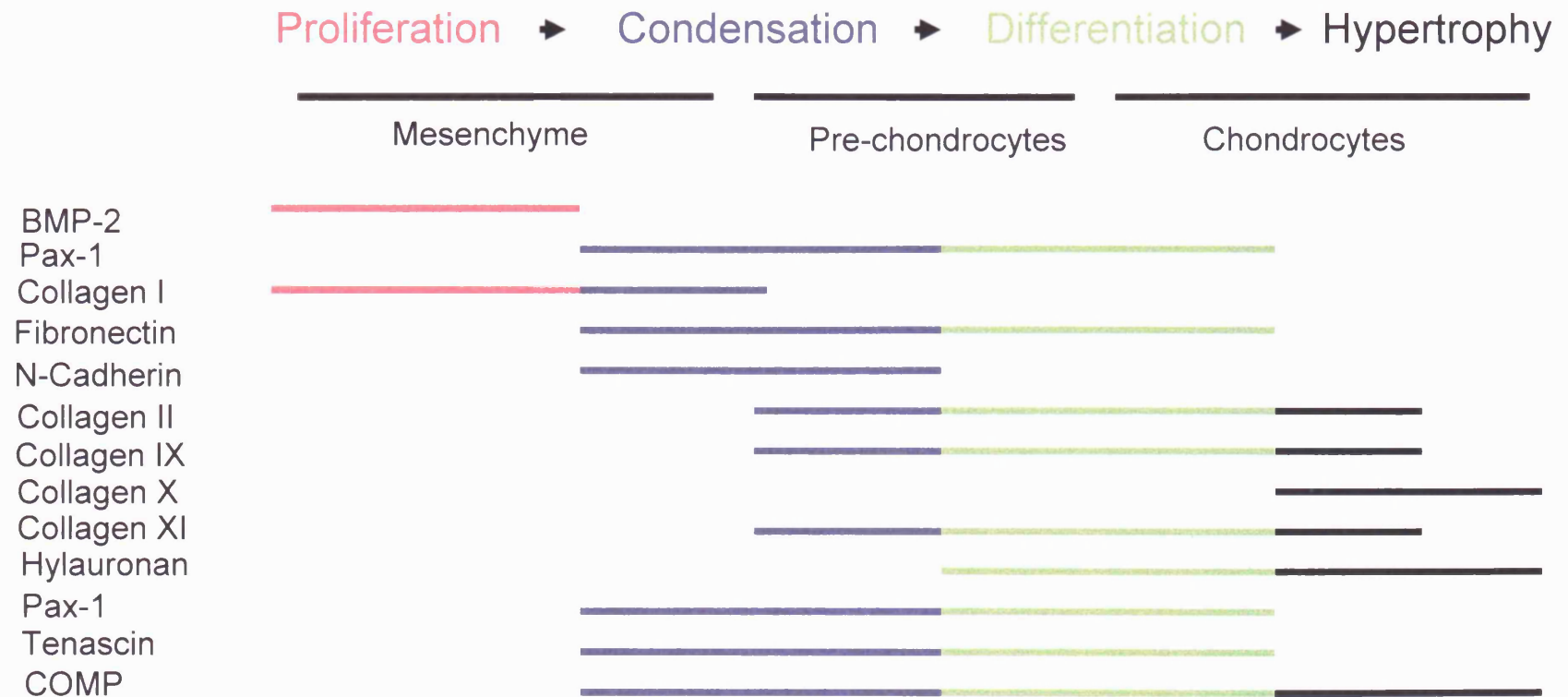


Figure 1.12. Schematic showing expression pattern of genes in chondrocyte proliferation, condensation, differentiation and hypertrophy.

cartilage and contributes towards its compression resistant nature. Mutations in *aggrecan* cause a form of SED known as the Kimberley type in which affected individuals have short trunk and limbs and develop early onset arthropathy (Gleghorn *et al.*, 2005). Decorin is a less abundant proteoglycan which coats collagen molecules and is thought to regulate their fibrillogenesis. Interestingly, irregularly arranged, thickened collagen fibrils have been described in DMC and are thought to be related to abnormality of the extracellular matrix make-up (Horton and Scott, 1982). Perlecan is a heparan sulphate proteoglycan found in cartilage which, if not secreted due to *heparan sulphate proteoglycan 2 (HSPG2)* gene mutations results in Silverman-Handmaker dyssegmental dysplasia, a lethal autosomal recessive disorder characterised by short ribs and an abnormal spine in addition to short long bones (Arikawa-Hirasawa *et al.*, 2001). The structure of a long bone is macroscopically divided into epiphysis, diaphysis and metaphysis as shown in Figure 1.1. The epiphysis is the end of the bone and is covered by hyaline cartilage forming an articular surface. During growth the epiphysis is separated from the metaphysis by the growth plate, replaced in adult life by a transverse bone plate. The diaphysis forms the shaft of the bone and is not seen in small bones such as the carpals and vertebrae. Bone is formed by the deposition of osteoid by osteoblasts followed by mineralization. Mineralisation in woven bone is initiated by spherical membrane bound vesicles known as matrix vesicles that originate from cytoplasmic processes of osteoblasts. The macroscopic and microscopic structure of bone is summarised pictorially in Figure 1.13 (Provot and Schipani, 2005).

1.8 Protein synthesis, transport and degradation

The finding of dilated rough endoplasmic reticulum (RER) and multiple vacuoles in cartilage and skin of DMC individuals has led to the hypothesis that the pathogenic mechanism underlying this skeletal dysplasia is one of abnormal storage within the ER (Horton and Scott, 1982, Nakamura *et al.*, 1997). The ER is a membrane system (illustrated in Figure 1.14) which is continuous with the nuclear envelope and both smooth and rough endoplasmic reticulum (Rutishauser and Spiess, 2002). The functions of the ER include the integration of membrane proteins in addition to the folding, modification and translocation of secretory proteins (Voeltz *et al.*, 2002). The ER also plays an essential role in the degradation of aberrant proteins by the quality control process of ER-associated degradation (ERAD), a process of export of aberrant proteins to the cytosol for degradation. This process requires chaperones, a family of conserved proteins which associate with unfolded or misfolded substrates, assisting in their correct folding and retaining them in the ER until the correct conformation has been achieved.

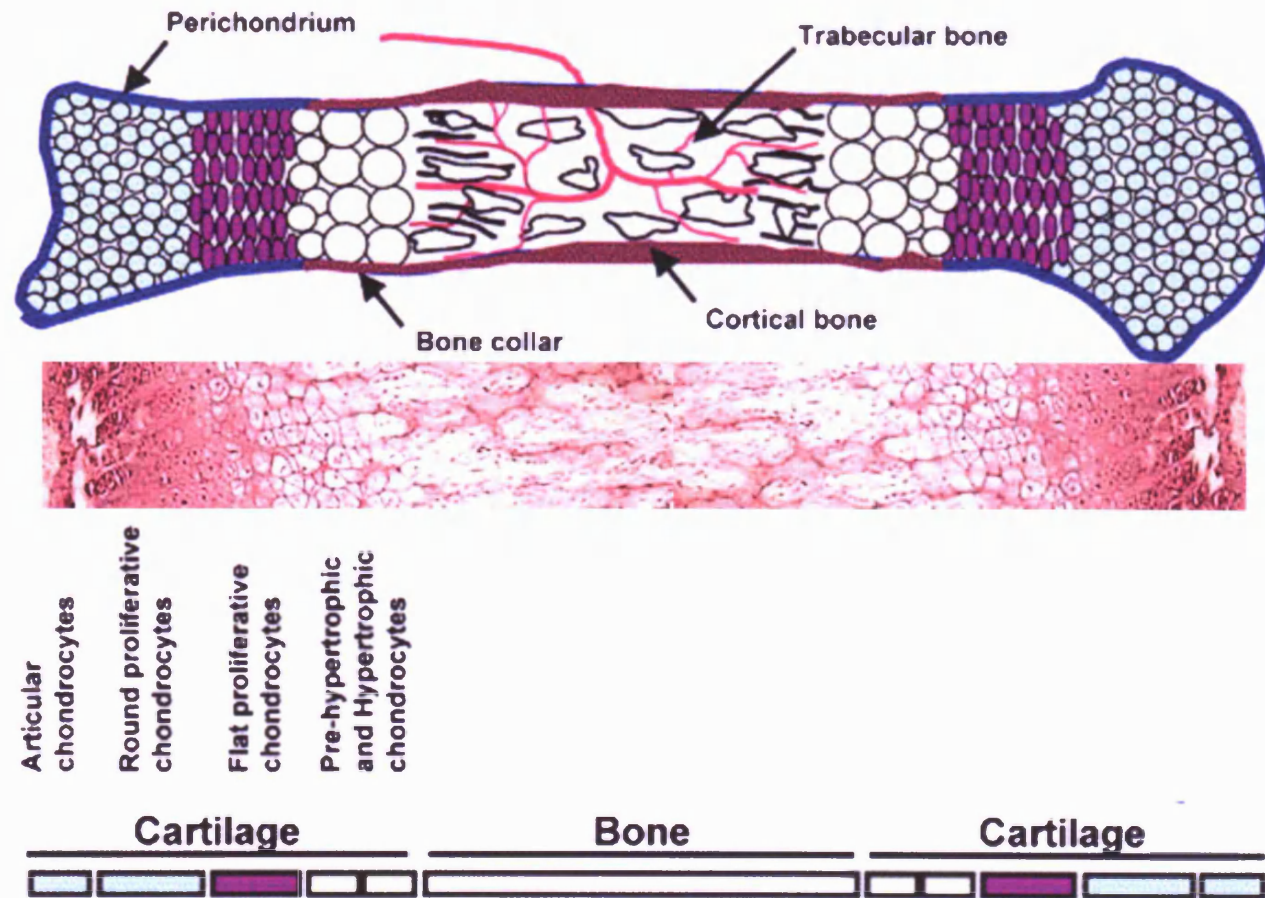


Figure 1.13. Macroscopic and microscopic histological structure of bone. Stages of chondrocyte differentiation are shown on the left with articular and round proliferating chondrocytes in blue, flat proliferating chondrocytes in purple and hypertrophic chondrocytes in cream. Histological image from www.bioscience.org. Figure adapted from Provot.

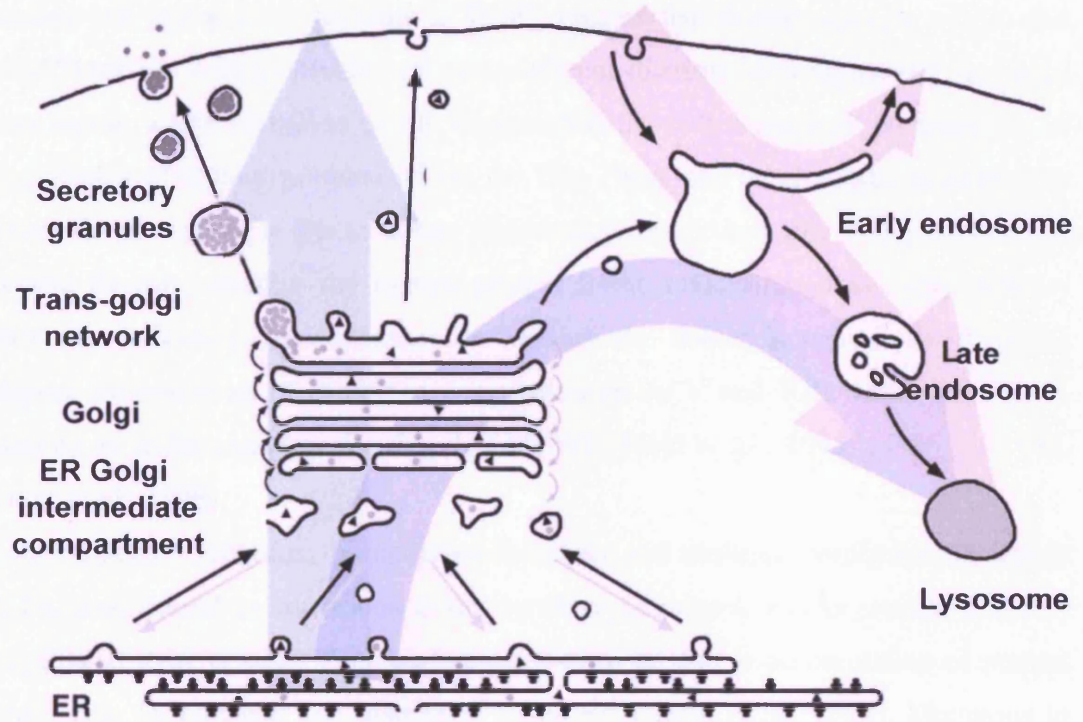


Figure 1.14. The endoplasmic reticulum transport system. The ER delivers proteins to the secretory and endocytic compartments of the cell. Adapted from Rutishauser *et al.*

Known disease causing mutant proteins which are ERAD substrates include the most common mutations in both cystic fibrosis and alpha-1-antitrypsin deficiency (Ward *et al.*, 1995, Le *et al.*, 1990). If a mutant protein fails to be recognised as such by ERAD, accumulation will occur in the ER leading to the impairment of cellular function and possibly cell death. This heterogeneous group of disorders are known as ER storage disorders and histological findings in DMC suggest that it may also lie within this group. There are a large number of very different diseases with apparently unrelated causes but in which dilatation of ER is characteristic with a common pathogenesis of accumulation of mutant proteins within the ER. These can be classified according to whether the phenotype is due to failure of secretion of a protein due to ER retention as in cystic fibrosis, whether the mutant protein forms toxic aggregates (eg. alpha-1-antitrypsin deficiency liver disease) or whether the defect is caused by defective transport machinery as in combined coagulation factor V and VIII deficiency due to mutations in *lectin mannose binding 1 (LMAN1)* (Ward *et al.*, 1995, Le *et al.*, 1990, Nichols *et al.*, 1998).

In the skeletal dysplasias, pseudoachondroplasia and multiple epiphyseal dysplasia (MED), both caused by mutations in the *cartilage oligomeric matrix protein (COMP)* gene, dilated RER is seen. This has been shown to be due to accumulation of mutant COMP, type IX collagen and matrilin 3 in the ER (Hecht *et al.*, 2004). Mutations in *MATN3* also causing MED have been shown to lead to Matrilin 3 accumulation in the ER (Cotterill *et al.*, 2005). Accumulation of matn3 within the RER of chondrocytes and subsequent reduced proliferation and dysregulation of apoptosis has been shown in a matn3 knock-in mouse model (Leighton *et al.*, 2007).

The wide range of clinical phenotypes affecting different body systems which can result from abnormal ER function highlight how critical this compartment is for normal cellular function. DMC may be another condition that lies within this spectrum.

1.9 Yeast two-hybrid system

The yeast two-hybrid system is an experimental means of identifying interacting proteins. It was developed following increased understanding of the role of eukaryotic transcription factors, specifically, the finding that transcription factors consist of at least two domains. These domains include a DNA binding domain (DBD) and a transcription activation domain (AD). Independently, the two domains are non-functional but when in close proximity can activate transcription. The yeast two-hybrid system exploits this fact by generating independent fusion proteins containing a single domain in addition to either a 'bait' or 'prey' protein. The protein of interest is

described as ‘bait’ and the proteins with which it interacts as ‘prey’. Interaction of the fusion proteins generates a functional transcription factor, which could, for example be a growth selective reporter gene thus allowing selection for only clones encoding interacting proteins. The system can be used to identify interacting proteins from a cDNA library encoding a large number of proteins. Initial experiments fused the DBD of one transcription factor to the AD of an unrelated transcription factor and assessed the ability of the combined protein to activate transcription in yeast (Brent and Ptashne, 1985). Subsequently this technique was modified in order to detect protein interactions. The yeast Gal4 transcription factor was split into separate DBD and AD and fused to two proteins known to interact. The plasmids encoding the fusion genes were co-transformed into yeast containing an engineered reporter gene (*lacZ*) and Gal4 transcription demonstrated (Fields and Song, 1989). Further modifications to the technique allowed the identification of new protein interactions by the expression of a ‘bait’ protein fused to a Gal4 DBD in yeast. A library of cDNAs were fused to the Gal4 AD and transformed into yeast which had previously been transformed with the bait fusion protein. A *LacZ* reporter gene, when activated led to a change in yeast colony colour to blue hence allowing the identification of colonies containing cDNA of interactors (Chien *et al.*, 1991). The method most commonly used now relies upon multiple reporter genes that are growth selective. Clones containing interacting proteins activate the transcription of auxotrophic genes thus allowing the clone to grow on nutritionally deficient media, hence allowing selection for clones containing interactors. The yeast two-hybrid technique is now a well established method of identifying interacting proteins and has been scaled up to array format in order to speed up the detection of large numbers of interacting proteins and to identify proteins acting within protein networks (Korf and Wiemann, 2005, Fields, 2005, Xia *et al.*, 2004). The technique has also been adopted to investigate whether proteins known to be mutated in similar conditions interact. In the human inherited ataxias, the yeast two-hybrid method was adopted to investigate 54 proteins known to be involved in 23 disorders of Purkinje cell degeneration. This was a highly successful approach, detecting 770 protein-protein interactions of which 83 % were confirmed in mammalian cells (Lim *et al.*, 2006). The use of both this and the array based technique are likely to vastly increase our understanding of protein interaction related to human disease in the future.

1.10 Summary

The studies described in this project aim to further understand the role of *DYM* through the identification of disease causing mutations, cellular and sub-cellular localisation

studies of wild-type and mutant transcripts and proteins and the identification of DYM interactors. Together, it is anticipated that these studies will clarify the role of the *DYM* gene product at a cellular level and improve our understanding of the underlying disease mechanism.

Chapter 2 – Materials and Methods

2.1 Materials

2.1.1 General reagents

Chemicals and reagents were purchased from Fisher Scientific (Loughborough, UK) or Sigma Aldrich (Poole, UK) except those listed below.

α [³² P]- dCTP	Amersham Bioscience (Chalfont St.Giles, UK)
Agarose	Cambrex BioScience (Rockland, ME, USA)
Aquamount	BDH (Poole, UK)
Dextran Sulphate	Amersham Bioscience (Chalfont St.Giles, UK)
dNTPs	Amersham Bioscience (Chalfont St.Giles, UK)
Zeta probe membrane	Biorad Laboratories (Herules, CA, USA)
IPTG	Melford Laboratories (Suffolk, UK)
L[³⁵ S]- methionine	Amersham Bioscience (Chalfont St.Giles, UK)
Lipofectamine	Invitrogen (Paisley, UK)
Long ranger acrylamide	Cambrex BioScience (Rockland, ME, USA)
Nonidet P-40	ICN Biomedicals (Irvine, CA, USA)
ProtoGel	National Diagnostics, (Hull, UK)
Random priming kit	Invitrogen (Paisley, UK)
Sequagel Complete/6	Flowgen (Ashby de la Zouch, UK)
YNB	MP Biomedicals, (Irvine, CA, USA)

2.1.2 Enzymes

Enzymes and buffers were purchased from New England Biolabs (Hitchin, UK) except those listed below.

Klenow DNA polymerase	Amersham Bioscience (Chalfont St.Giles, UK)
T4 ligase and buffer	Promega (Madison, WI, USA)
<i>Pfu</i> DNA polymerase	Statagene (La Jolla, CA, USA)

2.1.3 Cell culture media and reagents

DMEM (4500 mg/ml glucose), RPMI media and fetal calf serum were purchased from Gibco, Paisley, UK. Antibodies were purchased from Invitrogen, Paisley, UK unless otherwise stated.

2.1.4 Oligonucleotides

Oligonucleotides were purchased from Sigma Genosys (Cambridge, UK).

2.1.5 Molecular Biology kits

BigDye terminator	Applied Biosystems (Foster City, CA, USA)
cDNA multiple tissue panel	Clontech (Mountain View, CA, USA)
Dig RNA labelling kit	Roche (Basel, Switzerland)
ECL plus Western blotting detection system	Amersham Biosciences (Chalfont St.Giles, UK)
First strand cDNA synthesis kit	Roche (Basel, Switzerland)
Gel purification kit	Qiagen (Crawley, UK)
Miniprep kit	Qiagen (Crawley, UK)
Maxiprep kit	Qiagen (Crawley, UK)
Multiplex PCR master mix	Qiagen (Crawley, UK)
Paxgene blood RNA extraction kit	Qiagen (Crawley, UK)
PCR Purification kit	Qiagen (Crawley, UK)
Quikchange site directed mutagenesis	Statagene (La Jolla, CA, USA)
RNEasy kit	Qiagen (Crawley, UK)
TNT® Quick coupled transcription translation system	Promega (Madison, WI, USA)
TNT® T3 IVT kit	Promega (Madison, WI, USA)
Yeast plasmid isolation kit	Amersham Bioscience (Chalfont St.Giles, UK)

2.1.6 Patient samples

Venous blood samples (10 ml) were taken into EDTA tubes for DNA extraction or Lithium Heparin tubes for establishment of cell lines by the European Collection of Cell cultures, Porton Down, UK (ECACC).

2.1.7 Common solutions

Agarose loading dye

35 % (w/v) sucrose
6 x TBE
0.25% (w/v) bromophenol blue

Cell extraction buffer

50 mM Tris pH 7.4
100 mM NaCl
10 mM MgCl₂
0.1 % (w/v) SDS
1% (v/v) Triton x100
1 mM PMSF in DMSO
0.1 mg/ml DNase I
0.1 mg/ml RNase A

Cell lysis buffer

10 mM TRIS
10 mM EDTA
10 mM NaCl
0.5 % (w/v) SDS
100 µg/ml proteinase K

Church buffer

7 % (w/v) SDS
0.5 M Na₂HPO₄ pH 7.2
1 mM EDTA pH 8
1 % (w/v) BSA

Coomassie blue

0.5 % (w/v) coomassie blue
40 % (v/v) methanol

Formamide loading dye

5 parts deionised formamide
1 part EDTA
blue dextran to colour

Hybridisation solution

0.6 M NaCl
1 x PE
10% (w/v) dextran sulphate
50 % (v/v) formamide
Probe at required concentration
In DEPC dH₂O

Luria agar

1 % (w/v) tryptone
0.5 % (w/v) yeast extract
0.85 M NaCl
1.6 % (w/v) agar
2 mM NaOH

Luria broth

1 % (w/v) tryptone
0.5 % (w/v) yeast extract
0.85 M NaCl
2 mM NaOH

NET buffer

1 mM EDTA
20 mM TRIS HCl pH 8
100 mM NaCl

NETN buffer

1 mM EDTA
20 mM Tris-HCL pH 8
100 mM NaCl
0.5% (v/v) nonidet P40

1 x PBS

137 mM NaCl
2.7 mM KCl
1.4 mM KH₂PO₄
4.3 mM Na₂HPO₄.2H₂O

1 x PBS-tween

1x PBS +0.1% tween

10 x PCR buffer

45 mM Tris HCl pH8.8

11 mM (NH₄)₂ SO₄

45 mM MgCl₂

110 µg/ml BSA

6.7mM βME

4.4 µM EDTA

20 mM dATP

20 mM dCTP

20 mM dGTP

20 mM dTTP

Ponceau S

0.1% (w/v) ponceau S

5% (v/v) acetic acid

Prehybridisation solution

0.6 M NaCl

1 x PE

10% (w/v) dextran sulphate

50 % (v/v) formamide

In DEPC dH₂O

Protein digestion buffer

1 mM TRIS

1 mM EDTA

0.15 M NaCl

0.5 % (w/v) SDS

pH adjusted to 10.5 with NaOH

SDS-PAGE destain

40 % (v/v) methanol

10 % (v/v) acetic acid

2 x SDS-PAGE loading dye (Laemmli buffer)

62.5 mM TRIS-HCl pH 6.8

5 % (v/v)-β-ME

2 % (w/v) SDS

0.2 % bromophenol blue

10 % (v/v) glycerol

SDS-PAGE lower buffer

1.5 M TRIS-HCl pH 8.8

0.4 % (w/v) SDS

10 x SDS-PAGE running buffer

250 mM TRIS base

1.92 M glycine

1% (w/v) SDS

SDS-PAGE stain

0.025 % (w/v) coomassie blue
40% (v/v) methanol
10% (v/v) glacial acetic acid

SDS-PAGE upper buffer

0.5 M TRIS-HCl pH 6.8
0.4 % SDS

20 x SSC

3 M NaCl
0.3 M trisodium citrate
Adjust pH to 7 with HCl

Sucrose lysis buffer

11 % (w/v) sucrose
10 mM TRIS-HCl pH 7.5
5 mM MgCl₂

Synthetic complete dropout medium mix

2 g arginine
2 g isoleucine
2 g lysine
2 g methionine
3 g phenylalanine
6 g homoserine
3 g tryptophan
2 g tyrosine
9 g valine

Synthetic complete drop out medium

0.67 % (w/v) yeast nitrogen base
2 % (w/v) glucose
0.083 % (w/v) synthetic complete drop out mix
Adjust pH to 5.6 with NaOH
1.7 % (w/v) agar (for solid medium)

TE

10 mM TRIS-HCl pH 7.4
1 mM EDTA pH 8

Transfer buffer

25 mM TRIS base
192 mM glycine
20 % (v/v) methanol

Transformation buffer

15 mM CaCl₂
250 mM KCl
10mM PIPES
55mM MnCl₂

10 x Tris-borate buffer (TBE)

0.9 M tris

0.9 M boric acid

1 mM EDTA

Tris buffered saline (TBS)

0.05M tris

0.15M NaCl

Yeast extract, peptone and dextrose liquid media (YPD)

1 % (w/v) yeast extract

2 % (w/v) peptone

2% (w/v) glucose

0.01% (w/v) adenine sulphate

2.2 Methods

2.2.1 Multi-centre Research Ethical Committee (MREC) approval

Ethical approval was obtained for the study from the Multi-centre Research Ethics Committee (see Appendix A).

2.2.2 Family recruitment

Families were recruited following contact with clinical geneticists both in the United Kingdom and abroad. Clinical information was obtained in addition to blood or DNA samples and families were visited by the researcher where possible. Participation followed informed consent.

2.2.3 DNA extraction

2.2.3.1 Blood

Blood samples were received as 10 ml of whole blood in EDTA collection tubes. After chilling on ice for 30 min, ice cold sterile dH₂O was added to a volume of 50 ml and the sample mixed. Centrifugation was performed at 800 g, 4°C for 10 min. The supernatant was discarded, the pellet resuspended in 25 ml ice cold sucrose lysis buffer and incubated on ice for 10 min. Samples were centrifuged at 800 g, 4°C for 10 min, the supernatant discarded and the pellet resuspended in 3 ml of protein digestion buffer containing 0.2 mg/ml proteinase K. The samples were incubated at 37°C overnight. 2 ml of 6 M NaCl was added and the sample agitated for 20 sec and centrifuged at 1100 g, 4°C for 30 min. The supernatant was discarded, 10 ml ethanol added and the sample inverted until a DNA precipitate appeared. The precipitate was removed with a pipette tip and placed in 1 ml 70% ethanol. Centrifugation at 800g for 10 min was performed, the supernatant removed and the pellet air dried. Pellets were resuspended in 500 µl of dH₂O. A 500 fold dilution of the DNA was quantified at a wavelength of 260 nm and the concentration calculated using the equation 1 OD₂₆₀ equals 50 µg/ml.

2.2.3.2 Cell lines

EBV transformed lymphocyte cell lines were grown at 37 °C and 5% CO₂ in RPMI medium supplemented with 10% fetal calf serum and penicillin (50 U/ml) and streptomycin (50 µg/ml). DNA was extracted from 1 x10⁶ cells using the phenol-chloroform method as follows; cells were lysed with trypsin, washed in 5 ml PBS and centrifuged at 1100 g for 5 min. The cell pellet was resuspended in 2 ml cell lysis buffer

and incubated at 37 °C overnight, 2 ml phenol:chloroform (1:1) was added and centrifugation performed at 800 g for 15 min at room temperature. The aqueous phase was removed to a new tube and the step repeated, the aqueous phase was removed to a new tube and 4 ml chloroform added. One tenth of the sample volume of 4 M NaCl and 2 volumes of 100% ethanol were added and centrifugation performed at 800 g for 30 min at 4 °C. The supernatant was removed and the pellet washed in 2 ml 70 % ethanol and centrifugation repeated for 10 min. The supernatant was removed and the pellet air dried and resuspended in dH₂O.

2.2.4 RNA extraction

2.2.4.1 Blood

RNA was extracted from blood using PAXgene blood RNA collection tubes in association with PAXgene blood RNA extraction kit (Qiagen ,Crawley, UK), according to the manufacturer's instructions.

2.2.4.2 Cell lines

RNA was extracted from lymphocyte cell lines using RNEasy Midi kit (Qiagen, Crawley, UK), according to the manufacturer's instructions.

2.2.5 Oligonucleotide design

2.2.5.1 Genomic DNA

Polymorphic microsatellites surrounding *DYM* were selected from the Genethon linkage map. Oligonucleotide primer sequences for the markers were obtained from the NCBI database (www.ncbi.nlm.nih.gov). The 5' primer for each microsatellite pair was labelled with a fluorophore dye. PCR primers for amplifying DMC coding sequence were taken from El Ghouzzi *et al* 2003 (El Ghouzzi *et al.*, 2003).

2.2.5.2 cDNA

Primers to amplify *DYM* cDNA were designed using Primer 3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi).

2.2.5.3 Cloning oligonucleotides

To generate restriction enzyme sites within a cloned DNA sequence, oligonucleotides were designed for PCR with approximately 20 bases at the 3' end of the oligonucleotide

corresponding to the template DNA. A restriction enzyme site was inserted within the 5' end of the oligonucleotide, 4 nucleotides from the 5' terminal phosphate.

2.2.6 PCR

A general method for PCR was used based on the initial description of PCR (Saiki *et al.*, 1985).

PCR reactions were performed using the appropriate primers pairs as below.

1x PCR premix:

10 x PCR buffer	0.9 µl
Forward primer (10 µM)	0.5 µl
Reverse primer (10 µM)	0.5 µl
Taq	0.5 U
Genomic DNA	50 ng
dH ₂ O to 10 µl	

Cycling was performed in a DNA Engine Dyad/Tetrad thermal cycler (MJ Research, Waltham, USA) with cycling conditions of an initial denaturation step at 96 °C for 5 min followed by 35 cycles of 96 °C for 30 sec, annealing at a temperature appropriate to the primer pair for 30 sec and extension at 72 °C for 30 sec followed by a final extension step at 72° C for 10 min. Modifications to this PCR technique were made depending on requirements as below.

2.2.6.1 PCR of microsatellite markers

Microsatellite marker PCRs were performed using the fluorescently labelled oligonucleotides for each marker at the appropriate conditions as listed in Table 2.1. Fluorescent primers and PCR products were stored in the dark.

2.2.6.2 PCR for genomic sequencing

PCR for genomic sequencing of *DYM* was performed using the oligonucleotides for each exon at the appropriate conditions as listed in Table 2.2. (El Ghouzzi *et al.*, 2003).

2.2.6.3 PCR for cloning

PCR for cloning was scaled up to 100 µl using 1 ng of template plasmid DNA. Oligonucleotide primers used are listed in Table 2.3. The 72 °C extension step was modified to allow 2 min/kb of amplicon.

Primer name	Primer sequence and label	Annealing temperature	Product size (base pairs)
D18S1137	5' HEX- GGA CTT GCA CGC TAA TGA C 3'	55°C	189-195
UD18S1137	5' CTT GAC TTT TTC AGA AGT GGT 3'	55°C	
D18S1118	5' TET- TTT CCC AGA CCA CCA G 3'	55°C	172-192
UD18S1118	5' GCA AGC ATC AGA GAG TAA GG 3'	55°C	
D18S455	5' FAM- AGC CAT TTG TCA GGG TG 3'	55°C	136-150
UD18S455	5' GAT GTG TGC ATG TGT GTT CT 3'	55°C	
D18S450	5' TET- TGG GAC CAA CCT AAC TGT G 3'	55°C	226-278
UD18S450	5' TTG TAT GGT GCA TGA CCC T 3'	55°C	
D18S1126	5' HEX- CTT GCA GAC AAC ATA TCT TAC A 3'	55°C	244-274
UD18S1126	5' CTT GAC TTT TTC AGA AGT GGT 3'	55°C	
D18S470	5' FAM- AGC TTA CCA CAA GGC ATA ACT 3'	55°C	275-285
UD18S470	5' CCA TGA GGG TAG ACT GTT AAC 3'	55°C	
D18S472	5' HEX- GCC CAG CCA ACT TTT AAA TA 3'	55°C	149-163
UD18S472	5' GTG ACA ACA GTG AAA CTC CA 3'	55°C	
D18S363	5' FAM- TTG GGA ACT GCT CTA CAT TC 3'	55°C	177-247
UD18S363	5' GCT TCA TTC TCT CAC TGG AT 3'	55°C	
Dos Ex2F	5' FAM- GTT GTC TTT TGG AAA TGC AG 3'	60°C	254
Dos Ex14F	5' FAM- CCA CAG AAG ACA GAA AAG AAA 3'	60°C	237
Dos Ex15F	5' FAM- CTG TGT TGG GGT TTC TTT T 3'	60°C	280
Dos Ex15R	5' GGC TTA TTA AAC ATT CTC TGC T 3'	60°C	
Dos Alk1 6F	5' AGG CAG CGC AGC ATC AAG AT 3'	60°C	266
Dos Alk1 6R	5' AAA CTT GAG CCC TGA GTG CAG 3'	60°C	

Table 2.1. *DYM* microsatellite and fluorescent dosage PCR primers

Primer name	Primer sequence	Annealing temperature
DMC Ex2F	5' GGAAATGCAGGTTTAAGGAC 3'	56°C
DMC Ex2R	5' CATACCACACAAGTTGATATG 3'	56°C
DMC Ex3F	5' GTATTATAGTCATTAATTACCTTT 3'	56°C
DMC Ex3R	5' GTGAATGCCACAAATAGAAAAA 3'	56°C
DMC Ex4F	5' GCTCACTTTGATTTTAGTTCTT 3'	56°C
DMC Ex4R	5' CCATCAATAAATGAAAACATAAAAT 3'	56°C
DMC Ex5F	5' GAACTAGAAATAAACTATGGTG 3'	56°C
DMC Ex5R	5' ATCCTTAAATGTTACACATGATA 3'	56°C
DMC Ex6F	5' CAACTGTATTCTACTGTGTAC 3'	56°C
DMC Ex6R	5' GGACCCACAACTTCTCAAC 3'	56°C
DMC Ex7F	5' GTGACTGTTACTCTGTTTTTC 3'	58°C
DMC Ex7R	5' ATATGGGACGATACTCAGGT 3'	58°C
DMC Ex8F	5' GAATTAGATGACTCCCATAAAT 3'	58°C
DMC Ex8R	5' CACCAAAGAATTGAAAAAATCTG 3'	58°C
DMC Ex9F	5' CATTCACTCATTGGTGGCA 3'	58°C + 2% DMSO
DMC Ex9R	5' ACAAGCAATACTCTCCCAT 3'	58°C + 2% DMSO
DMC Ex10F	5' ATAAACCACAGCTAGAAATTA 3'	58°C
DMC Ex10R	5' CTGTGCCATGTACCCTAAAA 3'	58°C
DMC Ex11F	5' GCAAAGAGCACAAAATGAAAG 3'	55°C
DMC Ex11R	5' CCAGAAAGAAATGTAAAGACTA 3'	55°C
DMC Ex12F	5' GTGGACATGGAGTAATTCTTT 3'	58°C
DMC Ex12R	5' CTATCGTCTTAATTGTAAGTA 3'	58°C
DMC Ex13F	5' GAGAAGACTTGATTGTATGTTT 3'	58°C
DMC Ex13R	5' CCATCTTAATAGCTCAGCCA 3'	58°C
DMC Ex14F	5' CCACAGAAGACAGAAAAGAAA 3'	55°C
DMC Ex14R	5' GATTACAGGCAGGAGTACT 3'	55°C
DMC Ex15F	5' CAAGGAGGAAAAGACTGACT 3'	55°C
DMC Ex15R	5' CAGATTACCACAACTATTATAG 3'	56°C
DMC Ex16F	5' GTGAGAAAGGCGGCTACTA 3'	56°C
DMC Ex16R	5' CATTTACATCAAGGGACACG 3'	56°C
DMC Ex17F	5' CACCTGCAAACTCTACCA 3'	56°C
DMC Ex17R	5' CCTGTCTGTCTACTTCTGTT 3'	56°C

Table 2.2. *DYM* genomic sequencing PCR primers

Primer name	Primer sequence
DMC 5'MYC	5' CAT GAT ATC GCC ACC ATG GAG CAG AAG CTT ATT TCT GAG GAG GAC CTG AAC ATG GGA TCG AAT AGC AGC AGA 3'
DMC 508R	5'TGA TCA ACT GCA TCA AAC AGC AC 3'
ECORV 5'UTR	5' GAC GAT ATC CGA CTT CTC CTG 3'
5' GFP	5' GAA CCG CAT CGA GCT GAA GGG C 3'
GFPin frame	5' CGA TCA AGC TTC GAT GGG ATC GAA TAG CAG C 3'
GFPR	5' CTT TGG TTC TAG AAA GG 3'
R70XF	5' GGT CAT TAG TTG AAA ACA ATC CTT GAA CAG GAA ATC TTG GTG CAC 3'
R70XR	5' GTG CAC CAA GAT TTC CTG TTC AAG GAT TGT TTT CAA CTA ATG ACC 3'
E87KF	5' CTA GAA CCA AAA AAC TAA AAC TTT CAG CAG 3'
E87KR	5' CTG CTG AAA GTT TTA GTT TTT TGG TTC TAG 3'
R204XF	5' CAA CAT GAC TAG GAC ATG AGA CAA GTA CCT TCA CAC AAA TTG 3'
R204XR	5' CAA TTT GTG TGA AGG TAC TTG TCT CAT GTC CTA GTC ATG TTG
N469YF	5' GCA GCT TTA GCA TAT ATG TCG GCA CAG 3'
N469YR	5' CTG TGC CGA CAT ATA TGC TAA AGC TGC 3'
Q483XF	5' CAG TAT GCT GCC TAG AGG ATC ATC AG 3'
Q483XR	5' CTC ATC ATC CTC TAG GCA GCA TAC TG 3'
T3	5' ATT AAC CCT CAC TAA AG 3'
T7	5' TAA TAC GAC TCA CTA TAG G 3'
M13F	5'GTA AAA CGA CGG CCA GTG 3'
M13R	5' CAG GAA ACA GCT ATG AC 3'
DMCAb1F	5' GAG AGG ATC CGG SATC GAA TAG CAG CAG AAT C 3'
DMCAb2F	5' GAG AGG ATC CTA CAA CAT GAC TAG GAC ACG 3'
DMCAb1R	5' TCT CAA GCT TGT GGT TCT GAC ATT CTG CTG 3'
DMCAb2R	5' TCT CAA GCT TTC AGT CGG AAT CCA TGG TGA A 3'

Table 2.3. Cloning and vector sequencing primers

2.2.6.4 Fluorescent dosage PCR

DNA solutions were incubated at 55 °C for 2 hr and the DNA concentration determined and adjusted if necessary to give a concentration of 50 ng/μl. Ten control DNA samples were used in addition to patient samples. A multiplex PCR assay was performed in 20 μl reactions using multiplex PCR master mix (Qiagen Crawley, UK) to amplify *DYM* exons of interest and a control gene exon. Amplification of exons was performed using the reverse primers used for *DYM* gene amplification (see Table 2.2) in addition to forward primers labelled with the fluorescent phosphoramidite 6-FAM. Labelled primer sequences and the length of products produced are shown in Table 2.1. Cycling conditions consisted of an initial denaturation period of 15 min at 96 °C followed by 19 cycles of denaturation at 96 °C for 30 sec, annealing at 60 °C for 2 min, extension at 72 °C for 2 min followed by a final extension step of 10 min. 2 μl of PCR product was mixed with 1.8 μl formamide loading dye and 0.3 μl TAMRA size standard (Applied Biosystems Foster City, CA, USA). This mix was denatured at 96 °C for 5 min and electrophoresed in 1 x TBE on an ABI 377 DNA fragment analyser (Applied Biosystems Foster City, CA, USA). Gels were analysed on an Apple computer (Apple, Cupertino, USA) using the software packages Genescan v3.1 and Genotyper v2.0 (Applied Biosystems, Foster City, CA, USA).

2.2.7 Agarose gel electrophoresis

DNA samples were analysed by mixing 5 μl of sample with 1 μl sucrose loading dye (or scaled up for large scale cloning/ plasmid digests) and loading on a 2% agarose gel for amplicons <1kb or a 0.8% agarose gel for amplicons >1kb. 50 ng/ml of ethidium bromide was contained within the agarose gel. Gels were run at 10V/cm² in 1x TBE for 30-60 min and viewed using an ultraviolet trans-illuminator (Syngene, Cambridge, UK).

2.2.8 Genotyping of microsatellites

Microsatellite DNA markers surrounding the DMC disease gene were selected using information available at NCBI. Primers were labelled fluorescently such that PCR amplicons of different sizes and labels could be pooled together in order to increase throughput. PCR products were diluted 20 fold in dH₂O and 2 μl of each diluted product combined. 1.5 μl of formamide loading dye and 0.3 μl of TAMRA ® size standard were added to each sample. The samples were denatured at 96 °C for 3 min and kept on ice. Polyacrylamide gels were made using 12 cm well to read ABI Genescan plates. A 6% gel mix was made using 10 ml of SequaGel-6, 2.5 ml of SequaGel complete and 50 μl of 10% ammonium persulphate (APS). Fluorescent PCR samples were loaded on the gel

and electrophoresed in 1 x TBE on an ABI 377 DNA fragment analyser. Gels were analysed on an Apple computer (Apple, Cupertino, USA) using the software packages Genescan v3.1 and Genotyper v2.0. Cyrillic 2.0 software (Cyrillic Software, Wallingford, UK) was used to create haplotypes for the microsatellite allele results.

2.2.9 Fluorescent DNA sequencing

PCR amplification products were column purified using a PCR purification kit following the manufacturer's protocol (Qiagen, Crawley, UK). Purified PCR products were used for direct sequencing using the primers used to generate the PCR product as below:

Sequencing reaction:

BigDye terminator mix	2 µl
Primer (10 µM)	0.4 µl
PCR product/plasmid	2-3 µl
dH ₂ O to 10 µl	

Sequencing reactions were performed on a DNA Engine Dyad/Tetrad thermal cycler using 29 cycles of an initial denaturation step of 5 sec at 96 °C, annealing at 60 °C for 10 sec and extension at 72 °C for 4 min. The sequencing reactions were purified using DyeEx 2.0 purification columns following the manufacturer's protocol (Qiagen, Crawley, UK). The dried pellet was resuspended in 3 µl of formamide loading dye and denatured at 96 °C for 4 min and kept on ice until loading. Polyacrylamide gels were made using 36 cm well to read ABI plates. A 5% gel mix was made using 10.8 g urea, 3 ml of LongRanger® acrylamide, 3 ml 10 x TBE, 15.6 ml dH₂O, 150 µl 10% APS and 21 µl TEMED. Fluorescent sequencing samples were loaded on the gel and electrophoresed using 1 x TBE on an ABI 377 DNA fragment analyser. Gels were analysed on an Apple computer (Apple, Cupertino, USA) using Sequencing Analysis v3.4.1 software package (Applied Biosystems Foster City, CA, USA).

2.2.10 RT-PCR

Complementary DNA was synthesised by priming with oligodT in the presence of reverse transcriptase following the manufacturers protocol using First strand cDNA synthesis kit (Roche, Basel, Switzerland). PCR conditions were amended to include 30 PCR cycles. Primer sequences for RT-PCR are listed in Table 2.4.

cDNA Fragment	Primer names	Primer sequences	cDNA sequence amplified	Expected product size (base pairs)
DYM 1	5' UTR F FLJ 556R	5' GAC AGC GAC TTC TCC TGA CC 3' 5' CTG AGT CAG AAC TGT AAT TGC 3'	2-576	574
DYM 2	FLJ 1F FLJ 623R	5' GTC TTC CTT TCT AGA ACC AAA G 3' 5' TGA TCA ACT GCA TCA AAC AGC AC 3'	380-623	243
DYM 3	FLJ 522F FLJ 873R	5' CAA CTT CAT TTT ACT TAT GAA G 3' 5' GCA AGT CCA TAA AGC AGT CC 3'	522-893	371
DYM 4	FLJ 2F FLJ 3R	5' GGT CCA TGT CTT CCA TAC AC 3' 5' CAC AAA GAG CTG TGT ACA AAC 3'	756-1164	408
DYM 5	FLJ 1059F FLJ 4R	5' CAA GCC ATT ATG TCC TTC AAG 3' 5' CTC TGG GCA GCA TAC TGA TG 3'	1059-1595	536
DYM 6	FLJ 3F FLJ 4R	5' GCA ACT CTC CTC TTG TAT ACC 3' 5' CTC TGG GCA GCA TAC TGA TG 3'	1186-1595	409
DYM 7	FLJ 1458F FLJ2034R	5' CTG ATC CTG GTG GTA ATA AG 3' 5' CTC GGG CTG CTC CTC TTC 3'	1458-2052	594
DYM 8	FLJ 1024F 3' UTR R	5' GGA AAT CAT TAA GCA AGG CG 3' 5' TGC CAA GTA GCA CCA ATT CTC C 3'	1924-2327	403
DYM Int Ex2	Int Ex2 F Int Ex2 R	5' CGT TCT GGA ATC AGC TTC TC 3' 5' TTC TGC TGC TAT TCG ATC CC 3'	N/A	N/A
DYM Int Ex14	DMC Int Ex14F DMC Int Ex14R	5' CTA ATG ATG TTC CTC TAC CAG 3' 5' GTT GTG TTT TTT AGA CAG CAA AG 3'	N/A	N/A
BMPR2	BMPR2 F3F BMPR2 F2R	5' AGC CTC TCC ACC AAC ACA AC 3' 5' GGG TCT AGC TTG TTG GTT TCC 3'	2227-2417	191

Table 2.4. *DYM* RT-PCR primers

Where necessary, RT-PCR products were separated on 2% agarose gel and bands gel purified using a Gel purification kit following the manufacturer's protocol (Qiagen, Crawley, UK).

2.2.11 Fluorescent dosage PCR- calculation of dosage quotients

Genotyper v2.0 software was used to calculate the peak area for each exon. A mean peak area for each exon was calculated for 10 control samples and used to compare to patient samples. Dosage quotients were calculated using the following formula:

$$\text{DQ exon a/exon b} = \frac{\text{sample exon a peak area/sample exon b peak area}}{\text{control exon a peak area/control exon b peak area}}$$

2.2.12 Southern blot hybridisation

Restriction enzyme digestion of patient and control genomic DNA was performed as follows;

10 µg genomic DNA
5 µl appropriate restriction enzyme buffer
50 U enzyme
0.5 µl 10 x BSA or 5 µl 50mM spermidine
dH₂O to 50 µl

Samples were incubated for 16 hr at the temperature appropriate to the enzyme used. Digested products were run on 0.8% TBE agarose gel in 1 x TBE at 40V for 16 hr. Bands were visualised by soaking in TBE with ethidium bromide (100 ng/ml) for 30 min, washing in dH₂O for 30 min and transilluminating.

The gel was soaked in 0.25 M HCl for 10 min, and blotted onto Zeta probe membrane (Biorad Laboratories, Hercules, CA, USA) by overnight transfer using 0.4 M NaOH. Pre-hybridisation was performed in 12 ml Church buffer for 2 hr at 65 °C.

25 ng probe was incubated at 96 °C for 5 min and random priming reaction performed using a random priming kit according to the manufacturer's recommendations (Invitrogen, Paisley, UK). The reaction was incubated at room temperature for 1 hr. The probe was purified using G50 Sephadex beads according to the manufacturer's recommendations (Sigma, Cambridge, UK). The purified probe was added to the membrane in Church buffer and hybridised at 65 °C for 16 hr. Blots were washed twice in 2 x SSC, 0.1% SDS for 20 min at 65°C, twice in 0.2 x SSC, 0.1% SDS for 20 min at 65 °C and twice in 0.1 x SSC 0.1% SDS for 20 min at 65 °C or until the background count was 5-10 counts per second. The membrane was exposed to a Typhoon ® cassette

overnight and read on a Typhoon ® phospho-imager (Amersham Bioscience, Chalfont St.Giles, UK).

2.2.13 Cloning methods

2.2.13.1 Preparation of competent *E.coli*

E.coli DH5α or BL21 were streaked onto LB agar and grown overnight at 37 °C. A single colony was selected and inoculated into 2 ml LB and grown overnight at 37 °C. 1 ml of the overnight culture was inoculated into 500 ml LB/5mM MgCl₂ and grown at 37 °C until OD₆₀₀ was equal to 0.3-0.6. The culture was cooled rapidly on ice and centrifuged at 2500g for 10 min at 0 °C. The supernatant was discarded and the pellet resuspended in 150 ml ice cold transformation buffer. Centrifugation was repeated as above and the pellet resuspended in 40 ml ice cold transformation buffer. 3 ml DMSO was added and the cell suspension aliquotted and frozen in liquid nitrogen prior to storage at -80 °C.

2.2.13.2 Restriction enzyme digestion

PCR was performed as a 100 µl reaction and purified using a PCR column purification kit following the manufacturer's protocol (Qiagen Crawley, UK). PCR products or vectors were digested with the appropriate restriction enzymes using 2 µl of a restriction enzyme buffer (compatible for both enzymes), 5 U of each enzyme, 2 µg of plasmid or 5 µl PCR product and dH₂O to a final volume of 20 µl. Reactions were incubated at the temperature appropriate to the enzymes used for 2 hr and subsequently electrophoresed on an agarose gel. Insert and vector DNA were viewed on a trans-illuminator (Syngene, Cambridge, UK) and the bands excised with a scalpel and purified using a gel purification kit following the manufacturer's protocol (Qiagen, Crawley, UK).

2.2.13.3 DNA blunt ending

DNA was blunt ended by incubation at room temperature for 20 min of the following components;

25 µl PCR product
3 U Klenow DNA polymerase I
5 µl Klenow buffer
1 µl 2 mM dNTPS
18 µl dH₂O

Heat inactivation was performed by incubation at 70 °C for 10 min, followed by purification using Minelute PCR product purification kit (Qiagen, Crawley, UK).

2.2.13.4 DNA phosphorylation

DNA phosphorylation was performed by incubation at 37 °C for 30 min of the following components;

10 µl PCR product
3 µl kinase buffer
1 µl RATP
10 U T₄ polynucleotide kinase

Heat inactivation was performed by incubation at 70 °C for 10 min, followed by purification using Minelute PCR product purification kit (Qiagen, Crawley, UK).

2.2.13.5 DNA ligation

Ligation reactions were performed using approximately 100 ng of vector DNA and the formula:

$$\text{Amount of insert (ng)} = \frac{100 \times \text{size of insert (kb)}}{\text{size of vector (kb)}} \times \frac{3}{1}$$

Insert and vector were incubated at 16 °C for 16 hours with 1 µl 10 x T4 ligation buffer and 400 U T4 ligase (Promega, Madison, WI, USA) adjusted with dH₂O to a final volume of 10 µl.

2.2.13.6 Bacterial transformation

1 µl ligation reactions or 1 ng circular plasmids were transformed into chemically competent *E.coli* DH5α or BL21 (for protein expression). 100 µl aliquots of competent bacteria were thawed on ice and incubated with the plasmid for 30 min. Bacteria were 'heat shocked' at 42 °C for 90 sec and held on ice for 2 min, 1 ml LB was added and the mixture incubated at 37 °C for 1 hr shaking at 220 rpm. Centrifugation was performed at 3300 g for 3 min, the supernatant removed and the pellet resuspended in 100 µl LB and plated onto LB agar media containing the appropriate selective antibiotic (50 µg/ml ampicillin or 30 µg/ml kanamycin). The plates were incubated at 37 °C for 16 hr. Bacterial colonies were picked from the plates using a sterile toothpick and used to inoculate LB media. To screen recombinant plasmids, several colonies were picked and grown in LB media containing the appropriate antibiotic at 37 °C for 16 hr shaking at 220 rpm.

2.2.13.7 Storage and growth of bacteria

Both liquid bacterial cultures and bacteria plated onto LB agar were grown at 37 °C for approximately 16 hr. Long term storage was performed at –80 °C of a mixture of 800 µl of the overnight culture combined with 200 µl of 80% glycerol.

2.2.13.8 Preparation of plasmid DNA

2.2.13.8.1 Small scale

1.5 ml of overnight bacterial culture was centrifuged at 16 000g for 2 min at room temperature. The supernatant was discarded and the pellet used as per the manufacturer's protocol using a plasmid miniprep kit (Qiagen, Crawley, UK). Plasmid DNA was resuspended in 50 µl of dH₂O and stored at -20 °C.

2.2.13.8.2 Large scale

100 ml of overnight bacterial culture was centrifuged at 3300 g for 15 min at room temperature. The supernatant was discarded and the pellet used as per the manufacturer's protocol using a plasmid maxiprep kit (Qiagen Crawley, UK). Plasmid DNA was resuspended in 300 µl of dH₂O and stored at -20 °C. The DNA concentration was determined by quantifying a 100 fold dilution at a wavelength of 260 nm using the equation $1 \text{ OD}_{260} = 50 \text{ µg/ml}$.

All plasmids were confirmed by restriction enzyme digestion using the appropriate enzymes and direct DNA sequencing. Constructs produced using the cloning methods above are summarised in Appendix B.

2.2.14 Cell culture methods

2.2.14.1 Cell culture

Mouse and human cell lines were grown at 37 °C and 5% CO₂ in DMEM supplemented with 10% fetal calf serum and penicillin (50 U/ml) and streptomycin (50 µg/ml). Cells were maintained in flasks until 70% confluent and then split using the following procedure. Media was removed and the cells washed with 5 ml PBS to remove traces of serum. 2 ml trypsin was added and the cells incubated at 37 °C for 3 min. 5 ml medium was added and the cells pelleted by centrifugation at 135 g for 5 min. The supernatant was removed and the cells resuspended in fresh media and diluted 1/10. Media was changed every 3-4 days. When required for transfection studies, cells were seeded onto acid-treated glass coverslips at a density of 2×10^5 per 35 mm² dish and incubated as above for 24 hours prior to transfection.

2.2.14.2 Preparation of cell extracts

Cells were grown until confluent in 75 cm³ flasks and trypsinised as described above. The cell pellet was resuspended in 1 ml ice cold PBS/1mM PMSF. Centrifugation was performed at 800 g for 3 min at 4 °C. The cell pellet was resuspended in 200 µl ice cold PBS/1mM PMSF and incubated on ice for 30 min. The extract was then passed through a 27 g needle ten times. Centrifugation was performed at 16 000 g for 10 min at 4 °C. The supernatant was discarded and the pellet resuspended in 200 µl ice cold PBS/1mM PMSF. 200 µl of Laemmli buffer was added and the sample boiled for 10 min.

2.2.14.3 Transient transfection

Cells grown on coverslips were transfected using 1 µg plasmid DNA per 35 mm² dish and Lipofectamine reagent (Invitrogen, Paisley, UK), according to the manufacturer's recommendation. Cells were fixed in 2% paraformaldehyde, 0.1% glutaraldehyde for 30 min at room temperature 24 hr after transfection. Non-specific antibody cross reactivity was blocked by a 10 min pre-incubation in phosphate buffered saline (PBS) with 1% BSA. Antibodies were diluted to an appropriate dilution in 3% BSA in PBS. Cells were incubated at room temperature for 1 hr at 37 °C with the primary antibody. Following 3 x 5 min washes in 1 x PBS cells were incubated for one hr at 37 °C with the appropriate secondary antibody and hoescht (Invitrogen, Paisley, UK) at 1:500 dilution in 3% BSA in PBS. Following 3 x 5 min PBS washes, coverslips were mounted in 80% glycerol/3% n-propyl gallate (in PBS). Microscopy was performed using a Zeiss Axioskop 2 microscope (Carl Zeiss, Oberkochen, Germany) and images taken using Smartcapture software (Digital Scientific, Cambridge, UK) on an Apple computer (Apple, Cupertino, USA).

2.2.15 *In-vitro* translation of plasmids

Plasmid borne cDNA sequences downstream of a T7 promotor were *in-vitro* translated using the TNT® Quick Coupled transcription/translation system for T7. (Promega, Madison, WI, UK) 1 µg of plasmid was combined with 40 µl of TNT® master mix and 2 µl 35S methionine and the final volume adjusted to 50 µl with dH₂O. Plasmid borne cDNA sequences downstream of a T3 promotor were *in-vitro* translated using the TNT® T3 Rabbit reticulocyte lysate system according to the manufacturer's recommendations. Samples were incubated at 30 °C for 90 min and stored at -20 °C until required. Samples were run on an SDS-PAGE gel before exposing to autoradiography film and developing or exposing to a Typhoon cassette® overnight and reading on a Typhoon phospho-imager (Amersham Bioscience, Chalfont St.Giles, UK).

2.2.16 SDS-PAGE analysis of proteins

Protein samples were mixed with 2 x SDS-PAGE loading dye and heated at 96 °C for 5 min prior to gel electrophoresis. SDS polyacrylamide gels were poured in 10 x 7cm glass plates using the Protean III casting apparatus (BioRad, Hercules, CA, USA). Gels consisted of 2 sections, an upper 4% stacking gel and a lower section high percentage separating gel of variable concentration. The gel constituents are listed below.

	Stacking (4%)	Separating 7.5%	Separating 10%
dH ₂ O	1.5 ml	3 ml	1 ml
Protogel	325 µl	1.5 ml	2 ml
Upper/lower buffer	625 µl	1.5 ml	1.5 ml
10% APS	75 µl	75 µl	75 µl
TEMED	5 µl	5 µl	5 µl

Level setting was ensured by overlaying with isopropanol subsequently washed off with dH₂O. 20 µl of sample/loading dye mix was loaded and the gel electrophoresed at 150 V for 1 hr in 1 x SDS-PAGE running buffer. Gels were fixed using SDS-PAGE stain for 20 min with gentle agitation and washed over 30 min in SDS-PAGE destain. Gels were dried onto 3MM Whatman paper using a gel drying apparatus (BioRad, Hercules, CA, USA).

2.2.17 Western blot analysis

Protein gels were run as above but not fixed. Gels were soaked in transfer buffer for 5 min and a 10 x 6 cm piece of Protran® nitrocellulose membrane (Schleider and Schuell BioScience, Keene, NH, USA) soaked in transfer buffer for 10 min. The gel and membrane were sandwiched between 6 pieces of 3MM Whatman paper soaked in transfer buffer and run at 1 mAmp per cm² for 1 hr. The filter was washed in dH₂O and stained with Ponceau S for 1 min before washing again in dH₂O and allowing to dry. Nitrocellulose filters were blocked for 1 hr in 5% dried skimmed milk powder in 1 x PBS-0.1% Tween to prevent non-specific antibody cross reactivity. Antibodies were diluted in 5% dried skimmed milk powder in 1 x PBS-0.1% Tween. The filter was incubated with the primary antibody diluted to an appropriate concentration for 1 hr then 3 x 5 min washes in PBS-Tween were performed. The filter was incubated with the secondary antibody diluted to an appropriate concentration for 1 hr and 3 further washes in PBS-Tween were carried out over 30 min. The blot was developed using the ECL

plus Western blotting detection system (Amersham Biosciences, Chalfont St.Giles, UK) according to the manufacturer's protocol. The nitrocellulose membrane was exposed to autoradiography film and developed or exposed to a Typhoon ® cassette overnight and read on a Typhoon ® phospho-imager (Amersham Bioscience, Chalfont St.Giles, UK).

2.2.18 *In-situ* hybridisation

2.2.18.1 Probe generation

Primers were designed containing a T7 promotor sequence in either sense or antisense orientation in addition to 19 bases of *DYM* cDNA sequence probe generation using a T7 digoxigenin labelling kit (Roche, Basel, Switzerland). RT-PCR products were generated as previously described (see Section 2.2.10) using a primer containing a T7 promotor sequence in the forward direction (*1464FT7*) to generate a sense probe or in the reverse direction to generate an antisense probe (*1752RT7*). The complimentary primers (*1752R* and *1464F* respectively) contained no T7 promotor. An antisense β -*actin* probe was generated similarly and used as a positive control. Primer sequences are listed in Table 2.5. RT-PCR products were purified using Minelute PCR purification kit according to the manufacturer's recommendations (Qiagen Crawley, UK). Sense or antisense probes were generated by labelling with 11-dUTP digoxigenin using Dig RNA labelling kit (T7) according to the manufacturer's recommendations (Roche, Basel, Switzerland). The labelled probe was diluted in hybridisation buffer to a final concentration of 2 ng/ml.

2.2.18.2 Hybridisation

Tissue for use in situ-hybridisation studies was embedded in paraffin and sections cut at 4 μ n and mounted on Silane slides (Sigma-Aldrich, Cambridge, UK). Pretreatment of material was performed by immersion for 2 min in each the following solutions; xylene dewax, xylene, 99% IMS x 2, 95% IMS followed by immersion in DEPC dH₂O for 2 x 5 min.

Slides were placed in a humidified chamber and sections covered with the appropriate concentration of Proteinase K (2.5-10 μ g/ml) in 50 mM Tris-HCl, pH 7.6 and incubated at 37 °C for 1 hr. Slides were washed in DEPC dH₂O for 2 x 5 min. Prehybridisation was performed by incubation at 50 °C in 100 μ l prehybridisation solution for 1 hr. Prehybridisation solution was removed and 30 μ l hybridisation solution added and the section covered with a siliconised coverslip. Sections were incubated at 50 °C overnight. Coverslips were removed and slides washed at 50°C in 2 x SSC 50 % formamide preheated to 50 °C for 2 x 10 min. Slides were immersed in 3% BSA, 0.1%

Primer name	Primer sequence
1464FT7	5' CCAAGCTTCTAATACGACTCACTATAGGGAGACTGGTGGTAATAAGAACC 3'
1752RT7	5' GGCAGGAGTTGATGATCTTCTCCCTATAGTGAGTCGTATTAGAAGCTTGG 3'
1464F	5' GGAAATCATTAAAGCAAGGCG 3'
1752R	5' TGCCAAGTAGCACCAATTCTCC 3'
ActinF	5' CGTCGACAACGGCTCCGGC 3'
ActinRT7	5' GGTTGGCGCTCTTCTACTGTCTCCCTATAGTGAGTCGTATTAGAAGCTTGG 3'

Table 2.5. In-situ hybridisation probe primers

Triton in TBS for 5 min and then removed and covered with anti-digoxegenin alkaline phosphatase Fab conjugate (Roche, Basel, Switzerland) diluted 1 in 600 in TBS for 30 min at room temperature. Slides were rinsed in TBS pH 7.6 for 2 x 5 min and in dH₂O x 2. Slides were covered with alkaline phosphatase demonstration solution (1 NBT/BCIP tablet (Roche, Basel, Switzerland) dissolved in 10 ml dH₂O), covered with a siliconised coverslip and incubated in the dark until ready to be visualised by microscope. Slides were washed in running tap water and mounted in Aquamount (BDH, Poole, UK). Slides were visualised using a Zeiss microscope (Carl Zeiss, Oberkochen, Germany) and photographed using a Fujifilm CB300Z camera (Fujifilm, Beds, UK). Images were manipulated using Image Arithmetic version 2.5 software (Blackwell Publishing, Oxford, UK).

2.2.19 Yeast two-hybrid analysis

2.2.19.1 Growth and maintenance of yeast

The yeast host strain PJ69-4a was streaked onto YPD agar plates and incubated at 30 °C for 3-5 days (James *et al.*, 1996). Liquid yeast cultures were made by inoculating YPD media with a freshly streaked yeast colony and incubating at 30 °C shaking at 200 rpm for 16 hr. Yeast colonies were stored at 4 °C. Yeast transformed with plasmids were maintained on minimal media supplemented with 40 µg/ml uracil, 60µg/ml leucine, 20 µg/ml histidine and 40 µg/ml adenine. Uracil was omitted from media used to maintain yeast harbouring bait plasmid and uracil and leucine omitted from media used to maintain yeast harbouring both bait and prey plasmids. Yeast grown on minimal media agar were incubated at 30 °C for 5-7 days. Long term storage was performed at -80 °C of a mixture of 800 µl of the overnight culture combined with 200 µl of 80% glycerol.

2.2.19.2 Small scale yeast transformation

Yeast were transformed using the lithium acetate/ single stranded DNA/ polyethylene glycol method described by Agatep *et al* (Agatep *et al.*, 1998). Several colonies of yeast were picked with a sterile toothpick and inoculated into 5 ml liquid YPD and grown at 30 °C shaking at 220 rpm for 16 hr. The OD₆₀₀ of the culture was measured and 25 ml YPD inoculated with the overnight culture to give OD₆₀₀ of 0.25. Yeast was grown to an OD₆₀₀ of 0.7. The culture was centrifuged at 1000 g for 5 min at room temperature and the pellet resuspended in 12.5 ml sterile dH₂O. Centrifugation was performed at 1000 g for 5 min and the pellet resuspended in 0.5 ml 0.1 M lithium acetate. Centrifugation was performed for 30 sec at 16 000g and the pellet resuspended

in 200 µl 0.1 M LiAc .50 µl of the cell mixture was aliquotted for each transformation. The pellet was recentrifuged and the supernatant removed. To each pellet the following constituents were added in the order stated;

240 µl 50% PEG

36 µl 1 M LiAc

50 µl 2mg/ml sonicated boiled salmon sperm ssDNA

100 ng plasmid DNA

The cell pellet was resuspended by vortexing and the solution incubated at 30 °C for 30 min. The mixture was heat shocked at 42 °C for 30 min, centrifuged for 30 sec at 16 000 g and resuspended in 1 ml dH₂O. Transformed yeast were plated onto appropriate YPD agar selection media and grown at 30 °C for 3 days.

2.2.19.3 Large scale yeast transformation

100 ml culture of PJ69-4a previously transformed with bait plasmid was grown at 30°C overnight in –ura medium. The OD⁶⁰⁰ of the culture was measured and diluted in YPD to give an OD⁶⁰⁰ of 0.25 with a total volume of 300 ml. The culture was grown to an OD₆₀₀ of 0.7. The culture was centrifuged at 3000 g for 5 min, the supernatant removed and the pellet resuspended in 150 ml dH₂O. Centrifugation was repeated and the pellet resuspended in 6 ml 0.1 M LiAc and incubated for 15 min at 30°C. Centrifugation was repeated, the supernatant removed and the following added in the order stated;

14.4 ml 50% PEG

2.16 ml 1 M LiAc

3 ml salmon sperm DNA sonicated x 2 (boiled 5 min and kept on ice)

6 µg cDNA library

2.04 ml dH₂O

The mixture was vortexed vigorously until resuspended and incubated at 30°C for 30 min followed by heat shocking at 42°C for 60 min with inversion every 5 min. Centrifugation was performed at 3000 g for 5 min, the supernatant removed and the pellet resuspended in 12 ml dH₂O. 10 µl was removed and 240 µl dH₂O added. 10, 20 and 100 µl of this mixture were plated onto –ura/leu plates to calculate transformation efficiency and the remainder spread onto eight 23 cm x 23 cm plates (-ura/leu/his with 10 mM 3AT) and incubated at 30°C for 14 days.

2.2.19.4 Plasmid rescue

Yeast cells were grown in selective liquid media shaking at 220 rpm overnight at 30°C. 3 ml of culture was centrifuged at 600 g for 5 min and the supernatant removed. Plasmids were rescued from the yeast using a Zymoprep yeast plasmid rescue kit according to the manufacturer's instructions (Amersham Bioscience, Chalfont St.Giles, UK). Rescued plasmid samples were resuspended in 13 µl dH₂O and 1 µl used to transform *E.coli* DH5α.

2.2.20 MBP pulldown

2.2.20.1 MBP protein expression and purification

E. coli BL21 cells were transformed with pMALC2G expression plasmid or the engineered constructs pMALDMC1-98 or pMALDMC449-669. A single colony was inoculated into 5 ml LUB supplemented with ampicillin (50 µg/ml) and grown at 37 °C shaking 220 rpm for 16 hr. The overnight culture was diluted 1/20 in an appropriate volume of LUB with 0.2% glucose and ampicillin (50 µg/ml) and grown at 30 °C shaking 220 rpm for 2 hr. IPTG to a final concentration of 0.2 mM was added and the culture grown at 30 °C shaking 220 rpm for a further 2 hr. Centrifugation was performed at 3300 g for 15 min at 4 °C and the pellet resuspended in 2 ml cold NETN buffer and sonicated at 12 microns for 4 x 15 sec. The sonicated culture was centrifuged at 15 000 g for 10 min at 4 °C. The supernatant was removed and 40 µl 50 % amylose beads, (prepared following the manufacturer's recommendations) added. The solution was incubated for 1 hr at 4 °C with rotation and washed x 3 with 2 ml cold NETN buffer, centrifuging at 500 g for 5 min at 4 °C between washes.

2.2.20.2 MBP pulldown assay

The pelleted amylose beads coupled to MBP proteins were mixed with *in-vitro* translated 35S labelled protein. NETN buffer was added to give a total volume of 750 µl and the solution incubated on a rotating mixer for 90 min at 4 °C. Following incubation, the beads were washed x 3 with 1 ml cold NETN buffer centrifuging at 500 g for 5 min at 4 °C between washes. The supernatant was removed and the beads resuspended in 10 µl Laemmli buffer, boiled for 5 min and loaded onto a protein gel. The gel was stained and dried and exposed to a Typhoon cassette® overnight and read on a Typhoon phospho-imager (Amersham Bioscience, Chalfont St.Giles, UK).

2.2.21 Database analyses

GeneJockey software (Biosoft, Cambridge, UK) on an Apple G4 computer (Apple, Cupertino, USA) was used to analyse DNA sequence to identify restriction enzyme digestion sites and to translate DNA sequence into amino acid sequence. DNA and protein sequences were retrieved using databases available online from National Centre for Biotechnology information (NCBI, www.ncbi.nlm.gov). Sequence comparison was carried out using the basic local alignment and search tool (BLAST) nucleotide and protein servers also available at NCBI. Protein prediction tools used were available online from the Expert Protein Analysis System proteomics server of the Swiss Institute of Bioinformatics (www.expasy.org).

Chapter 3 - Identification and characterisation of mutant *DYM* alleles

3.1 Introduction

Preliminary work leading up to this study mapped the DMC locus to chromosome 18q12-21 using an autozygosity mapping approach (Thauvin-Robinet *et al.*, 2002). Simultaneously, a further group mapped both DMC and SMC to the same locus, providing support for the previous suggestion that DMC and SMC are allelic disorders (Ehtesham *et al.*, 2002). A schematic of these marker data is shown in Figure 3.1. A positional cloning approach was adopted to identify the DMC gene within the linked region. Direct sequencing of the coding regions and exon-intron boundaries of the eight known genes within the region failed to detect any deleterious sequence variants, however, mutations were identified in affected individuals within the predicted transcript known as *FLJ20071*. Initial sequence analysis suggested a gene composed of 9 exons over 190 Kb of genomic sequence. However by BLAST analysis of predicted exons against EST databases, additional expressed sequences were identified. The additional ESTs were aligned to genomic sequence and this resulted in the identification of a total of 17 exons over 250 Kb of genomic sequence as illustrated in Figure 3.2 (El Ghouzzi *et al.*, 2003). A total of 23 different mutations in the DMC gene have been reported, including four novel and one previously reported mutation occurring within four families included in this study, as previously summarised in Table 1.3 (Cohn *et al.*, 2003, El Ghouzzi *et al.*, 2003, Neumann *et al.*, 2006, Paupe *et al.*, 2004, Girisha *et al.*, 2007). The majority (13/18) of previously reported mutations predict a truncated protein product hence it has been suggested that DMC is a loss of function phenotype resulting from a reduction in available biologically active protein.

In a number of inherited disorders, the identification of mutations enables correlation to be made between genotype and phenotype. However, in disorders as rare as DMC and SMC, this may prove problematic as only a small number of mutations are identified and reported. The identification of missense mutations may allow identification of critical domains exposing discrete functions for a protein. To date, only one missense mutation has been identified in DMC and another in SMC, and even these are observed as compound heterozygote in association with truncating mutations thus limiting insight gained from clinical correlation (Cohn *et al.*, 2003).

In order to recruit DMC individuals for this study, contact was made with Clinical Geneticists in the UK, Europe, and the United Arab Emirates in addition to the national radiological expert on skeletal dysplasias (Professor Christine Hall, Great Ormond Street Hospital). This resulted in the recruitment of eight families with ten DMC

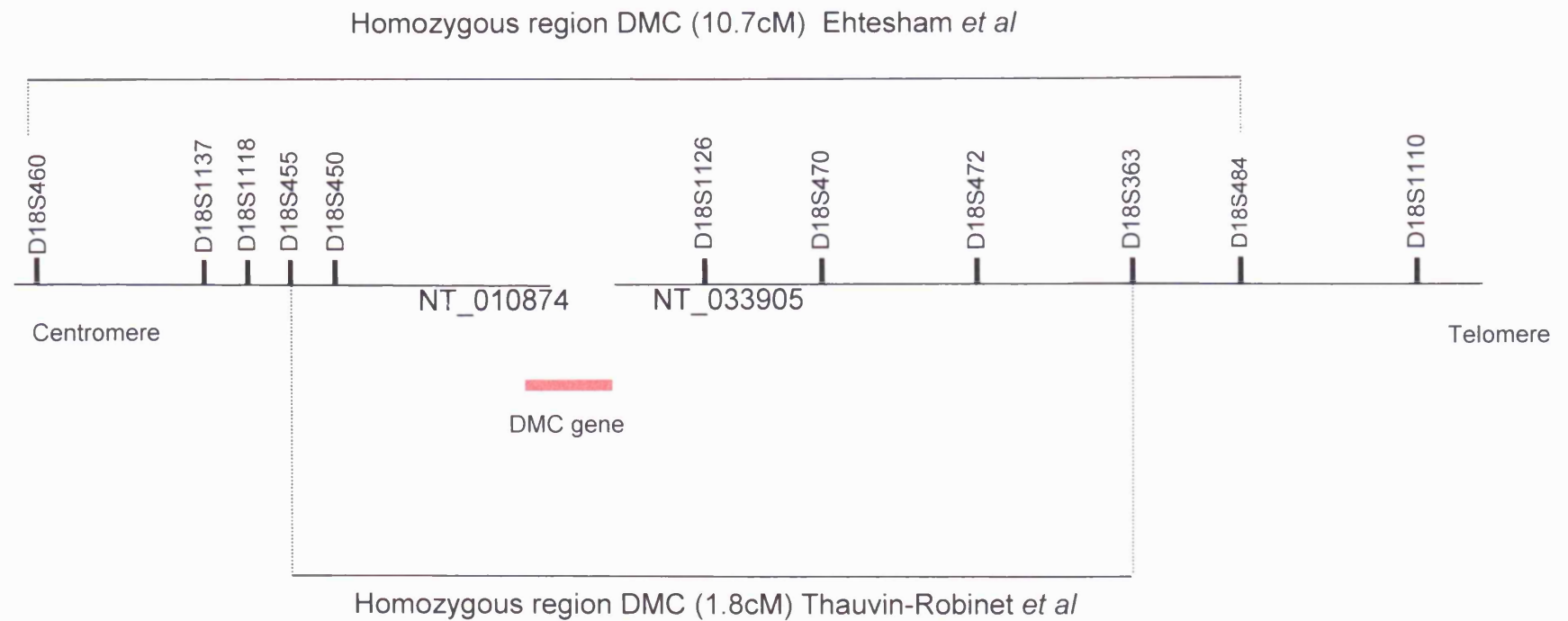


Figure 3.1. Schematic illustrating chromosomal position of microsatellite markers and DMC gene. Overlapping regions of homozygosity identified by autozygosity mapping studies by Ehtesham *et al* and Thauvin Robinet *et al* are also shown.

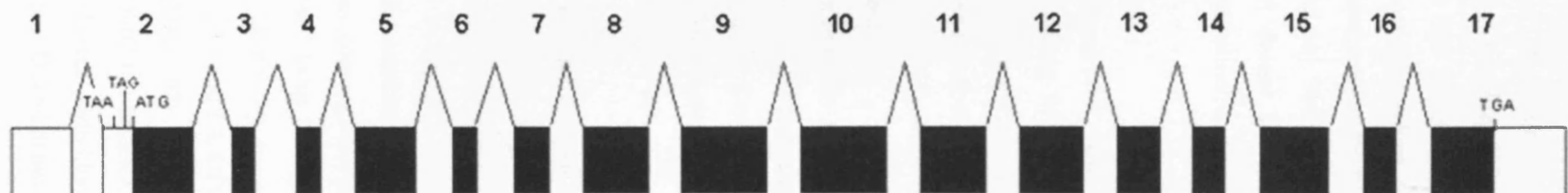


Figure 3.2. Schematic illustrating the structure of the *DYM* gene which is composed of 17 exons, shown as black blocks drawn approximately to scale. Introns are indicated by lines connecting exons and are not drawn to scale. The ATG initiator codon is located 53 bp downstream of the beginning of exon 2 flanked by a Kozak consensus sequence and the TGA termination codon is located in exon 17. Adapted from El Ghouzzi *et al.*

affected individuals. The initial method of mutation detection employed was direct sequencing of coding exons and intron-exon boundaries rather than a screening method such as denaturing high performance liquid chromatography (DHPLC) which may be more appropriate in a large cohort. Direct sequencing revealed mutations in two of the eight probands. Subsequent analysis included cDNA amplification in order to look for changes at an RNA level. The finding of exon duplication/repetition in cDNA led to the use of fluorescent dosage PCR and Southern blot analysis in order to further characterise the duplication/repetition events.

3.2 Family resource

Ten affected individuals from eight independently ascertained pedigrees were recruited into the study following Multi-centre Research Ethical Committee approval (see Appendix A). The pedigrees are illustrated in Figures 3.3 and 3.4, with four of the families being consanguineous (first cousin parents in families DMC 01, 02 and 03 and 2nd cousin parents in family DMC 04). All patients were diagnosed with DMC/SMC following examination by a Clinical Geneticist and had characteristic findings on skeletal survey, confirmed by Professor Christine Hall where possible. Clinical features of the 10 patients who all had characteristic clinical and radiological findings are summarised in Table 3.1. Four families (DMC 02, 03, 06 and 08) were visited by the researcher who performed clinical examination on affected individuals and reviewed hospital notes where available.

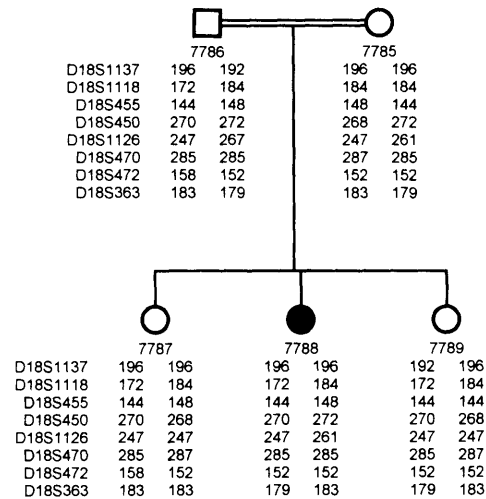
3.3 Microsatellite marker analysis

All family members from whom DNA samples were obtained were typed for microsatellite markers surrounding the DMC locus on chromosome 18q12-21, using markers illustrated in Figure 3.1 with *DYM* being located between markers D18S450 and D18S1126. Microsatellite marker analysis was performed based on the assumption that offspring from consanguineous pedigrees would be homozygous for markers surrounding the DMC locus having inherited two copies of the mutant allele from a common ancestor (identity by descent). Results of microsatellite marker analysis are illustrated in Figures 3.3 and 3.4. Of the four consanguineous families, only the affected individual 7345 (DMC 02) is homozygous for the entire region. DMC affected individual 7788 (DMC 01) is heterozygous for the *DYM* flanking markers D18S450 and D18S1126, whereas individuals 7175 (DMC 04) and 7758 (DMC 03) are homozygous at marker D18S450 but heterozygous at D18S1126.

DNA number	7170	7175	7263	7264	7329	7332	7345	7758	7788	7815
Clinical Features	DMC 05	DMC 04	DMC 06	DMC 06	DMC 07	DMC 07	DMC 02	DMC 03	DMC 01	DMC 08
Ethnic origin	Lebanese	Palestinian	Gujerati	Gujerati	Italian	Italian	Lebanese	Pakistani	Tamil	British
Consanguinity	N	1 st cousins	N	N	N	N	1 st cousins	1 st cousins	2 nd cousins	N
Height (cm)	102	100	100	93	134	101	137	124	87	103
OFC (cm)	Unknown	Unknown	48	47.5	52	47.6	54	45	46	49.5
Age (yr)	8	11	12	5	22	12	18	22	4	17
Retardation	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
Dysmorphism	N	N	Y	Y	Y	Y	Y	Y	N	Y
Short trunk	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Barrel chest	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Rhizomelia	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Scoliosis	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Vertebral platyspondyly	Y	Y	Y	Y	Y	Y	Y	Unknown	Y	Y
Vertebral notching	Y	Y	Y	Y	Y	Y	Y	Unknown	Y	N
Lacy iliac crests	Y	Y	Y	Y	Y	Y	Y	Unknown	Y	Y
Metaphyseal dysplasia	Y	Y	Y	Y	Y	Y	Unknown	Unknown	Y	Y
Epiphyseal dysplasia	Y	Y	Y	Y	Y	Y	Unknown	Unknown	Y	Y
Odontoid hypoplasia	Y	Y	Y	Y	Y	Y	Y	Unknown	Unknown	Y
White cell enzymes	Not done	Not done	Normal	Normal	Normal	Normal	Unknown	Unknown	Not done	Normal
Urinary GAGS	Normal	Normal	Normal	Normal	Normal	Normal	Unknown	Unknown	Normal	Normal
Mutation identified	Y	N	Y	Y	N	N	Y	N	Y	N
Nucleotide change	IVS 11 c.1252-1G>A		Exon 14 repetition	Exon 14 repetition			Exon 2 duplication		c.208C>T c.1363C>T	

Table 3.1. Clinical features of study subjects.

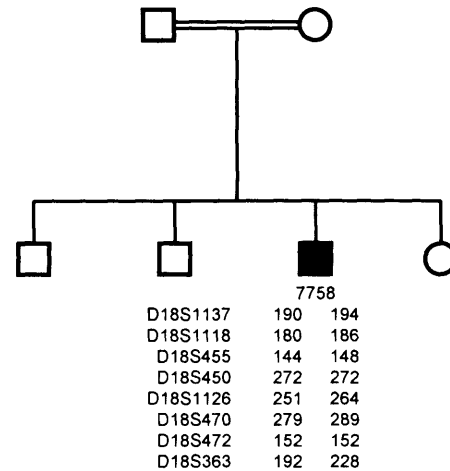
DMC 01



DMC 02



DMC 03



DMC 04

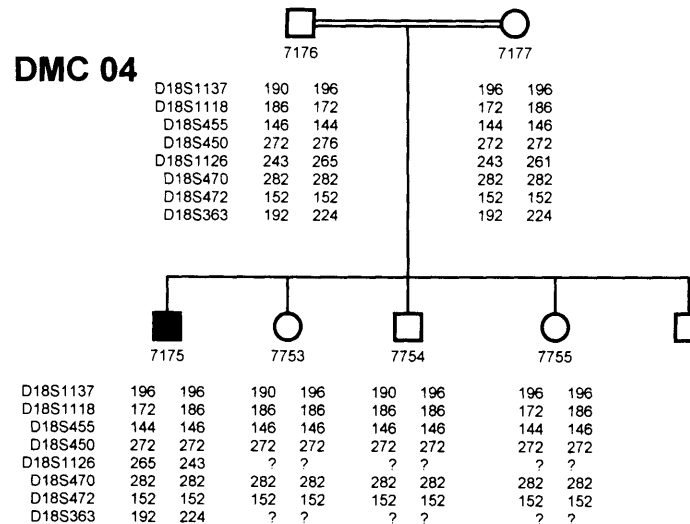


Figure 3.3. DMC consanguineous pedigrees and microsatellite marker results. Families are referred to by a number commencing with DMC and individuals assigned a 4 digit number.

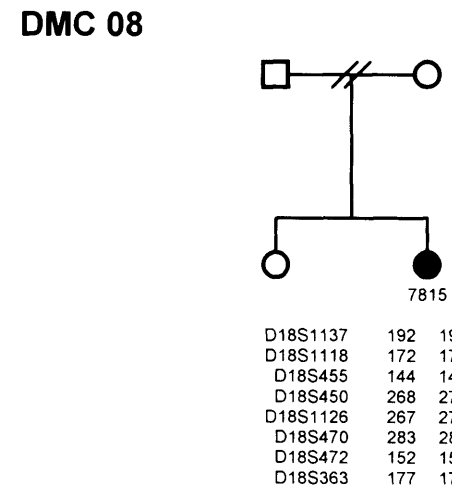
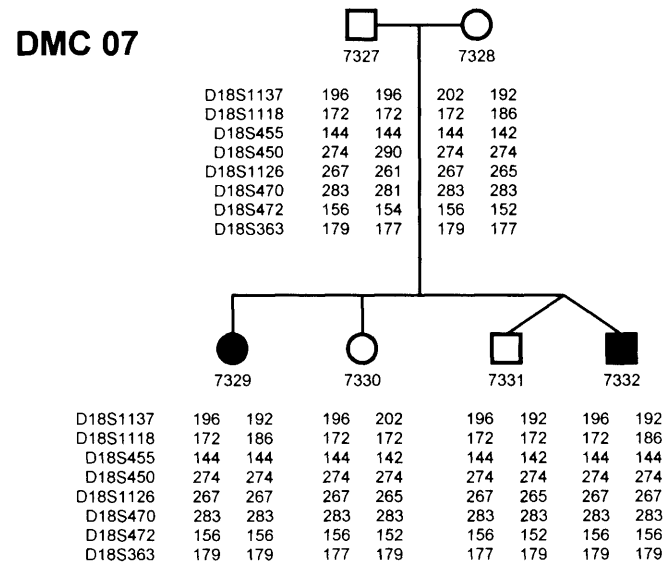
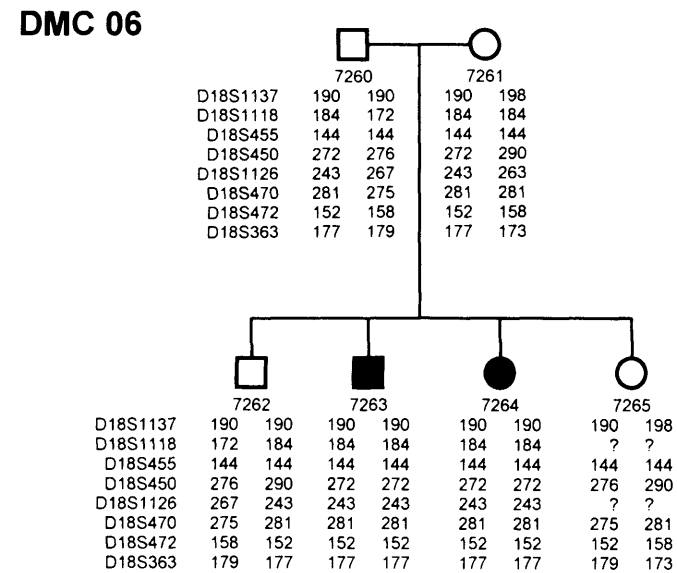
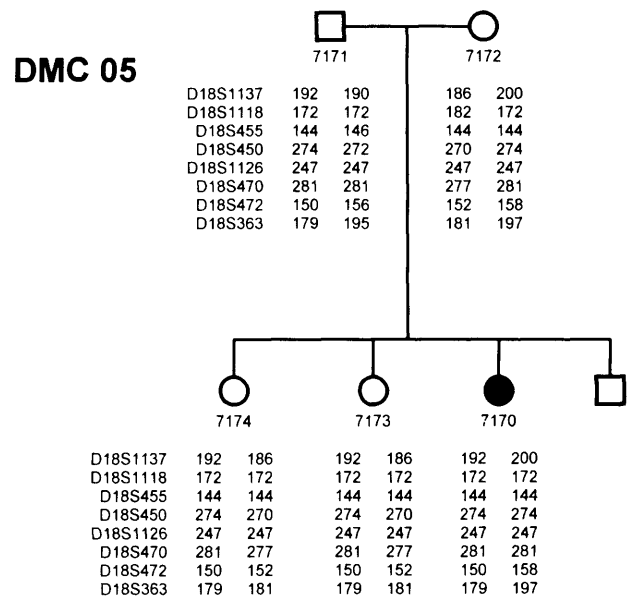


Figure 3.4. DMC non-consanguineous pedigrees and microsatellite marker results. Families are referred to by a number commencing with DMC and individuals assigned a 4 digit number.

The affected siblings 7329 and 7332 (DMC 07) are homozygous for markers D18S455 to D18S363. Their parents are said to originate from the same small village hence may have a common ancestor from whom they have both inherited the same mutant *DYM* allele. The affected siblings 7263 and 7264 (DMC 06) are homozygous for all markers in the region but have parents who are not known to be consanguineous. The affected individual 7170 (DMC 05) is homozygous for markers D18S1118 to D18S470 and also has parents who originate from the same village and religious sect. The subsequent finding of homozygous mutations in siblings 7263 and 7264 and individual 7170 support the hypothesis that two copies of the same mutant allele may have been inherited from a common ancestor.

Family studies did not identify any common disease bearing haplotypes but all families were of different ethnicity except families DMC 02 and 05 both of Lebanese origin.

3.4 Identification of *DYM* mutations by direct sequencing

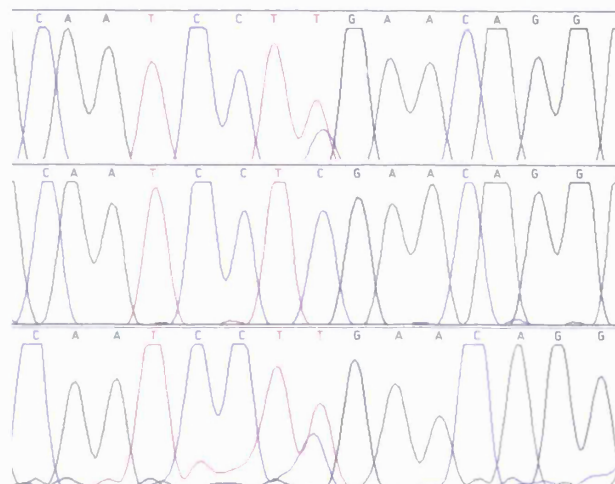
The *DYM* gene comprises 17 exons of which 16 are coding as illustrated in Figure 3.2. Initially, a proband from each of the eight affected families was sequenced for all 16 coding exons and exon-intron boundaries of *DYM* using the primers and PCR conditions previously listed in Table 2.2. Direct sequencing of the *DYM* gene revealed mutations in two affected individuals from two independent pedigrees. The proband in family DMC 05 (7170) was found to be homozygous for the splice site mutation IVS 11 c.1252-1G>A. The affected individual in family DMC 04 (7788) was found to be compound heterozygote for the truncating mutations c.208C>T (R70X) and c.1363C>T (R455X) in keeping with the results of heterozygosity for microsatellite markers. The nucleotide changes and predicted protein alterations produced are summarised in Table 3.2. Following identification of mutations in the proband, sequencing was performed on parents and other affected individuals to confirm segregation within the family. Chromatograms illustrating the mutations in family DMC 04 are shown in Figure 3.5.

3.5 RT-PCR of patient cDNA

Direct sequencing of coding exons and exon-intron boundaries may not detect abnormalities of genomic DNA such as deletions, duplications and rearrangements. In order to look for abnormalities of cDNA which may reflect changes of genomic DNA which would not have been identified by sequencing, *DYM* cDNA from patient samples was amplified in nine overlapping fragments as described in section 2.2.10 using

Family and DNA number	Nucleotide change	Affected exon	Protein change
DMC 05 (7170)	Homozygous IVS 11 c.1252-1G >A	Exon 12 (5' end) splice acceptor	
DMC 04 (7788)	c.208C>T	4	R70X
	c.1363C>T	14	R455X

Table 3.2. *DYM* Mutations identified in study subjects by direct sequencing.

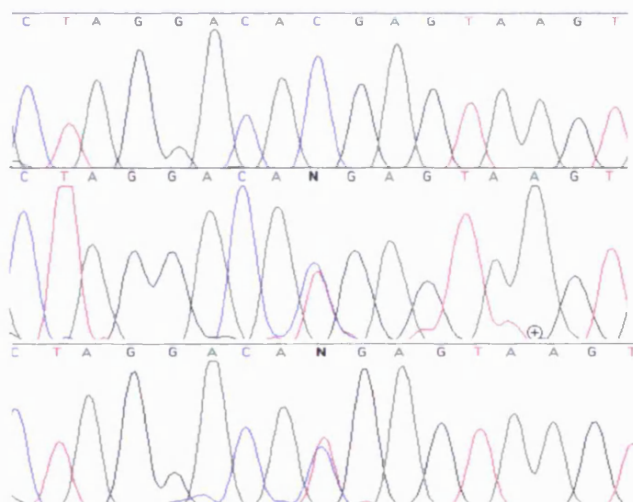


Mother

Father

Proband (7788)

c.208C>T
R70X



Mother

Father

Proband (7788)

c.1363C>T
R455X

Figure 3.5. Chromatograms from family DMC 04 demonstrating segregation of 2 truncating mutations through the family. The mother is heterozygous for the mutation c.208C>T leading to a premature stop codon (R70X). The father is heterozygous for the mutation c.1363C>T leading to a premature stop codon (R455X). Both mutations are seen in the proband 7788.

primers listed in Table 2.4. Expected RT-PCR fragment sizes using these primers are listed in Table 3.3. Aberrantly sized bands were sequenced as described in section 2.2.9. RT-PCR was undertaken on the affected individual from DMC 05 (7170) in order to further characterise the splice site mutation IVS11 c.1252-1G>A identified on sequencing of genomic DNA. Amplification of fragment DYM 6 using primers *FLJ3F* and *FLJ4R* which should produce a product of 409 bp (see Table 3.3), resulted in a faint, marginally smaller band (compared to that produced from control cDNA). Following gel purification, sequencing revealed the splice site mutation to result in the skipping of the first 28 bases of exon 12 and pick up at a cryptic splice site (AG..GCAGCT) within exon 12 (see Figure 3.6). This is predicted to result in a frame shift and a termination codon after three aberrant amino acids that would result in an open reading frame encoding a protein of 420 amino acids.

RT-PCR of cDNA from the two affected siblings (7263 and 7264) from family DMC 06 revealed the product produced with primers *FLJ 1458F* and *FLJ 2034R* (DYM 7, see Table 3.3) to be aberrantly sized with an approximate size of 1000 bp compared to the expected size of 594 bp seen in controls (see Figure 3.7a). Sequencing of this band revealed four copies of exon 14 in series consistent with the size difference detected by gel electrophoresis. This predicts the amino acid sequence following the first copy of exon 14 to consist of an aberrant six amino acids followed by a termination codon as shown in Figure 3.7b.

RT-PCR in the DMC 02 proband (7345) demonstrated the RT-PCR product resulting from amplification using primers *5'UTR F* and *FLJ 556R* to be of an aberrant size of approximately 800 bp compared an expected fragment of 574 bp seen in controls (see Figure 3.7c). Sequencing of this band revealed a duplication consisting of nucleotide 90 of the 5'UTR to the end of exon 2, consistent with the size difference detected by gel electrophoresis. This predicts the amino acid sequence following the first copy of exon 2 to consist of an aberrant 22 amino acids followed by a termination codon as shown in Figure 3.7b. Hence duplication/repetition in both cases is predicted to result in a truncated protein.

To confirm these findings, RT-PCR was performed using primer pairs (*Int Ex2F* and *Int Ex2R*, *Int Ex14F* and *IntEx14R*) each predicted to generate a product specific to the aberrant repeated cDNA sequences but not in the normal sequence (see Figure 3.7b and Tables 3.3 and 2.4). Bands of the sizes expected (122 bp in 7345; 79, 182 and 285 bp in 7263 and 7264) were seen in affected individuals and parents from whom RNA was available (see Figure 3.7b) and were not detected in unaffected control cDNA. Positive

cDNA fragment name	cDNA sequence amplified	expected product size (base pairs)
DYM 1	2-576	574
DYM 2	380-623	243
DYM 3	522-893	371
DYM 4	756-1164	408
DYM 5	1059-1595	536
DYM 6	1186-1595	409
DYM 7	1458-2052	594
DYM 8	1924-2327	403
DYM Int Ex2	N/A	N/A
DYM Int Ex14	N/A	N/A
BMPR2	2227-2417	191

Table 3.3. *DYM* cDNA fragments amplified by RT-PCR

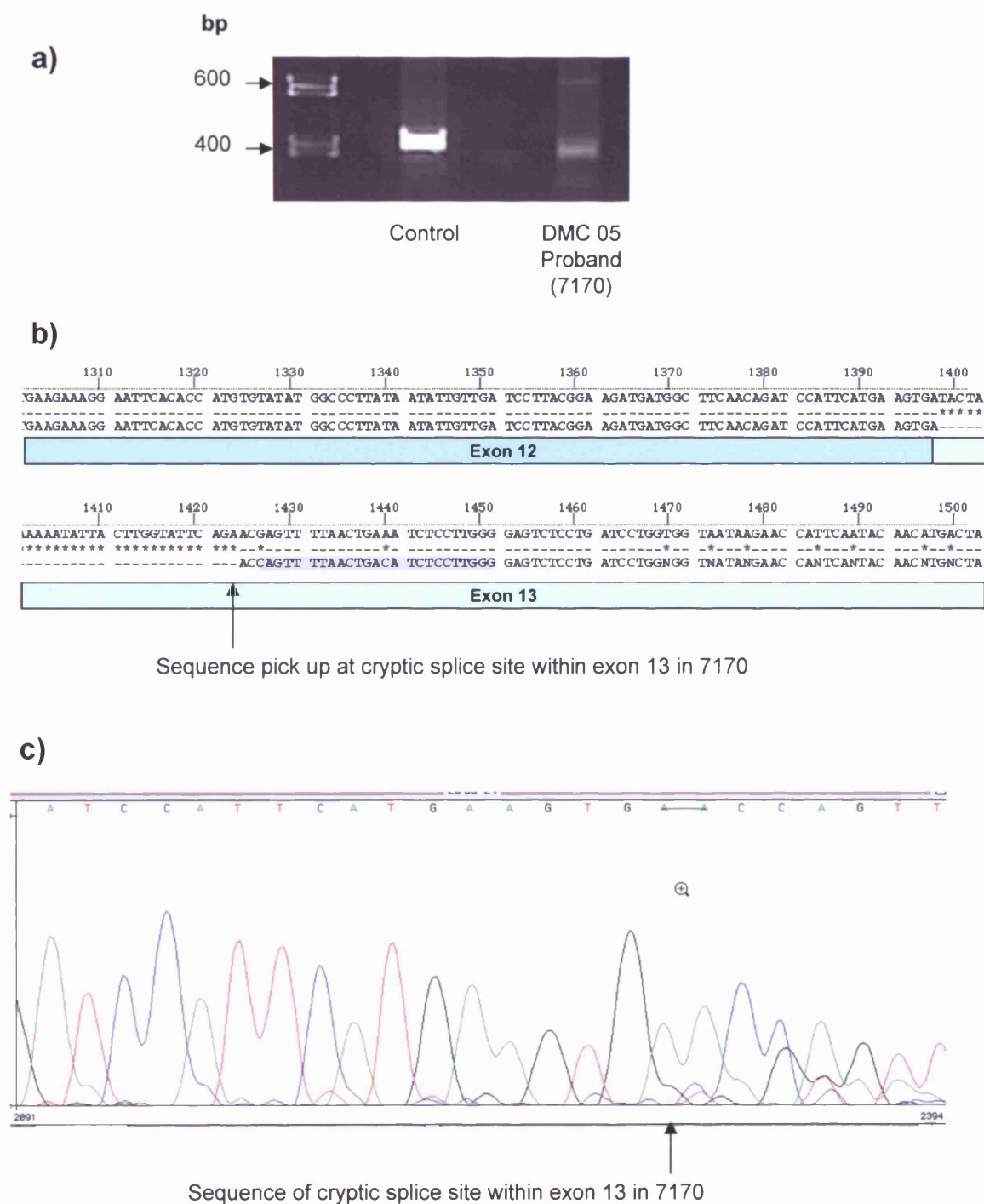


Figure 3.6.

- RT-PCR using primers *FLJ3F* and *FLJ4R* . Faint, slightly smaller band seen in 7170 compared to control.
- 7170 sequence aligned with *DYM* cDNA showing sequence of cryptic splice site within exon 13.
- Chromatogram of 7170 sequence showing sequence of cryptic splice site within exon 13.

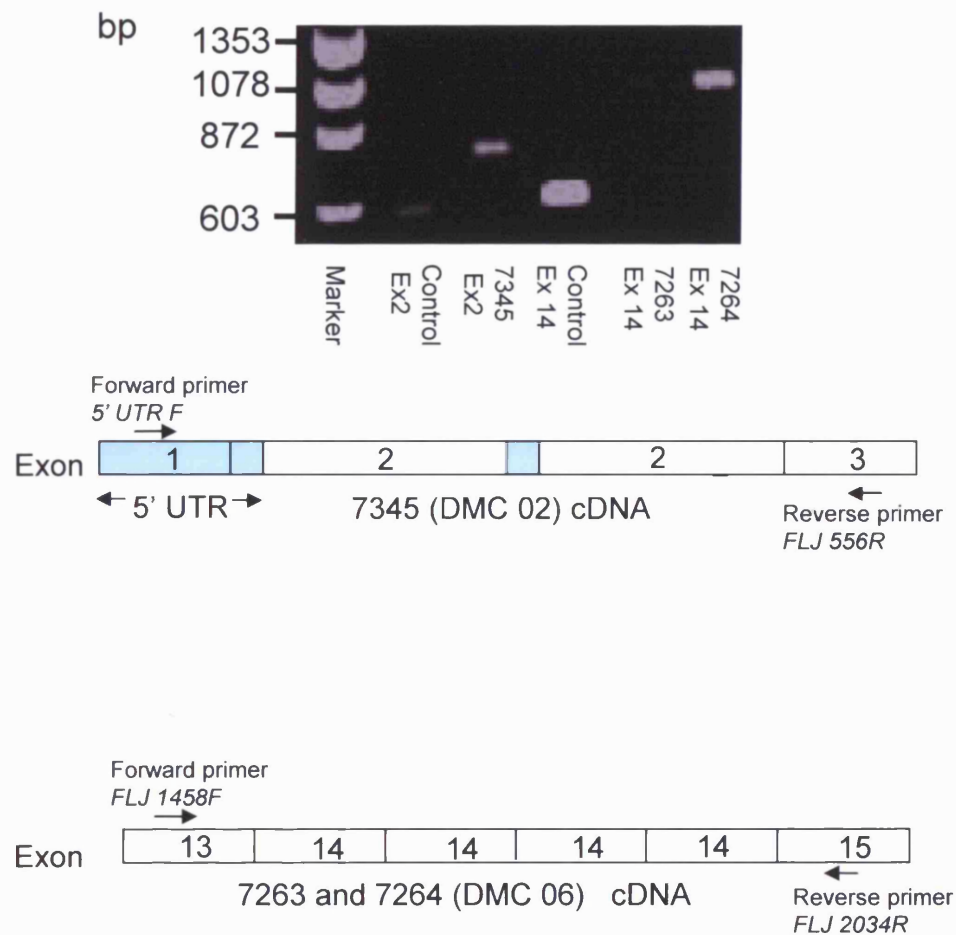


Figure 3.7a. RT-PCR of exons 1-3 using a forward primer in exon 1 and a reverse primer in exon 3 gives a product size of 574 bp in unaffected control compared to ~800 bp in 7345 as seen on the gel image. RT-PCR of exons 13-15 using a forward primer in exon 13 and a reverse primer in exon 15 gives a product size of 594 bp in control compared to ~1000 bp in 7263 and 7264 also seen in the gel image. schematic shows position of primers and aberrant cDNA structure in affected individuals, as revealed by sequencing.

i)

AGT TTA TTT TCT TTG CTG TCT AAA AAA CAC AAC AAA GTT CTG GAA CAA GCC ACA
Ser Leu Phe Ser Leu Leu Ser Lys Lys His Asn Lys Val Leu Glu Gln Ala Thr
CAG TCC TTG AGA GGT TCG CTG AGT TCT AAT GAT GTT CCT CTA CCA GAT TAT GCA
Gln Ser Leu Arg Gly Ser Leu Ser Ser Asn Asp Val Pro Leu Pro Asp Tyr Ala
CAA GAC CTA AAT GTC ATT GAA GAA GTG ATT CGA ATG ATG TTA GAG ATC ATC AAC
Gln Asp Leu Asn Val Ile Glu Glu Val Ile Arg Met Met Leu Glu Ile Ile Asn

ii)

AGT TTA TTT TCT TTG CTG TCT AAA AAA CAC AAC AAA GTT CTG GAA CAA GCC ACA
Ser Leu Phe Ser Leu Leu Ser Lys Lys His Asn Lys Val Leu Glu Gln Ala Thr
CAG TCC TTG AGA GGT TCG CTG AGT TCT AAT GAT GTT CCT CTA CCA GAT TAT TTT
Gln Ser Leu Arg Gly Ser Leu Ser Ser Asn Asp Val Pro Leu Pro Asp Tyr Phe
ATT TTC TTT GCT GTC TAA
Ile Phe Phe Ala Val Stop

iii)

GTG TCT TTT TGG AAA TGC AGG TTT AAG GAC AAA ATT ATC TGC TTA AGC TAG AAG
ATG GGA TCG AAT AGC AGC AGA ATC GGC GAT CTT CCT AAA AAT GAG TAC TTG AAA
Met Gly Ser Asn Ser Ser Arg Ile Gly Asp Leu Pro Lys Asn Glu Tyr Leu Lys
AAG TTA TCA GGC ACG GAA TCT ATC TCT GAG AAT GAC CCG TTC TGG AAT CAG CTT
Lys Leu Ser Gly Thr Glu Ser Ile Ser Glu Asn Asp Pro Phe Trp Asn Gln Leu
CTC TCA TTT TCT TTC CCT GCA CCA ACT AGC AGT AGT GAG TTG AAA CTC TTG GAG
Leu Ser Phe Ser Phe Pro Ala Pro Thr Ser Ser Ser Glu Leu Lys Leu Leu Glu

iv)

GTG TCT TTT TGG AAA TGC AGG TTT AAG GAC AAA ATT ATC TGC TTA AGC TAG AAG
ATG GGA TCG AAT AGC AGC AGA ATC GGC GAT CTT CCT AAA AAT GAG TAC TTG AAA
Met Gly Ser Asn Ser Ser Arg Ile Gly Asp Leu Pro Lys Asn Glu Tyr Leu Lys
AAG TTA TCA GGC ACG GAA TCT ATC TCT GAG AAT GAC CCG TTC TGG AAT CAG CTT
Lys Leu Ser Gly Thr Glu Ser Ile Ser Glu Asn Asp Pro Phe Trp Asn Gln Leu
CTC TCA TTT TCT TTC CCT GCA CCA ACT AGC AGC TCC GCA CCG ACG CGG CCC GGG
Leu Ser Phe Ser Phe Pro Ala Pro Thr Ser Ser Ser Ala Pro Thr Arg Pro Gly
CTG GAG CCG AGC CGG GGC CGA GCT GCA GGC CGG ACC GGA GCC GGA TGA
Leu Glu Pro Ser Arg Gly Arg Ala Ala Gly Arg Thr Gly Ala Gly Stop

Figure 3.7b. (i) Wild type exon 14 cDNA sequence is shown in bold with flanking exon 13 and exon 15 cDNA sequence and translated amino acid sequence below. **(ii)** Mutant cDNA sequence in patients 7263 and 7264 with 2nd copy of exon 14 shown in italics is predicted to encode 6 aberrant amino acids followed by a stop codon. **(iii)** Wild type coding exon 2 sequence is shown in bold with flanking non-coding exon 2 and exon 3 sequence. Translated amino acid sequence is shown below. **(iv)** Mutant cDNA sequence in patient 7345 with 2nd copy of 5'UTR from nucleotide 90 shown in italics which is predicted to encode 22 aberrant amino acids followed by a stop codon.

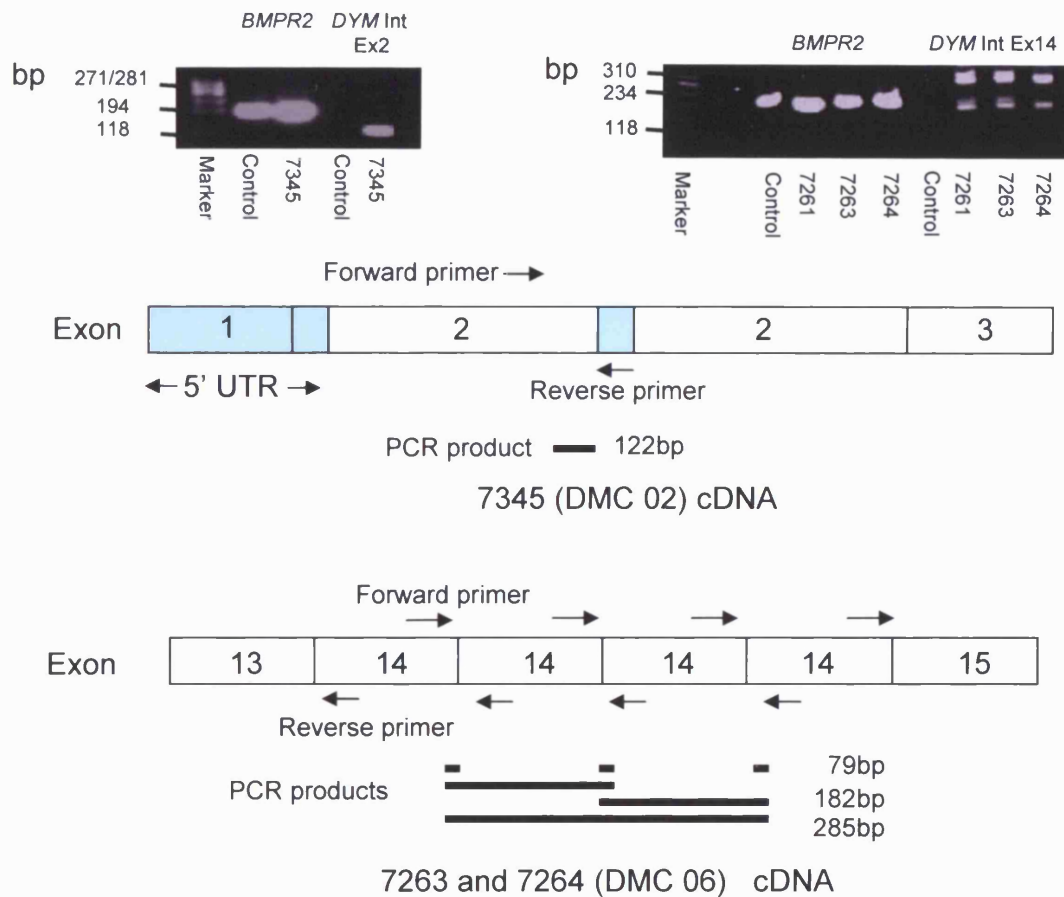


Figure 3.7c. RT-PCR using a forward primer in *DYM* Exon 2 and a reverse primer in the 5' UTR does not produce a band in the unaffected control but produces a band of the size expected (122 bp) in 7345. RT-PCR using a forward primer at the 3' end of exon 14 and a reverse primer at the 5' end of exon 14 does not produce a band in the unaffected control but produces bands of the sizes expected (79, 182 and 285 bp) in 7263 and 7264. Positive control RT-PCR amplification of a 191 bp product of *BMPR2* (chromosome 2q36-37) revealed an identical product in all samples tested.

control RT-PCR amplification of a 191 bp product of *BMPR2* (chromosome 2q36-37) was performed and revealed an identical product in all samples tested.

3.6 Fluorescent dosage PCR

Fluorescent dosage PCR is a method of comparing exon copy number by the quantitative amplification of exons of interest with fluorescently labelled primers. Primers are designed such that PCR product sizes for different exons vary in size and can be separated on acrylamide gel. After separation by size, the amount of fluorescence for each PCR product is measured by laser and can be compared within a sample for several different exons. The ratio of fluorescence levels for different exons can be calculated in individuals and compared in order to compare exon copy number. Fluorescent dosage PCR was performed in order to confirm that the duplications identified in cDNA from patients 7345, 7263 and 7264 were present at a genomic level. *DYM* exons 2, 14 and 15 were amplified in addition to *ALK1* exon 2 (control exon) for genomic DNA from the three patients and for ten control samples using the fluorescently labelled forward primers *Dos Ex2F*, *Dos Ex14F*, *Dos Ex15F*, *Dos Alk1 6F* and the reverse primers *DMC Ex 2R*, *DMC Ex14R*, *Dos Ex15R* and *Dos Alk1 6F* respectively (see Tables 2.1 and 2.2). Examples of dosage plots obtained are shown in Figure 3.8b and the dosage quotients calculated for patient samples (as described in Section 2.2.11) are shown in Figure 3.8a. If the copy number of each individual exon is the same, a dosage quotient of 1 would be expected. The two affected siblings (7263 and 7264) from family DMC 06 had dosage quotients of approximately 4 for *DYM* exon 14 (compared with *DYM* exons 2 and 15 and *ALK1* exon 6) when exon 14 was used as the numerator and approximately 0.25 when used as the denominator. This is consistent with previous findings on RT-PCR and sequencing revealing the presence of 4 copies of exon 14. The proband (7345) from family DMC 02 had a dosage quotient of approximately 2 for *DYM* exon 2 when exon 2 was used as the numerator and approximately 0.5 when used as the denominator, again consistent with previous findings of a duplication of exon 2.

3.7 Southern blot analysis

Southern blot analysis was carried out in order to provide additional confirmation of the genomic duplication but also to begin to initiate mapping of the extent of the duplication. *DYM* exon 2 probe corresponds to 298 base pairs (bp) of *DYM* cDNA including all of exon 2 and the *DYM* exon 14 probe corresponds to 1040 bp of *DYM* genomic DNA including all of exon 14 as illustrated in Figure 3.9a. Probes were

a)

Peak area			Dosage quotient			
			7345			
Exon	7345	control	DYM2	DYM14	DYM15	ALK16
DYM2	9082	21308		0.39	0.53	0.45
DYM14	716	4318	2.2		0.73	1.14
DYM15	5204	22960	1.9	0.79		0.85
ALK16	2289	12084	2.3	0.93	0.84	
			7263			
DYM2	7263	control	DYM2	DYM14	DYM15	ALK16
DYM2	808	7433		4	1.01	0.88
DYM14	713	1617	0.25		0.25	0.22
DYM15	1104	9966	0.97	4		0.87
ALK16	452	4708	1.14	4.65	1.16	
			7264			
DYM2	7264	control	DYM2	DYM14	DYM15	ALK16
DYM2	2870	7433		3.68	0.94	0.95
DYM14	2342	1617	0.26		0.25	0.25
DYM15	3639	9966	1.05	4		1.02
ALK16	1736	4708	1.05	3.97	1	

b)

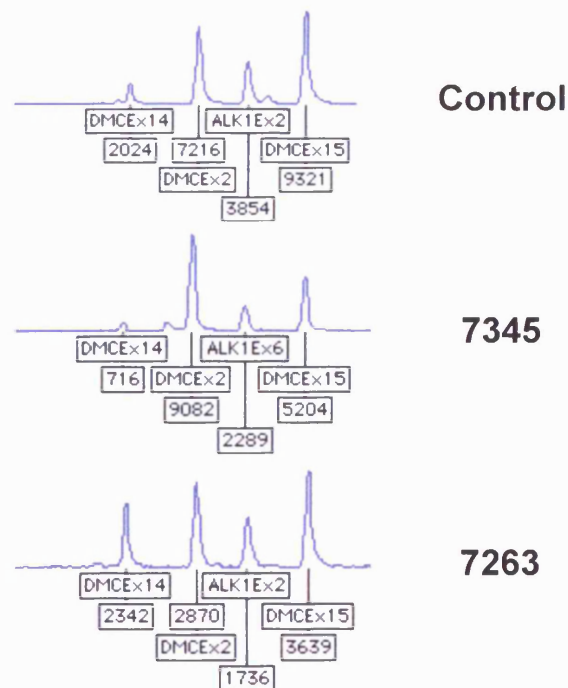
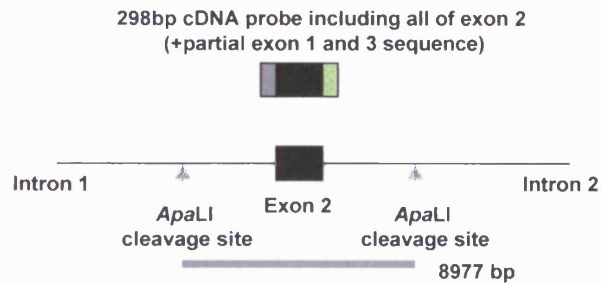
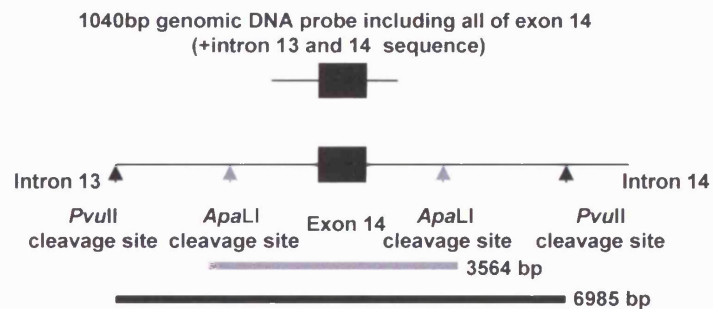


Figure 3.8.

- a) Dosage quotients calculated for affected individuals from families DMC 02 (7345) and 06 (7263 and 7264). Peak areas are shown in the left hand column. The average peak area was calculated for 10 control samples. Dosage quotients reflecting duplication/repetition are highlighted in blue.
- b) Examples of dosage plots generated for patients and controls. Each peak is labelled with the exon represented and the peak area. Note the different peak patterns in control compared to patients.



*Apa*LI cleavage sites in wild type sequence produces fragment of 8977 bp



*Pvu*II and *Apa*LI cleavage sites in wild type sequence produce fragments of 6985 and 3564 bp respectively

Figure 3.9a. Exon 2 probe corresponds to 298 base pairs (bp) of *DYM* cDNA including all of exon 2 and following digestion of genomic DNA with the restriction enzyme *Apa*LI will result in a fragment of 8977 bp. *DYM* exon 14 probe corresponds to 1040 base pairs of *DYM* genomic DNA including all of exon 14. Digestion of genomic DNA with the restriction enzymes *Pvu*II and *Apa*LI will result in fragments of 6985 and 3564 bp respectively.

generated by initial PCR as described in section 2.2.6 using cDNA primers *5'UTR F* (within 5'UTR) and *FLJ1R* (within exon 3) for the exon 2 probe and intronic primers *Ex14 Gen South F* and *Ex14 Gen South R* for exon 14 probe. Primer sequences are listed in Table 3.4. DMC WT pCDNA3 and normal control genomic DNA were used respectively as templates. Following PCR product purification, blunt ending and phosphorylation were performed as described in sections 2.2.13.3 and 2.2.13.4. The insert was ligated into the vector pBKS using the restriction enzyme site *EcoRV*. The plasmid was transformed into *E. coli* DH5 α in order to allow easy amplification of plasmid DNA. Following extraction of plasmid DNA, exon 2 probe sequence was removed from the vector by digestion with *HindIII* and *XbaI* followed by separation on agarose gel and gel purification. Exon 14 probe sequence was similarly isolated from the vector by digestion with *SmaI* and *ApaI* with subsequent electrophoresis and gel purification.

Following digestion of high molecular weight genomic DNA with *ApaI* and sequential probing with *DYM* Exon 2 and Exon 14 probes, bands of the expected sizes (8977 and 3564 bp respectively) were seen both in control samples, family DMC 02 proband (7345) and his father (7344) as illustrated in Figure 3.9b. A maternal DNA sample was not available. ImageQuant (Amersham Bioscience, Chalfont St.Giles, UK) software analysis confirmed the ratio of exon 2 and exon 14 band densities as 2:1 in 7345 (affected) and 3:2 in 7344 (father) compared to 1:1 in control samples. Following digestion of high molecular weight DNA with the restriction enzyme *PvuII* and Southern analysis with a *DYM* exon 14 probe, a band of the expected size (6985 bp) was seen both in control samples and family DMC 06 affected individuals (7263,7264) and parents (7260,7261). However, an additional fragment of approximately 5500 bp was observed in each family member tested, as shown in Figure 3.9c. Analysis following digestion with *ApaI* revealed restriction fragments of the expected size only.

3.8 Discussion

3.8.1.1 *DYM* mutations

This study has identified five novel *DYM* mutations. These include a splice site mutation, two nonsense mutations and two genomic duplication/repetition events both resulting in exon duplication and predicted to cause a truncated protein product. The majority (13/18) of mutations identified as causative of DMC are predicted to result in a truncated protein product. It is likely that nonsense mediated mRNA decay (NMD) would result in the destruction of mutant transcripts.

Primer name	Primer sequence
5' UTR F	5' GAC AGC GAC TTC TCC TGA CC 3'
FLJ1R	5' GAG TTT CAA CTC ACT ACT GC 3'
Ex 14 Gen South F	5' GTC TTG CTC TGC ACA TTA CA 3'
Ex 14 Gen South R	5' CCT GGC TGA AAG GAA AAT 3'

Table 3.4. *DYM* Southern blot probe primers

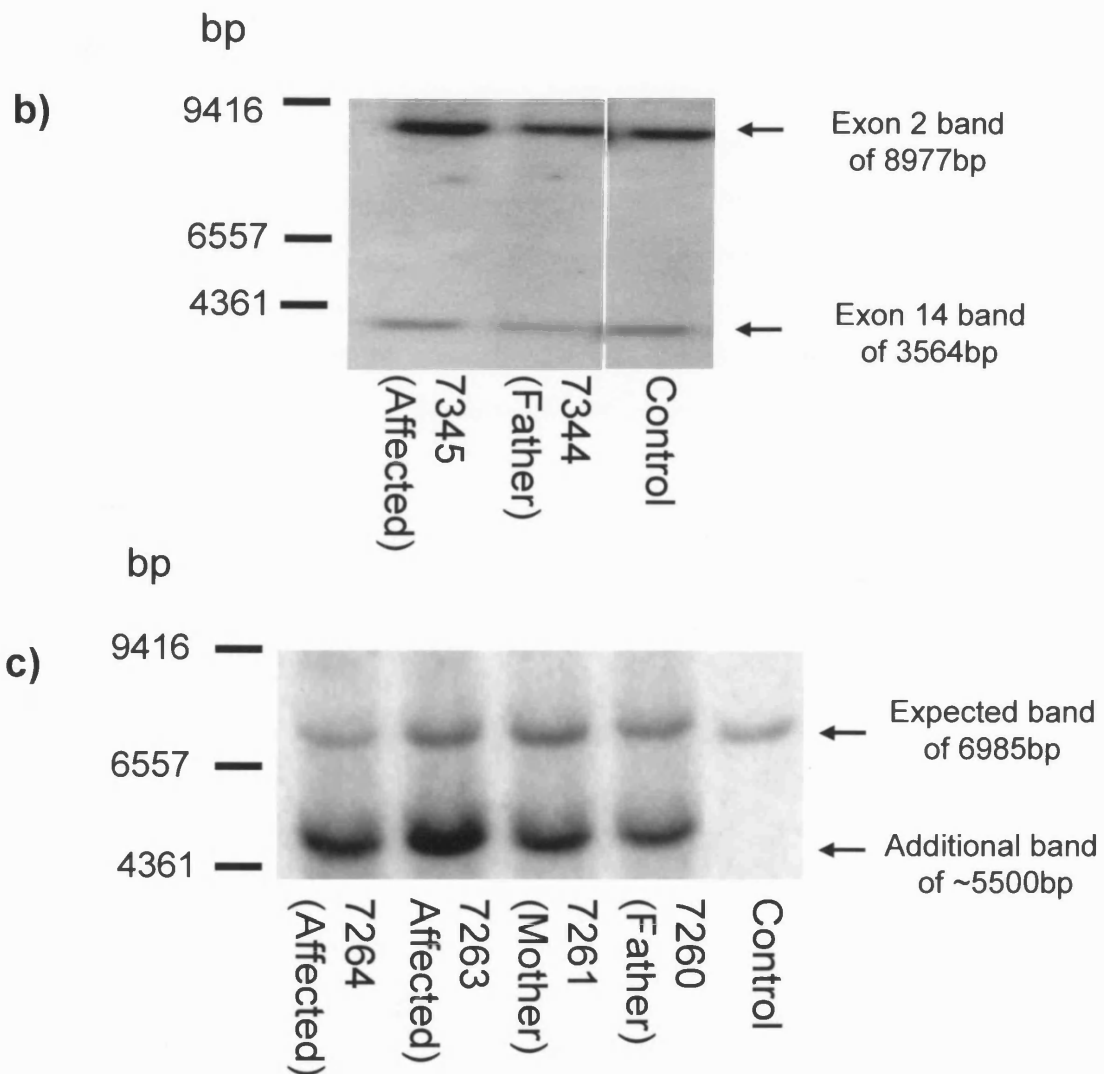


Figure 3.9 b) and c).

b) Southern blot analysis following digestion of genomic DNA with *Apa*LI using *DYM* exon 2 and 14 probes results in bands of the expected sizes (8977 and 3564 bp respectively) in normal control, affected 7345 and father of 7345.

c) Southern blot analysis following digestion of genomic DNA with *Pvu*II using *DYM* exon 14 probe results in a band of the expected size (6985 bp) in all those tested. An additional band of ~5500 bp is also seen in affected siblings 7263 and 7264 and both parents.

3.8.1.2 Missense mutations

The site of missense mutations within a gene may give clues as to domains within the protein necessary for normal function. There have been reports of two missense mutations (E87K and N469Y) in the *DYM* gene, both mutations occurring as compound heterozygotes, one in an individual with DMC in association with a truncating mutation and the other in two patients affected by SMC who also had splice site mutations (Cohn *et al.*, 2003). No other mutations have been reported in SMC. The SMC splice site mutation (c.IVS 7-1 T>G) is predicted to lead to skipping of exon 8, a shift in the reading frame after amino acid 207 and a subsequent termination codon 62 amino acids upstream. Studies of RNA from cartilage of this SMC patient demonstrated there to be normal sized cDNA transcripts in addition to transcripts of a size compatible with exon 8 skipping (Cohn *et al.*, 2003). This has led to the hypothesis that the absence of mental retardation in SMC is due to the presence of some functional DYM with the brain requiring lower levels compared to cartilage in order to function normally. No detailed clinical information was reported on the DMC patient with the N469Y missense mutation but it would be interesting to know the extent of mental retardation in this individual, which may be less marked if some DYM function remains. Both missense mutations occur at amino acids conserved through evolution (see Figure 3.10) hence are likely to be located in regions of the protein with functional importance. N469Y affects an amino acid predicted to be a glycosylation site which may be critical for normal protein functioning.

3.8.1.3 Duplication mutations

The exon 2 duplication and exon 14 repetition predict the insertion of 22 and six amino acids respectively followed by a premature termination codon. It is expected that the mutant transcripts would be destroyed by NMD with subsequent failure to produce a protein product. The ATG initiator codon is located in exon 2 with the duplicated sequence containing a 2nd repeat ATG initiation site. If translation were initiated in the 2nd (repeated) copy of exon 2 a normal transcript would be produced. However, exon 2 contains only 53 bp of untranslated sequence which may not be sufficient to promote use of this initiation site and hence translation may preferentially occur at the first ATG. Genomic duplications which result in exon repetition/duplication at a cDNA level have been identified in individuals 7345, 7263 and 7264. This is the first report of such a mutation mechanism being detected in a recessive disorder and has implications for mutation screening in other recessive disorders. Genomic duplication mutations would

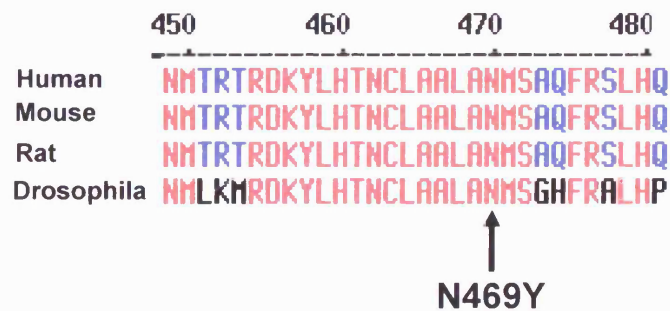
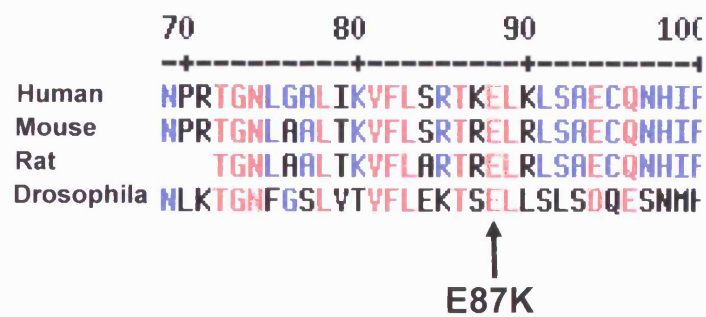


Figure 3.10. *DYM* missense mutations E87K and N469Y occur at amino acids conserved through evolution. Human, mouse, rat and drosophila amino acid sequence is shown with most conserved amino acids in red and least conserved amino acids in black. Moderately conserved residues are illustrated in blue.

not be identified by direct sequencing. In many cases, the only technique used to look for pathogenic mutations. Southern blot analysis has demonstrated the genomic architecture surrounding exon 14 in individuals 7263 and 7264 to be disturbed hence leading to an aberrantly sized fragment after probing with an exon 14 probe following *Apa*LI digestion. The presence of a band of the expected size in addition to an aberrantly sized band suggests that at least one of the external copies has been unaffected by the duplication. The exact nature and extent of the duplication has not been identified but these results suggest the duplicated sequence is likely to commence between the *Apa*LI and *Pvu*II restriction sites (see Figure 3.9a) as only a band of the expected size was observed following *Pvu*II digestion of genomic DNA. Whilst the real extent of genomic duplication has not been identified, implicit in these observations is the assumption that the duplication events encompass sufficient intronic sequence to enable normal splicing.

3.8.1.4 Limitations of techniques used for mutation identification

The precise ascertainment of clinical phenotype is critical in the study of any genetic disorder but particularly in the skeletal dysplasias in which there is often significant overlap between conditions. The diagnosis of DMC is made following a combination of clinical examination, skeletal survey and the exclusion of the main differential diagnosis, Morquio's disease (a mucopolysaccharidosis) by urinary sulphate and white cell enzyme analysis. Steps which have been undertaken to ensure only the recruitment of individuals affected with DMC into this study have included the collection of extensive clinical data including review of hospital notes, review of X-rays by an expert in skeletal dysplasias (Professor Christine Hall, Great Ormond Street Hospital, London) wherever possible and clinical examination by the researcher in those patients in whom this was feasible.

Despite this, of the eight families recruited, mutations were identified by direct sequencing in two families and duplication/repetition identified in a further two families. The initial screening technique used for mutation analysis included direct sequencing of exons and exon-intron boundaries but would not identify mutations in promoter regions or further into introns, gene deletion or exon rearrangements, deletions or duplications. The subsequent identification of exon duplication/repetition was made initially following RT-PCR of patient cDNA. However, RNA was not available from all patients hence it is possible that other patients may also have duplications/deletions. This could be investigated in the absence of RNA samples by the use of a DNA based technique such as MAPH capable of detecting exon duplication/deletion.

The gene causative of DMC was identified using an autozygosity mapping approach (Ehtesham *et al.*, 2002, Thauvin-Robinet *et al.*, 2002). Results of microsatellite marker analysis of the *DYM* locus in patients included in this study highlight the limitations of this technique. Compound heterozygosity for *DYM* mutations (family DMC 04) has been demonstrated in the offspring of a consanguineous partnership. Genotyping data in this individual would have falsely suggested exclusion of this chromosomal region assuming identity by descent. Heterozygosity around the *DYM* locus found in the offspring of other consanguineous parents could also be explained by compound heterozygosity for *DYM* gene mutations which have not been identified. However, it must also be considered that DMC is a heterogeneous disorder and there may be other, as yet undiscovered loci. There is no published evidence to confirm this although it has been suggested by Neumann *et al* who failed to identify mutations in a patient with SMC (Neumann *et al.*, 2006).

3.8.2 Exon duplication/repetition as a disease mechanism

RT-PCR and fluorescent dosage PCR have demonstrated duplication of exon 2 and part of the 5' UTR in one individual and repetition of four copies of exon 14 in two affected siblings. Both duplication/repetition events result in a frameshift and a premature termination codon suggesting loss of protein function as a disease mechanism. The existence of exon repetition in mRNA in the absence of duplication in the genome has been reported and hence this possibility must be considered in our cases (Frantz *et al.*, 1999). However, data from fluorescent dosage PCR performed on genomic DNA was consistent with exon duplication/repetition in the genome demonstrating that the duplication/repetition is not only present in RNA. Furthermore, Southern blot analysis also provided independent confirmation that the duplications are present at a genomic level.

The mechanisms which may underlie the duplication events must be considered and include non-allelic homologous recombination (NAHR), non-homologous end-joining (NHEJ) and replication slippage, all of which occur as errors during meiotic recombination. NAHR is the recombination between repeat regions on opposite alleles and can result in either duplication or deletion of the intervening DNA. NHEJ is the joining of DNA breakpoints which are not homologous from the same or different chromosomes and may also result in duplication or deletion (Lupski, 1998). The final mechanism to be considered is replication slippage which occurs through misalignment between regions of repeat sequence leading to insertion of the DNA sequence between

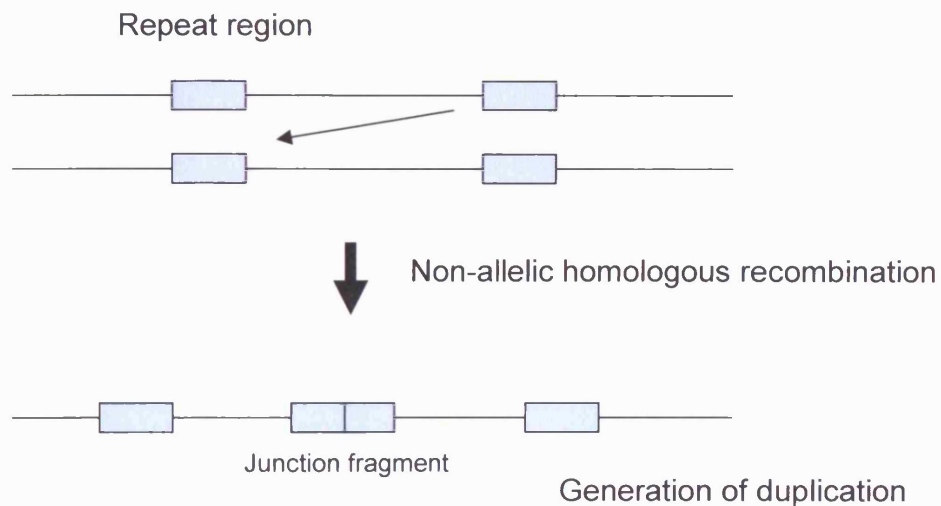
the repeat regions and subsequent duplication. NAHR, NHEJ and replication slippage are depicted pictorially in Figure 3.11.

BLAST analysis of *DYM* intronic sequence surrounding exons 2 and 14 was used to search for regions of homology where non-allelic homologous recombination or replication slippage may have arisen. This revealed multiple Alu repeat sequences in the intronic sequence both up and downstream of exons 2 and 14, including several in the region between *PvuII* and *ApaLI* cut sites surrounding exon 14. Southern blot analysis was consistent with one of these repeats forming the site of initiation of duplication in DMC 06 family members, as analysis following digestion with *PvuII* revealed a band of the appropriate size whereas analysis following *ApaLI* digestion revealed an aberrantly sized band. Repeat regions of approximately 1 kb in length were identified 15 kb downstream of exon 14 and both 36 and 64 kb upstream of exon 14 which could also have provided potential sites for NAHR or replication slippage to occur.

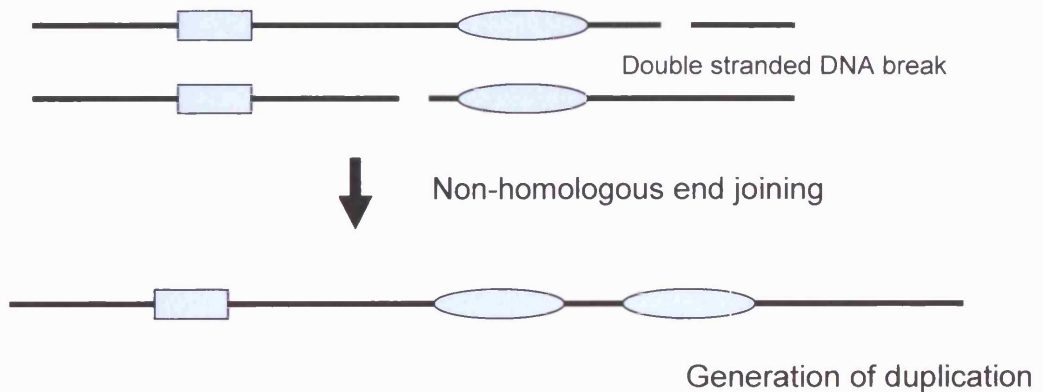
NAHR at Alu repeat elements may have occurred in families DMC 02 and DMC 06 as illustrated in Figure 3.12. Rearrangements arising at Alu elements have been reported in several conditions including Duchenne muscular dystrophy and hereditary liability to neuropathy and pressure palsies (HNPP) and have been estimated to account for 0.3% of human genetic diseases (Deininger and Batzer, 1999). Recent estimates from computational data with experimental verification have identified nearly 500 Alu mediated deletions within the human genome, of which approximately 60% lie within known or predicted genes, suggesting that this mechanism is an important contributor to deletion/duplication events (Sen *et al.*, 2006).

Low copy repeat (LCR) structures are highly homologous regions of the genome which have also been implicated in NAHR resulting in gene deletion, duplication or rearrangement. NAHR at LCRs has been identified as a causative mechanism in human genomic disorders such as Charcot-Marie-Tooth disease, Williams syndrome and Smith-Magenis syndrome (Shaw and Lupski, 2004). In contrast to the 1 kb repeat regions within the *DYM* intronic sequence, LCRs are usually significantly longer at 10-500 kb. However, studies of recurrent deletions in Smith-Magenis syndrome have demonstrated that recombination repeatedly occurs within a ~500 bp region of the entire 124 kb homologous region (Shaw *et al.*, 2004). It is therefore plausible that a smaller repeat region such as the repeat identified in *DYM* intronic sequence surrounding exon 14 may be sufficient to lead to recombination by NAHR. Moreover, the minimum length of sequence required for strand exchange during homologous recombination

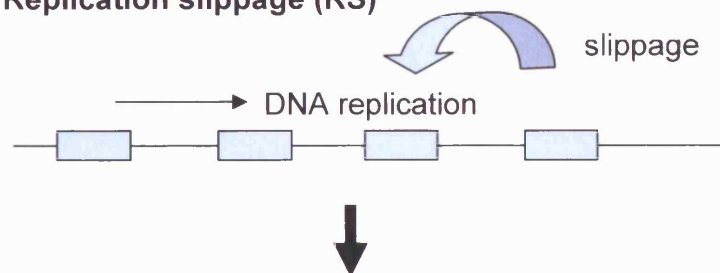
Non-allelic homologous recombination (NAHR)



Non-homologous end joining (NHEJ)



Replication slippage (RS)



Slippage during replication leads to introduction of additional sequence



Figure 3.11. Mechanisms of generation of repeat sequences during meiosis. NAHR arises through recombination between repeat sequences on opposite alleles whereas NHEJ comes about through the joining of non-homologous DNA breakpoints. RS occurs through the insertion of sequence lying between repeat regions.

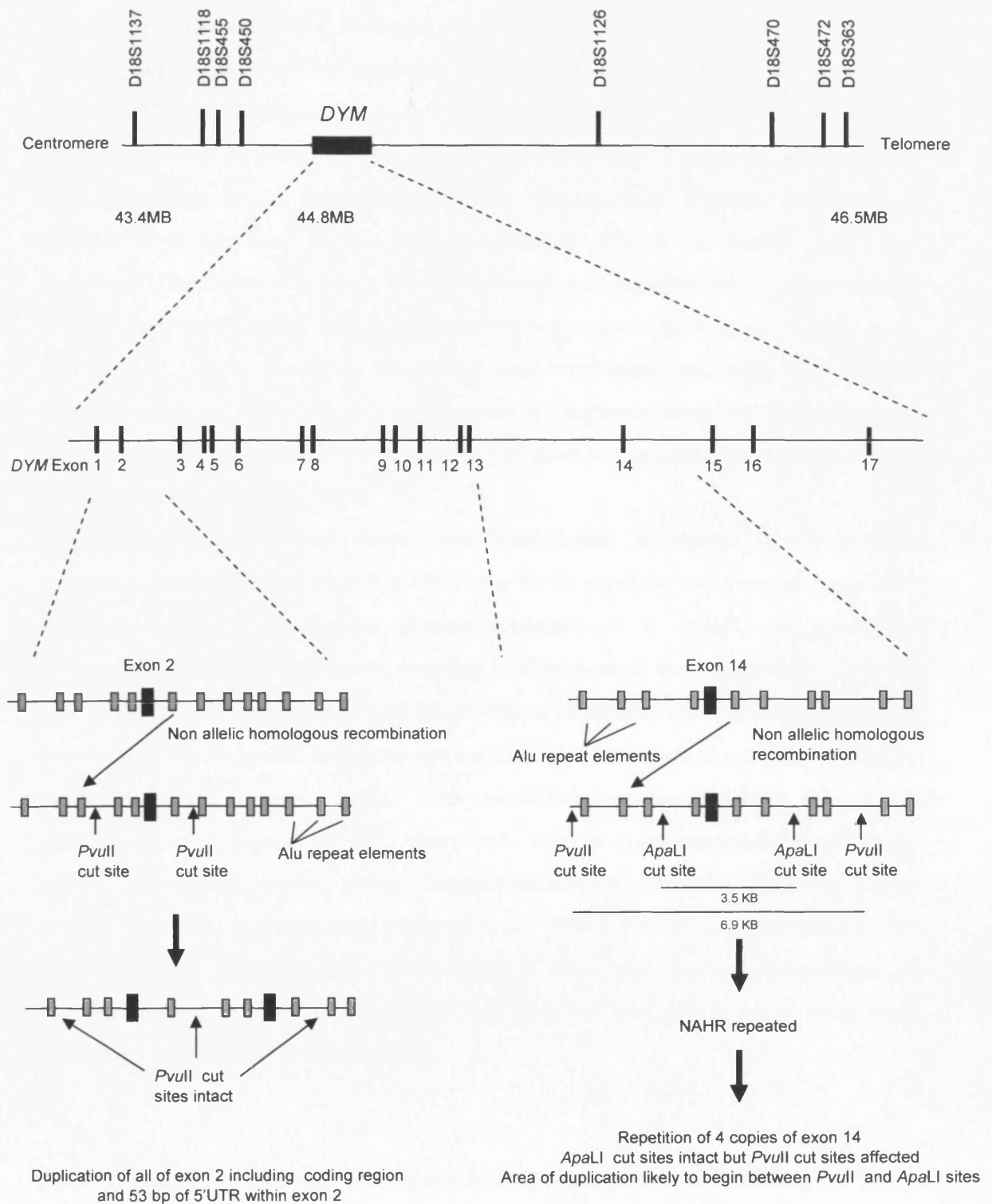


Figure 3.12. Multiple Alu repeat elements are located in the intronic sequence both 5' and 3' to exons 2 and 14. If NAHR occurred at Alu sites external to *PvuII* cut sites around exon 2, exon duplication could occur not affecting the size of fragment produced by *PvuII* digestion. NAHR occurring between *PvuII* and *ApaLI* cut sites around exon 14 would result in an aberrantly sized fragment after digestion with *PvuII* but a normally sized fragment after digestion with *ApaLI* as suggested by Southern blotting.

(minimal efficient processing segment, MEPS) during meiosis has been demonstrated to be between 300 and 500 bp (Reiter *et al.*, 1998)

Studies have highlighted the existence of variation in copy number (copy number variation, CNV) of large segments of the human genome and suggested that this phenomenon may be more common than previously thought. Estimates using genome-wide fluorescent *in-situ* hybridisation (FISH) with bacterial artificial chromosomes (BACS) as probes have shown that approximately 5% of the human genome is duplicated (Cheung *et al.*, 2001). Genomic imbalance was detected in approximately 50% of individuals tested in one study, including regions containing known genes (Iafrate *et al.*, 2004). However, the sample sizes were small and using mathematical estimates based on deletions and duplications in Duchenne muscular dystrophy as a model, others have suggested 1 in 8 humans will have a segmental deletion and 1 in 50 a duplication (van Ommen, 2005).

Interestingly, increased gene density has been found in regions of the genome containing duplicons suggesting that CNV may be an important mechanism underlying mutations causing many human diseases (Fredman *et al.*, 2004). An underlying genomic deletion/duplication event resulting in disruption of the normal exon structure may account for a proportion of individuals whose phenotype appears to be linked to a disease gene but in whom mutations are not identified by standard mutation screening techniques. More recent analyses have identified over-representation of protein encoding genes in regions of CNV. Conversely, there is under-representation of genes known to be disease causing, perhaps because paralogues within the duplicated CNVs provide functional compensation (Nguyen *et al.*, 2006). As our understanding of the repetitive nature of the genome increases, it is likely that duplication/repetition or deletion events such as those identified in this study will be found to be the underlying mutational mechanism in other disorders.

3.8.3 Premature termination codons and nonsense mediated mRNA decay

NMD is a cellular surveillance system which recognises aberrant mRNAs and degrades them such that abnormal and potentially deleterious proteins are not produced. NMD machinery recognises premature termination codons occurring more than 50 nucleotides upstream of an exon-exon junction created by splicing. A protein complex known as the exon-exon junction complex, (EJC) localises upstream of the 5' splice site following splicing and should be removed on initiation of translation. However, if an aberrant premature termination codon is present, the EJC remains on the mRNA, marking it for

NMD. Subsequently the aberrant transcript is degraded by decapping and digestion by exonucleases (Baker and Parker, 2004). The five mutations/duplication events identified in this study all predict a truncated protein product. Given that the majority of *DYM* mutations predict a truncated protein product, it is expected that NMD would destroy mutant transcripts and hence the phenotype is likely due to the absence of any functioning Dymeclin.

3.8.4 Implications for other recessive disorders

The term genomic disorder has been used to describe diseases caused by loss, gain or disruption of genes by changes to the genomic structure. The number of diseases which fall within this group is rapidly increasing. The finding of genomic duplication as causative in DMC suggests that this mechanism should be considered in other recessive disorders and has implications for mutation screening in such cases. Advances in array comparative genomic hybridisation (array CGH), a method of identifying genomic copy number now allow resolution at a single exon level, and as this technique becomes cheaper and more widely available will begin to address this problem (Dhami *et al.*, 2005).

3.8.5 Summary

This study has identified five mutations in the *DYM* gene in affected individuals, of which four are novel. Of these, two are nonsense, c.208C>T and c.1363C>T predicting R70X and R455X respectively, one splice site (IVS 11 c.1252-1G >A) and two complex genomic repetition events each predicted to result in a frameshift and subsequent premature termination codon. This is the first report (see Appendix C) of a genomic repetition event leading to a pathogenic disruption in the exonic structure of a gene causative of an autosomal recessive disorder (Kinning *et al.*, 2005). This type of mutational event cannot be detected by the most commonly employed methods of mutation detection and has implications for mutation screening in a wide range of genetic disorders.

Chapter 4 – Characterisation of the *DYM* gene product

4.1 Introduction

The DMC phenotype suggests that the *DYM* gene product (Dymeclin) is essential in the development and homeostasis of cartilage and bone and in critical neuronal function. Over the last few years, a significant amount of information has become available through online databases relating to genes, genomic organisation, transcripts and protein products. Such data enables comparisons that offer insight as to possible function of, for example, novel proteins by homology. The amino acid sequence of the protein of interest can be compared to those available in the databases, in order to identify members of protein families and detect functional motifs. However, online databases only provide predictions and are subject to limiting factors including the quality and quantity of data available and the extent of homology between proteins within the databases and the novel protein of interest. It is essential that experimental data is also obtained to confirm or refute the predictions. Bioinformatic analysis was undertaken with the aim of obtaining insight as to the putative function of *DYM*. In addition, experiments were performed to determine the tissue expression patterns at an RNA and protein level, whilst transfection studies provided evidence for its sub-cellular location.

4.2 *In-silico* bioinformatic analysis

DYM genomic sequence annotation (NCBI Gene ID: 54808, Ensembl ENSG00000141627) reveals a gene comprising 17 exons over 400 kb with the ATG initiator located in exon 2. However alternative splice variants have been suggested (Aceview, NCBI) and are supported by expressed sequence tag (EST) evidence. Indeed, published Northern blot analysis demonstrates 2 differentially sized transcripts of approximately 3 kb and 6 kb with variable expression levels in different tissues as shown in Figure 4.1 (El Ghouzzi *et al.*, 2003). However, there is no further published evidence to confirm the existence of alternatively spliced gene products.

DYM encodes a 669 amino acid protein predicted to have a molecular weight of 75.9 kD (Compute pI/MW). Database analysis reveals no homology to any known human protein but orthologues appear to be present in many species with the C-terminal region being highly conserved through species as illustrated in Figure 4.2. This suggests that the C-terminus has an important conserved function through evolution. The presence of transmembrane domains can be predicted from regions of hydrophobic amino acids which are likely to be located within a membrane. Examples of hydropathy plots

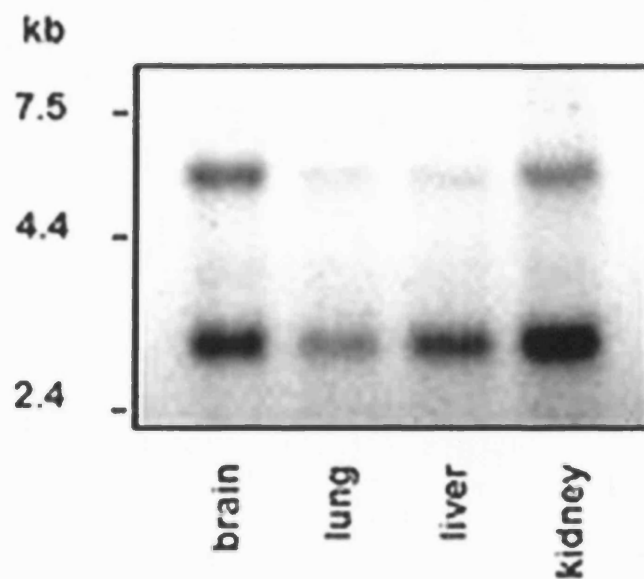


Figure 4.1. Northern blot analysis of *DYM*. RNA transcripts of approximately 3 kb and 6 kb are demonstrated in all four tissues tested with differing levels of expression. From El Ghouzzi *et al.*



Figure 4.2. Multiple alignment of DYM amino acid sequence across human, mouse, rat and drosophila with consensus sequence also shown. Most highly conserved amino acids are illustrated in red with least conserved residues in black (Multalign).

predicting transmembrane domains within Dymeclin are illustrated in Figure 4.3 suggesting that Dymeclin has between 3 and 6 transmembrane domains, depending on the stringency of the prediction (TopPred, HMMTop). Analysis looking for distinctive protein domains/motifs provides no evidence for a signal sequence indicative of a protein that is secreted. However, there is strong support for the presence of a myristoyl domain and 17 dileucine motifs (InterProScan, ScanProsite, PRINTS, SMART, PSORTII). Myristoyl domains allow a protein to be acylated by the addition of myristate (a C14-saturated fatty acid) to the N-terminal residue, a protein modification step which has been shown to assist in the association of a protein with a membrane (Boutin, 1997). Dileucine motifs comprise two sequential leucine molecules and are thought to be involved in the targeting of transmembrane proteins to the endosomal-lysosomal system (Bonifacino and Traub, 2003). There are also 3 dilysine (KKXX) motifs present, an ER retrieval motif which interacts with coat protein I (COPI) vesicles which are required for retrograde protein trafficking from the Golgi to the ER (Cosson and Letourneur, 1997). Proteins harbouring this motif are mostly found in the ER although may undergo cycling between the Golgi and ER (Duden, 2003). The presence of two O-glycosylation and one N-glycosylation sites suggest that the protein may undergo post-translational modification by glycosylation, a mechanism of addition of a glycosyl group to a protein. This modification step takes place in the ER and is thought to stabilise the protein and prevent its degradation. Several consensus sites of phosphorylation are also present implying that phosphorylation may be used as a mechanism to activate or inactivate the protein. A schematic illustrating the predicted transmembrane domains, myristoyl, dileucine and dilysine motifs is shown in Figure 4.4. The N-terminus of Dymeclin is predicted to be cytoplasmic (MTOP) and the protein is predicted to be most likely to localise to the endoplasmic reticulum although there is no ER retention motif (PSORTII). These *in-silico* predictions suggest that Dymeclin is likely to be a transmembrane protein involved in protein sorting/targeting within the cell.

4.3 RT-PCR of cDNA panel

In order to determine the expression patterns of *DYM* transcripts, RT-PCR was performed as described in section 2.2.10. A variety of cDNAs from human adult tissues (brain, heart, lung, liver, skeletal muscle, kidney, pancreas and placenta) (BD Multiple Human Tissue cDNA panel, Clontech, Mountain View, CA, USA) were used as template with a forward primer in *DYM* exon 7 (*INTF7*) and a reverse primer in *DYM* exon 9 (*INTR9*) to produce a product of 310 bp. A 431 bp fragment of the house keeping gene *GAPDH* was also amplified using the primers *GAPDHF* and *GAPDHR* to

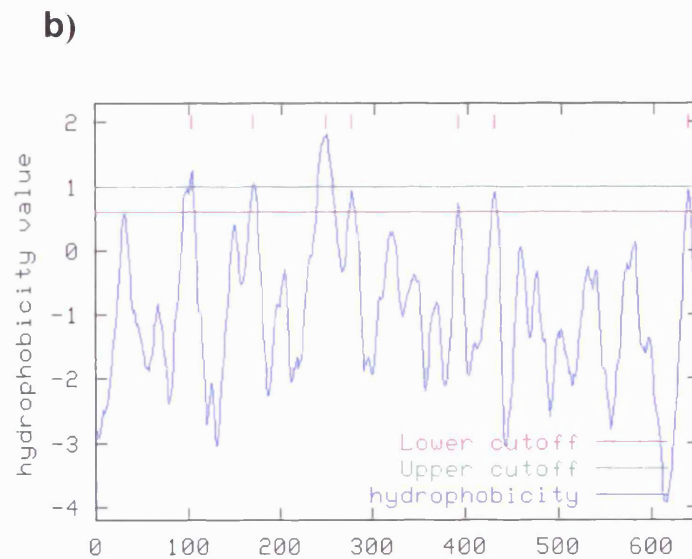
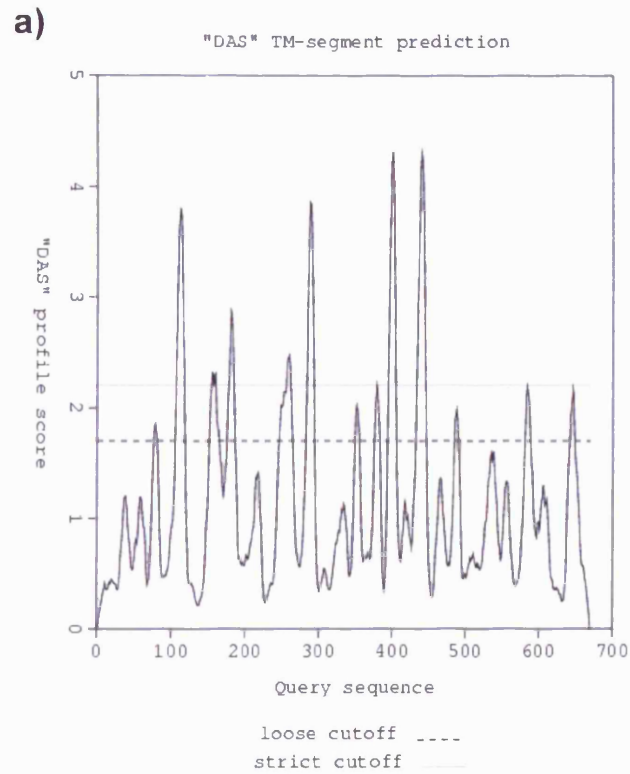


Figure 4.3.

- a) Hydropathy plot produced by HMMTop predicting 6 transmembrane domains taking the higher cut off level indicated by the unbroken line.
- b) Hydropathy plot produced by TopPred predicting 3 transmembrane domains taking the higher cut off level indicated by the green line.

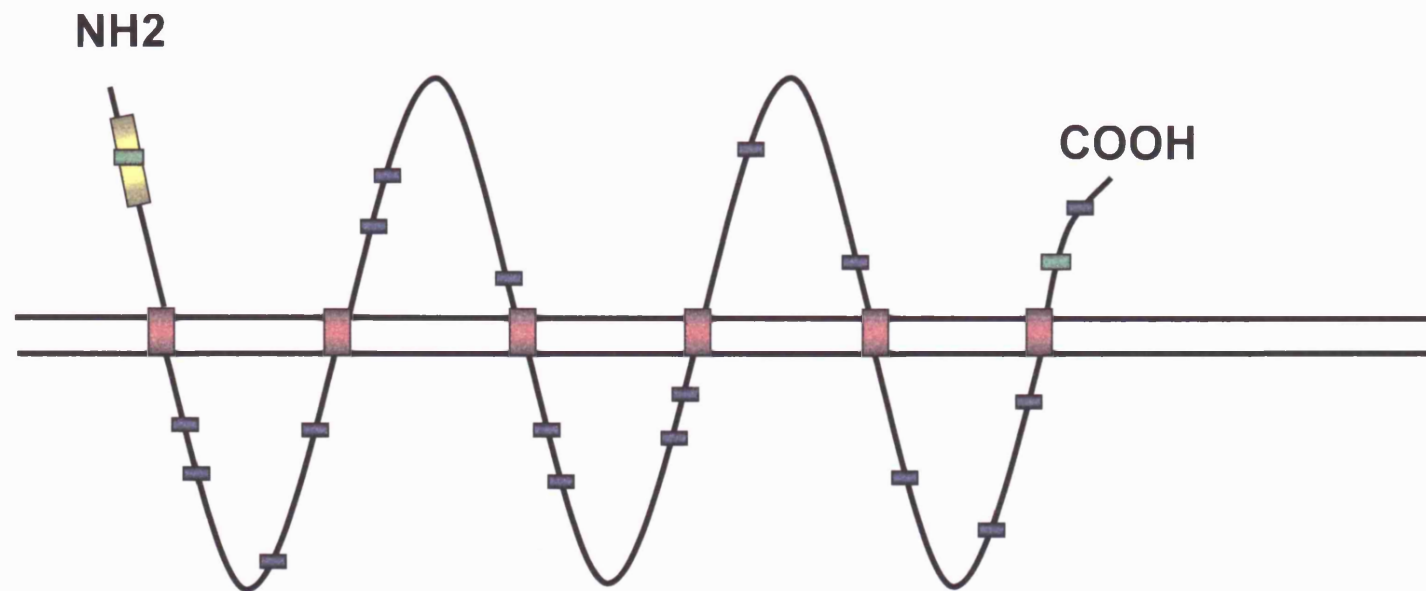


Figure 4.4. Schematic illustrating DYM predicted structure. Transmembrane domains (red), myristoylation site (yellow), di-leucine (blue) and di-lysine motifs (green) are shown. This is a schematic only and is not drawn to scale. Predictions from Prosite, SOSUI, TMPred, PSORTII.

act as a positive control. Primer sequences are listed in Table 4.1. A *DYM* transcript was demonstrated in all tissues tested as illustrated in Figure 4.5a. Published data illustrated in Figure 4.5b has shown a transcript to be present in a selection of tissue types relevant to the DMC phenotype including chondrocytes and fetal brain (El Ghouzzi *et al.*, 2003). These results demonstrate that *DYM* is widely expressed with a transcript seen in all tissues tested.

4.4 *In-vitro* translation

In-vitro translation of *DYM* wild type cDNA in the expression vector pCDNA3 produced as described in section 2.2.15 resulted in a protein product of 75 kD as illustrated in Figure 4.6. This is consistent with the predicted molecular weight of 75.9 kD. Translation of *DYM* mutant cDNAs encoding truncating mutants R70X, R204X and Q483X produced smaller protein products of, respectively <10, ~20 and ~45 kD. The protein product produced by cDNA encoding the missense mutations E87K and N469Y produced bands of similar size to the wild type product.

4.5 Transient transfection studies

5' myc and GFP-tagged *DYM* wild type constructs and mutants R70X, E87K, N469Y, Q483X and K626N produced as described in sections 4.13.2 and 4.13.3 were used for transient transfection studies using mouse fibroblasts (NIH3T3) as described in section 2.2.14.1. Transient transfection of both myc and GFP-tagged *DYM* wild type constructs into mouse fibroblasts demonstrated localisation of the tagged protein largely to the cytoplasm in a reticular distribution, consistent with Dymeclin being within the endoplasmic reticulum. The missense mutations E87K and N469Y resulted in aberrant localisation in approximately 50% of transfected cells with accumulation in the cytoplasm in a punctate pattern. A similar pattern was seen in the nonsense mutations R70X and Q483X and the frameshift mutation K626N (which predicts an aberrantly extended protein). This distribution was seen both in myc and GFP-tagged proteins. Subcellular localisation images are illustrated in Figure 4.7.

In order to confirm that Dymeclin localises to the endoplasmic reticulum, co-localisation studies were performed. Mouse fibroblasts were transiently transfected with myc-tagged wild type Dymeclin as previously described. Cells were fixed as described in section 2.2.14.3 after either 24, 48, 72, or 96 hr. Myc-tagged proteins were detected using the primary antibody mouse anti-c myc (Invitrogen, Paisley, UK) diluted 1/500. In order to show endoplasmic reticulum localisation, cells were incubated as described

Primer name	Primer sequence	Annealing temperature
INT F7	5' AAG AAG TTT TGC GAC AGA GC 3'	58°C
INT R9	5' GGC CAG AGG GGA AGA AAG 3'	58°C
GAPDHF	5' AGA ACA TCA TCC CTG CCT C 3'	58°C
GAPDHR	5' GCC AAA TTC GTT GTC GTC ATA CC 3'	58°C

Table 4.1. RT-PCR primer sequences.

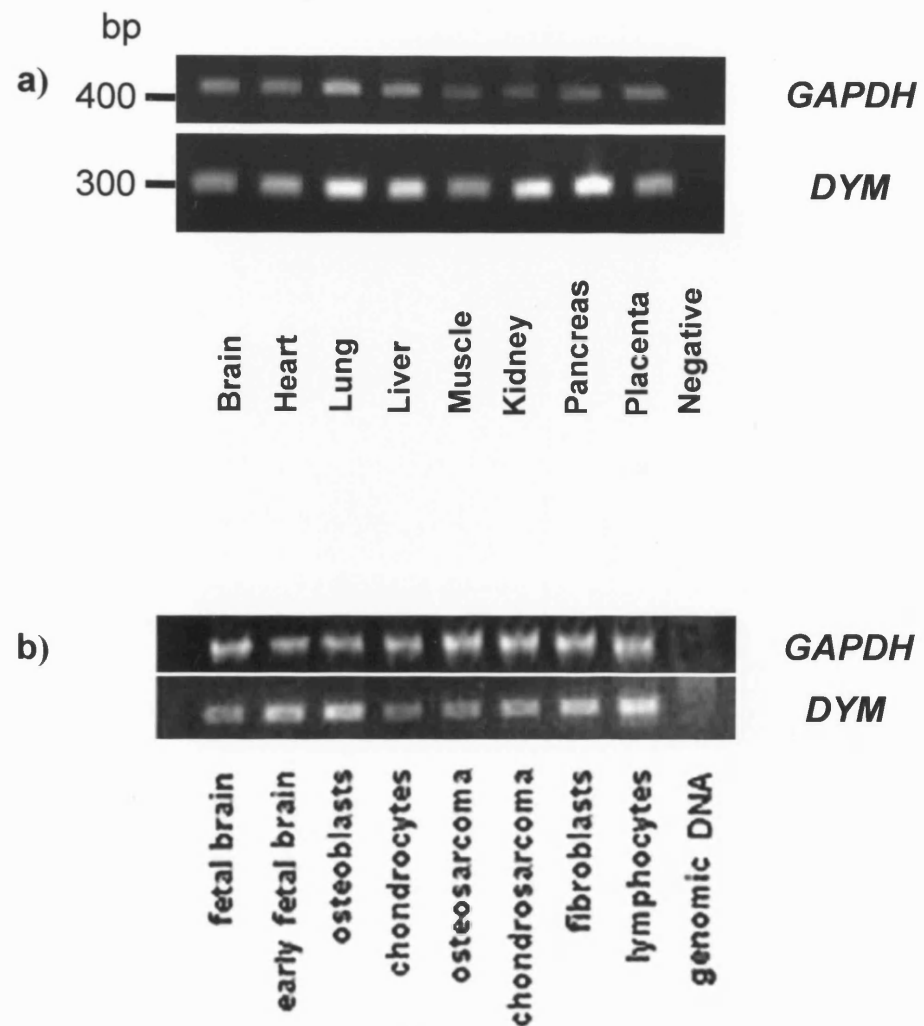


Figure 4.5.

a) RT-PCR on a variety of tissues using *DYM* primers *INTF7* and *INTR9* and *GAPDH* primers *GAPDHF* and *GAPDHR* demonstrating a transcript in all tissues tested.

b) RT-PCR on a variety of tissues relevant to the phenotype of DMC again showing a transcript in all tissues tested.

From El Ghouzzi *et al.*

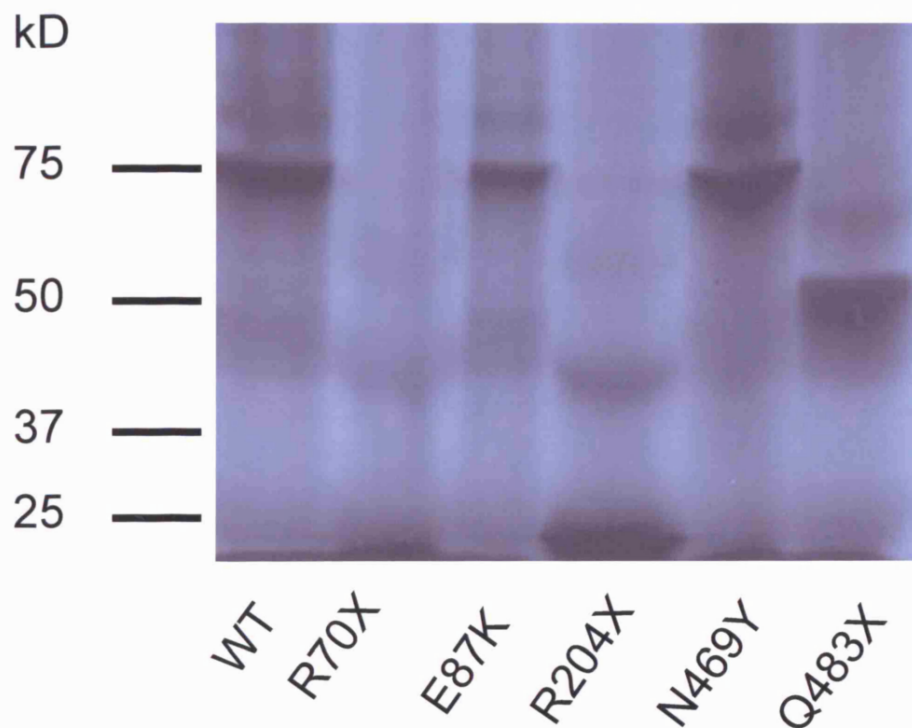
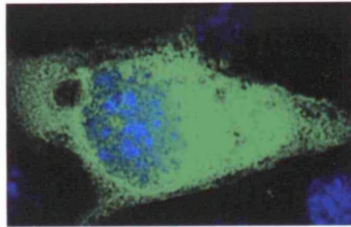
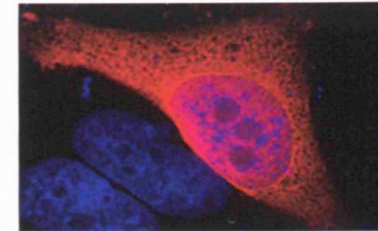


Figure 4.6. In-vitro translation (IVT) of *DYM* wild type and mutant constructs. IVT using the expression vector pCDNA3 results in a protein product of ~ 75 kD for the wild type protein. Truncating mutations R70X, R204X and Q483X produce smaller protein products of, respectively <10, ~20 and ~45 kD. The protein product produced by cDNA containing missense mutations E87K and N469Y produces bands of similar size to the wild type product.

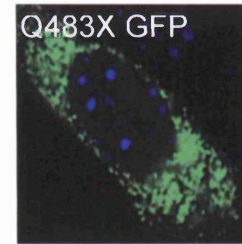
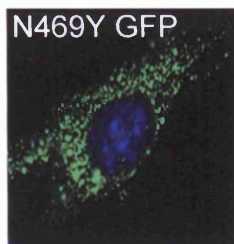
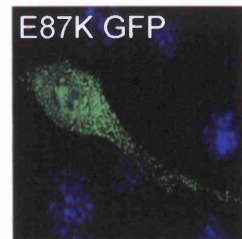
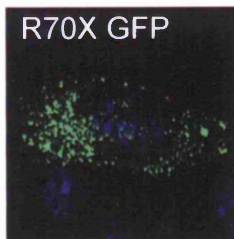
a) Wild type GFP



b) Wild type Myc



c)



d)

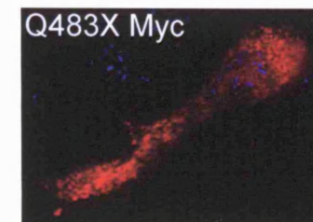
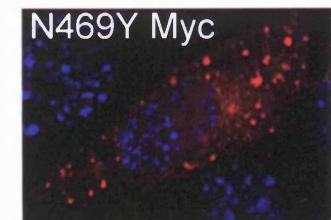
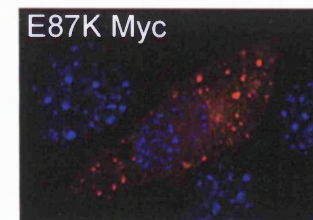


Figure 4.7. Sub-cellular localisation by transient transfection of NIH3T3 cells. a) and b) show cells transfected with wild type Green fluorescent protein and Myc tagged *DYM* constructs respectively. c) and d) show cells transfected with a range of mutant Green fluorescent protein and Myc tagged *DYM* constructs respectively.

in section 2.2.14.3 with a rabbit primary antibody to GRP78 BiP, a resident ER protein (Invitrogen, Paisley, UK) at a dilution of 1/500. Alexa Fluor® 594 donkey anti-mouse IgG and Alexa Fluor® 488 goat anti-rabbit IgG, (Invitrogen, Paisley, UK) both at a dilution of 1/500 were used as the secondary antibodies respectively. DNA was stained with Hoescht at a dilution of 1/500 (Invitrogen, Paisley, UK). Imaging was performed as described in section 2.2.14.3. Myc-tagged Dymeclin showed co-localisation to the endoplasmic reticulum as illustrated in Figure 4.8. Co-localisation studies confirmed the localisation of Dymeclin to the ER.

4.6 Western blot analysis

In order to confirm that transfected cells were producing myc-tagged Dymeclin wild type or mutant proteins, cell extracts were made from cells transfected with myc-tagged wild type and mutant Dymeclin constructs as described in section 2.2.14.2 and separated by SDS-PAGE as described in section 2.2.16. Following transfer to a nitrocellulose membrane, Western blot analysis was performed using an antibody against the myc-tag (Invitrogen, Paisley, UK) as described in section 2.2.17. This resulted in bands of ~75 kD corresponding to proteins produced from constructs encoding wild type Dymeclin and missense mutations E87K and N469Y, a band of ~50 kD corresponding to the truncating mutant Q483X and no band was seen corresponding to the truncating mutant R70X, likely because the protein product is too small to be resolved on the concentration of gel used (see Figure 4.9). Western blotting has shown transfected cell extracts to contain wild type and missense myc-tagged Dymeclin of ~75 kD, comparable to the predicted molecular weight. Expression of the truncating mutation Q483X resulted in a smaller tagged protein, as would be expected.

In order to confirm that Dymeclin is a membrane protein, Western blot analysis was performed using cell extracts made following digestion in a variety of buffers. Cell extracts were prepared as described in section 2.2.14.2 but supplemented with either 7 M urea or 1% Triton x 100 in addition to 50 or 500 mM NaCl. Extracts were incubated on ice for 30 min and then centrifuged at $16\,000 \times g$ for 10 min to obtain soluble and insoluble fractions. Digestion in Laemmli buffer alone resulted in the tagged protein being found within the cell pellet suggesting that Dymeclin is in an insoluble cell compartment. Digestion in buffer containing Triton x 100, a detergent which solubilises all cell membranes resulted in a band representing Dymeclin being found within the supernatant, confirming that Dymeclin is likely to be within a cell membrane. Digestion with urea increases protein solubility by disrupting non-covalent bonds, but does not disrupt membranes. Following incubation in a 7 M urea solution,

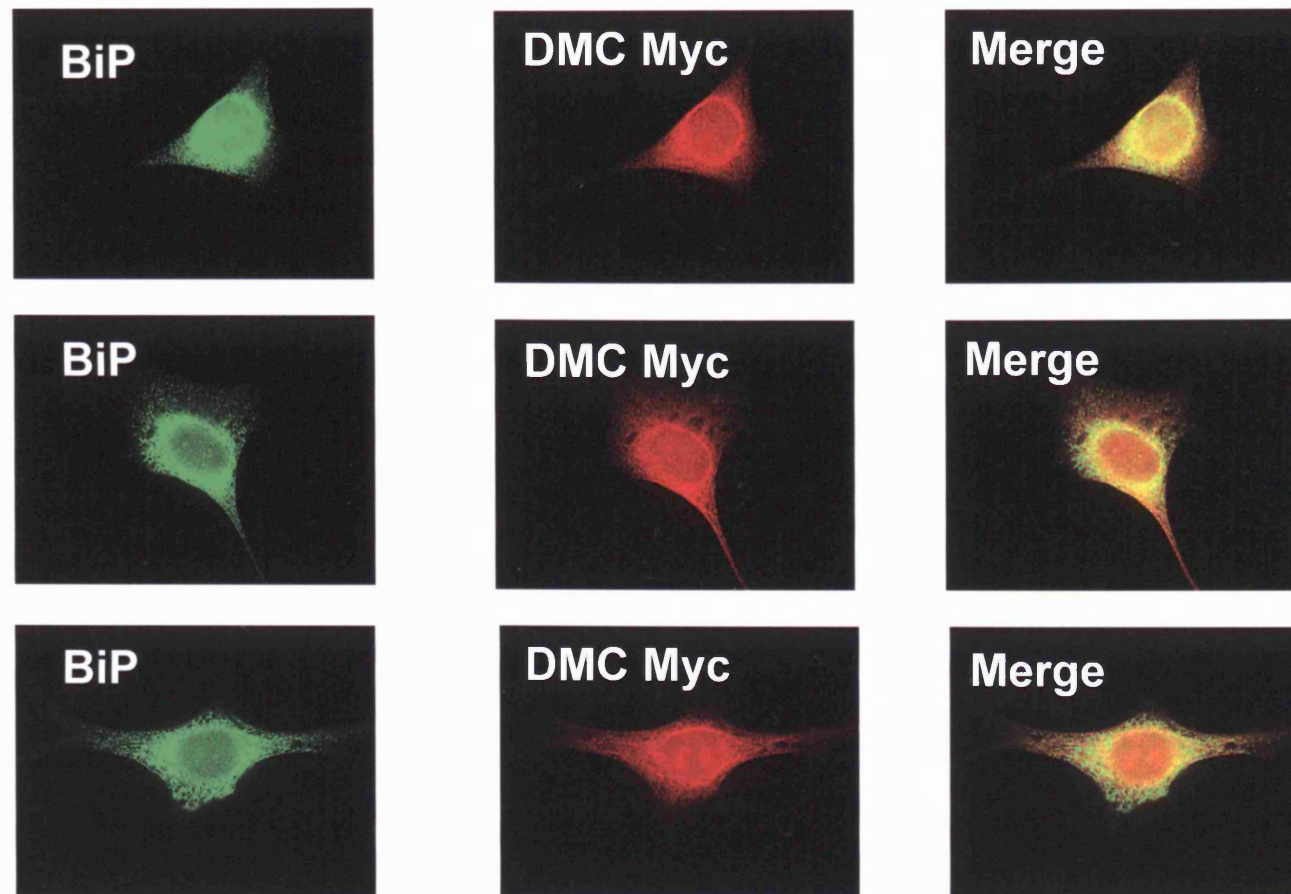


Figure 4.8. Sub-cellular co-localisation by transient transfection of NIH3T3 cells. BiP (ER marker) localisation is shown in green and DYM myc-tagged protein in red . Merged pictures in the right panel demonstrate co-localisation in yellow. Cells were fixed 96 hours post-transfection.

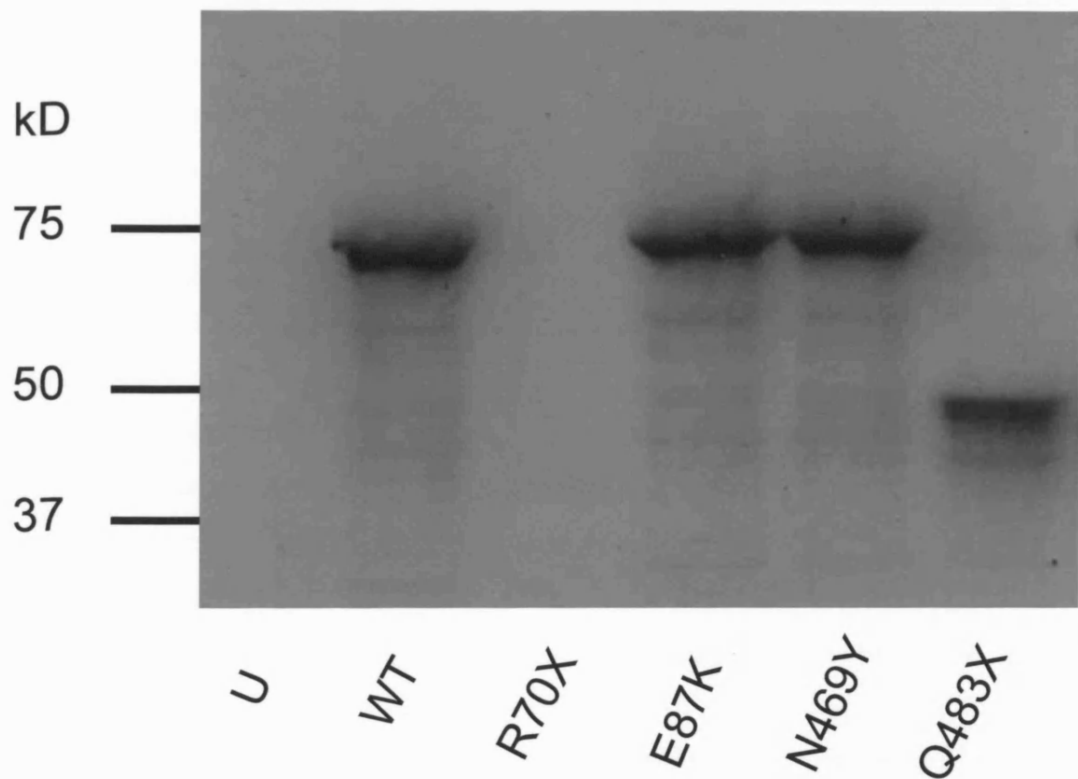


Figure 4.9 Western blot of mouse fibroblast (NIH3T3) cell extracts transfected with myc-tagged *DYM* wild type and mutant constructs using an antibody against the myc-tag (Invitrogen, Paisley, UK). This revealed bands of ~75 kD corresponding to wild type *DYM* and missense mutations E87K and N469Y, ~50 kD corresponding to the truncating mutant Q483X and no band corresponding to the truncating mutant R70X (likely because this band is too small to be seen on this gel). U corresponds to untransfected cell extracts, used as a negative control.

tagged Dymeclin was found within the cell pellet suggesting that Dymeclin is located within a membrane in the cell (see Figure 4.10). These results confirm that Dymeclin is a membrane associated protein.

4.7 Discussion

In-silico predictions suggest that Dymeclin is a transmembrane protein of approximately 75 kD which contains a myristoylation site, multiple dileucine and two dilysine motifs in addition to N- and O-glycosylation and phosphorylation sites. It has no homology to any known protein but orthologues in many other species (illustrated in Figure 4.2) suggest a conserved function through evolution. Dymeclin is predicted to lie within a membrane and indeed contains a myristoylation site which has previously been shown to assist in protein-membrane association (Boutin, 1997). Dileucine and dilysine motifs imply a function in protein sorting and targeting within the cell, particularly through the ER and endosomal/lysosomal system (Bonifacino and Traub, 2003). The presence of glycosylation sites suggests that Dymeclin undergoes post-translational modification and in addition, phosphorylation sites imply activation/inactivation by this mechanism. Phosphorylation is known to act as an important regulatory mechanism in the kinase mediated control of endosomal trafficking through phosphoinositol-III kinases (Lindmo and Stenmark, 2006). It is possible that a similar regulatory process controls the trafficking function of Dymeclin.

In-vitro translation and Western blot analysis of myc-tagged *DYM* transfected cell extracts have confirmed the molecular weight of Dymeclin to be 75 kD, corresponding to *in-silico* prediction. Truncated mutant *DYM* proteins were shown to be of a lower molecular weight as expected and missense mutants of a similar size to the wild type protein.

Sub-cellular co-localisation studies have demonstrated that Dymeclin lies within the endoplasmic reticulum and given that Dymeclin has been shown to be a transmembrane protein it is likely to lie within the endoplasmic reticulum membrane. Furthermore, Western blot studies with various solubilisation conditions have confirmed that Dymeclin is likely to be a membrane associated protein. Mutant Dymeclin appears to localise aberrantly within the cell, accumulating in the cytoplasm in a punctuate distribution. This distribution would be consistent with mutant Dymeclin being within lysosomes or ER exit sites although in order to determine the exact cellular localisation of the mutant protein, it would be necessary to perform further co-localisation studies and it is likely that *in-vivo* NMD would act to destroy mutant transcripts hence preventing their translation.

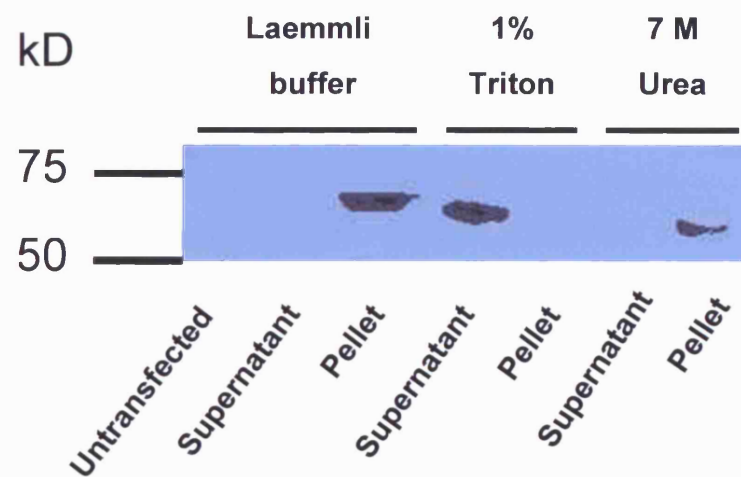


Figure 4.10.

Western blot of mouse fibroblast (NIH3T3) cell extracts transfected with myc-tagged *DYM* wild type construct using an antibody against the myc tag (Invitrogen, Paisley, UK). Cell extracts were incubated in Laemmli buffer supplemented with 50 mM NaCl and 1 % Triton, or 7M Urea.

The potential role of Dymeclin in protein transport within the cell must be considered. Localisation to the ER suggests involvement in the transport of substances through this compartment, and the presence of dileucine sites implies that transport may be to the endosomal/lysosomal system. Given that Dymeclin appears to lie mainly within the ER, it is likely to be acting as a transporter rather than being transported itself. The finding of dilated RER in the cartilage of affected individuals suggests that there is accumulation of either transported substrates or mutant Dymeclin. Abnormal folding and hence aberrant localisation may result in the case of mutant Dymeclin if NMD does not occur (Engfeldt *et al.*, 1983, Nakamura *et al.*, 1997). Unusually thick collagen fibres have been reported in DMC, and also seen in cases of known extracellular matrix deficiency (Horton and Scott, 1982). The regulation of collagen fibrils relies upon normal matrix glycoproteins (Chandrasekhar *et al.*, 1984, Vogel and Trotter, 1987, Kuijer *et al.*, 1988) suggesting that there may be a co-existing matrix abnormality in DMC. Given that Dymeclin is thought to be involved in protein transport through the ER, the hypothesis that Dymeclin is transporting matrix proteins is attractive. Brain and cartilage, the tissues affected in DMC, have high proteoglycan requirements, which may explain the tissue specific nature of the condition in contrast to the widespread expression of *DYM* transcripts. Cartilage matrix vesicles are large and thus more difficult to transport, thus possibly leaving the tissue more sensitive to a failure in transport by the absence of functioning Dymeclin.

The function of Dymeclin and its orthologues have not been demonstrated other than in *C.elegans* where a role of the orthologue *HID-1* in the regulation of synaptic vesicle release has been demonstrated (personal communication, Professor Chris Ponting, Comparative Genomics Group, MRC Functional Genetics Unit, University of Oxford) (Ailion and Thomas, 2003). Given the neurological phenotype of mental retardation in DMC, a role for Dymeclin in synaptic protein transport is attractive.

In summary, *DYM* transcripts have been shown to be widespread in contrast to the tissue specific phenotype of the condition. Dymeclin has been shown to lie within the ER membrane and is predicted to function in protein transport through this compartment, likely to the endosomal/lysosomal system. It is hypothesised that Dymeclin is transporting extracellular matrix substances with the tissue specific phenotype of DMC accounted for by the high requirement of cartilage and brain tissue for ECM components.

Chapter 5 - *In-situ* hybridisation studies

5.1 Introduction

The techniques which can be adopted to determine the expression pattern of gene transcripts and their protein products within tissues include *in-situ* hybridisation and immunohistochemistry. *In-situ* hybridisation is a method of detecting a specific DNA or RNA sequence within a tissue whereas immunohistochemistry detects the protein product using an antibody generated against an amino acid sequence specific to the protein of interest. Antibodies against Dymeclin were generated as part of this project to use for immunohistochemistry and Western blotting. An antibody was generated in rat against the C-terminal 98 amino acids of Dymeclin in addition to a peptide antibody generated in rabbit against Dymeclin amino acid residues 9-23. Despite extensive manipulation of experimental conditions, it was not possible to obtain consistent results either for immunohistochemistry or Western blotting hence experiments using these antibodies will not be discussed further. Subsequently, commercial antibodies against Dymeclin have become available but again, their experimental use has not been successful (personal communication, Dr Carolyn Dent, Department of Molecular and Medical Genetics, King's College, London).

In order to characterise the expression pattern of the *DYM* gene at a tissue level both during fetal development and adult life, *in-situ* hybridisation studies were performed on fetal and adult mouse tissues. The aim of these studies was to relate the pattern of *DYM* expression to the clinical phenotype seen in DMC. *In-situ* hybridization is a method of detecting specific DNA or RNA sequences in morphologically preserved tissue sections by hybridising the complementary strand of a nucleotide probe to the sequence of interest. The advantage of *in-situ* hybridisation is that it allows expression to be visualised in different cell types making up tissues. If RNA is extracted from tissues for Northern blotting or RT-PCR, then it is not possible to determine the exact cellular origin of the RNA. By *in-situ* hybridisation, a histological image of the expression pattern can be produced and changing expression patterns through development can be visualised. However, it is important to note that the expression pattern of RNA may not reflect that of the protein it encodes as RNA undergoes post-transcriptional regulation. Immunohistochemistry would be necessary to determine if the protein expression pattern is the same but would require an appropriate antibody.

To allow the probe to penetrate the tissue, the permeability of the cell has to be increased whilst at the same time maintaining the structural integrity of the tissue. This is achieved by digestion of the tissue section in a solution of the enzyme proteinase K

which acts to break down proteins within the tissue and allow probe penetration. RNA *in-situ* hybridisation was used for this study because, although RNA is more prone to degradation, it is present at higher levels than its homologous DNA sequence and is thus easier to detect. A single stranded oligonucleotide probe was produced because it can be easily labelled by *in-vitro* transcription, forms stable hybrids with mRNA and should give a specific signal with low background staining if washes are stringent and steps are included to digest RNAases. Single stranded RNA probes were produced by *in-vitro* transcription from a promoter sequence ligated to the 5' end of the PCR fragment by incorporation within the oligonucleotide primer sequence. In addition to producing an antisense *DYM* probe, an antisense β -actin probe was generated as a positive control and a sense *DYM* probe for use as a negative control.

A variety of labelling techniques can be used to visualise gene expression patterns by *in-situ* hybridisation. Radioactive labelling is a commonly used method but the safety considerations related to the use of radioactivity must be borne in mind. Digoxigenin labelling was used for this study, a method which employs digoxigenin, a small steroid easily attached to RNA nucleotides (usually deoxy-uridine triphosphate (dUTP)). Subsequent detection employs an anti-digoxigenin antibody and alkaline phosphatase based detection system by incubation with 5-bromo-4 chloro-3-indolyl phosphate (BCIP) and Nitro blue tetrazolium (NBT).

Initial attempts at *in-situ* hybridisation were unsuccessful hence the parameters influencing the technique will be discussed as these had to be manipulated in order to achieve the results presented in this chapter. If an RNA probe is to be used, it is essential that the tissue has been prepared appropriately in order to preserve RNA which may easily be degraded by RNAases. All reagents and equipment must be treated to ensure they are RNase free and tissues optimally fixed. Initial experiments using tissues already available in the laboratory were not successful hence further tissue was obtained to ensure adequate fixation with preservation of target nucleic acids in addition to cell and tissue morphology.

The ability of the target RNA sequence to bind to the labelled probe is dependent upon the complementarity and length of the probe, hybridisation temperature, pH, solvent concentration and the presence of monovalent cations (eg. sodium). Maximum hybridisation rate can be achieved with a longer probe but the probe needs to be small enough to diffuse into tissues. Monovalent cations interact electrostatically with nucleotides such that electrostatic repulsion between the two strands decreases with increasing salt concentration hence increasing the stability of the duplex. Organic solvents (such as formamide) reduce the thermal stability of RNA thus allowing lower

hybridisation temperatures to be used. Dextran sulphate can be used to increase the apparent probe concentration as it is strongly hydrated in aqueous solution thus limiting the availability of water to other macromolecules.

Parameters which were varied before optimal conditions were obtained included concentration of proteinase K used for predigestion of tissues, choice of probe, hybridisation temperature, wash stringency (temperature, salt and solvent concentration) and dextran sulphate concentration.

5.2 Preparation of mouse tissue

Mice on a C57BL/6 genetic background were used for this study. Embryos were prepared on the desired day of pregnancy after sacrificing the mother by cervical dislocation. Day 20 embryos were decapitated prior to fixing. Adult mouse tissues were decalcified prior to fixing by immersion in a solution of 10 % EDTA for 24 hr. Mouse fetuses of 16 and 20 days gestation and adult mouse brain, muscle, paw and lower limb were transected and fixed in 10 % formalin in saline for 24-48 hours. Dehydration was performed overnight in a series of solutions of increasing ethanol concentration. Tissue was embedded in paraffin before being sectioned at 4 μ n onto silane slides (Sigma-Aldrich, Cambridge, UK).

5.3 Probe design

Riboprobe is the term used to describe an RNA sequence complementary to the nucleotide sequence to which it will bind as a probe. Riboprobes were designed against the 3' region of the *DYM* transcript which is most conserved through species as illustrated in Figure 5.1. This region shows a high level of homology between human and mouse sequences (89 % sequence identity). The region of highest homology was chosen so the probe could be used on both mouse and human tissues, although only mouse tissues were used in this study. Riboprobe generation is described in section 2.2.18.1. In order to detect a signal from *in-situ* probing, the mRNA of the gene of interest must be abundant and the *DYM* transcript has been shown to be strongly expressed in all tissues tested (see section 4.3). In addition to the anti-sense probe, a sense probe was used as a negative control and a β -actin probe as a positive control. *In-situ* hybridisation was carried out as described in section 2.2.18.2.

```

ACGTGACCACTGGCTGCGGCGCGCGCTGGGGTCTGGTCCCTAGCCGC Human
-----CCAC Mouse
**
GTCTCGGTGGCTTCCCTCTTCTCTCGGCGCGCTGGGGTCCGCGCGA 100
G----CGTCCGCTTCCCGCTTCTGACCCCAACTCCT--CCTGGCTCTCG 48
**
CGCGCGCGCGCGCTGGAGCGAGCGCGGGCTCGAGCTGCAGCGCGGACCGGA 150
CAGCGAC-----AGCGG--CTGGGCTTGGAACTCGGCTGGGAACCGGA 58
**
CGCGGATCTGTACCGCTGAGAGCTGGAAACATGAGGCGCTGAGCGCGTG 200
GTCT-----CTGAGCTACGGAAACATGAGGCGCTTAAAGCGCGCG 126
**
TGCGGCACTGGGCTGGGCGCGGACAGCGAGCTTCTCTGACCGCTCTGC 250
TGCAAT-ACCGCGCGCGCTGTG--CGCGGCGG--CTCGGCTTCCCT-TGC 149
**
CAGCTCTGCATCGCTCGCGCGGCTCGGTGGAGCTGGAGCAGATCCCGAGC 300
TGCTTC-----CGCGCGCGGCT--GAAAGCGGAGT-----CGCGG 292
**
CGGCTGAGAGAGTTGTCTTTTGGAAATCGAGGTTTAAAGGACAAATTATC 350
CGGCTGAGAGAGTTGTCTTTTGGAAATCGAGGTTTAAAGGAGAGGCTGTC 293
**
TGCTTAAGCTAGAGATGGGATCGAATAGAGAGAGAGTGGGATCTTCC 400
TACC--ACGTAGAGATGGGATCGAATAGAGAGAGAGATGAGTATCTTCC 301
**
TAAAAATTA TACTTAAAAAATTATTA TAA TAA TAA TAA TAA TAA 450
GAAAAACGAGTACTTAAAAAATTGCTATGCGCTGAGTGTATATCTGAAA 351
**
ATGAGCGCTTGGGAATGAGTTTCTTATTTTCTTCTTCTTCTTCTTCTT 500
AGGAGCGCTTCTGGGAATGAGTTTCTTATTTCTTCTTCTTCTTCTTCTT 401
**
AGCATAGTGAATTAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTT 550
AGCATAGTGAATTAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTT 451
**
GTATTAGTAAAAAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTT 600
GTATTAGTAAAAAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTT 501
**
AGGTCTTCTTCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTT 650
AGGTCTTCTTCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTT 551
**
AAGCAGATCTTCAATTTGGGAGAGACAGCAATGCTTTGTTATTATTGCTG 700
AAGCAGATCTTCAATTTGGGAGAGAGACAGCAATGCTTTGTTATTATTG 601
**
TTTCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCT 750
TTTCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCT 651
**
ATTATATATGAAGAAAAATTTCTTCTTGAAGATCTTGAAGATCTTGAAG 800
ATTATATATGAAGAAAAATTTCTTCTTGAAGATCTTGAAGATCTTGAAG 701
**
GATCTTTTGAAGAAAAATTTCTTCTTGAAGATCTTGAAGATCTTGAAGAT 850
GATCTTTTGAAGAAAAATTTCTTCTTGAAGATCTTGAAGATCTTGAAGAT 751
**
TGCATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGAT 900
TGCATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGAT 801
**
TGGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 950
TGGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 851
**
AGCATCAGGATCAGAGATTTGATGCGAGGTCATGCTCTCTCTATACAGAG 1000
AGCATCAGGATCAGAGATTTGATGCGAGGTCATGCTCTCTCTATACAGAG 901
**
GAAATCTGTGAAGAGCTTATATATATATATATATATATATATATATAT 1050
TAAATCTGTGAAGAGCTTATATATATATATATATATATATATATATAT 951
**
CTCTTCAGGCGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1100
CTCTTCAGGCGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1001
**
CTTTATGGAATCTGATCAGGAGTAGCAACAGGAGCTTCTGAGCTGTCTT 1150
CTGTATGGAATCTGAGCAGGAGTAGCAACAGGAGCTTCTGAGCTGTCTT 1051
**
ACTAGGTGCTGTGGGCGAGCAAGCGGCTGCTCTCTCTCTCTCTCTCTCT 1200
GCTAGGCGGCGCGAGGAGCAAGCGGCTGCTCTCTCTCTCTCTCTCTCT 1101
**
CTCTGCGCAACAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1250
CTCTGCGCAACAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1151
**
GATGCTCTCAATGCGCAACAGGCTTCTCTCTCTCTCTCTCTCTCTCTCT 1300
GATGCTCTCAATGCGCAACAGGCTTCTCTCTCTCTCTCTCTCTCTCTCT 1201
**
GAAGACACAGATAGAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1350
GAAGACACAGATAGAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1251

```

Figure 5.1. Alignment of *DYM* cDNA sequence from H.sapiens and M. musculus. Consensus sequence is indicated by stars. The riboprobe sequence used is in bold, within the most conserved region of the sequence (89 % homology between human and mouse). Produced using ClustalW (www.ebi.ac.uk/Tools/clustalw2).

5.4 Embryology and histology of the skeletal system

In order to interpret *in-situ* hybridisation studies of the *DYM* transcript in fetal mouse tissues, it is necessary to understand the processes involved in development of the skeletal and nervous systems, the phenotypically relevant tissues.

The skeleton originates from three different embryonic tissues; somites, mesoderm and neural crest. The somites, transient developmental structures of the embryo located on either side of the neural tube, give rise to the sclerotome which forms the axial skeleton (vertebrae and ribs). Mesenchymal cells from the lateral plate mesoderm form the appendicular skeleton (limbs). Skeletal development can be divided into three stages; formation of the template which will go on to form the skeletal elements is followed by condensation then differentiation of mesenchymal precursor cells into prechondrocytes. Precartilaginous condensations of mesenchyme form at future sites of bone from 13 days post-conception in the mouse embryo (Theiler stage 21). The cranial vault and facial bones undergo intramembranous ossification by direct differentiation of mesenchymal cells into osteoblasts and bone matrix.

Precartilaginous condensations are replaced by cartilage by 14 days post-conception (Theiler stage 22) except for the more distal limb elements (metacarpals, metatarsals and phalanges) which remain as mesenchymal condensations. Centres of ossification can be visualised from 15 days post-conception (Theiler stage 23) in the cranial vault and long bones and gradually increase in size through development. By 16 days post-conception, the earliest gestation at which *in-situ* hybridisation was performed for this study, ossification has commenced in all bones except the phalanges which ossify by day 18. We would thus expect all bones except the phalanges to have undergone ossification in the fetal mouse tissues used in this study.

Cartilage can be classified based on its morphologic appearance into hyaline, elastic and fibrocartilage. Hyaline cartilage is the most common type and forms the temporary skeleton in the embryo as well as the epiphyseal plate. At a histological level, cartilage can be seen to comprise extracellular matrix containing lacunae within which are chondrocytes whereas chondroblasts are located in the periphery. Both chondroblasts and chondrocytes secrete the extracellular matrix. High levels of extracellular matrix are seen surrounding chondrocytes and form the territorial matrix which shows strong basophilic staining with haematoxylin and eosin. Between the chondrocytes is the interterritorial matrix which stains less strongly due to its lower proteoglycan level.

In bone, the extracellular matrix is calcified and contains osteocytes (a fully differentiated cell formed from an osteoblast) within lacunae. Endochondral ossification occurs at the growth plate within long bones and contains chondrocytes at different

stages of differentiation. Quiescent chondrocytes form the resting zone, where there is a high ratio of extracellular matrix to cell volume. As the cells proliferate, they become flattened and organised into columns before progressing to hypertrophy with a large cell volume. At the stage of terminal differentiation, matrix vesicles bud off from the plasma membrane of the chondrocyte and initiate mineralisation. Blood vessel invasion occurs, the extracellular matrix becomes ossified and the bone marrow cavity is formed. Chondrocyte differentiation in the growth plate is illustrated in Figure 5.2 and has been shown to be abnormal in DMC patients with chondrocytes failing to form columns but arranged irregularly with some seen to degenerate as previously shown in Figure 1.6 (Engfeldt *et al.*, 1983, Nakamura *et al.*, 1997, Horton and Scott, 1982). The expression pattern of *DYM* in the cartilage and in particular, the growth plate was therefore studied in depth in order to relate expression patterns to the histological phenotype.

In early human postnatal life, chondrocytes in the epiphysis undergo a similar differentiation process and are replaced by osteoblasts and invading vessels which form the secondary ossification centre which separates the articular cartilage at the surface of the bone and the growth plate cartilage. However in mice, articular cartilage begins to form during embryogenesis at synovial joints. Articular cartilage comprises four zones- the superficial or tangential zone, intermediate or transitional zone, radial zone and the zone of calcified cartilage, as depicted in Figure 5.3.

5.5 Embryology and histology of the nervous system

The mouse nervous system develops from the neural tube (derived from ectoderm) which at around nine days post-conception closes and simultaneously three primary brain vesicles form, of which the walls are termed prosencephalon, mesencephalon and telencephalon. They go on to form the primitive forebrain, midbrain and hindbrain respectively. The cerebellum becomes obvious at around 14 days post-conception and by 15 days the cerebral hemispheres form which can be seen to form different layers by 17 days. Within the cerebral hemispheres at this stage are the lateral ventricles subdivided into anterior and posterior horns. The primitive hindbrain forms the pons, cerebellum and medulla oblongata.

Histologically, the cells of the central nervous system can be broadly divided into neurons and glial or supporting cells. Neurons contain large amounts of endoplasmic reticulum and ribosomes because they are metabolically very active hence may be sensitive to lack of Dymeclin if it is acting as a substrate transporter through the ER. Neurons can be subdivided and include, for example, pyramidal cells in the cerebral cortex and Purkinje cells within the cerebellum. Glial cells themselves can be

Resting
chondrocytes

Proliferating
chondrocytes

Hypertrophic
chondrocytes

Zone of
ossification

Zone of
vascular
invasion

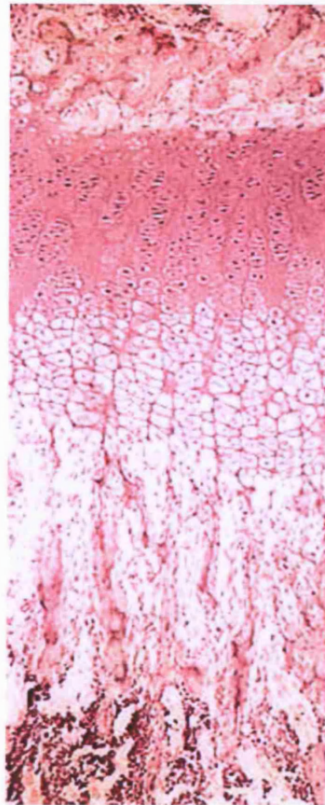


Figure 5.2. Light microscope image of the growth plate. Chondrocytes progress through stages from resting to proliferating, before becoming hypertrophic. Image from www.bioscience.org.

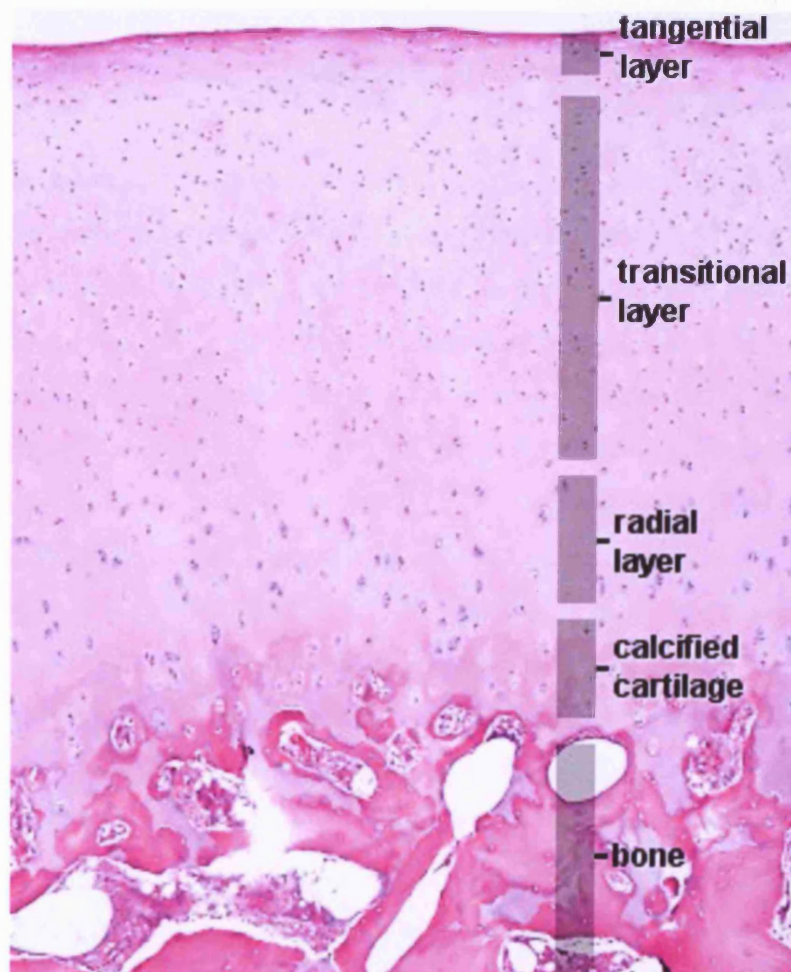


Figure 5.3. Light microscopic image of articular cartilage showing the different layers. Image from www.lab.anhb.uwa.edu.au.

subdivided into astrocytes, oligodendrocytes and ependymal cells which line the ventricles. Individuals with DMC have microcephaly with a small brain which has been shown by computerised tomography to be structurally normal (Girisha *et al.*, 2007). There are no reports in the literature of post-mortem findings on a DMC affected patient hence it is not known if there are any histological abnormalities of the central nervous system. However, mental retardation of variable severity and unknown cause is seen in DMC although not in the allelic disorder, SMC.

5.6 *DYM* expression in fetal mouse tissues

Day 16 mouse embryos stained with haematoxylin and eosin and *in-situ* hybridised antisense β -actin, *DYM* sense and antisense probes are shown in Figures 5.4, 5.5 and 5.6. These images show that the *DYM* transcript is expressed widely in many different tissue types throughout the embryo but most strongly expressed in cartilage, bone, brain, liver and epithelium. Very high expression levels are seen in the cells surrounding the ventricles in the brain, in all bony structures and in epithelium. Figures 5.5 and 5.6 illustrate a longitudinal section through the hind limb of a day 20 mouse embryo with *DYM* expression shown in Figures 5.5d and 5.6d. *DYM* expression is highest in the epithelium and bony structures in comparison to β -actin which shows maximal expression in muscle. Although the skin appears grossly normal in DMC affected individuals, abnormalities have been shown at a cellular level which include dilated RER and multiple vacuoles (El Ghouzzi *et al.*, 2003). High expression levels of *DYM* are seen within phenotypically relevant tissues (cartilage and brain) as well as in the epithelium which has been shown to be abnormal histologically. High expression levels have been demonstrated within the fetal liver despite the fact that there is no known structural or functional hepatic abnormality in DMC.

5.7 *DYM* expression in adult mouse tissues

Sagittal sections through adult mouse brain in Figure 5.7 and 5.8 show *DYM* RNA expression throughout the brain but with particularly high levels in the molecular layer of the cerebellum and surrounding the ventricles in the ependymal cells. Expression of *DYM* within the cerebellum is highest within the Purkinje cell layer with a lower level of expression in the granular layer (see Figure 5.9), although this may represent to some extent the difference in cell density.

Transverse sections through adult mouse paw (see Figure 5.10) show *DYM* expression in epithelium, muscle, articular cartilage and in osteocytes within lacunae in bone. Figure 5.11 shows sections of muscle from adult mouse, with *DYM* seen to be expressed

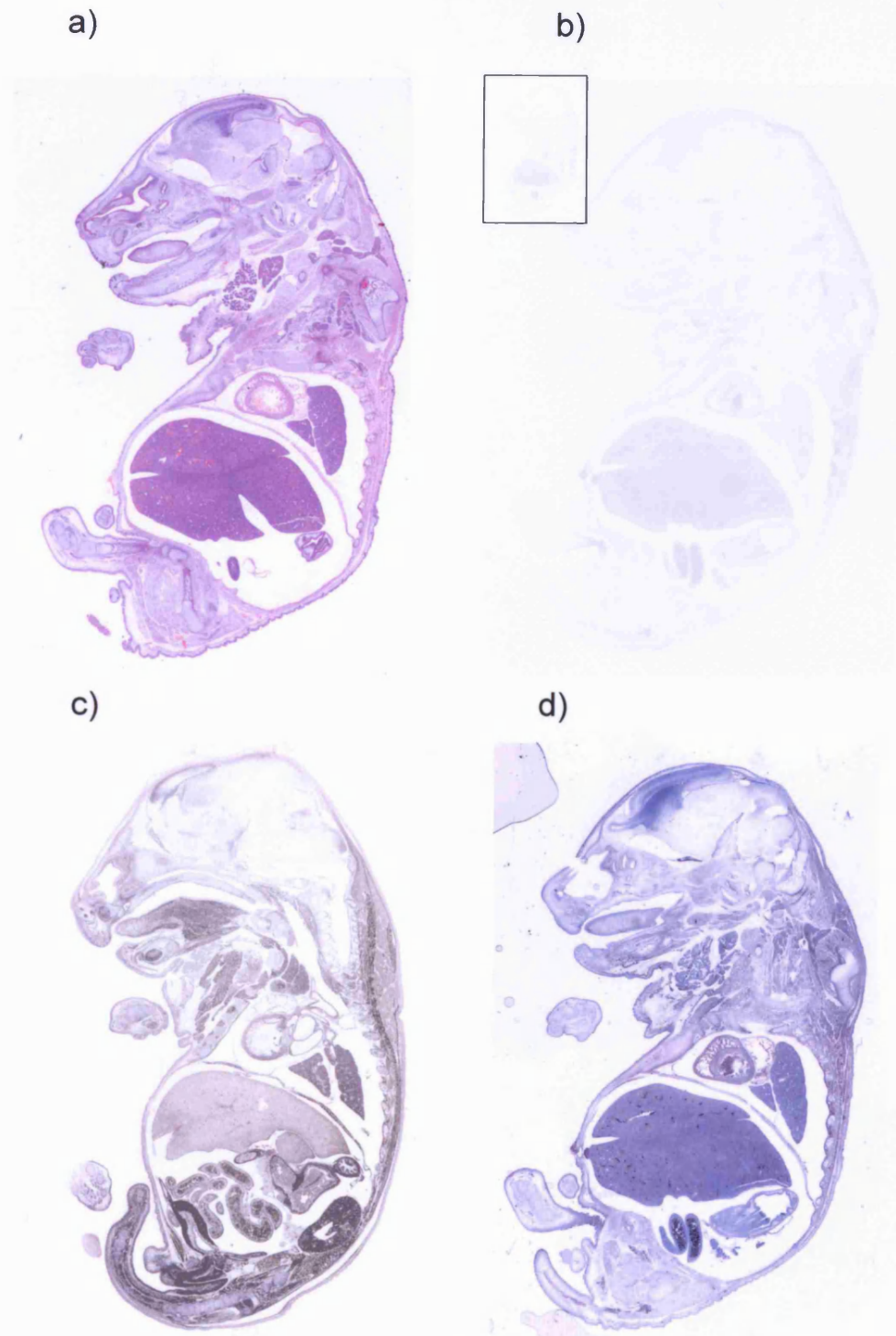


Figure 5.4. Day 16 mouse embryos, sagittal section, scanned image x 6.

- a) Haematoxylin and eosin.
- b) In-situ hybridisation – *DYM* sense probe illustrating low background level of staining. Brightness level has been reduced to allow structures to be seen. Inset picture shows original image.
- c) In-situ hybridisation – *Actin* antisense probe- expression is widespread but maximal in muscle.
- d) In-situ hybridisation – *DYM* antisense probe- maximal expression levels are seen in bone, brain, epithelium and liver.

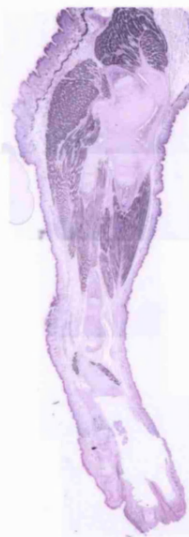
a)



b)



c)



d)



Figure 5.5. Day 20 mouse embryo limb, sagittal section, x 2.5.

- a) Haematoxylin and eosin.
- b) In-situ hybridisation – *DYM* sense probe -illustrating low background level of staining. Brightness level has been reduced to allow structures to be seen. Inset picture shows original image.
- c) In-situ hybridisation –*Actin* antisense probe- expression is widespread but maximal in muscle.
- d) In-situ hybridisation – *DYM* antisense probe- maximal expression levels are seen in bone and epithelium.

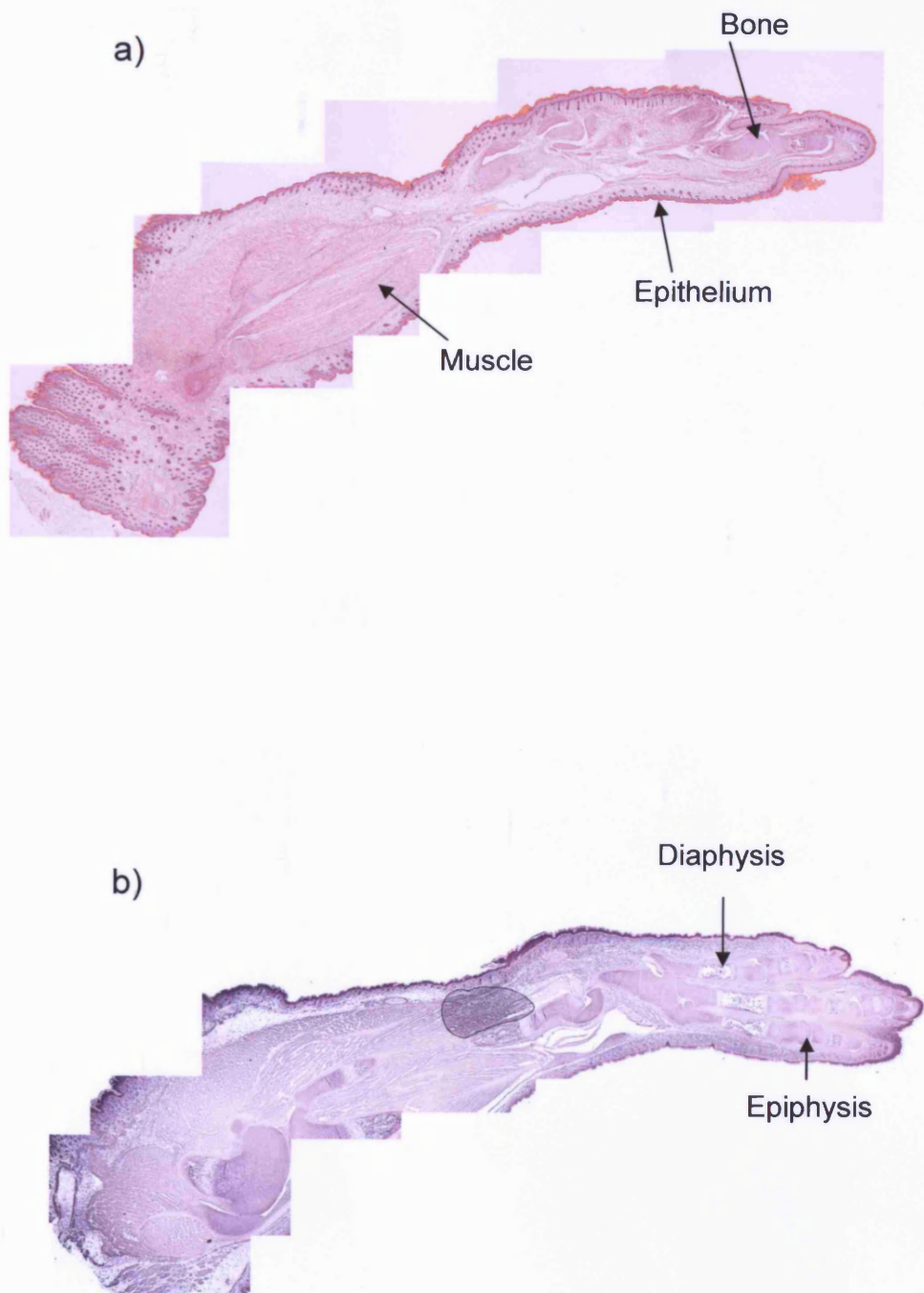


Figure 5.6. Day 20 mouse embryo, limb section, x 4.

- a) Haematoxylin and eosin. Muscle and epithelium are labelled.
- b) In-situ hybridisation – *DYM* antisense probe. The diaphysis and epiphysis of the developing bone are identified. High expression levels are seen in bone and epithelium.

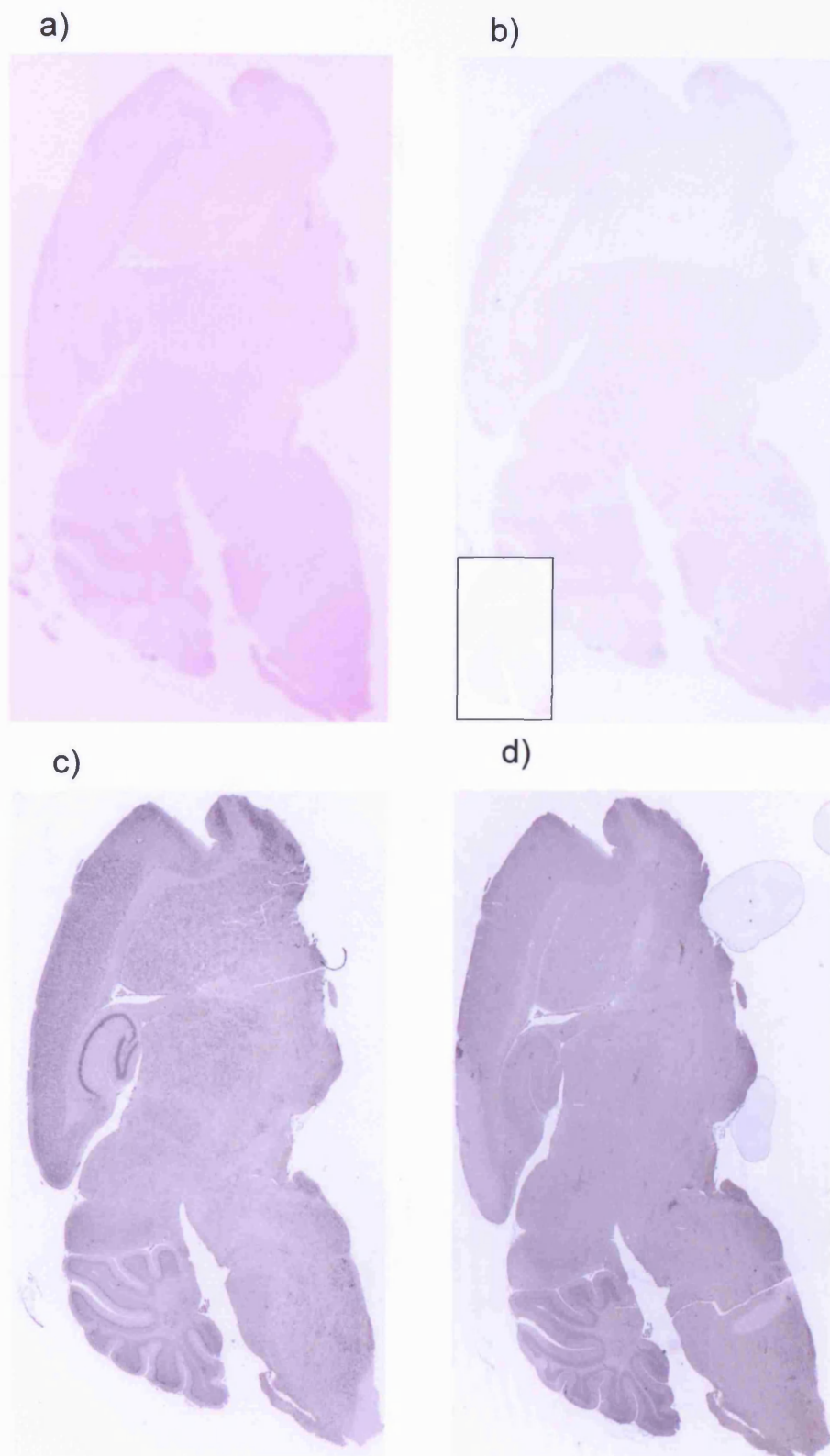


Figure 5.7. Adult mouse brain sagittal section, scanned image x 6.

- a) Haematoxylin and eosin.
- b) In-situ hybridisation – *DYM* sense probe-illustrating low background level of staining. Brightness level has been reduced to allow structures to be seen. Inset picture shows original image.
- c) In-situ hybridisation –*Actin* antisense probe.
- d) In-situ hybridisation – *DYM* antisense probe.

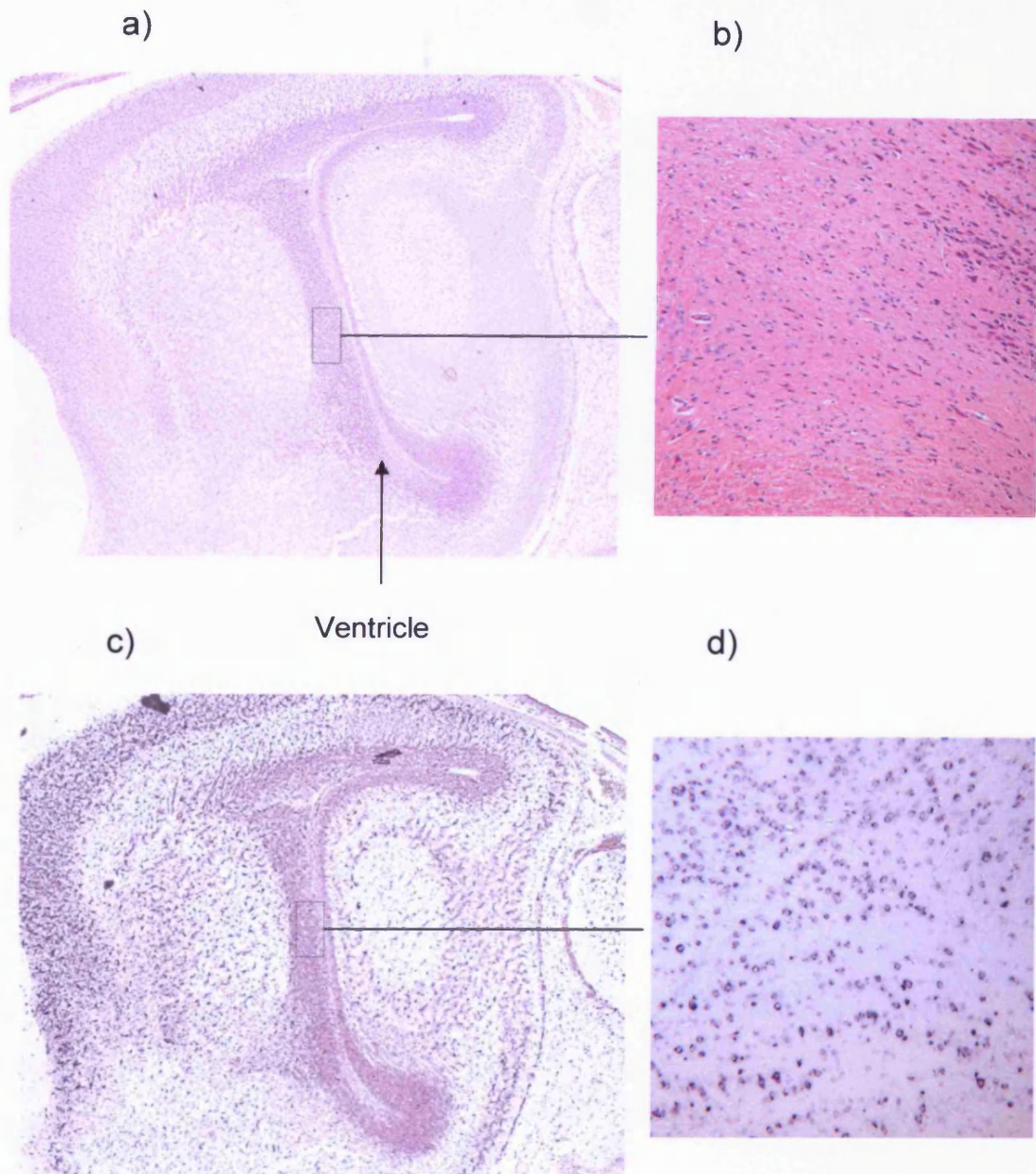


Figure 5.8. Adult mouse, sagittal section of brain.

- a) Haematoxylin and eosin x 2.5, ventricle is indicated.
- b) Haematoxylin and eosin x 10.
- c) In-situ hybridisation – *DYM* antisense probe x 2.5, high levels of expression are seen surrounding the ventricles.
- d) In-situ hybridisation – *DYM* antisense probe x 10, ependymal cells show high expression.

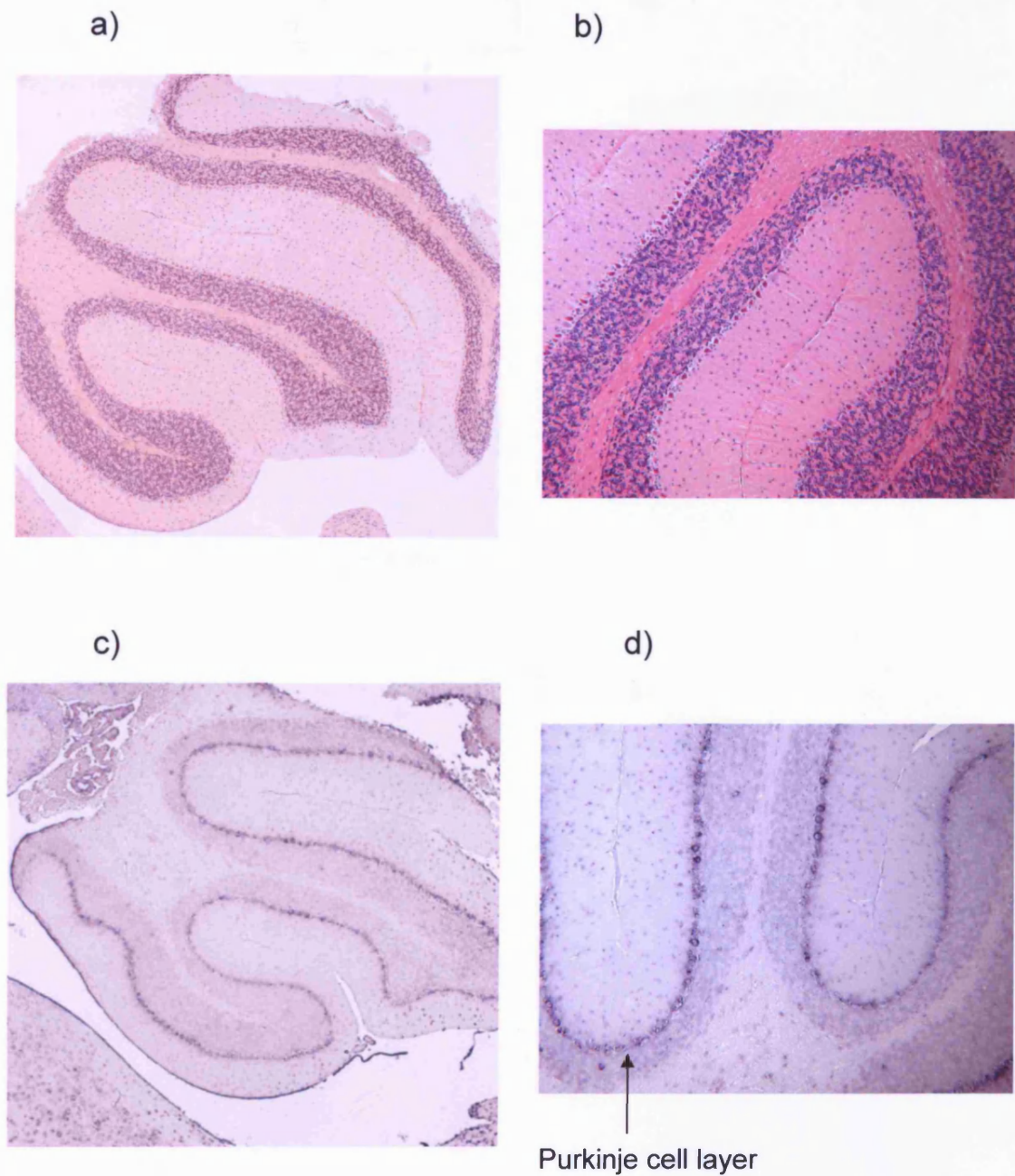


Figure 5.9. Adult mouse brain, sagittal section of cerebellum.

- a) Haematoxylin and eosin x 2.5.
- b) Haematoxylin and eosin x 10.
- c) In-situ hybridisation – *DYM* antisense probe x 2.5 showing high level of expression in molecular and Purkinje cell layer.
- d) In-situ hybridisation – *DYM* antisense probe x 10 demonstrating high level of expression in Purkinje cell layer.

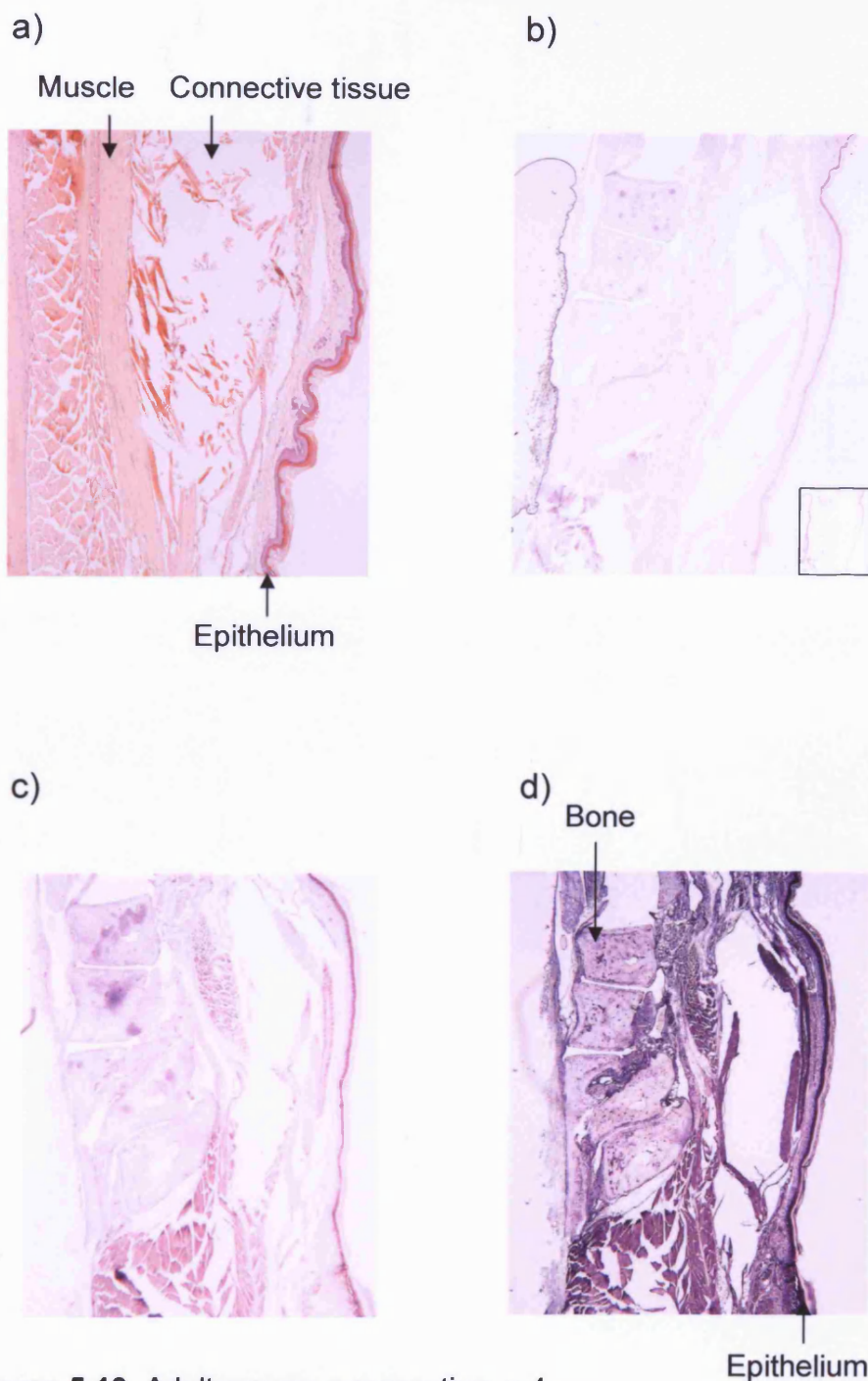


Figure 5.10. Adult mouse, paw section, x 4.

- a) Haematoxylin and eosin.
- b) In-situ hybridisation – *DYM* sense probe, illustrating low background level of staining. Brightness level has been reduced to allow structures to be seen. Inset picture shows original image.
- c) In-situ hybridisation - *Actin* antisense probe.
- d) In-situ hybridisation – *DYM* antisense probe showing high expression in epithelium, articular cartilage and within lacunae in bone.

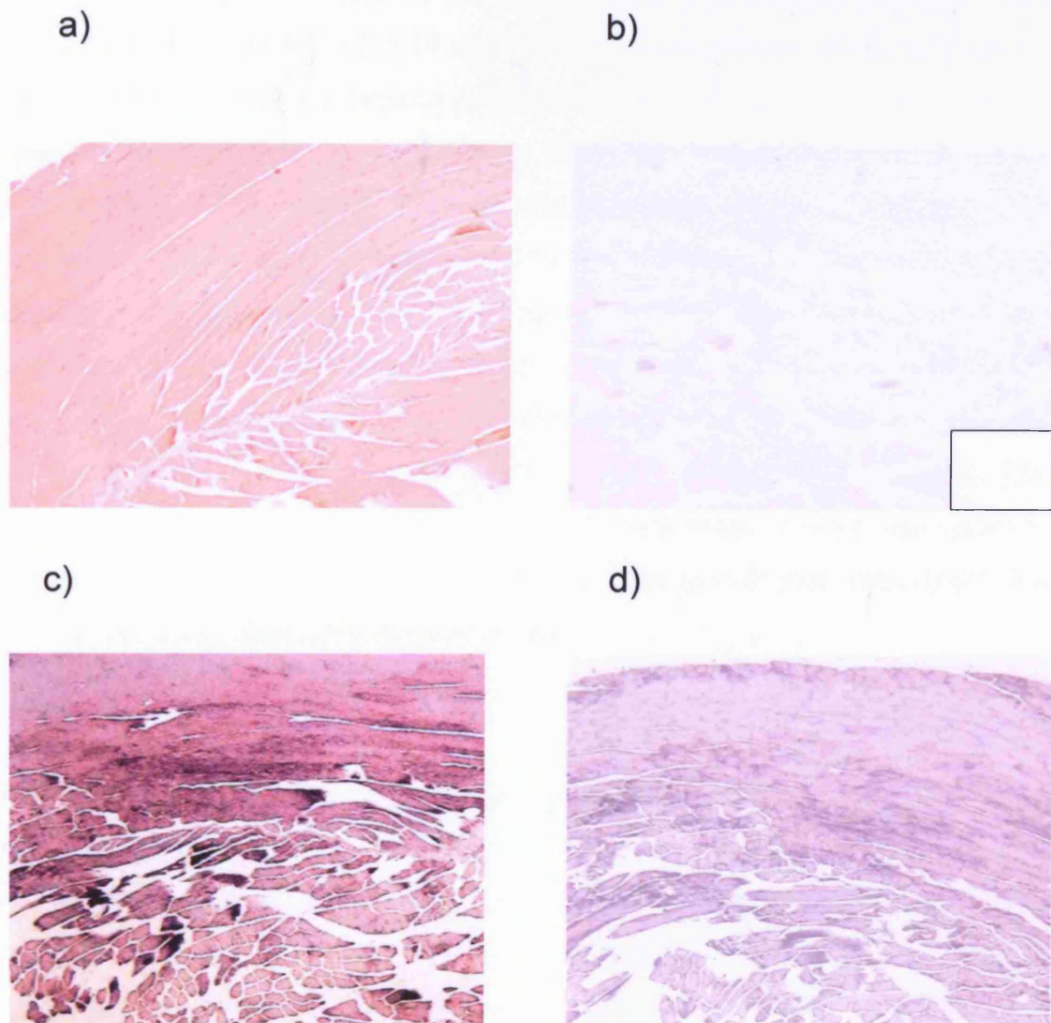


Figure 5.11. Adult mouse, muscle, x 4.

- a) Haematoxylin and eosin
- b) In-situ hybridisation – *DYM* sense probe illustrating low background level of staining. Brightness level has been reduced to allow structures to be seen. Inset picture shows original image.
- c) In-situ hybridisation – *Actin* antisense probe, showing high expression levels in muscle.
- d) In-situ hybridisation – *DYM* antisense probe, some expression is seen in muscle.

at lower levels than *β-actin*. Figure 5.12 illustrates sections through an adult mouse knee joint with *DYM* expression seen in the growth plate and articular cartilage. Higher magnification (see Figures 5.13, 5.14 and 5.15) shows expression levels are greatest in resting chondrocytes and the zone of ossification within the growth plate, in articular cartilage in the tangential and calcification layers and within chondrocytes lying in lacunae within cartilage matrix. There is also expression within macrophages in the bone marrow. The developing femur of a D16 mouse embryo (see Figure 5.16) shows expression of *DYM* in chondrocytes at all stages of differentiation with highest levels in resting and hypertrophic chondrocytes. Figures 5.17 and 5.18 show images which have been manipulated using Image Arithmetic software such that the image is reversed and areas with high levels of the RNA of interest are seen as light rather than dark. These images illustrate clearly the high levels of *DYM* expression in bone and epithelium within the tangential layer of articular cartilage and the growth plate compared to high expression in muscle seen using the *β-actin* probe.

5.8 Discussion

In-situ hybridisation studies have shown that *DYM* is most strongly expressed in cartilage and bone, particularly in resting and hypertrophic chondrocytes in the growth plate, in the tangential and calcification layers in articular cartilage and within chondrocytes within lacunae embedded in the territorial matrix. Expression levels are also high in the brain, particularly in the ependymal cells surrounding the ventricles and in the Purkinje cell layer of the cerebellum. Fetal liver also demonstrated high *DYM* expression at day 16 although by day 20, expression had decreased (data not shown). High Dymeclin expression levels have been shown in cartilage and epithelium of the skin, both tissues which have been shown to be abnormal at a cellular level. Given that no histological studies have been performed on DMC affected brain, it is difficult to correlate these findings with the disease phenotype of mental retardation.

Images of *in-situ* hybridisation using a *DYM* gene probe in mouse brain from the Allen atlas of mouse brain (www.brain-map.org) are shown in Appendix D, Figures D (i) - (iii). These images are taken from 56 day old adult male mice and show detailed expression patterns by comparison to the annotated diagrams of brain structure provided. These images corroborate the finding of high *DYM* expression in the Purkinje layer of the cerebellum and give high resolution images of *DYM* expression throughout the brain.

The expression patterns of *EGF-containing fibulin-like extracellular matrix protein*

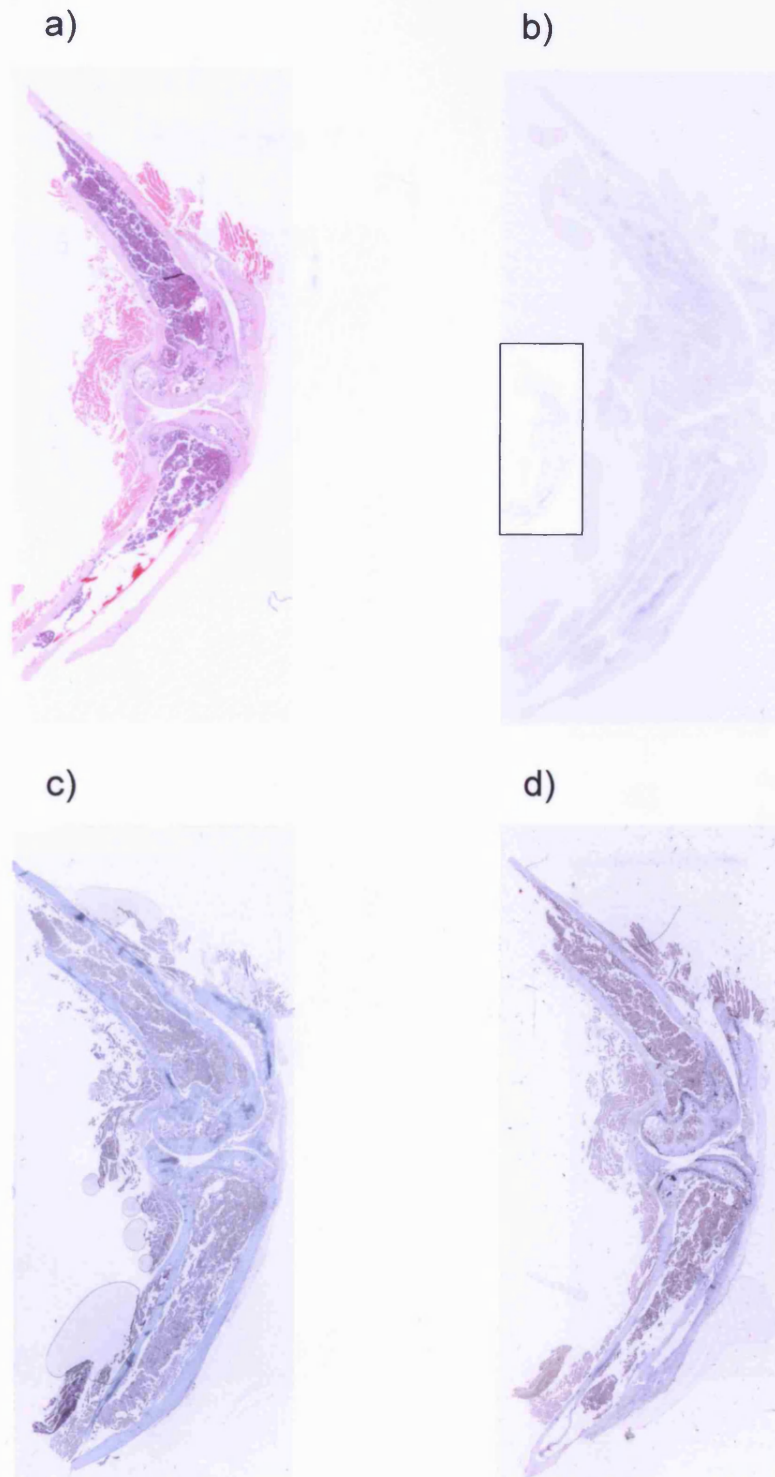


Figure 5.12. Adult mouse, sagittal joint section, scanned image x 6.

- a) Haematoxylin and eosin.
- b) In-situ hybridisation – *DYM* sense probe. Brightness level has been reduced to allow structures to be seen. Inset picture shows original image.
- c) In-situ hybridisation – *Actin* antisense probe.
- d) In-situ hybridisation – *DYM* antisense probe. High expression is seen in the growth plate and articular cartilage.

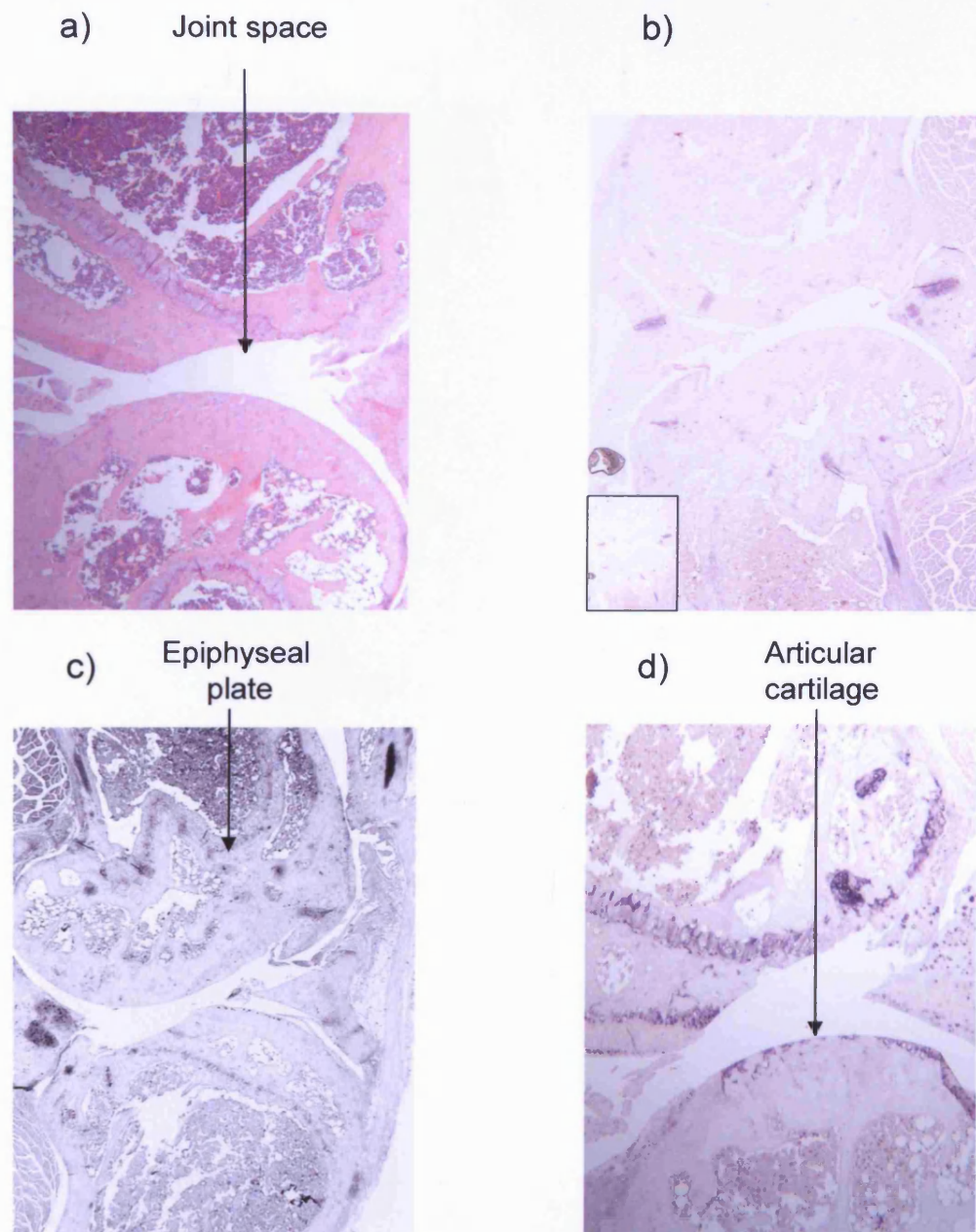


Figure 5.13. Adult mouse, sagittal joint section, x 2.5.

- a) Haematoxylin and eosin.
- b) In-situ hybridisation – *DYM* sense probe. Brightness level has been reduced to allow structures to be seen. Inset picture shows original image.
- c) In-situ hybridisation – *Actin* antisense probe, expression is highest in the bone marrow and muscle with some expression in cartilage.
- d) In-situ hybridisation – *DYM* antisense probe. High expression levels are seen in the epiphyseal plate and articular cartilage.

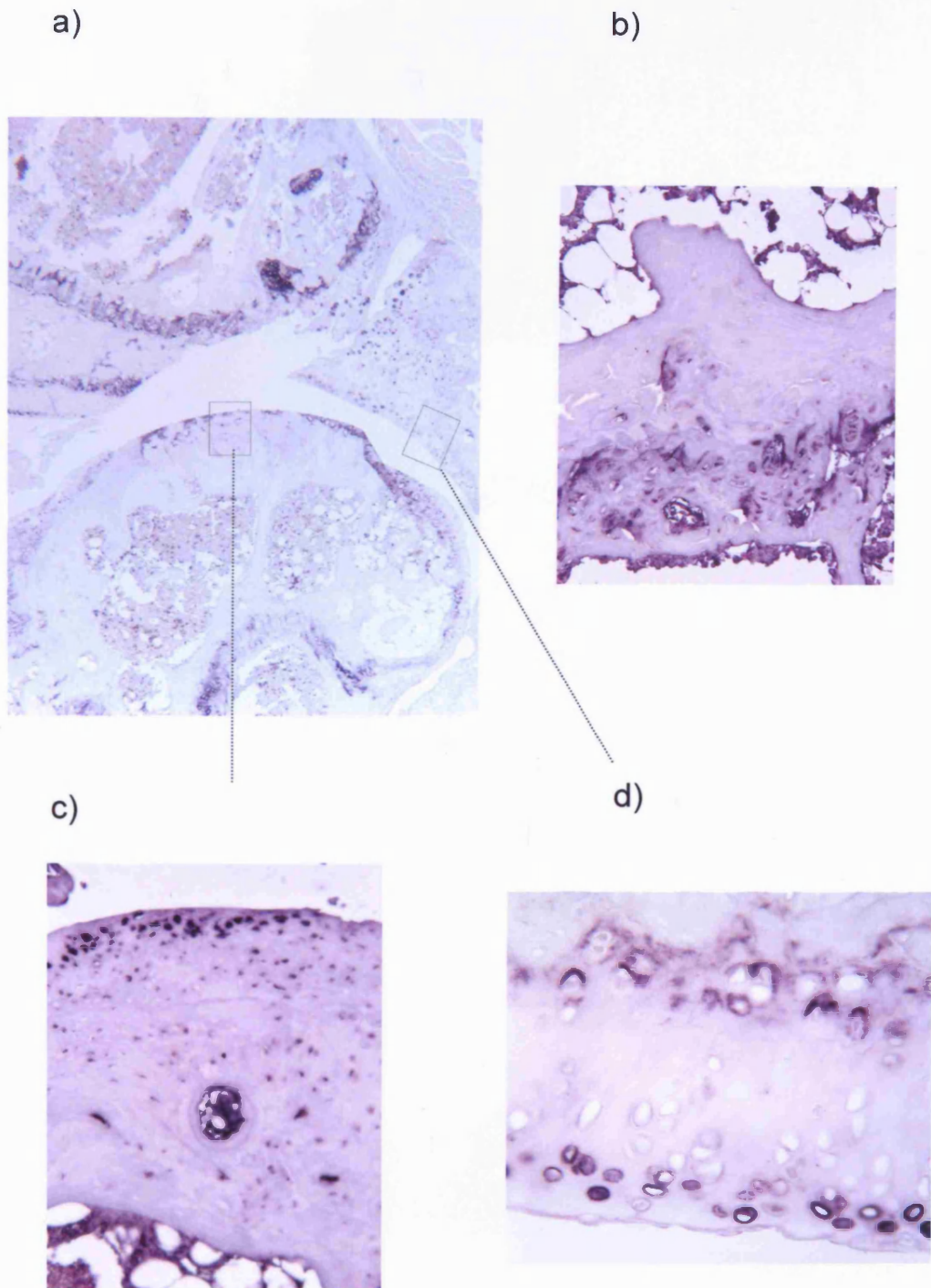
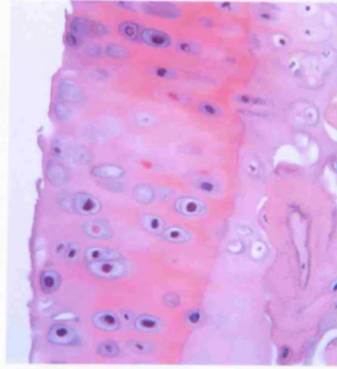


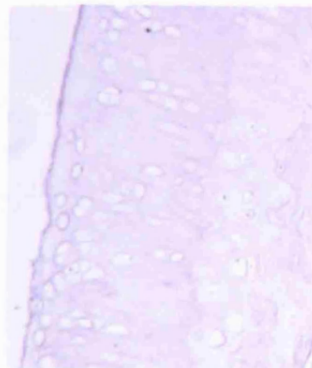
Figure 5.14. Adult mouse, knee joint.

- a) In-situ hybridisation – *DYM* antisense probe, x 2.5.
- b) In-situ hybridisation – *DYM* antisense probe, growth plate x 10, there is widespread expression throughout the growth plate.
- c) In-situ hybridisation – *DYM* antisense probe, articular cartilage x 10, expression is highest in the tangential layer.
- d) In-situ hybridisation – *DYM* antisense probe, articular cartilage x 40, expression is seen in the tangential and radial layers.

a)



b)



c)

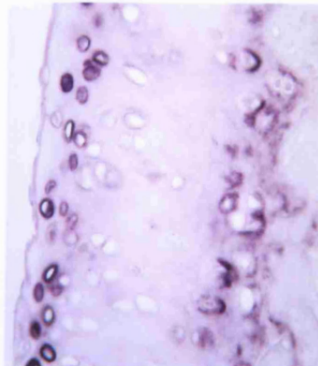


Figure 5.15. Adult mouse articular cartilage, x 40.

- a) Haematoxylin and eosin.
- b) In-situ hybridisation – *DYM* sense probe.
- c) In-situ hybridisation – *DYM* antisense probe. High expression is seen in tangential and radial layers.

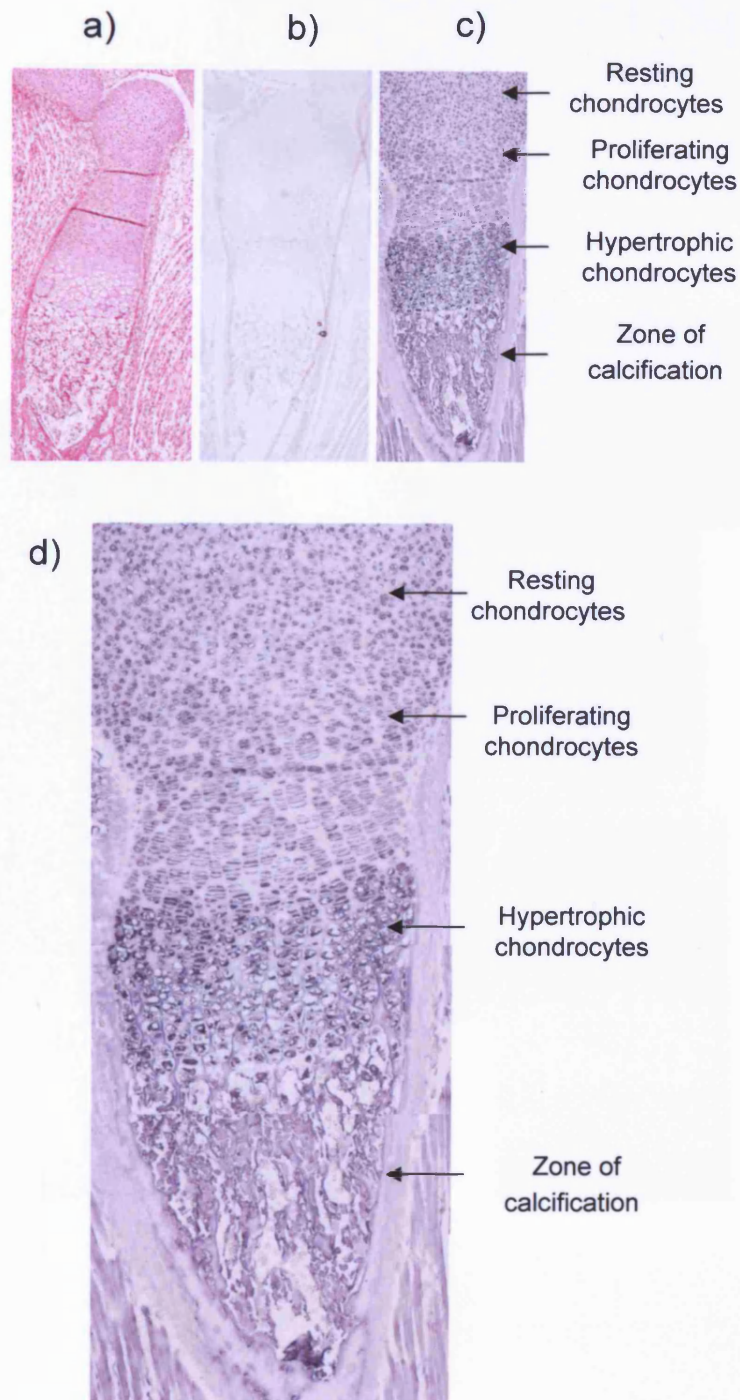


Figure 5.16. Day 16 mouse embryo femur, x 4.

- a) Haematoxylin and eosin.
- b) In-situ hybridisation – *DYM* sense probe showing low background level of staining.
- c) In-situ hybridisation – *DYM* antisense probe.
- d) In-situ hybridisation – *DYM* antisense probe. Image magnified x 2.5. *DYM* is expressed throughout the bone but at highest levels in hypertrophic chondrocytes.

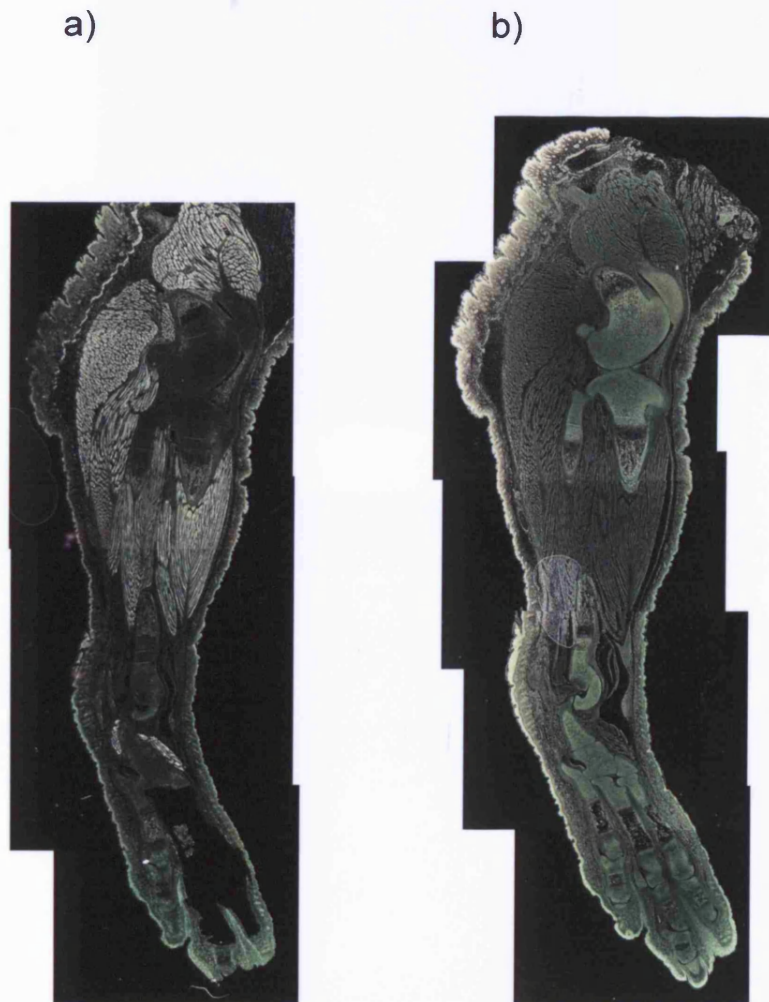


Figure 5.17. D20 mouse embryo limb, sagittal section, x 2.5
Images manipulated using Image arithmetic subtract function with
lightest areas representing highest level of expression.

- a) In-situ hybridisation – *Actin* antisense probe showing maximal expression in muscle.
- b) In-situ hybridisation – *DYM* antisense probe showing maximal expression in bone and epithelium.

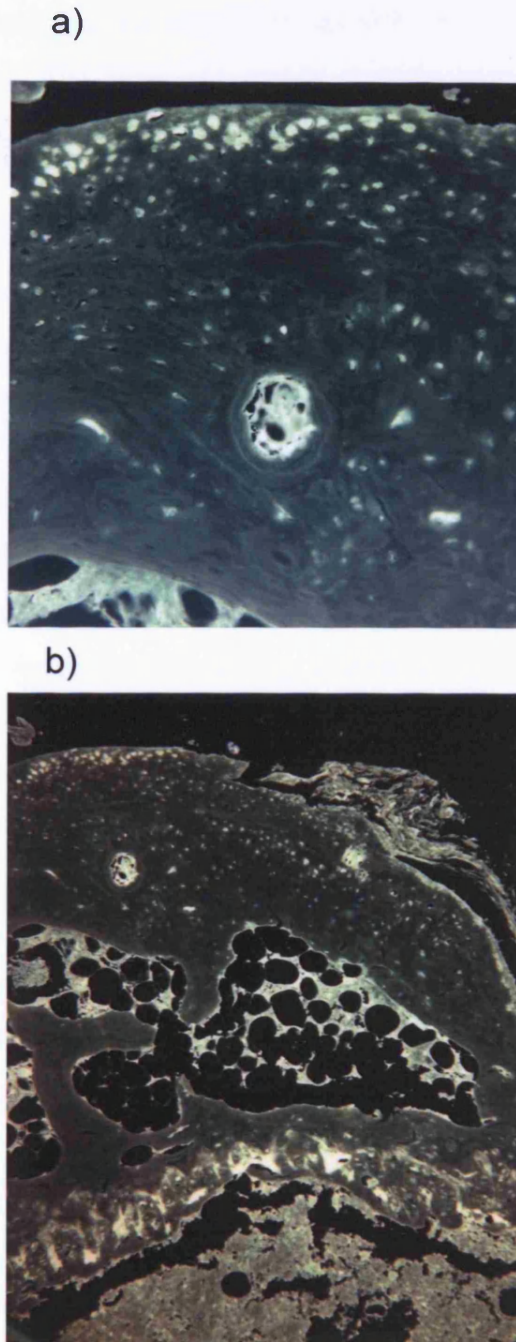


Figure 5.18. Adult mouse femur.
Images manipulated using Image arithmetic subtract function with lightest areas representing highest level of expression.

- a) Articular cartilage x 10. In-situ hybridisation – *DYM* antisense probe showing maximal expression in chondrocyte lacunae and tangential layer.
- b) Growth plate and articular cartilage x 4. In-situ hybridisation – *DYM* antisense probe showing maximal expression in growth plate, chondrocyte lacunae and tangential layer of cartilage.

(*EFEMP1*) and *EAP20* within mouse brain are illustrated in Appendix D, Figures D (iv) - (v). Yeast-two hybrid analysis identified both *EFEMP1* and *EAP20* as Dymeclin interactors (see Chapter 6). The expression pattern of all three genes is similar although this may reflect cell density to some degree. *EFEMP1* encodes the extracellular matrix protein fibulin-3 and has been shown to be expressed in condensing mesenchyme and subsequently in the cartilage primordia of craniofacial bones, limb buds and vertebrae in the mouse embryo but not in the developing brain. It has thus been suggested that *EFEMP1* has either a mechanical role in the developing skeleton or a regulatory role given that it contains epidermal growth factor (EGF) repeats which may bind to EGF receptors and modulate chondrocyte proliferation and differentiation by promoting protein-protein interactions (Ehlermann *et al.*, 2003). It is not known to have a role in brain development despite being expressed in the adult mouse brain (McLaughlin *et al.*, 2007). *EAP20* is the human homologue of the yeast protein, vacuolar sorting protein 25 (*VPS25*) and is required for endosomal sorting, a method of transporting proteins for recycling or degradation. The finding of similar expression patterns within the mouse brain of interacting proteins *DYM* and *EAP20*, both thought to be acting in protein transport and *EFEMP1*, an extracellular matrix protein may indicate that have a shared function, perhaps in the transport of extracellular matrix components.

Chondrocytes within the growth plate undergo a complex differentiation process, progressing through resting, proliferative and hypertrophic stages before becoming apoptotic. Many genes are required for this process to proceed normally and if mutated, cause abnormal proliferation or differentiation of chondrocytes and subsequent skeletal dysplasia. *DYM* has been shown to be most highly expressed in resting and hypertrophic chondrocytes. Other genes which are known to be expressed at high levels in resting chondrocytes include *COL2A1*, *SOX* transcription factors 5, 6, and 9 and *Fgdl* (Kornak and Mundlos, 2003, Gorski *et al.*, 2000). Mutations in *COL2A1* can cause a variety of conditions ranging in severity from lethal Type II achondrogenesis to Stickler syndrome which in addition to abnormalities of the joints, has severe myopia and cleft palate as clinical features. *SOX9* acts to regulate *COL2A1* expression and is required for the differentiation of mesenchymal cells into chondrocytes whereas *SOX5* and *SOX6* are required for the differentiation of chondrocytes themselves (Akiyama *et al.*, 2002). *Fgdl*, the mouse homologue of the gene mutated in faciogenital dysplasia (FGD) has also been shown to be expressed in resting chondrocytes. FGD is characterised by short stature, abnormal facial bones and spinal abnormalities as well as urogenital defects (Gorski *et al.*, 2000). The variety of syndromes which have been

discussed here highlight the broad range of phenotypes which can result from genes known to be expressed at high levels within resting chondrocytes.

Genes which are known to be expressed in hypertrophic chondrocytes include *COL10A1*, *RUNX2*, *osteoclast differentiation factor* (also known as *osteoprotegerin ligand*) and *osteoprotegerin (OPG)* (Kornak and Mundlos, 2003). Mutations in *COL10A1* cause the dominant condition Schmid metaphyseal chondrodysplasia, a relatively mild disorder characterised by short limbs, lumbar lordosis and abnormal epiphyses and metaphyses radiologically. Dilated RER has been shown in chondrocytes in Schmid metaphyseal chondrodysplasia as in DMC (Cooper *et al.*, 1973). *Osteoclast differentiation factor*, as its name suggests is required for osteoclast differentiation and has been found to cause a rare recessive form of osteopetrosis in which increased bone lay down results in bone marrow obliteration. *Osteoprotegerin* suppresses bone resorption and recessive mutations results in Juvenile Pagets disease (OMIM 239000) with increased bone turnover and bone fragility. The disorders discussed here caused by genes expressed within hypertrophic chondrocytes result in largely skeletal phenotypes compared to conditions caused by defects in genes expressed in resting chondrocytes. Given that several of the genes expressed within resting chondrocytes are transcription factors, this is not surprising.

Although a number of novel genes involved in chondrocyte differentiation and endochondral ossification have been identified over the years, there are likely many more genes yet to be discovered which are essential for this complex process. A large-scale gene expression profile study aimed to identify genes expressed during skeletal development using ESTs. Approximately 200 genes were found to be overexpressed in cartilage of which 10% encoded proteins already known to be involved in cartilage and/or bone development or homeostasis. Variable expression levels were shown at different stages of development suggesting that different genes are active at various stages in bone/cartilage development. This approach may lead to the identification of candidates for the many skeletal dysplasias for which causative genes have not yet been found (Tagariello *et al.*, 2005).

RT-PCR and Northern blot analysis have shown that *DYM* transcripts are found in a broad range of tissues in contrast to *in-situ* hybridisation findings of *DYM* expression largely in brain and cartilage. If Dymeclin is acting in substrate transport through the ER, a cellular mechanism necessary in all cell types, it must be considered how the tissue specific phenotype results. Brain, epithelium and cartilage are all tissues with high requirements for extracellular matrix and in particular, proteoglycans. Brain is a metabolically active tissue as neurotransmitters must be transported through synapses.

Cartilage requires ongoing excretion of extracellular matrix components. Expression of *DYM* in fetal liver was also shown although liver function in DMC patients has not been shown to be abnormal. Similarly, liver is a metabolically active tissue with high requirement for protein transport within cells. It is likely that the brain and cartilage specific phenotype occurs because these tissues are exquisitely sensitive to functional Dymeclin levels. Thus *DYM* mutation leads to failure of protein trafficking and subsequent substrate accumulation with possible deleterious effect.

Taken together, the findings of dilated RER and multiple vacuoles in chondrocytes of affected individuals, localisation of Dymeclin to the ER and *DYM* expression in cartilage, epithelium and brain, further support the hypothesis that Dymeclin is involved in protein trafficking through the ER of substances critical for chondrocyte differentiation and neuronal function.

Chapter 6 - Identification of Dymeclin interactors by yeast two-hybrid analysis

6.1 Introduction

The identification of mutations in *DYM* as causative of the DMC phenotype highlights the crucial role of Dymeclin in cartilage, bone and brain development and normal function. The finding of abnormal chondrocyte columnarisation with degenerating chondrocytes in cartilage of affected individuals indicates that loss of function of Dymeclin disrupts chondrocyte differentiation. The observation of dilated rough endoplasmic reticulum and multiple vacuoles in both chondrocytes and fibroblasts implies that there may be accumulation of as yet unidentified substances within DMC cells (Engfeldt *et al.*, 1983, Nakamura *et al.*, 1997). Dymeclin is a novel protein with no homology to any other known protein and its biological properties have yet to be elucidated. *In-silico* predictions suggest that Dymeclin is a transmembrane protein with motifs that may indicate a role in protein sorting/targeting. This, taken together with the findings in this study of localisation to the ER, have led to the hypothesis that Dymeclin is involved in transport through the ER, of proteins required for normal chondrocyte function. A yeast two-hybrid screen was therefore performed in an attempt to identify proteins which interact with Dymeclin and hence gain greater insight as to its function. These interactors may be either primarily involved in transport through the ER, or themselves represent transport cargo.

A yeast two-hybrid screen is capable of detecting the interaction of two fusion proteins, the first containing the gene under investigation (bait) fused to the DNA binding domain (DBD) of a transcription factor. The second protein is one of a library of transcripts fused to the activation domain of the same transcription factor. Reconstitution of the transcription factor will drive the expression of two transgenes required for yeast survival under limiting conditions, thus permitting selection of clones containing interacting proteins. If the bait protein interacts with a library transcript, transcription is activated and the prey (library) protein can subsequently be identified. The yeast two-hybrid technique is composed of several stages as discussed in Section 2.2.19 and is summarised in Figure 6.1.

6.2 Characterisation of a prey cDNA library

The yeast two-hybrid system used relies on a yeast strain (PJ69-4A) which is auxotrophic for four amino acids, uracil, leucine, histidine and adenine and requires these amino acids to be supplied in the media or synthesised from an exogenous source

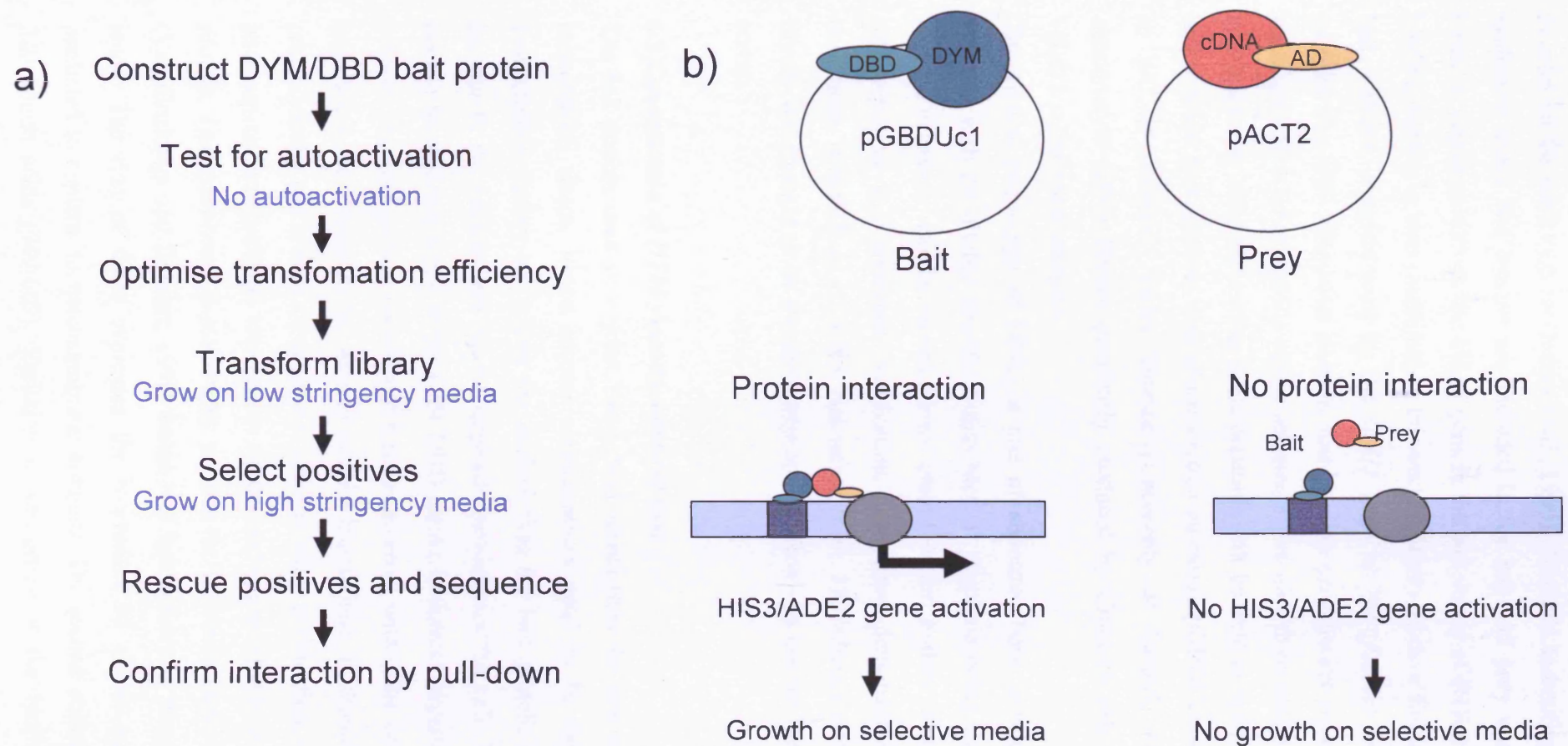


Figure 6.1. Overview of the Yeast two-hybrid strategy.

a) Summary of steps in yeast two-hybrid analysis.

b) Diagram illustrates the bait vector containing the gene encoding the protein of interest fused to the DNA binding domain (DBD) of a transcription factor. The prey vector comprises a library of cDNAs fused to the transcription factor activation domain (AD). If the protein products interact, the transcription factor will be activated when the DBD and AD come together and HIS3/ADE2 genes will be activated hence growth will occur on selective media.

in order for the yeast to grow (James *et al.*, 1996). Plasmids harbouring genes which can synthesise uracil and leucine are included in the bait and prey vectors respectively. Uracil is synthesised from the *URA3* gene in the bait vector pGBDUc1. The Gal4 DNA binding domain is also contained in this vector, expressed as a fusion protein with the bait. Leucine is synthesised by the *LEU2* gene on the plasmid pACT2 which also contains the Gal4 activation domain fused to the prey library transcript. Yeast cells containing these two plasmids can be selected for by growth on media deficient in uracil and leucine as cells harbouring these plasmids will be able to synthesise the deficient amino acids. Bait proteins were constructed in the vector pGBDUc1 and the prey library in pACT2 obtained. Victor Duance (University of Cardiff) kindly provided a chondrocyte cDNA library previously produced by Clontech (Mountain View, CA, USA) but since discontinued.

The clinical phenotype of DMC is one of abnormal bone development and brain function with the allelic disorder Smith McCort dysplasia being characterised by the absence of mental retardation and normal brain function. Failure of normal chondrocyte differentiation and abnormal ossification have been demonstrated in both these conditions, (Engfeldt *et al.*, 1983, Nakamura *et al.*, 1997) hence a chondrocyte cDNA library was thought to be the most appropriate library to use for the yeast two-hybrid screen.

6.3 Construction of *DYM* yeast two-hybrid bait

The bait protein used in a yeast two-hybrid screen must be able to act as part of a transcription factor, hence several considerations need to be taken into account. Transcription factors act within the nucleus thus the bait protein should be able to localise to the nucleus and not be targeted elsewhere in the cell. The presence of a nuclear localisation signal in the Gal4 DBD assists in correct targeting of the chimeric protein. The bait protein should not form complexes within the cell or require post-translational modification for normal protein interactions. It should not be degraded proteolytically and it must be soluble in order to allow purification. Dymeclin does not have any cell compartment localisation signals and is predicted to be a transmembrane protein. Transmembrane domains may prevent the bait from localising to the nucleus (Van Crielinge and Beyaert, 1999) hence two bait constructs were designed for this study. The first of these expresses the N-terminal 98 amino acids of Dymeclin, predicted to contain no transmembrane domains. The second contains the C-terminal 220 amino acids (449-669). Variation in stringency of the bioinformatic structure prediction programmes used has resulted in two alternative protein models, one

predicted to contain three transmembrane domains within this region but the other not thought to have any transmembrane regions within this part of the protein.

Plasmids expressing the Dymeclin bait proteins, termed pGBDUDYM1-98 (N-terminal 98 amino acids) and pGBDUDYM449-669 (C-terminal 220 amino acids) were cloned using the methods described in section 2.2.13. Briefly, the vector was derived from pGDBU-c1LIPA221, containing the insert LIPA 221 encoding 221 amino acids of Lamin interacting protein (provided by Dr Sue Shackleton), which was digested with *EcoRI* and *BamHI* to remove the insert. The vector band was gel purified according to the manufacturer's recommendations (Qiagen, Crawley, UK). PCR primers were designed to amplify the sequence encoding either amino acids 1-98 (*EcoRI* Y2H1-98F and *BamHI* Y2H1-98R) or 449-669 (*EcoRI* Y2H449-669F and *BamHI* Y2H449-669R) of *DYM* whilst at the same time introducing a 5' *EcoRI* site and a 3' *BamHI* site for cloning. Primer sequences are listed in Table 6.1. PCR reactions were performed as previously described using DMCWTpCDNA3 as template. The PCR product was purified using a PCR purification kit according to the manufacturer's recommendations (Qiagen Crawley, UK). This DNA fragment was subsequently digested with *EcoRI* and *BamHI* and ligated into the previously digested and purified pGDBU-C1 vector. Ligations were transformed into *E.coli* DH5 α and restriction digest of purified DNA was performed to confirm the presence of an insert. Clones were then sequenced to confirm the absence of PCR errors and to ensure the insert was cloned in frame.

The plasmids pGBDUDYM1-98 and 449-669 were used to transform the yeast strain PJ69-4a. Transformed yeast were maintained on media deficient in uracil, as transformed yeast would contain the plasmid containing the *URA3* gene and thus synthesise uracil.

6.4 Verification of DYM1-98 and 449-669 as suitable bait

If the reporter genes (histidine and adenine in the yeast strain PJ69-4a) are activated by the bait fusion protein in the absence of an interacting protein (auto-activation), no meaningful results would be obtained and the screen would not be valid. Approximately 5% of bait fusion proteins are thought to cause auto-activation (Van Crielinge and Beyaert, 1999) which occurs when the Gal4 DBD functions correctly but motifs within the prey protein are able to recruit RNA polymerases directly and transformation of the prey protein alone into yeast results in activation of the transcription factor. The subsequent transformation of library proteins hence cannot be relied upon as false positives would be obtained in the absence of library transformation.

Primer name	Primer sequence
EcoRI Y2H1-98F	5' GAG AGA ATT CAT GGG ATC GAA TAG CAG CAG 3'
BamHI Y2H1-98R	5' TCT CGG ATC CGA TGA ACA CTT TCA GCA AAC 3'
EcoRI Y2H449-669F	5' GAG AGA ATT CTA CAA CAT GAC TAG GAC ACG 3'
BamHI Y2H449-669R	5' TCT CGG ATC CTC AGT CGG AAT CCA TGG TGA A
DMC MBP1-98F	5' GAG AGG ATC CGG ATC GAA TAG CAG CAG AAT C 3'
DMC MBP1-98R	5' TCT CAA GCT TGT GGT TCT GAC ATT CTG CTG 3'
DMC MBP449-669F	5' GAG AGG ATC CTA CAA CAT GAC TAG GAC ACG 3'
DMC MBP449-669R	5' TTC ACC ATG GAT TCC GAC TGA AAG CTT GAG A 3'
pMALF	5' GGT CGT CAG ACT GTC GAT GAA GCC 3'
pMALR	5' CGC CAG GGT TTT CCC AGT CAC GAC 3'
GAL4ADF	5' TAC CAC TAC AAT GGA TG 3'
pACTR	5' CAG TTG AAG TGA ACT TGC GG 3'
EFEMPF1	5' CAA CCA CTA TGG AGG ATA CCT CTG C 3'
EFEMP1R	5' GGG GGC ACA TAC ATG AGA AT 3'
EFEMPF2	5' TGT GAG ACC ACA AAT GAA TGC 3'
EFEMPR2	5' TCC ATT TTC TTA CAA AGC AG 3'
VPSF1	5' GCT CGC TGG TCC TGT CCT TCT GC 3'
VPSR1	5' AGC CAG GGC TTC CCC CTA CCC CAT 3'

Table 6.1. Yeast two-hybrid cloning and sequencing primers.

The yeast two-hybrid system used in this study contains two selective reporter genes, histidine and adenine (James *et al.*, 1996). An additional reporter was introduced because histidine is known to be a 'leaky' reporter with transcription from the Gal4-histidine promoter activated at low levels in the absence of a binding protein. Yeast cells can therefore survive on media deficient in histidine in the presence of bait protein alone. This is overcome by the addition of 3-aminotriazole (3-AT) to the media which, when at a low level, inhibits the histidine gene product. However, at high levels 3-AT can interfere with protein interaction or kill the yeast consequently it is important to determine the appropriate level of 3-AT required to inhibit auto-activation.

In order to determine if pGBDUDYM1-98 caused auto-activation, the plasmid was transformed into PJ69-4a. Plates were made deficient in histidine with increasing concentrations of 3-AT (1-50 mM), a diluted yeast culture spread on them and incubated for several days. A level of 10 mM 3-AT was deemed adequate to inhibit background histidine production without suppressing weak protein interactions as 10 mM 3-AT permitted growth whereas 25 mM 3-AT completely obliterated all growth.

6.5 Optimisation of transformation efficiency

Transformation of bait and prey plasmid into yeast using lithium acetate is a method which has previously been well validated (Agatep *et al.*, 1998) and has been used successfully in many yeast two-hybrid screens (Haque *et al.*, 2006, Lloyd *et al.*, 2002). However, transformation efficiency can be variable depending on the following factors; the size of plasmids used, the quality of DNA, the strain of yeast and its stage in growth and the quality of carrier DNA used (Gietz, 2006). A yeast two-hybrid screen aims to test the interaction of every cDNA in the library with the bait protein. The chondrocyte cDNA library used is predicted to contain 3.6×10^6 transformants. and given that some library cDNAs may be represented at a higher level in the library than others, it has been suggested that the number of transformants screened should be between three and six times the number in the library (Gietz, 2006). Hence a minimum total of 1.1×10^7 transformants should be screened in order to provide complete coverage of the library. In order to optimise the transformation efficiency, small scale transformations of PJ69-4a containing the bait expression construct were performed using variable amounts of cDNA. Graphs illustrating the results are shown in Figure 6.2. Higher transformation efficiencies were obtained using smaller amounts of cDNA but the total number of transformants generated was lower hence it was decided to use an intermediate amount of cDNA in order to obtain an adequate number of transformants without wasting library cDNA. The reaction was scaled up 150 times using 15 μ g of library cDNA in

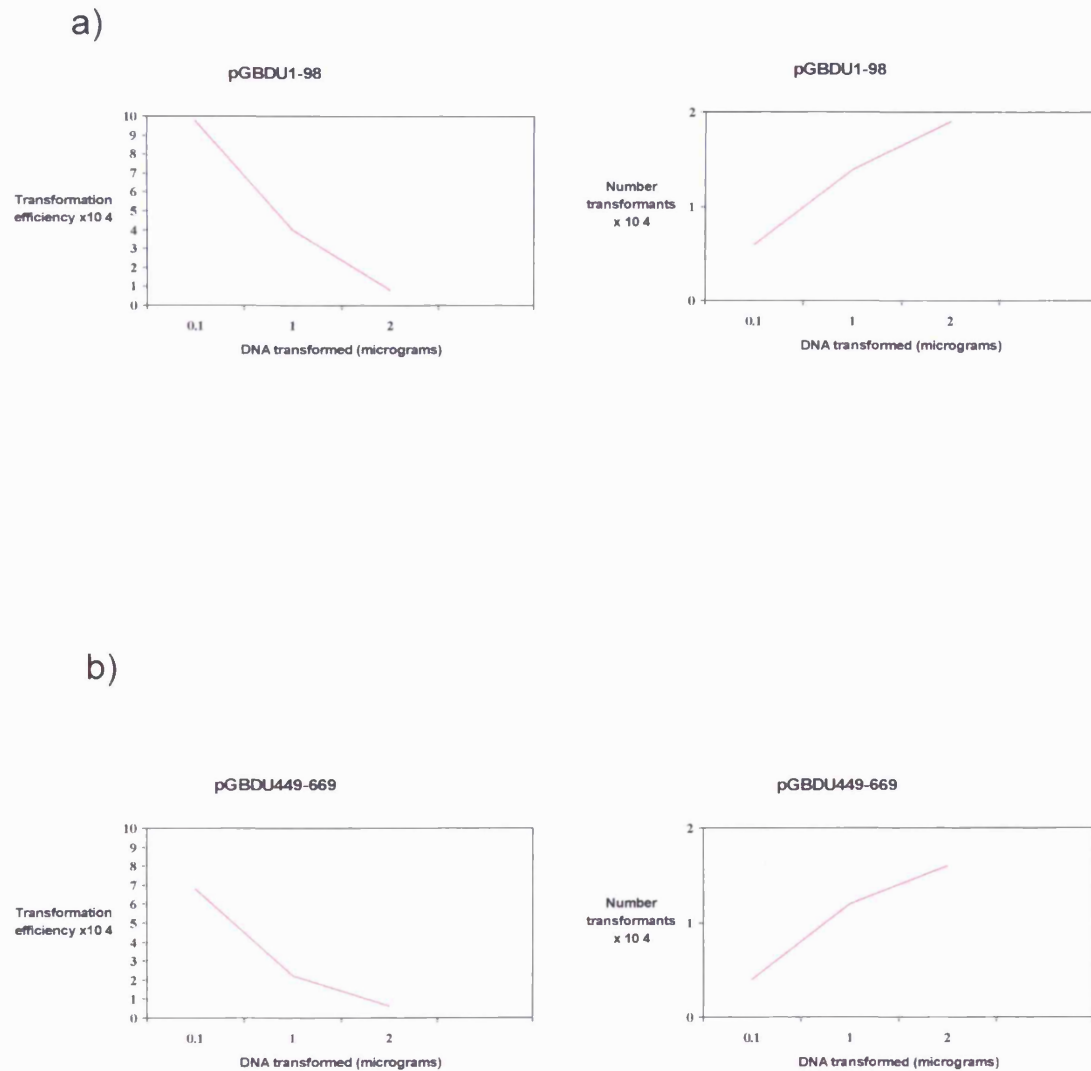


Figure 6.2. Graphs demonstrating yeast transformation efficiency and number of transformants per DNA transformed.

a) represents pGBDUDYM1-98, and b) pGBDUDYM449-669.

The transformation efficiency is higher with smaller amounts of DNA but the number of transformants lower.

order to obtain a total number of transformants in excess of the total required (1.1×10^7).

6.6 Library transformation and analysis of positive clones

The chondrocyte cDNA library in the vector pACT2 was transformed into PJ69-4a pre-transformed with the either bait plasmid pGBDUDYM1-98 or pGBDUDYM449-669. Transformants were plated onto media deficient in uracil, leucine, histidine and supplemented with 10 mM 3-AT. Plates were incubated for 14 days at 30 °C. If an interaction occurred between the bait fusion protein and the prey fusion protein, transcription would be activated from the histidine promoter and a yeast colony would grow. The transformation efficiency of the scale up reaction was estimated to be 5×10^3 / μg DNA for DYM1-98 bait and 6.5×10^2 / μg DNA for DYM449-669 bait providing total numbers of transformants of 7.5×10^4 and 9.8×10^3 respectively. This was significantly lower than previous transformations and not high enough to provide complete library coverage. However, 45 yeast colonies were identified using pGBDUDYM1-98 bait and six yeast colonies using pGBDUDYM449-669.

The histidine reporter is known to be less stringent than the adenine reporter and may allow clones with weak interactions to grow. As the yeast two-hybrid system used contains a second reporter (adenine), putative interacting clones were confirmed using the adenine reporter system. Positive clones were re-streaked onto media lacking uracil, leucine and histidine supplemented with 10 mM 3-AT and in duplicate onto media additionally lacking adenine to act as a further selection step. Of the initial 45 positive clones using DYM1-98 bait, 16 positive clones (colonies 3, 4, 5, 6, 12, 18, 19, 22, 23, 25, 26, 27, 34, 35, 39 and 43) were confirmed and of the initial 6 colonies using DYM449-669 bait, 4 positive clones (colonies 1, 2, 3 and 5) were confirmed. Positive clones are shown in Figure 6.3. Plasmid rescue was achieved by extraction of plasmid DNA from the yeast strain, transformation into *E.coli* DH5 α and subsequent sequencing as described in methods sections 2.2.19.4, 2.2.13.6 and 2.2.9 respectively. Transformation into competent cells allows the DNA to be easily manipulated and analysed. Insert sequence data was obtained for 12 of the positive clones with DYM1-98 bait and 3 of the positive clones using DYM449-669 bait. These results are summarised in Table 6.2.

6.7 Sequence identification of interacting clones

Primers were designed that flank the cloning site of pACT2 to sequence the cDNA insert. Primer *GAL4ADF* is on the sense strand of the Gal4 activation domain and

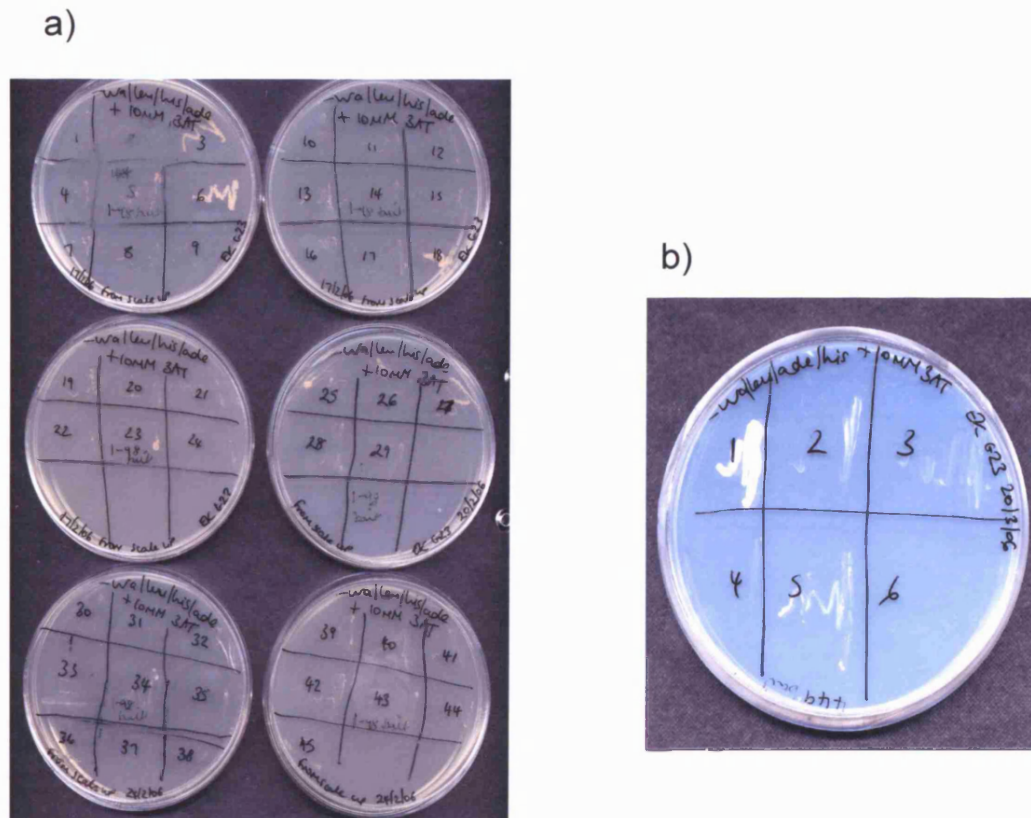


Figure 6.3. Confirmation of positive interacting clones. The initial 45 (using 1-98 bait) and 6 (using 449 bait) positive clones identified from their growth on media deficient in uracil, leucine and histidine and supplemented with 3-AT were streaked onto media additionally deficient in adenine. Growth on this media confirms the 16 positive clones using 1-98 bait (illustrated in a) and 4 positive clones using 449 bait (illustrated in b) which were able to strongly activate the reporter gene

Bait	Colony number	Plasmid rescued	Sequenced successfully	Sequence result
1-98	3	Yes	No	
1-98	4	Yes	No	
1-98	5	Yes	Yes	ENC1 (partial)
1-98	6	Yes	Yes	EFEMP
1-98	12	No		
1-98	18	Yes	Yes	DYM and vector
1-98	19	Yes	Yes	VPS25(EAP20)
1-98	22	Yes	Yes	Zinc CSL
1-98	23	Yes	Yes	DYM and vector
1-98	25	Yes	Yes	vector
1-98	26	No		
1-98	27	Yes	Yes	DYM and vector
1-98	29	Yes	Yes	DYM and vector
1-98	34	No		
1-98	35	No		
1-98	39	No		
1-98	43	No		
449-669	1	Yes	Yes	DYM
449-669	2	Yes	Yes	DYM
449-669	3	Yes	Yes	DYM
449-669	5	No		

Table 6.2. Identification of the 16 positive clones using DYM1-98 bait and the 4 positive clones using DYM449-669 bait after screening on adenine depleted media. Clones were numbered when identified from the initial screens (51 positives in total).

would reveal the coding strand of the cDNA insert. Primer *pACT2R* is on the antisense strand after the cloning site and would produce antisense sequence of the cDNA insert. Primer sequences are listed in Table 6.1. The 12 plasmids which were successfully rescued into bacteria were sequenced using both primers *GAL4ADF* and *pACT2R* as described in section 2.2.9. DNA sequence obtained was used in BLAST search of known gene sequences to determine the identity of the cDNA insert. Sequence analysis of the 12 interacting clones revealed one to comprise vector sequence alone, and four contained both *DYM* and vector sequence. An additional three clones contained *DYM* sequence only. The remaining four were identified as *EGF-containing fibulin-like extracellular matrix protein 1* (*EFEMP1*, Accession number NM_001039349), *Homo sapiens vacuolar protein sorting protein 25 (VPS25) homologue*, (known in humans as *EAP20*, NM_032353), *Ectodermal neural cortex 1* (*ENC1*, NM_003633) and *Zinc finger CSL domain containing type 3* (*ZCSL3*, also known as *DPH4 homologue*, NM_181706). Sequencing of both sense and antisense strands confirmed these findings. Sequencing of the *EFEMP1* insert revealed that the first 61 out of 494 amino acids were missing but that the 3' region of the insert included the polyA. *EAP20* sequence contained the entire cDNA and all 176 amino acids of the protein would be produced in the chimeric protein. Only 101 bp of the 5' UTR of *ENC1* cDNA matched the insert sequenced with the remaining sequence not having a good BLAST match hence it was decided not to pursue investigation of this clone further as interaction was not occurring with protein coding sequence. Sequencing of *ZCSL3* showed that the sequence encoding the initial 73 amino acids was missing but that the insert extended to the polyA. Insert sequences are shown in Figures 6.4, 6.5 and 6.6. Nested primers were designed within the sequence of *EAP20* and *EFEMP1* in order to sequence the whole of the insert as sequencing using primers *GAL4ADF* and *pACT2R* had revealed the 5' and 3' regions of the cDNA alone. Sequencing using the internal primers *EFEMPF1*, *EFEMPR1*, *EFEMPF2*, *EFEMPR2*, *EAP20F1* and *EAP20R1* confirmed the full length clone sequences to match published *EFEMP* and *EAP20* cDNA sequence. Primer sequences are shown in Table 6.1.

6.8 Confirmation of the interaction of DYM with EAP20 and EFEMP1

To investigate the validity of the interactions between Dymeclin and the putative interacting proteins, MBP pull down assays were performed as described in section 2.2.20. The pull down assay requires the production of two proteins, the protein of interest (in this case, the original bait protein) with an MBP tag and the interacting protein produced by *in-vitro* translation and [³⁵S] methionine labelled. A fusion gene

GAATTCCGGGCCGC Val Asn His Tyr Gly Gly Tyr 68

484 CCA GAC GCT TGT AAA GGT GGA ATG AAG TGT GTC AAC CAC TAT GGA GGA TAC 85
 Leu Cys Leu Pro Lys Thr Ala Gln Ile Ile Val Asn Asn Glu Gln Pro Gln

535 CTC TGC CTT CCG AAA ACA GCC CAG ATT ATT GTC AAT AAT GAA CAG CCT CAG 102
 Gln Glu Thr Gln Pro Ala Glu Gly Thr Ser Gly Ala Thr Thr Gly Val Val

586 CAG GAA ACA CAA CCA GCA GAA GGA ACC TCA GGG GCA ACC ACC GGG GTT GTA 119
 Ala Ala Ser Ser Met Ala Thr Ser Gly Val Leu Pro Gly Gly Phe Val

637 GCT GCC AGC AGC ATG GCA ACC AGT GGA GTG TTG CCC GGG GGT GGT TTT GTG 136
 Ala Ser Ala Ala Ala Val Ala Gly Pro Glu Met Gln Thr Gly Arg Asn Asn

688 GCC AGT GCT GCT GCA GTC GCA GGC CCT GAA ATG CAG ACT GGC CGA AAT AAC 153
 Phe Val Ile Arg Arg Asn Pro Ala Asp Pro Gln Arg Ile Pro Ser Asn Pro

739 TTT GTC ATC CGG CGG AAC CCA GCT GAC CCT CAG CGC ATT CCC TCC AAC CCT 170
 Ser His Arg Ile Gln Cys Ala Ala Gly Tyr Glu Gln Ser Glu His Asn Val

790 TCC CAC CGT ATC CAG TGT GCA GCA GGC TAC GAG CAA AGT GAA CAC AAC GTG 187
 Cys Gln Asp Ile Asp Glu Cys Thr Ala Gly Thr His Asn Cys Arg Ala Asp

841 TGC CAA GAC ATA GAC GAG TGC ACT GCA GGG ACG CAC AAC TGT AGA GCA GAC 204
 Gln Val Cys Ile Asn Leu Arg Gly Ser Phe Ala Cys Gln Cys Pro Pro Gly

892 CAA GTG TGC ATC AAT TTA CGG GGA TCC TTT GCA TGT CAG TGC CCT CCT GGA 221
 Tyr Gln Lys Arg Gly Glu Gln Cys Val Asp Ile Asp Glu Cys Thr Ile Pro

943 TAT CAG AAG CGA GGG GAG CAG TGC GTA GAC ATA GAT GAA TGT ACC ATC CCT 238
 Pro Tyr Cys His Gln Arg Cys Val Asn Thr Pro Gly Ser Phe Tyr Cys Gln

994 CCA TAT TGC CAC CAA AGA TGC GTG AAT ACA CCA GGC TCA TTT TAT TGC CAG 255
 Cys Ser Pro Gly Phe Gln Leu Ala Ala Asn Asn Tyr Thr Cys Val Asp Ile

1045 TGC AGT CCT GGG TTT CAA TTG GCA GCA AAC AAC TAT ACC TGC GTA GAT ATA 272
 Asn Glu Cys Asp Ala Ser Asn Gln Cys Ala Gln Gln Cys Tyr Asn Ile Leu

1096 AAT GAA TGT GAT GCC AGC AAT CAA TGT GCT CAG CAG TGC TAC AAC ATT CTT 289
 Gly Ser Phe Ile Cys Gln Cys Asn Gln Gly Tyr Glu Leu Ser Ser Asp Arg

1147 GGT TCA TTC ATC TGT CAG TGC AAT CAA GGA TAT GAG CTA AGC AGT GAC AGG 306
 Leu Asn Cys Glu Asp Ile Asp Glu Cys Arg Thr Ser Ser Tyr Leu Cys Gln

1198 CTC AAC TGT GAA GAC ATT GAT GAA TGC AGA ACC TCA AGC TAC CTG TGT CAA 323
 Tyr Gln Cys Val Asn Glu Pro Gly Lys Phe Ser Cys Met Cys Pro Gln Gly

1249 TAT CAA TGT GTC AAT GAA CCT GGG AAA TTC TCA TGT ATG TGC CCC CAG GGA 340
 Tyr Gln Val Val Arg Ser Arg Thr Cys Gln Asp Ile Asn Glu Cys Glu Thr

1300 TAC CAA GTG GTG AGA AGT AGA ACA TGT CAA GAT ATA AAT GAG TGT GAG ACC 357
 Thr Asn Glu Cys Arg Glu Asp Glu Met Cys Trp Asn Tyr His Gly Gly Phe

1351 ACA AAT GAA TGC CGG GAG GAT GAA ATG TGT TGG AAT TAT CAT GGC GGC TTC 374
 Arg Cys Tyr Pro Arg Asn Pro Cys Gln Asp Pro Tyr Ile Leu Thr Pro Glu

1402 CGT TGT TAT CCA CGA AAT CCT TGT CAA GAT CCC TAC ATT CTA ACA CCA GAG 391
 Asn Arg Cys Val Cys Pro Val Ser Asn Ala Met Cys Arg Glu Leu Pro Gln

1453 AAC CGA TGT GTT TGC CCA GTC TCA AAT GCC ATG TGC CGA GAA CTG CCC CAG 408
 Ser Ile Val Tyr Lys Tyr Met Ser Ile Arg Ser Asp Arg Ser Val Pro Ser

1504 TCA ATA GTC TAC AAA TAC ATG AGC ATC CGA TCT GAT AGG TCT GTG CCA TCA 425
 Asp Ile Phe Gln Ile Gln Ala Thr Thr Ile Tyr Ala Asn Thr Ile Asn Thr

1555 GAC ATC TTC CAG ATA CAG GCC ACA ACT ATT TAT GCC AAC ACC ATC AAT ACT 442
 Phe Arg Ile Lys Ser Gly Asn Glu Asn Gly Glu Phe Tyr Leu Arg Gln Thr

1606 TTT CGG ATT AAA TCT GGA AAT GAA AAT GGA GAG TTC TAC CTA CGA CAA ACA 459
 Ser Pro Val Ser Ala Met Leu Val Leu Val Lys Ser Leu Ser Gly Pro Arg

1657 AGT CCT GTA AGT GCA ATG CTT GTG CTC GTG AAG TCA TTA TCA GGA CCA AGA 476
 Glu His Ile Val Asp Leu Glu Met Leu Thr Val Ser Ser Ile Gly Thr Phe

1708 GAA CAT ATC GTG GAC CTG GAG ATG CTG ACA GTC AGC AGT ATA GGG ACC TTC 493
 Arg Thr Ser Ser Val Leu Arg Leu Thr Ile Ile Val Gly Pro Phe Ser Phe

1759 CGC ACA AGC TCT GTG TTA AGA TGA ATA ATA GTG GGG CCA TTT TCA TTT
 Stop

1810 TAG TCTTTTCTAAGAGTCAACCACAGGCATTTAAGTCAGCCAAAGAATATTGTTACCTTAAAGCAC
 1876 TATTTTATTTATAGATATATCTAGTGCATCTACATCTCTATACTGTACACTCACCCATAATTCAAAC
 1943 AATTACACCATGGTATAAAGTGGGCATTTAATATGTAAAGATTCAAAGTTTGTCTTTATCTATATAT
 2010 GTAAATTAGACATTAACTCACTAAACTGGTCTTCTCAAGAGAGCTAAGTATACACTATCTGGTGAA
 2077 ACTTGGATTCTTTCTATAAAGTGGGACCAAGCAATGATGATCTTCTGTGGTGCTTAAGGAACTT
 2144 ACTAGAGCTCCACTAACAGTCTCATAAGGAGGCAGCCATCATAACCATTGAATAGCATGCAAGGGTA
 2211 AGAATGAGTTTTTAACGTCTTGTAAAGAAAATGGAAAAGGTCAATAAAGATATATTTCTTTGAAAA
 AAAAAAAAAAAAAAAAAA***CTCGAG***

Figure 6.5. cDNA sequence of EFEMP1 identified by sequencing of Y2H bait 1-98 clone 6. Cloning sites are shown in bold and italic. cDNA sequence is shown in codons with the corresponding amino acid above. cDNA sequence is numbered on the left and amino acids on the right. 5' and 3'UTR are indicated by continuous sequence. The first 61 amino acids are not encoded by the insert sequence.

GAATTCCGGCGCCGCGTCGACCAA Lys Arg Glu Tyr Asp Leu Gln Arg Cys Glu Asp 85
 265 CTA GGA AAT GAA GAG ACA AAA AGA GAG TAT GAC CTG CAG CGG TGT GAA GAT
 Asp Leu Arg Asn Val Gly Pro Val Asp Ala Gln Val Tyr Leu Glu Glu Met 102
 316 GAT CTA AGA AAT GTA GGA CCA GTA GAT GCT CAA GTA TAT CTT GAA GAA ATG
 Ser Trp Asn Glu Gly Asp His Ser Phe Tyr Leu Ser Cys Arg Cys Gly Gly 119
 367 TCT TGG AAT GAA GGT GAT CAC TCT TTT TAT CTG AGT TGC AGA TGT GGT GGA
 Lys Tyr Ser Val Ser Lys Asp Glu Ala Glu Glu Val Ser Leu Ile Ser Cys 136
 418 AAA TAC AGT GTT TCC AAG GAT GAA GCG GAA GAA GTT AGC CTG ATT TCT TGT
 Asp Thr Cys Ser Leu Ile Ile Glu Leu Leu His Tyr Asn Stop 149
 469 GAT ACA TGT TCA CTA ATT ATA GAA CTC CTT CAT TAT AAC TAA AATTGTTTCA
 522 ACTTGAATGCTTTAACTGTGGTATTGAGACATGATGAGAAGCCGTTGAGCTTTGTCCATTCAAGGA
 589 AATGGATTATTTGTCAGCCCGATTATTTGCAAAGAAAATATACAATTATCAAGCAGAGACCACCTCA
 656 GATTAATAGAGAAATGAATATAATTATTTATTTGTATTTATAATTTATATTACAAGTTGTCCAAAA
 723 TCTAGTTTGATTATTTTGGTGAATCCCATCTGAGAACTTTATATATGTCATTTTTTCCATGACAC
 790 AAAATAAGAAATTTTCTTAATATGAGCATGTCTCCTCTTTTCAGTGGGTAAACAGTATCATCATATTCC
 857 TTTATAGTTACTAAGTTTCTGGTCTAACCATAACGTTTCTAGTTGCTAAAGTTATGGGGCTGTTGT
 924 ACACCTAGAATCTTTGTGAATAATGTGAGGCCAGTTCTTCCATAAGGAAGGCTGGTTATGGATATTC
 991 ATAAGGTTATTTCAAAGTTAATAAAGACAAAGTGGCAACTGTAGAAAAGTTGCTCCCAATCTGGT
 1058 CCGTATTTCCAAAGCTTGTAAACACTGCCATCTTCAAATTTAAATGTAATTTAGGAAATGGAAGCTT
 1125 TTAATTTTAATTTCTTTTTTCTAAGACAGAGTCTCACTCTGTGCGCCAGCTGGAGTGCAGTGCCAT
 1192 GATCTCAGTCTGCTGTAACCTCCGCTCTCCTGGGTCAAGTGATTCTCCAGCCTCAGCCTCCCGAATA
 1259 GCTGGGATTACAGTGGCCACCACCATGCCAGCATGGTGAAACCCCATATCTACTAAAAATAATTT
 1326 TTTTAAAAATTAGCCAGTCACGTTGGTGTCTGCCTGTGGTCCCAGCTACTCGGGAGGCTGAGGCACA
 1393 AGAATCACTTGAACCCAGGAGGCAGAGTGAGCCGAGGTTGTGCCACTGCACTGCAGCCTGGGCTACA
 1460 GAGAGCAAGACTCTGTCTCAAAAAAATAAAATAAAATAAAATAAAAGATATAGTATTAATGCCT
 1527 GTATTTATGAAGACAGGGCATTGCACTGCTGCTATGCTGAAATGAACTTTATAAAGTATTATTGA
 1594 CTATAGCCAACTACAGATGGCATATGCCTTAAAGATGCATTTTCTATGTATTTAATATTTGTGTA
 1661 CAGTAAAAGTCAGAACTCATTTAAGAACTTTAAGGGATCCTTTAGAAACATTTTACTTAATTTCA
 1728 AATGAGTGGGCTGAGCATGTAGACTAGCAGCTTTTATCAAACCTGGCATTGTCAGGTTGAACCTTAT
 1795 AAATACGAATAGTTACTACTCTGACTTGTAAATGTAGCTGCTAAAGTATGAGGTATATTTAAGAAGA
 1862 ATAGATTAGTTTCTATATACCTATAAGAACAAGAATTCAAAGTGAGGTAGTCATCCAGGTTCCCCC
 1929 TCCCACAATATTTTGTATCATATAAATTTAGATGACTTTTTTGGGTTACATTTTATCCATATATTT
 1996 TGAACCTAACAGAACTTGGCAGAAATGTGTGTTGAATATTCTTTACATTTAACAATTAACAACAACTAA
 2063 AACTGAATAGCAAATAGTTTATTGGTAAGTACACGGTTTCAACGGGAGTAATAAATTCACATGAAAA
 2130 GGAGACAATAATCAAGTCAAAAGAATAAATGCTTACTAATCATCAGAAAATCTGTGGCCATTAGGGC
 2197 TGGCACGTAAAAATCCAAATCACTCAGAGGCCAAATCTGTAAAATCAAATGAGGTTGAAGACAAT
 2264 TAGTGAAAGTAATGTTATAGTATCATCATATTCCTTTTGCTATCTGTCCCTTTTCTCTATGCCAGA
 2331 TAAACACTTTCTCTTACTCATCAATCAAGAATAAAATTTCTACATCAAAGTAAGACTTGAGAAAA
 2398 TAAACTTATGTGTATATATGTTTAAACTCCAAAGTTTGGGACATTCAAACAATGTACATGAAATTAT
 2465 ATGTGTGTGGTGTGTGAATAAATATGTATGAATATAAAAAAAAAAAAAAAAA***CTCGAG***

Figure 6.6. cDNA sequence of ZCSL identified by sequencing of Y2H bait1-98 clone 22. Cloning sites are shown in bold and italic. cDNA sequence is shown in codons with the corresponding amino acid above. cDNA sequence is numbered on the left and amino acids on the right. 5' and 3'UTR are indicated by continuous sequence. The first 73 amino acids are not encoded by the insert sequence.

consisting of MBP and the bait protein is created within a plasmid with a chemically inducible bacterial promoter. MBP binds to amylose beads hence the fusion protein can be purified by binding to amylose beads followed by centrifugation and washing. The second protein (potential interactor) is produced by *in-vitro* translation using a transcription-translation kit and [³⁵S] methionine radiolabelled as described in section 2.2.15. The two proteins are incubated at conditions which will promote protein interactions. If the two proteins interact, a complex will be produced consisting of the MBP fusion protein bound to amylose beads and the radiolabelled *in-vitro* translated protein. Stringent washes and centrifugation will isolate the protein interaction complexes but if the two proteins of interest do not interact, the radiolabelled protein will remain in the supernatant and will be washed away. The protein samples bound to the amylose beads are subjected to SDS-PAGE and subsequent Typhoon® phosphor-imager analysis. Maltose binding protein bound to amylose beads is used as a negative control to ensure that maltose binding protein itself does not interact with the *in-vitro* translated protein. The pull-down technique is summarised in Figure 6.7.

cDNAs encoding Dymeclin amino acids 1-98 (DYM1-98) or 449-669 (DYM449-669) were cloned into the vector pMALC2G (which encodes MBP) such that the inserted sequence would be in-frame with the MBP sequence and a chimeric protein produced. Primers *EcoRI* Y2H1-98F and *BamHI* Y2H1-98R for DYM1-98, and *EcoRI* Y2H449-669F and *BamHI* Y2H449-669R for DYM449-669 were used to amplify *DYM* sequence from the template vector DMCFLpCDNA3. Primers were designed such that *EcoRI* and *BamHI* restriction enzyme sites would be introduced at the 5' and 3' ends of the PCR product respectively. Both the vector pMALC2G and the PCR products were digested with *EcoRI* and *BamHI* and the products purified and ligated as described in sections 2.2.9 and 2.2.13.5. Primer sequences are listed in Table 6.1.

The cDNA sequences of the two interactors as illustrated in Figures 6.4 and 6.5 were cloned into pCMVTag3C vector with a T3 promotor by digestion with the enzymes *EcoRI* and *XhoI*. As *XhoI* was found to have a cut site within the *EAP20* insert at cDNA position 300 as illustrated in Figure 6.4, the vector was digested for varying lengths of time to produce a partial cut and the band of the size appropriate to the full length sequence was gel purified and ligated into pCMVTag3C. *In-vitro* translations were carried using pCMVTag3C containing inserts *EFEMP1* and *EAP20*, termed pCMVTag3CEFEMP1 and pCMVTag3CEAP20. Expression levels of the two *in-vitro* translated proteins varied such that a 50 µl reaction was performed for *EFEMP1* whereas a 300 µl reaction was required for *EAP20*. Transformation into *E.coli* BL-21 was performed for MBP constructs and the fusion proteins induced. Culture volumes of

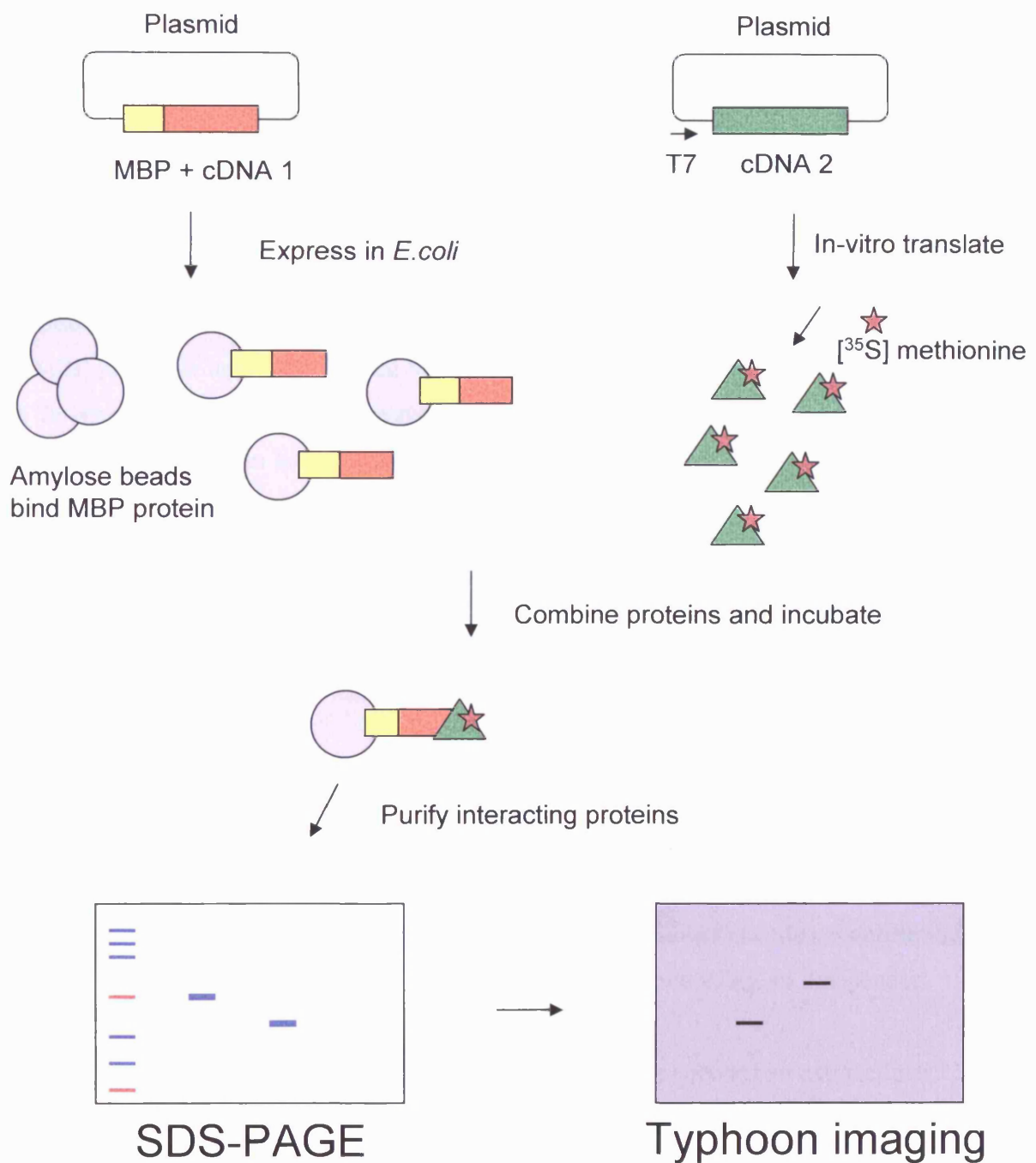


Figure 6.7. Diagram illustrating the maltose binding protein (MBP) pulldown technique. A plasmid is produced containing an MBP tagged cDNA, and another with a cDNA downstream of a T3 promoter. The MBP fusion protein is expressed in *E. coli* and bound to amylose beads. The 2nd protein is *in-vitro* translated with [³⁵S] methionine (indicated by red star) and combined with the MBP fusion protein. Following incubation and washing, if the proteins interact, they are co-purified. The protein samples are resolved by SDS-PAGE and Typhoon imaging used to detect the radiolabelled protein. MBP only protein is used as a negative control.

1, 2 and 3ml were used for each pulldown experiment with MBP, MBPDYM1-98 and MBPDYM449-669 respectively due to their variable expression levels. Induced proteins, including MBP alone were bound to 40 µl amylose beads. Amylose beads previously bound to induced proteins were incubated with *in-vitro* translated proteins as described in section 2.2.20. The resulting proteins were analysed by SDS-PAGE and autoradiography. SDS-PAGE analysis with Coomassie blue staining revealed bands of approximately 50, 55 and 70 kDa corresponding respectively to MBP alone, MBP-DYM1-98 and MBP-DYM449-669. Radioactive proteins of approximately 50 and 25 kD respectively were detected, corresponding to *EFEMP1* and *EAP20* *in-vitro* translation products. The predicted molecular weights are 54 and 20 kD respectively. The MBP pulldown assays performed with MBP-DYM1-98 detected interactions with both *in-vitro* translated proteins however, similar assays performed with MBP-DYM 449-669 did not detect an interaction (see Figure 6.8). Thus we may conclude that the N-terminal 98 amino acids of Dymeclin interact with EFEMP1 and EAP20 whereas the 220 C-terminal amino acids do not.

6.9 Discussion

6.9.1 Dymeclin interactors

EFEMP1 and EAP20 have been identified from this yeast two-hybrid screen as Dymeclin interacting proteins. MBP pulldown assays have confirmed that both EFEMP1 and EAP20 interact with the N-terminal 98 amino acids of Dymeclin but not with the C-terminal 220 amino acids. The yeast two-hybrid screen also identified ZCSL3 as a Dymeclin interactor but this has yet to be confirmed using an independent experimental method.

EFEMP1 is a secreted protein with high homology to fibulin. It encodes an extracellular matrix glycoprotein of 493 amino acids with a predicted molecular weight of 54.6 kD. The fibulin protein family are extracellular matrix proteins containing calcium binding epidermal growth factor like domains and are thought to function as intramolecular bridges in the stabilisation of supramolecular extracellular matrix structures (Argaves *et al.*, 2003). EFEMP1 contains six calcium binding epidermal growth factor like domains. Calcium binding growth factor like domains are required for calcium binding in order to permit protein-protein interactions and are also present in proteins involved in skeletal development including cartilage oligomeric matrix protein and bone morphogenetic protein 1 (Online database PROSITE). Fibulins are thought to be important in fetal development with *Fibulins 1* and *2* shown to be expressed in

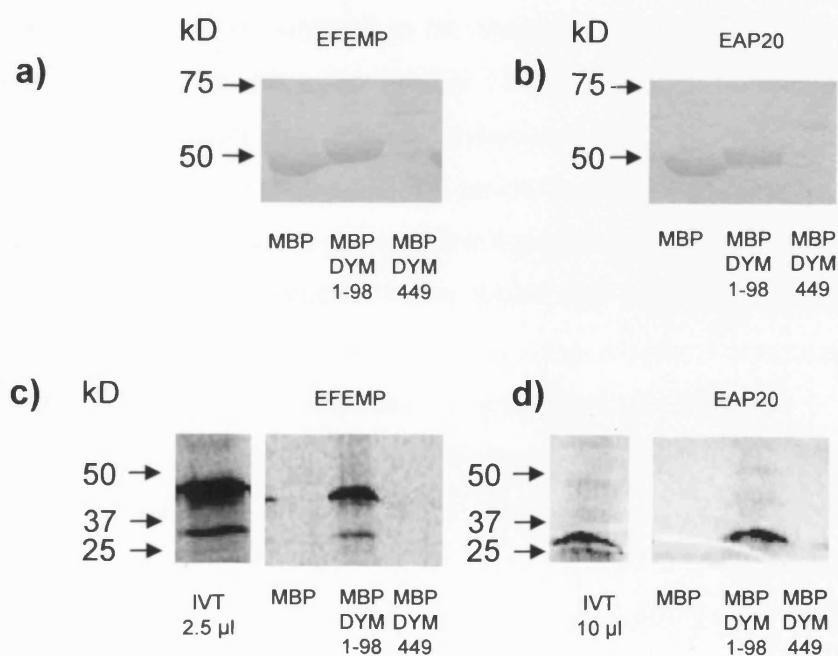


Figure 6.8. In-vitro translation of EFEMP and EAP20 and MBP pulldown. SDS-PAGE with Coomassie blue staining shown in;
a) with bands of ~50,55 and 70kD respectively representing induced MBP, MBP-DYM1-98 and MBP-DYM449-669 proteins for pulldowns with EFEMP.
b) shows induced MBP, MBP-DYM1-98 and MBP-DYM449-669 proteins for pulldowns with EAP20.
c) and d) show Typhoon phosphor-imaging of 35 S-methionine labelled *in-vitro* translated EFEMP and EAP20.

developing vascular, skeletal and neuronal structures. Specifically, human embryos between 6 and 8 weeks show expression of *Fibulins 1* and *2* in all perichondrial structures although not in the chondrocytes themselves. By gestational week 10, the interterritorial matrix surrounding the hypertrophic chondrocytes expresses *Fibulins 1* and *2* (Miosge *et al.*, 1996). Interestingly, disruption of the *Fibulin-1* gene caused by a chromosomal translocation has been reported in synpolydactyly type 2 (OMIM 608180) in which affected individuals have metacarpal and metatarsal fusion and synpolydactyly (Debeer *et al.*, 2002). Mouse *in-situ* hybridisation studies have shown *EFEMP1* (encoding the protein also known as fibulin 3) to be expressed in early development (day 9.5 post-conception) in the condensing mesenchyme which gives rise to bone and cartilage and subsequently (day 14.5 post-conception) in the developing bone of the axial and cranial skeleton (Ehlermann *et al.*, 2003). The *EFEMP1* knockout mouse model develops multiple herniae and pelvic organ prolapse and has a decreased lifespan with signs of premature aging. Histologically, there is evidence of reduced elastic fibre networks in connective tissue fascia and fragmented elastic fibres on electron microscopy. It has thus been hypothesised that *EFEMP1* acts to protect elastic fibres from degradation by ECM enzymes (McLaughlin *et al.*, 2007).

Heterozygous mutations in the *EFEMP1* gene have been reported to cause the human retinal disorders Malattia Leventinese (ML) and Doyme honeycomb retinal dystrophy (DHRD) (Michaelides *et al.*, 2006, Stone *et al.*, 1999) with accumulation of the abnormally folded mutant protein in the endoplasmic reticulum of retinal pigment epithelial cells demonstrated to be the underlying pathogenic mechanism (Marmorstein *et al.*, 2002). However, no retinal abnormalities were seen in the *EFEMP1* knockout mouse model (McLaughlin *et al.*, 2007) although a knock-in mouse with a heterozygous missense mutation in the *EFEMP1* gene had similar eye changes to those found in families with ML/DHRD (Marmorstein *et al.*, 2007, Fu *et al.*, 2007).

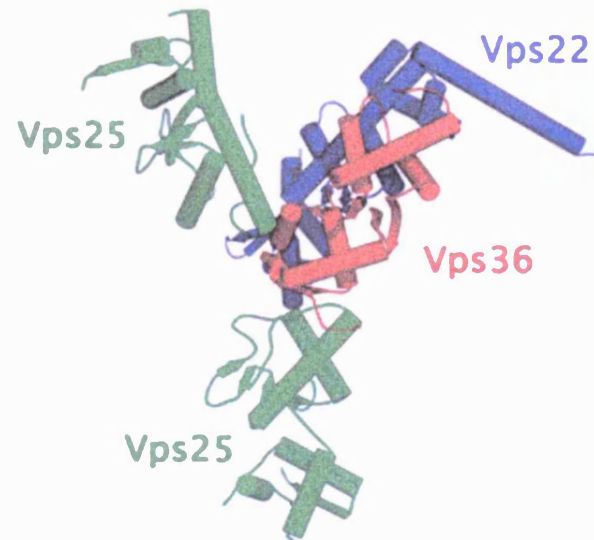
The functional relevance of the interaction between Dymeclin and EFEMP1 has not yet been established. EFEMP1 is a secreted extracellular matrix protein, and given the potential role of Dymeclin in protein transport within the cell, Dymeclin may be transporting EFEMP1. The hypothesis that the extracellular matrix in DMC individuals is deficient in EFEMP1 because of the absence of functioning Dymeclin to transport EFEMP1 is attractive and could be demonstrated by immunohistochemistry of affected tissue. A role in skeletal development for fibulins 1 and 2 is supported by their expression in developing cartilage and bone taken together with the disruption of the fibulin 1 gene in a type of synpolydactyly. This role may be shared by EFEMP1 (fibulin 3) given the homology amongst this protein family.

Dymeclin has also been shown to interact with EAP20, a class E VPS protein (vacuolar sorting proteins involved in the endosomal pathway are known as Class E VPSs) which is predicted to comprise 178 amino acids and have a molecular weight of 20.7 kD. EAP20 is a component of the endosomal sorting complex required for transport II (ESCRT-II), which is a protein complex consisting of 2 molecules of EAP20 in addition to one each of VPS36 and VPS22 (Hierro *et al.*, 2004). ESCRT-II is required for the formation of multivesicular bodies (MVBs) and sorting of endosomal cargo proteins into MVBs. Multivesicular bodies are a type of late endosome which contain multiple vesicles, the purpose of which is to allow storage of transmembrane proteins for subsequent transport elsewhere within the cell. The MVB pathway transports transmembrane proteins either into the lumen of the lysosome for degradation or to the trans-golgi network or cell membrane for recycling. The ESCRT-II complex transiently associates with the MVB in order to allow the transport of cargo proteins into the MVB lumen. This transport process additionally requires the protein complexes ESCRT-I (comprising VPS 23, 28 and 37) and ESCRT-III (comprising VPS 2, 20 and 24 and Snf 7). Cargo proteins are first ubiquitinated which allows them to be recognised and bound by ESCRT-I which subsequently recruits ESCRT-II. This interaction initiates the formation of ESCRT-III which itself transports cargo proteins into the MVB (Hurley and Emr, 2006). The endosomal sorting pathway is not only responsible for the transport of proteins into MVBs. Endosomal sorting is a two-way process and can either transport proteins into the cell from the cell membrane or export them from the trans-golgi network. Proteins can be sent for degradation to the lysosome or recycled and reused either as cell membrane or trans-golgi proteins. Figure 6.9 illustrates the ESCRT-II complex and the endosomal sorting pathway (Hierro *et al.*, 2004).

The only human disorder currently attributed to a Class E vacuolar sorting protein defect is a form of Fronto-temporal dementia which has been shown to be caused by mutations in *VPS2* (Skibinski *et al.*, 2005). Class E *VPS* mouse knockout models are also limited in number. The *VPS54* knockout mouse has a neurological phenotype with ataxia and tremor progressing to paralysis and early death by 3 months of age (Schmitt-John *et al.*, 2005). No human disease has been attributed to a defect in *VPS54* likely because a defect is embryonically lethal. However, mutations in *VPS33*, a class C VPS protein have been shown to cause the arthrogryposis-renal dysfunction-cholestasis (ARC) syndrome, a recessive multi-system disorder caused by defective membrane fusion mechanisms in involved organs (Gissen *et al.*, 2005).

The N-terminal domain of Dymeclin is predicted to be cytoplasmic (MTOPI prediction programme) and has been shown to interact with EAP20 and EFEMP1. Given that

a)



b)

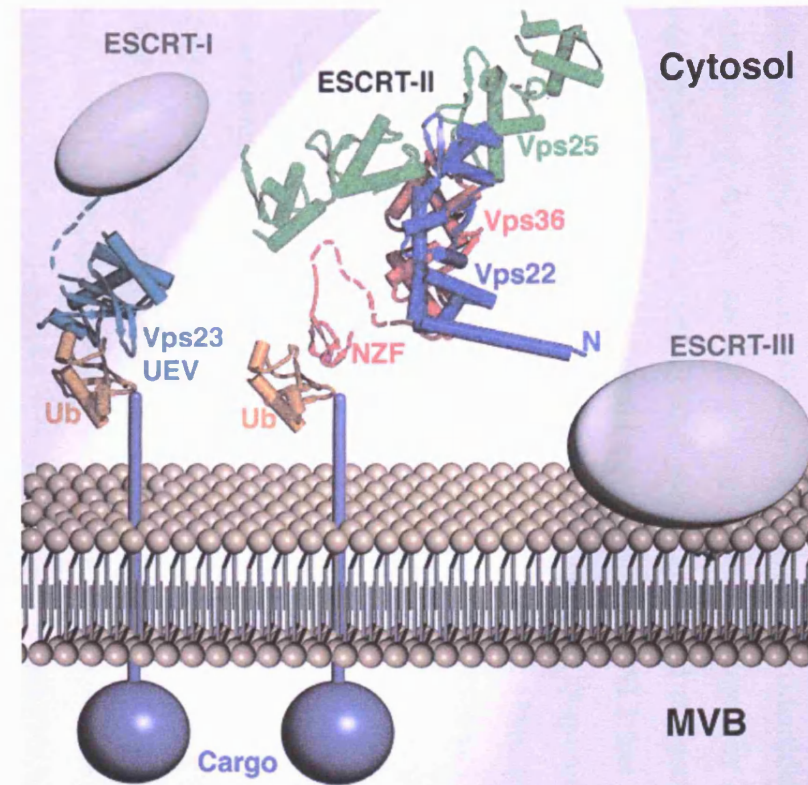


Figure 6.9. a) Structure of the ESCRT-II complex, which comprises two VPS25 (human homologue known as EAP20) molecules in addition to one each of VPS36 and VPS22.
b) ESCRT-II interacts with ESCRT-I and ESCRT-III in order to allow the transport of cargo proteins to the MVB. From Hierro *et al.*

Dymeclin is interacting with both an endosomal sorting component and an extracellular matrix protein, it is hypothesised that Dymeclin is required for cargo transport through the endosomal compartment. Yeast two-hybrid analysis also identified *ZCSL3* as a Dymeclin interactor, however this has not yet been confirmed by the use of an additional experimental method. The *ZCSL3* gene is predicted to encode a protein of 148 amino acids with a molecular weight of 17 kD. *ZCSL3* has been shown to stimulate ATPase activity of Heat shock protein-70 type chaperones, which are responsible for aiding the folding of newly synthesised proteins (online database PROSITE). If Dymeclin is acting as a cargo transporter within the cell, it may interact with chaperone proteins, themselves required for correct protein folding. Dymeclin was found to interact with itself although this interaction has not been independently confirmed. Dymeclin may form dimers or more complex structures, but given that it is a transmembrane protein, this would be unusual.

In light of these findings, it is necessary to consider how the bone and brain specific nature of DMC may arise given that the *DYM* transcript is widely expressed (El Ghouzzi *et al.*, 2003). Cartilage and brain are both tissues which require large amounts of ECM to maintain their integrity, and thus may be more susceptible to abnormality if cells are less able to secrete ECM components. It is possible that chondrocytes are more sensitive to transport defects due to the large size of collagen complexes which are transported to form the extracellular matrix. Neurons secrete many neurotransmitters which have to be transported over long distances in order to traverse the long neuronic processes and thus may also be more sensitive to deficiencies in protein transport.

6.9.2 Other skeletal disorders attributed to defects in ER transport

The skeletal disorder Cranio-lenticulo-sutural dysplasia (CLSD) has been attributed to a defect in a component of the ER to Golgi transport system. CLSD is an autosomal recessive disorder characterized by delay in the closure of skull fontanelles, cataracts, facial dysmorphism and skeletal defects (posterior wedging of the vertebral bodies and narrow iliac wings). Mutations in *SEC23A*, a component of COPII-coated vesicles which transport secretory proteins from the ER to the Golgi were found to be causative of this disorder (Boyadjiev *et al.*, 2006). Microscopic analysis of fibroblasts shows dilatation of the endoplasmic reticulum in individuals affected with CLSD, a feature also seen in DMC. A disruption in ER transport of secretory proteins has been shown to account for the CLSD phenotype with extracellular matrix proteins accumulating in the ER. The zebrafish *crusher* mutant, due to a similar defect in *Sec23a* shows abnormal

craniofacial chondrocyte and bone formation. The ER of chondrocytes becomes distended with accumulating proteins and cartilage ECM deposits, including type II collagen are much reduced, findings which have also been seen in DMC (Lang *et al.*, 2006, Engfeldt *et al.*, 1983). A number of human disorders have been linked to mutations in genes encoding proteins essential for membrane trafficking within the cell. For example, X-linked spondyloepiphyseal dysplasia tarda is a skeletal dysplasia characterised by bony abnormalities of the vertebrae and hips (Gedeon *et al.*, 1999) caused by mutations in a *sedlin*, a component of transport protein particle (TRAPP). TRAPP is a large multiprotein complex shown in yeast to be required for the targeting and fusion of ER-to-Golgi transport vesicles. Mutant *sedlin* is thought to disrupt the secretion of type II collagen and other cartilage components (Oka and Krieger, 2005).

6.9.3 Experimental limitations

The large scale transformation of yeast with prey cDNA library resulted in a transformation efficiency much lower than in previous small scale transformations. This was not sufficiently high to give complete coverage of the cDNA library and hence some interactors may have been missed in this screen. The fact that no two positive clones contained the same cDNA sequence suggests that the screen was not completely saturated and also implies that potential interactors may not have been ascertained. Yeast transformation efficiency is known to be particularly sensitive, with contributory factors including the quality of the carrier DNA, poly-ethylene glycol concentration, the quality and concentration of the cDNA library and the temperature and length of heat-shock (Gietz, 2006). Experimental reagents remained the same throughout with a longer period of heat-shocking used for the large scale transformation, as recommended in all yeast two-hybrid protocols. It is possible that this variable was not optimal which may have resulted in a lower transformation efficiency.

Extraction of prey plasmids from transformed yeast proved experimentally difficult with some positive clones not being successfully rescued and sequenced. Many different experimental methods were used including sodium hydroxide extraction, phenol chloroform extraction and commercially available kits (Qiagen, Crawley, UK) in addition to the use of glass beads in order to break down the yeast cell membrane. Each yeast cell only contains a very small amount of prey vector in addition to the bait vector making extraction difficult. Growth of the yeast strain in –ura/leu medium was performed in order to try and eliminate the prey vector and overcome this problem. However the failure to successfully extract all prey vectors means there may be other functionally interesting interactors which are yet to be identified. The bait protein MBP-

DYM449 expressed at much lower levels on protein induction possibly because it does contain transmembrane domains. This may decrease expression and account for the lack of observed interaction in MBP pulldowns.

6.9.4 Summary

Yeast two-hybrid analysis has identified and confirmed the interaction of Dymeclin with EAP20, a component of the endosomal sorting pathway, and EFEMP1, an extracellular matrix protein. Given that Dymeclin has been shown to localise to the endoplasmic reticulum and affected individuals have been shown to have dilated RER, these findings support the role of Dymeclin in protein transport through the ER.

Chapter 7 - General discussion

DMC is a spondylometaepiphyseal dysplasia associated with mental retardation. At a cellular level chondrocytes fail to columnarise normally leading to irregular ossification (Engfeldt *et al.*, 1983). Dilated RER and multiple vacuoles are seen in DMC affected chondrocytes and fibroblasts leading to the hypothesis that the condition is a storage disorder (El Ghouzzi *et al.*, 2003). The gene causative of DMC was identified by autozygosity mapping as part of the preliminary work leading up to this project and is located on chromosome 18q12-21. It encodes a 669 amino acid protein with homology to no other known proteins but orthologues in many species suggesting a conserved function through evolution. *In-silico* analyses have predicted that Dymeclin lies within a membrane and contains myristoyl, dileucine and dilysine motifs. These motifs have been shown to assist in protein-membrane association, in the targeting of transmembrane proteins to the endosomal-lysosomal system and in Golgi/ER protein trafficking (Cosson and Letourneur, 1997, Boutin, 1997, Bonifacino and Traub, 2003).

This study has identified novel mutations in *DYM*, including two complex genomic duplication/repetition events, which may have arisen as a result of non-allelic homologous recombination. This is the first report of genomic duplication/repetition mutations occurring in a recessive disorder and an important finding because conventional mutation screening by direct sequencing would fail to detect such mutations. Approximately 5% of the human genome is made up of repeat elements hence it is likely that similar mutation mechanisms may result in gene disruption giving rise to other genetic disorders (Cheung *et al.*, 2001). This may account for some individuals in whom linkage is demonstrated to a disease locus but in whom no causative mutation can be identified. The more frequent use of techniques such as multiple amplifiable probe hybridisation (MAPH) to detect exon deletion/duplication and the use of high resolution CGH may in future allow genomic duplications to be detected on a routine basis and could be employed to study further DMC individuals in whom causative mutations have not yet been identified.

Subcellular localisation studies have shown that wild type Dymeclin localises to the ER yet mutant proteins show aberrant localisation in a punctuate distribution within the cytoplasm. This pattern is consistent with localisation within ER exit sites or lysosomes and could be confirmed by further co-localisation studies. *In-vivo*, it is anticipated that NMD would destroy transcripts encoding a truncated protein thus preventing translation

of the protein product. Abnormal cellular localisation was also demonstrated for *DYM* missense mutants whose transcripts would not undergo NMD thus abnormal Dymeclin distribution within the cell may reflect the *in-vivo* situation in patients with such mutations. Only two *DYM* missense mutations have been reported to date, one each occurring in DMC and SMC phenotypes and both as compound heterozygote. The SMC missense mutation (E87K) was found in association with a splice site mutation predicted to result in exon skipping. RT-PCR demonstrated fragments of a length consistent with exon skipping in addition to normally sized fragments not containing the missense mutation indicating the production of some normal transcript (Cohn *et al.*, 2003). Some residual Dymeclin function may account for the SMC phenotype without mental retardation, perhaps because a lower level of functional Dymeclin is required within brain compared to cartilage.

Yeast two-hybrid analysis identified an extracellular matrix protein (EFEMP1) and a component of the endosomal sorting complex (EAP20) as Dymeclin interactors. Dymeclin may be acting as a matrix protein transporter with EFEMP1 as cargo. EAP20 is a component of the endosomal sorting complex required for transport necessary for multivesicular body formation (MVB) and the sorting of endosomal cargo proteins. The MVB pathway transports transmembrane proteins either to lysosomes for degradation or to the trans-golgi network or cell membrane for recycling. This interaction of Dymeclin with EAP20 provides further support that Dymeclin is required for protein trafficking through the ER, although the nature of its cargo is not yet known. The hypothesis that Dymeclin is transporting extracellular matrix components is attractive given that irregularly arranged, thickened collagen fibrils have been described in DMC and are thought to be related to an abnormality of the extracellular matrix make-up (Horton and Scott, 1982). We do not yet know whether the dilated RER and vacuoles seen in DMC cells contain accumulated cargo, mutant Dymeclin or both. Studies to identify the contents of dilated RER in pseudoachondroplasia have found accumulation of cartilage oligomeric matrix protein and type IX collagen associated with compromised chondrocyte function and cell death. It has been suggested that mutant COMP is disrupting the processing of other proteins leading to their accumulation (Maddox *et al.*, 1997, Delot *et al.*, 1998, Hecht *et al.*, 1995). Subsequently, analysis of a mouse model with a missense *COMP* mutation has been shown to have aberrantly localised mutant cartilage oligomeric matrix protein, reduced chondrocyte proliferation and increased apoptosis (Pirog-Garcia *et al.*, 2007). Similar studies in addition to mass spectroscopy could be undertaken on DMC affected tissue in order to determine the

nature of the accumulation and to examine the make-up of the extracellular matrix in order to detect any abnormality.

These findings lead the researcher to question why a protein involved in trafficking through the ER might result in a bone and brain specific phenotype when protein transport is necessary in all cell types. *DYM* RNA species have been identified in a broad range of tissues by RT-PCR and Northern blot analysis. *In-situ* hybridisation has shown that *DYM* RNA is expressed at high levels in fetal and adult mouse cartilage and brain. Unfortunately, it has not been possible to demonstrate Dymeclin protein expression within tissues as experiments using antibodies against Dymeclin were not successful. However, the specific phenotypically relevant findings of Dymeclin expression at high levels within resting and hypertrophic cartilage and within the brain provide compelling evidence that the phenotype is at least in part due to variable expression in different tissues. A further hypothesis is that Dymeclin is acting to transport components of the extracellular matrix as both brain and cartilage have a high requirement for extracellular matrix thus will be more sensitive to abnormalities of their transport.

The use of model systems is a very useful technique in the elucidation of the pathogenetic mechanisms underlying human genetic disorders. Mouse models are probably most commonly used with successful models within the skeletal dysplasia group including the *diastrophic dysplasia sulphate transporter* which causes a number of autosomal recessive phenotypes in humans including achondrogenesis type Ib and diastrophic dysplasia and *perlecan* in Schwartz-Jampel syndrome (Forlino et al., 2005, Rodgers *et al.*, 2006). The zebrafish (*Danio rerio*) has also been used and is an ideal organism to look at skeletal development as the embryo is translucent hence the developing skeleton can be visualised (Mari-Beffa *et al.*, 2007). The zebrafish is evolutionarily apart from human but has advantages over mouse models in that it can be bred quickly in large numbers, takes up relatively little space and costs of upkeep are lower. Basic skeletal development follows the same pattern thus the organism has been used as a model successfully in, for example, the *sec23a* mutant in cranio-lenticulosutural dysplasia, a disorder of abnormal facial bone development (Boyadjiev, S. A. *et al.*, 2006, Lang *et al.*, 2006). A Zebrafish model is currently under development for DMC and it is hoped that this will elucidate further the function of *DYM*.

In summary, this study has identified novel mutations in the *DYM* gene including complex genomic rearrangements and is the first report of this mutation mechanism in an autosomal recessive disorder. Dymeclin has been shown to be an endoplasmic

reticulum protein involved in trafficking through the endosomal compartment. Inadequate levels of functioning Dymeclin result in abnormal chondrocyte differentiation and subsequent phenotypic abnormality.

Appendix A

MULTI-CENTRE RESEARCH ETHICS COMMITTEE RESPONSE FORM

DETAILS OF APPLICANT

1. Name and address of Principal Researcher:

**Professor R C Trembath
Professor of Medical Genetics
University of Leicester
Leicester Royal Infirmary
Department of Genetics
Adrian Building
University Road
Leicester LE1 7RH**

2. Title of Project

**A clinical and molecular genetic Study of the Skeletal Dysplasia
Dyggve Melchior Clausen syndrome**

3. Name and address of Sponsor:

**Medical Research Council
20 Park Crescent
London W1B 1AL**

DETAILS OF MREC:

4. Name and address of MREC:

**South East MREC
Room 76 B Block
40 Eastbourne Terrace
London W2 3QR**

5. MREC Reference Number:

Listed below is a complete record of the review undertaken by the MREC with the decisions made, dates of decisions and the requirement at each stage of the review:

14th January 2004

Approved subject to amendment

- a) Further information is required regarding the assent process.
- b) Further information is required as to how subjects are likely to be recruited to the study.
- c) Clarification is required with regard to comparison of the results from subjects in the study with results from a normal population.
- d) Reassurance is required that an objection from a child to enter the study will not be overridden by consent from their parent(s). This should be stated in the information sheets.
- e) A separate consent forms is needed for the blood test, skin biopsy and bone samples.

Subject Information Sheet

- f) Travel expenses should be paid to subjects.
- g) The Committee questions whether the abbreviation for Smith McCort dysplasia is correct on the information sheets.
- h) The Committee agreed that both conditions being studied should be detailed on the information sheets
- i) The information sheet for the parent should be amended to refer to "your child" instead of "you".
- j) "Principle" is incorrectly spelt at the end of the information sheets and requires amending.
- k) It should be stated that the South East MREC has reviewed and approved the study.
- l) Jargon should be removed from the information sheets.
- m) The results and information relating to new developments should be offered to subjects.
- n) Subjects should be informed of the size of the piece of bone being removed.

THE FINAL DOCUMENTS AND ARRANGEMENTS APPROVED BY THE MREC

The following items have been approved by the South East MREC:

Application form dated 03 12 03
Protocol version 2 January 2004
Professor R C Tremath's curriculum vitae
Sample letter to invitation/introduction version 1 December 2003
Return slip version 1 December 2003
GP/Consultant letter version 1 December 2003
Information for Children version 1 December 2003
Information sheet version 1 December 2003
Adult information sheet version 2 January 2004
Parent information sheet re. Dyggve Melchior Clausen Syndrome and Smith
Mccort Dysplasia – Version 2 January 2004
Consent form for individuals under the age of 18 – version 2 January 2004
Consent form for individuals under the age of 18 for blood, skin and bone version
2 January 2004
Consent form for individuals under the age of 18 for blood version 2 January
2004
Consent form for individuals under the age of 18 for bone version 2 January 2004
Consent form for individuals under the age of 18 for skin biopsy version 2
January 2004
Consent form adult patient – version 2 January 2004
Consent form for adult patient for blood version 2 January 2004
Consent form for adult patient for bone version 2 January 2004
Consent form for adult patient for skin biopsy version 2 January 2004
Consent form parent – version 2 January 2004
Consent form for parent for blood version 2 January 2004
Methods of initial recruitment of study -√
Compensation arrangements for subjects - √
Payments to researcher -√
Provision of expenses for subjects - √

Date of approval **29th January 2004**

Signature of Chairman/Administrator: 

Date: **29th January 2004**

Name: **MRS JANE MARTIN**

Appendix B

Construct Name	Vector	Insert	Insert 5' primer	Insert 3' primer	5'enzyme	3'enzyme
IMAGE 3462520	pCMVSPORT6	3' terminal 1770 bp DYM	N/A	N/A	<i>EcorRV</i>	<i>XhoI</i>
DYMWT+5'UTR pCMVSPORT6	pCMVSPORT6	Full length DYM cDNA inc 5'UTR	EcoRV5'UTR	DMC508R	<i>EcorRV</i>	<i>XhoI</i>
DYMWT+5'UTR pCDNA3	pCDNA3	Full length DYM cDNA inc 5'UTR	N/A	N/A	<i>EcorRV</i>	<i>XhoI</i>
DYMWT 5'myc pCMVSPORT6	pCMVSPORT6	Full length DYM cDNA	DMC 5'myc	DMC508R	<i>EcorRV</i>	<i>XhoI</i>
DYMWT 5'myc pCDNA3	pCMVSPORT6	Full length DYM cDNA	N/A	N/A	<i>EcorRV</i>	<i>XhoI</i>
DYMWT+5'UTR pEGFPC1	pEGFPC1	Full length DYM cDNA inc 5'UTR	N/A	N/A	<i>HindIII</i>	<i>XhoI</i>
DYMWTpEGFPC1	pEGFPC1	Full length DYM cDNA	GFPinframe	GFPR	<i>HindIII</i>	<i>XhoI</i>
DYM1-98 pMALC2G	pMALC2G	DYM codons 1-98	DMCAb1F	DMCAb1R	<i>BamHI</i>	<i>HindIII</i>
DYM449-669 pMALC2G	pMALC2G	DYM codons 449-669	DMCAb2F	DMCAb2R	<i>BamHI</i>	<i>HindIII</i>
DYM1-98pGBDUc1	pGBDUc1	DYM codons 1-98	BamHIY2H1-98F	EcoRIY2H1-98R	<i>BamHI</i>	<i>EcoRI</i>
DYM449-669pGBDUc1	pGBDUc1	DYM codons 449-669	BamHIY2H449F	EcoRIY2H449R	<i>BamHI</i>	<i>EcoRI</i>

Appendix B. Table illustrating *DYM* constructs produced by cloning.

Appendix C

ONLINE MUTATION REPORT

Genomic duplication in Dyggve Melchior Clausen syndrome, a novel mutation mechanism in an autosomal recessive disorder

E Kinning, C Tufarelli, W S Winship, M A Aldred, R C Trembath

J Med Genet 2005;42:e70 (<http://www.jmedgenet.com/cgi/content/full/42/12/e70>). doi: 10.1136/jmg.2005.033829

Background: Dyggve Melchior Clausen syndrome (DMC) is a severe autosomal recessive skeletal dysplasia associated with mental retardation. Direct sequencing of genomic DNA has identified causative mutations in the gene *Dymeclin* (chromosome 18q12–21), with the majority predicting the generation of a truncated protein product.

Objective: To carry out molecular genetic studies in three DMC kindreds.

Results: Two novel nonsense mutations and two complex genomic duplication events resulting in exon repetition were identified.

Conclusions: Exon dosage assessment or mRNA analysis, in addition to direct genomic DNA sequencing, should be employed in the investigation of DMC affected individuals. Genomic duplication may be the causative mutation mechanism in other autosomal recessive disorders.

The skeletal dysplasias are a heterogeneous group of at least 200 distinct clinical disorders affecting approximately 1 in 4000 individuals in European populations.¹ Dyggve Melchior Clausen syndrome (DMC, OMIM 223800), an autosomal recessive spondylo-epimetaphyseal dysplasia, is characterised by short trunk dwarfism, microcephaly, a coarse facial appearance, and mental retardation.² Radiographs show multiple abnormalities including vertebral platyspondyly, lacy iliac wings, laterally displaced irregularly ossified femoral heads, and a hypoplastic odontoid.³ Aspects of the DMC phenotype mirror the clinical features seen in the mucopolysaccharidosis known as Morquio's disease; hence chondrocyte dysfunction was postulated to be a result of abnormal metabolite storage.² However, biochemical studies have been inconsistent, such that pathogenesis of DMC remains unclear. Autozygosity mapping led to the localisation of a DMC disease gene to chromosome 18q12–21, with no evidence of locus heterogeneity.^{4–5} A few mutations were detected in a gene encoding the predicted transcript termed FLJ 20071.^{6–7} The DMC gene spans approximately 400 kb and is composed of at least 17 exons.^{6–7} *Dymeclin*, the protein product of the DMC gene, is widely expressed and several splicing isoforms have been predicted (Aceview, NCBI), and differentially sized transcripts seen on northern blot.⁸ The longest open reading frame encodes a putative transmembrane protein of 669 amino acids (75 kDa), but to date the exact nature and function of the protein is unknown. The analysis of naturally occurring mutations is instrumental in gaining functional insights and more generally in understanding the molecular mechanisms causing human genetic disease.

For a better understanding of the allelic architecture of DMC mutations, we undertook a molecular screen of affected subjects

from three kindreds. We now report the identification and characterisation of four novel disease alleles including complex genomic rearrangements which, when expressed, result in exon duplication or repetition. Recent studies have suggested that regions of the human genome may be subject to regional duplication/deletion on a large scale, which we show may lead to gene disruption and subsequent disease.⁹

METHODS

Genotyping and DMC gene mutation analysis

Four subjects from three independently ascertained kindreds were investigated. Multicentre research ethics approval was given for the study (South East MREC, London). Pedigrees are illustrated in fig 1. Genomic DNA was extracted from peripheral blood leucocytes according to standard protocols. Microsatellite markers extending over a 3.1 Mb distance surrounding the DMC locus (fig 1 and supplementary tables 1 and 2*) were amplified in multiplex polymerase chain reactions (PCR) and the products separated by electrophoresis on an ABI 377 DNA fragment analyser. Alleles were scored manually using Genescan v3.1 and Genotyper v2.0 software (Applied Biosystems, Foster City, California, USA).

All 16 coding exons of the DMC gene, together with the exon–intron boundaries, were amplified as previously described.⁶ Products were purified and sequenced on an ABI 377 automatic sequencer using the fluorescent dideoxy-terminator method. Primer sequences are available as supplementary material.

Reverse transcriptase PCR and fluorescent dosage PCR

RNA was extracted either from peripheral blood collected into Paxgene® collection tubes (Qiagen, Valencia, California, USA) following the manufacturer's instructions or from EBV transformed lymphocyte cell lines (subjects 7261 and 7345) using RNeasy kit (Qiagen). Complementary DNA was synthesised by priming with oligo-dT using First Strand cDNA synthesis kit (Roche Diagnostics, Lewes, UK). DMC cDNA was amplified in eight overlapping PCR products and size separated on 2% agarose gel (supplementary table 3). Products of an aberrant size were extracted using a Gel purification kit (Qiagen) and purified DNA sequenced as described above.

Fluorescent dosage PCR to quantify locus specific genomic DNA was carried out as a multiplex assay using the primers listed in supplementary table 4 and Qiagen multiplex PCR master mix. DMC exons 2, 14, and 15 were amplified together with *ALK1* exon 2 (control), samples run on an ABI

Abbreviations: DMC, Dyggve Melchior Clausen syndrome; DQ, dosage quotient; STR, simple tandem repeat

*The supplementary tables can be viewed on the journal website (www.jmedgenet.com/supplemental).

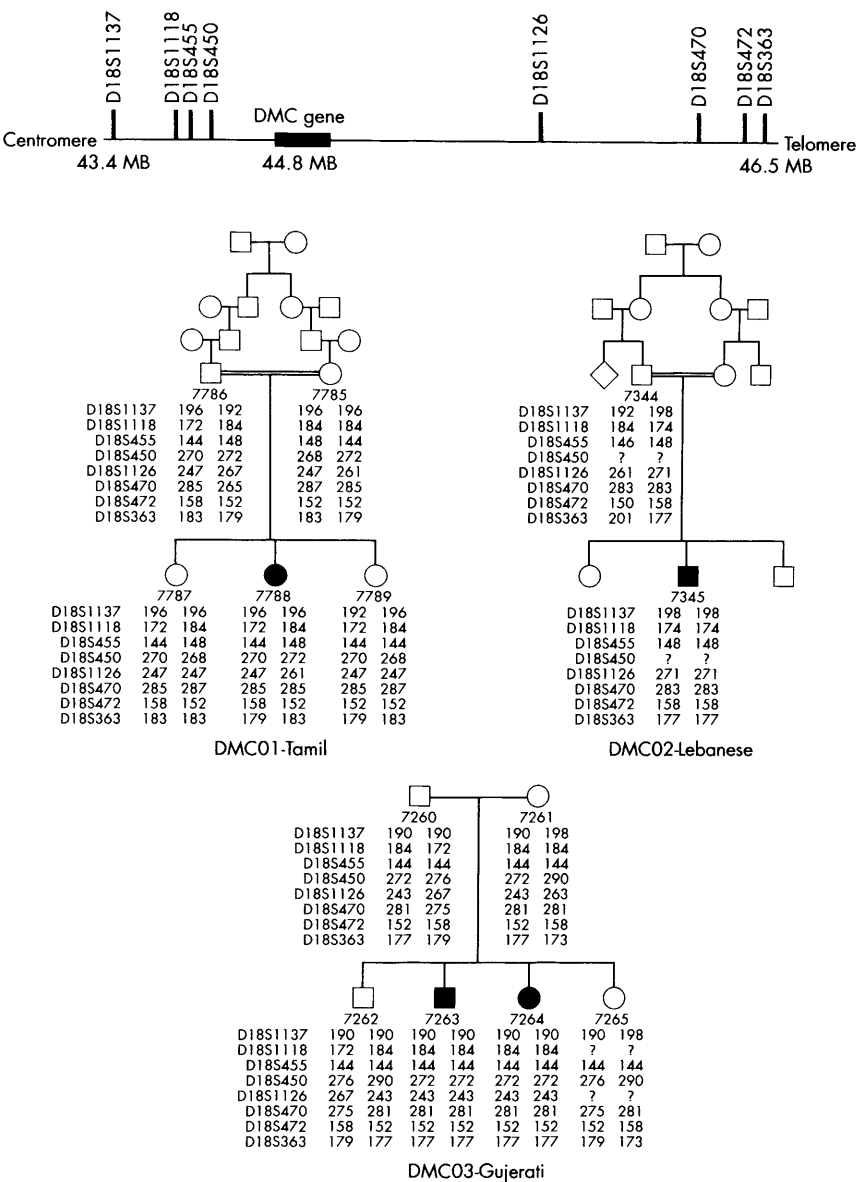


Figure 1 Position of simple tandem repeat (STR) markers around DMC locus. Pedigrees DMC01-03 including ethnic origin and STR marker results.

377 DNA fragment analyser, and gels analysed using the software packages Genescan v3.1 and Genotyper v2.0. Three independent experiments were done and used for comparison. The mean peak area for each exon was calculated for 10 control samples and used to compare with patient samples. Dosage quotients (DQ) were calculated for the exon of interest in relation to two other DMC exons or ALK1 exon 2 using the following formula, as previously described⁹: $DQ \text{ exon } a/\text{exon } b = [\text{sample exon } a \text{ peak area}/\text{sample exon } b \text{ peak area}] \div [\text{control exon } a \text{ peak area}/\text{control exon } b \text{ peak area}]$.

Southern blot analysis

A total of 10 µg of genomic DNA was digested with the restriction enzyme *PvuII* or *ApaI* and separated on 0.8% agarose gel. After electrophoresis DNA was transferred to Zeta probe membrane (Biorad) by alkaline blotting according to the manufacturer's instructions. A 25 ng sample of probe DNA was ³²P labelled with 15 µCi [α -³²P]dCTP (NEN Life Science, Boston, Massachusetts, USA) in a standard random primed

reaction (Invitrogen, San Diego, California, USA). Hybridisation was undertaken at 65°C for 16 hours, membranes were washed twice at 65°C in 2×SSC/0.1% SDS (sodium dodecyl sulphate) for 20 minutes, then twice in 0.1×SSC/0.1% SDS at 65°C for 20 minutes. The filters were exposed to a Storage Phosphor screen for 16 hours and analysed on a Typhoon[®] phosphorimager using ImageQuant software (Amersham Biosciences, Amersham, UK). DMC exon 2 and exon 14 probes correspond, respectively, to 298 base pairs (bp) of DMC cDNA including all of exon 2 and 1040 bp of DMC genomic DNA including all of exon 14, as illustrated in fig 3A.

RESULTS

Identification of new mutations in the DMC gene

In kindred DMC01, the proband (7758) was heterozygous for all informative genotypes at surrounding polymorphic markers (fig 1). Direct sequencing showed a paternally inherited C→T transition at nucleotide position 208 (exon 4) leading to the introduction of a premature stop codon

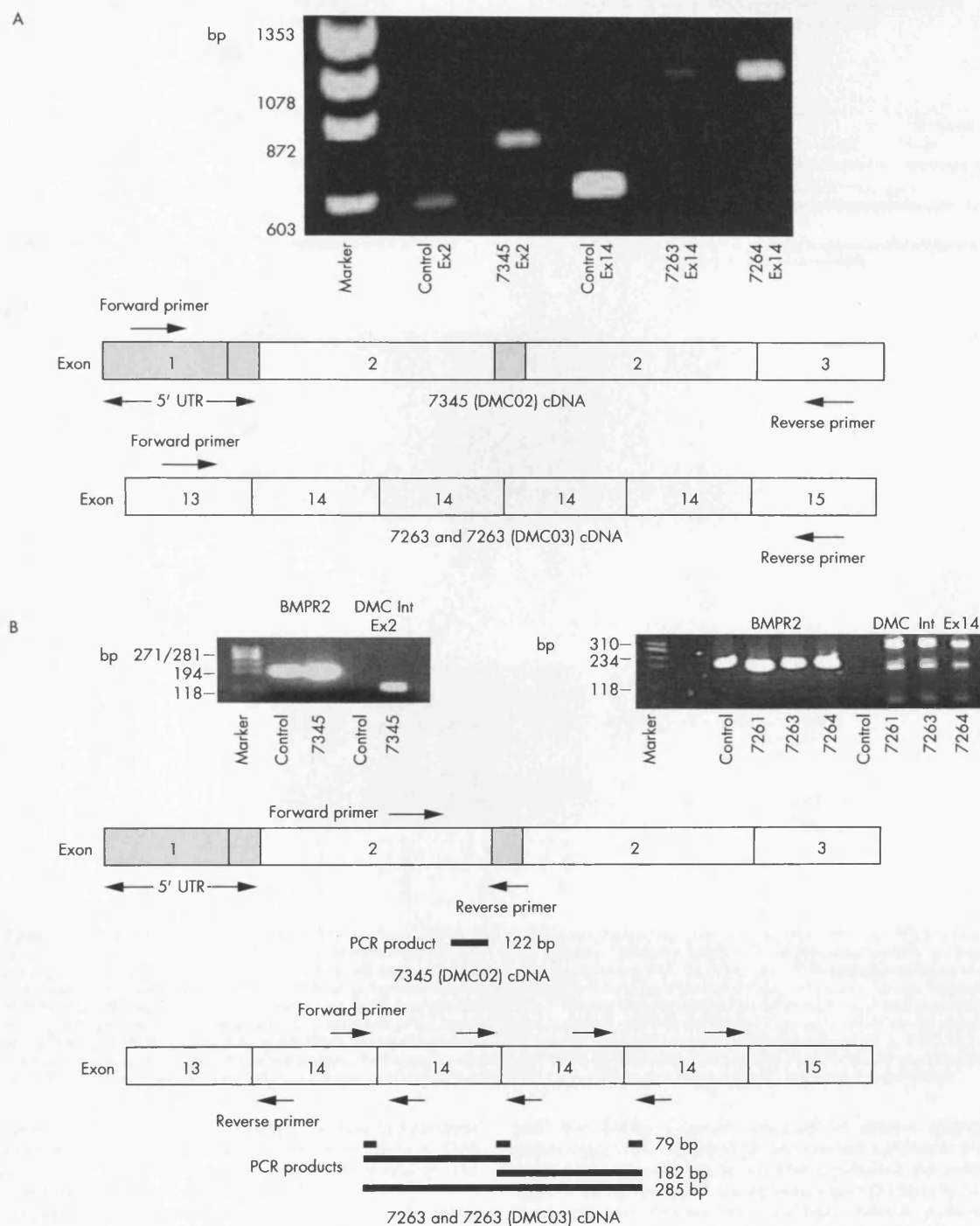


Figure 2 (A) Reverse transcriptase polymerase chain reaction (RT-PCR) gel image and schematic of cDNA amplified. Dyggve Melchior Clausen syndrome (DMC) cDNA amplified with forward primer in exon 1 (5' untranslated region) and reverse primer in exon 3 produces product of 574 base pairs (bp) in control compared with approximately 800 bp in 7345. DMC cDNA amplified with forward primer in exon 13 and reverse primer in exon 15 produces product of 594 bp in control compared with approximately 1000 bp in 7263 and 7264. (B) RT-PCR gel image and schematic of aberrant subject cDNA amplified. Forward primer at the end of exon 2 and reverse primer at the beginning of exon 2 do not produce a product in control cDNA as expected. A product of 122 bp is shown in 7345. Forward primer at the end of exon 14 and reverse primer at the beginning of exon 14 do not produce a product in control cDNA but products of 79 182 and 285 bp are shown in 7261, 7263, and 7264. A BMPR2 amplicon of 191 bp is produced for all samples.

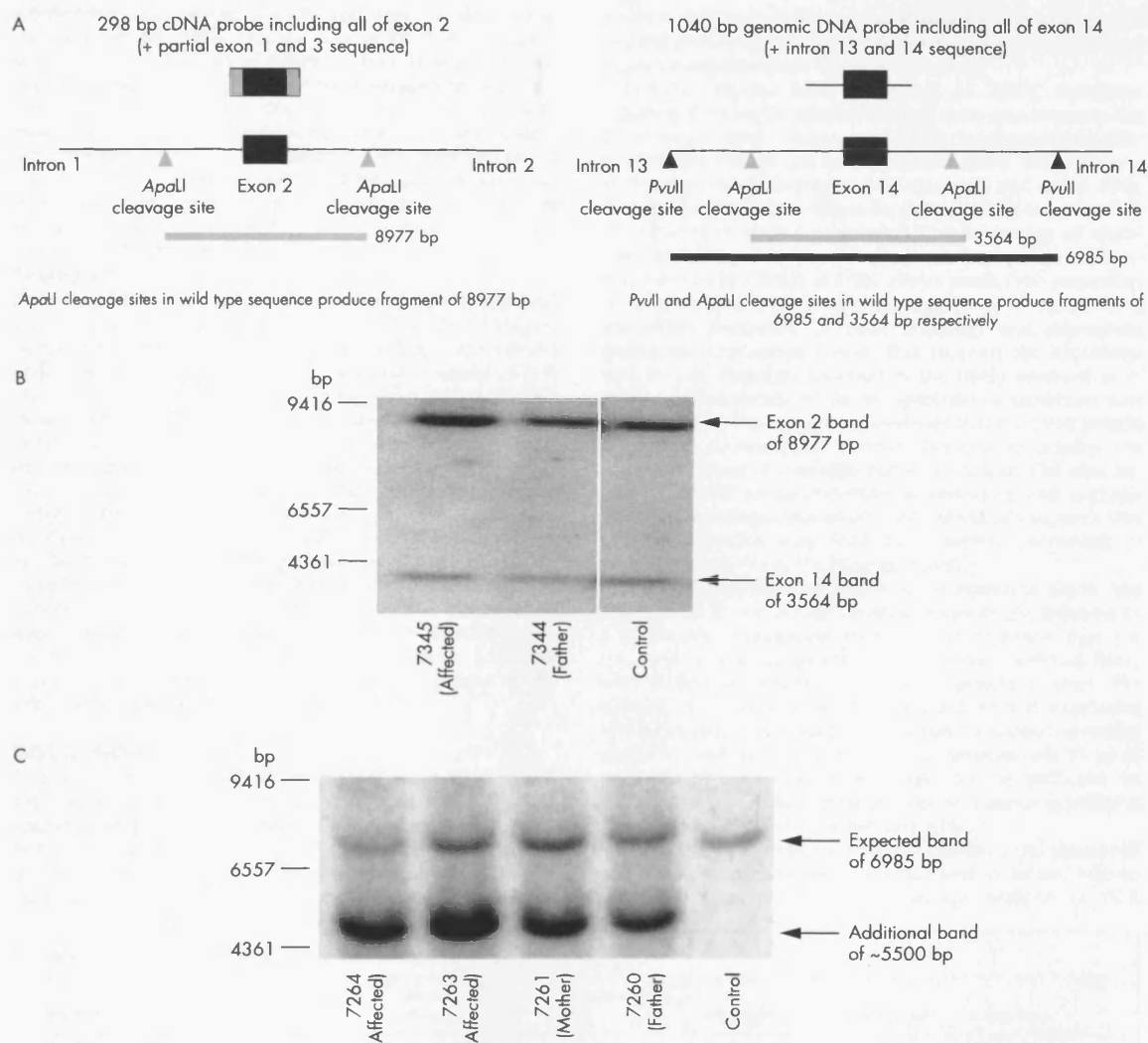


Figure 3 (A) Schematic illustrating probes used for Southern blotting, and restriction enzyme cleavage sites. Exon 2 probe comprises 298 base pairs (bp) of cDNA sequence including all of exon 2 in addition to partial exon 1 and 3 sequence. Following digestion of wild type sequence with *ApalI* and probing with exon 2 probe, a fragment of 8977 bp will be produced. Exon 14 probe comprises 1040 bp of genomic DNA sequence including all of exon 14 and surrounding intronic sequence. Following digestion of wild type sequence with *ApalI* or *PvuII* and probing with exon 14 probe, fragments of 3564 and 6985 bp, respectively, will be produced. (B) Southern blot analysis. Following digestion of genomic DNA with *ApalI* and sequential probing with DMC exon 2 and exon 14 probes, fragments of the expected sizes, respectively 8977 and 3564 bp, were seen both in control samples and in DMC02 proband (7345) and father (7344). Densitometry of exon 2 to exon 14 restriction fragments showed a ratio of 2:1 in 7345, 3:2 in 7344, and 1:1 in control. (C) Southern blot analysis. *PvuII* digestion of genomic DNA in DMC03 subjects (7263 and 7264) and both parents (7260 and 7261) revealed a smaller aberrantly sized fragment of approximately 5500 bp in addition to a fragment of the expected size (6985 bp).

replacing an arginine residue (R70X). The maternal chromosome harboured a mutation at nucleotide position 1363 (exon 12), also a C→T transition, which results in the introduction of a premature stop codon (R455X).

Homozygosity for all eight simple tandem repeat (STR) markers across the DMC locus at chromosome 18 was observed for each of the three affected subjects from kindred DMC02 (7345) and DMC03 (7263, 7264). Family studies identified distinct putative disease bearing haplotypes not shared by either family (fig 1). Direct sequencing of genomic DNA in each of the affected subjects failed to reveal any deleterious sequence variants. Reverse transcriptase polymerase chain reaction (RT-PCR) analysis of lymphocyte derived DMC exon 1 to exon 3 cDNA from subject 7345 (DMC02) revealed an aberrantly sized product of approximately 800 bp compared

with the 574 bp fragment observed in control subjects. Sequencing of the gel purified product revealed a 193 nucleotide duplication encompassing all of exon 2 including the coding region (140 bp) and the 5' untranslated region (53 bp) (fig 2A).

Amplification of exons 13 to 15 of DMC cDNA of probands in kindred DMC03 identified approximately 1000 bp fragments, significantly larger than the 594 bp amplification seen in control subjects, and sequence analysis revealed tandem repeats of four copies of exon 14 (fig 2A). To confirm these findings, RT-PCR was carried out using primer pairs, each predicted to generate a product specific to the aberrant repeated cDNA sequences but not in the normal sequence (fig 2B and supplementary table 3). Bands of the sizes expected (122 bp in 7345 and 79, 182, and 285 bp in 7263 and 7264) were seen in affected individuals and parents from

whom RNA was available (fig 2B) and were not detected in control (normal) cDNA. Positive control RT-PCR amplification of a 191 bp product of BMPR2 (chromosome 2q36–37) was done and revealed a product in all samples tested.

Analysis by fluorescent dosage PCR revealed DQs for DMC exon 2 compared with either DMC exons 14, 15, or ALK1 exon 6 (chromosome 12q11–14) in individual 7345 (DMC02) of ~2 (exon 2 as the numerator). DQs for DMC exon 14 compared with either DMC exons 2, 15, or ALK1 exon 6 in affected subjects 7263 and 7264 (DMC03) were ~4 (data not shown).

Southern blot analysis

Following digestion of genomic DNA with *Apa*LI and sequential probing with DMC exon 2 and exon 14 probes, bands of the expected sizes (8977 and 3564 bp, respectively) were seen in control samples and DMC02 proband (7345) and father (7344), as illustrated in fig 3B. A maternal DNA sample was not available. ImageQuant software analysis confirmed the ratio of exon 2 and exon 14 band density as 2:1 in 7345 (affected) and 3:2 in 7344 (father) normalised to control samples. Following digestion of high molecular weight DNA with the restriction enzyme *Pvu*II and Southern blot analysis with a DMC exon 14 probe, a band of the expected size (6985 bp) was seen both in control samples and DMC03 affected individuals (7263, 7264) and parents (7260, 7261). However, an additional fragment of approximately 5500 bp was observed in each DMC03 family member tested, as shown in fig 3C. Analysis following digestion with *Apa*LI revealed restriction fragments of the expected size only.

DISCUSSION

In a screen of subjects affected with DMC we have identified four novel DMC mutations including complex genomic rearrangements that when transcribed result in exon duplication or repetition. We unexpectedly identified two nonsense mutant alleles inherited as compound heterozygote in the offspring of cousin parents (DMC01). These findings highlight a

potential limitation to the method of autozygosity mapping (the method used to identify the DMC disease gene), widely adopted in gene mapping of rare recessive disorders.

Previous reports have identified 16 DMC mutations distributed among 21 kindreds (table 1) by direct sequencing of genomic DNA; hence gross deletions/rearrangements/duplications would not have been identified. These disease alleles were seen to segregate with classical DMC or the allelic disorder Smith McCort dysplasia, distinguished from DMC by the absence of mental retardation.^{6–7,10} Reviewing all mutations including those in this report, it is apparent that the majority (14/20 (70%)) of DMC alleles predict the generation of a truncated protein product, some as a result of a frameshift (including by exon skipping) and subsequent premature termination codon. This supports the hypothesis that loss of dymeclin function is the likely mechanism of disease pathogenesis. While the spectrum of mutations now identified provides compelling evidence that the DMC protein product is necessary for normal chondrocyte activity, the precise function of dymeclin remains unclear. The observation of dilated rough endoplasmic reticulum and multiple vacuoles in chondrocytes of affected individuals suggests that lack of dymeclin may lead to abnormal processing or defective synthesis of cartilage proteins.¹¹

The exon 2 duplication and exon 14 repetition predict the insertion of 22 and six amino acids, respectively, followed by a premature termination codon. It is expected that the transcript would be destroyed by nonsense mediated decay with subsequent failure to produce a protein product. The ATG initiator codon is located in exon 2; thus if translation was initiated in the second copy of exon 2 a normal transcript would be produced. However, exon 2 contains only 53 bp of untranslated sequence, which may not be sufficient to promote the use of this initiation site and hence translation may preferentially occur at the first ATG.

The existence of exon repetition of mRNA in the absence of genome duplication has been reported both in rat and human genes.^{12–13} However, fluorescent dosage analysis of PCR

Table 1 Dyygve Melchior Clausen syndrome (DMC) mutations

Nucleotide change	DMC exon	Amino acid change	Consequence on protein	Phenotype	Ethnic origin	Reference
G→C 34 bp 3' of exon 1	Intron 1		?	DMC	Portuguese	10
Duplication exon 2	Exon 2	S47R +FS	Frameshift and truncation	DMC	Lebanese	This report
C48G	2	Y16X	Truncated	DMC	Dominican	7
C208T	4	R70X	Truncated	DMC	Tamil	This report
IVS 3 194-1G→A	4 (5' end) splice acceptor		Skipping exon 3	DMC	Lebanese	10
G259A	4	E87K	Missense	SMC	Guamese	7
IVS 4 288-2A→G	5 (5' end) splice acceptor		Skipping exon 4	DMC	Spanish	10
T396A	5	Y132X	Truncated	DMC	Not stated	7
C580T	7	R194X	Truncated	DMC	Tunisian	6
C610T	7	R204X	Truncated	DMC	Moroccan	6
T656G	8	R219X	Truncated	DMC	Moroccan	6
763delA	8	T254Q	Frameshift and truncation	DMC	Not stated	7
IVS 7-1 T→G	8 (5' end) splice acceptor		Skipping exon 7	SMC	Guamese	7
IVS 10 1125-1G→T	10 (3' end) splice donor		Skipping exon 11 (in frame)	DMC	Moroccan	6
C1363T	12	R455X	Truncation	DMC	Tamil	This report
IVS 11 1252-1G→A	12 (5' end) splice acceptor		Skipping exon 11 (in frame)	DMC	Lebanese	6
A1405T	13	N469Y	Missense	DMC	Not stated	7
C1447T	13	Q483X	Truncated	DMC	Moroccan	6
Repetition 4 copies exon 14	14	A525F+FS	Frameshift and truncation	DMC	Gujerati	This report
1877delA	17	K626N+FS	Frameshift and extension	DMC	Moroccan	6

amplified lymphocyte derived genomic DNA provided independent confirmation of the genomic rearrangements. Southern blot analysis in subjects from kindred DMC02 did not reveal any novel restriction fragments, and a doubling of exon 2 band density in the proband (compared with exon 14 and normalised to control) suggests that the duplicated sequence includes the entire sequence between *Apa*LI restriction sites (fig 3, panels A and B). In contrast, hybridisation with the DMC exon 14 probe after *Pvu*II digestion of genomic DNA from DMC03 subjects revealed an aberrantly sized fragment of approximately 5500 bp in addition to a fragment of the expected size (6985 bp), as illustrated in fig 3C. *Apa*LI digestion showed only fragments of the expected size. We interpret the above data to suggest that at least one of the external copies has been unaffected by the duplication, and hence a fragment of the expected size is produced. The exact nature of the duplication is not yet known but our results suggest the duplicated sequence is likely to start between the *Apa*LI and *Pvu*II restriction sites (fig 3A).

We used BLAST analysis of the DMC intronic sequence surrounding exons 2 and 14 to search for regions of homology where non-allelic homologous recombination or replication slippage may have arisen. This revealed multiple Alu repeat sequences, including several in the region between *Pvu*II and *Apa*LI cut sites surrounding exon 14 which Southern blot analysis indicated may be the site of initiation of duplication. Rearrangements arising at Alu elements have been reported in several conditions including Duchenne muscular dystrophy, and are estimated to account for 0.3% of human genetic diseases.¹⁴ Recent studies have highlighted the existence of variation in copy number of large segments of the human genome and suggested that this phenomenon may be more common than previously thought. Genomic imbalance was detected in approximately 50% of individuals tested in one study, including regions containing known genes.¹⁵ However, the sample sizes were small and using mathematical estimates based on Duchenne muscular dystrophy as a model, others have suggested that one in eight humans will have a segmental deletion and one in 50 a duplication.⁸ Interestingly, increased gene density has been found in duplications, suggesting that this phenomenon may be an important mechanism underlying mutations causing many human diseases.¹⁶ It may even account for a proportion of individuals whose phenotype appears to be linked to a disease gene but in whom mutations are not identified by standard mutation screening techniques.

While we were unable to characterise exactly the real extent of genomic duplication, implicit in our observations is the assumption that the duplication events encompass sufficient intronic sequence to enable normal splicing. We are now seeking to characterise in more detail the duplicated intronic sequence in both cases, which may provide further insight into the underlying mechanism.

Conclusions

This study has identified novel mutations including complex genomic rearrangements of the DMC gene, each likely to disrupt the normal function of the protein product significantly. Genomic duplication mutations have not previously been reported in DMC and this study highlights the merit of using additional methods of mutation analysis in this severe skeletal dysplasia. Genomic duplication leading to gene disruption may be an important mechanism underlying other human genetic disorders.

ACKNOWLEDGEMENTS

We would like to thank the families who participated in this study, Shanta Patel, Nabitha Chabilal, and Dr Ashok Vellodi for assistance

with patient recruitment, and Professor Christine Hall (Great Ormond Street Hospital, London) for review of radiographs. This work was funded by the Medical Research Council through a clinical training fellowship to EK.

ELECTRONIC DATABASES



The supplementary tables can be viewed on the journal website (www.jmedgenet.com/supplemental).

Authors' affiliations

E Kinning, C Tufarelli, M A Aldred, R C Trembath, Division of Medical Genetics, Departments of Genetics and Cardiovascular Science, University of Leicester, Leicester, UK

E Kinning, R C Trembath, Leicestershire Genetics Centre, Leicester Royal Infirmary

W S Winship, Department of Paediatrics, University of Natal, South Africa

Competing interests: none declared

Correspondence to: Professor Richard Trembath, Division of Medical Genetics, Department of Genetics and Cardiovascular Science, University of Leicester, University Road, Leicester LE1 7RH, UK; richard.trembath@genetics.kcl.ac.uk

Received 8 April 2005

Revised version received 26 May 2005

Accepted for publication 30 May 2005

REFERENCES

- 1 Stoll C DB, Roth MP. Birth prevalence rates of skeletal dysplasias. *Clin Genet* 1989;**35**:88–92.
- 2 Beighton P. Dyggve-Melchior-Clausen syndrome. *J Med Genet* 1990;**27**:512–15.
- 3 Spranger J, Maroteaux P, Der Kaloustian VM. The Dyggve-Melchior-Clausen syndrome. *Radiology* 1975;**114**:415–21.
- 4 Ehteshami N, Cantor RM, King LM, Reinker K, Powell BR, Shanske A, Unger S, Rimoin DL, Cohn DH. Evidence that Smith-McCort dysplasia and Dyggve-Melchior-Clausen dysplasia are allelic disorders that result from mutations in a gene on chromosome 18q12. *Am J Hum Genet* 2002;**71**:947–51.
- 5 Thauvin-Robinet C, El Ghouzzi V, Chemaitilly W, Dagoneau N, Boute O, Viot G, Megarbane A, Sefiani A, Munnich A, Le Merrer M, Cormier-Daire V. Homozygosity mapping of a Dyggve-Melchior-Clausen syndrome gene to chromosome 18q21.1. *J Med Genet* 2002;**39**:714–17.
- 6 El Ghouzzi V, Dagoneau N, Kinning E, Thauvin-Robinet C, Chemaitilly W, Prost-Squarcioni C, Al Gazali L, Verloes A, Le Merrer M, Munnich A, Trembath RC, Cormier-Daire V. Mutations in a novel gene *Dymeclin* (FU20071) are responsible for Dyggve-Melchior-Clausen syndrome. *Hum Mol Genet* 2003;**12**:357–64.
- 7 Cohn DH, Ehteshami N, Krakow D, Unger S, Shanske A, Reinker K, Powell BR, Rimoin DL. Mental retardation and abnormal skeletal development (Dyggve-Melchior-Clausen dysplasia) due to mutations in a novel, evolutionarily conserved gene. *Am J Hum Genet* 2003;**72**:419–28.
- 8 Van Ommen GB. Frequency of new copy number variation in humans. *Nat Genet* 2004;**37**:333–4.
- 9 Yau SC BW, Mathew CG, Abbs SJ. Accurate diagnosis of carriers of deletions and duplications in Duchenne/Becker muscular dystrophy by fluorescent dosage analysis. *J Med Genet* 1996;**33**:550–9.
- 10 Paupe V, Gilbert T, Le Merrer M, Munnich A, Cormier-Daire V, El-Ghouzzi V. Recent advances in Dyggve-Melchior-Clausen syndrome. *Mol Genet Metab* 2004;**83**:51–9.
- 11 Engfeldt B, Bui TH, Eklof O, Hjerpe A, Reinholdt FP, Ritzen EM, Wikstrom B. Dyggve-Melchior-Clausen dysplasia. Morphological and biochemical findings in cartilage growth zones. *Acta Paediatr Scand* 1983;**72**:269–74.
- 12 Frantz SA, Thiara AS, Lodwick D, Ng LL, Eperon IC, Samani NJ. Exon repetition in mRNA. *Proc Natl Acad Sci USA* 1999;**96**:5400–5.
- 13 Rigatti R, Jia JH, Samani NJ, Eperon IC. Exon repetition: a major pathway for processing mRNA of some genes is allele specific. *Nucleic Acids Res* 2004;**32**:441–6.
- 14 Deininger PL, Batzer MA. Alu repeats and Human Disease. *Mol Genet Metab* 1999;**67**:183–93.
- 15 Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Ying Q, Schere SW, Lee C. Detection of large-scale variation in the human genome. *Nat Genet* 2004;**36**:949–51.
- 16 Fredman D, White SJ, Potter S, Eichler EE, Den Dunstan JT, Brookes AJ. Complex SNP-related sequence variation in segmental genome duplications. *Nat Genet* 2004;**36**:861–6.

Appendix D

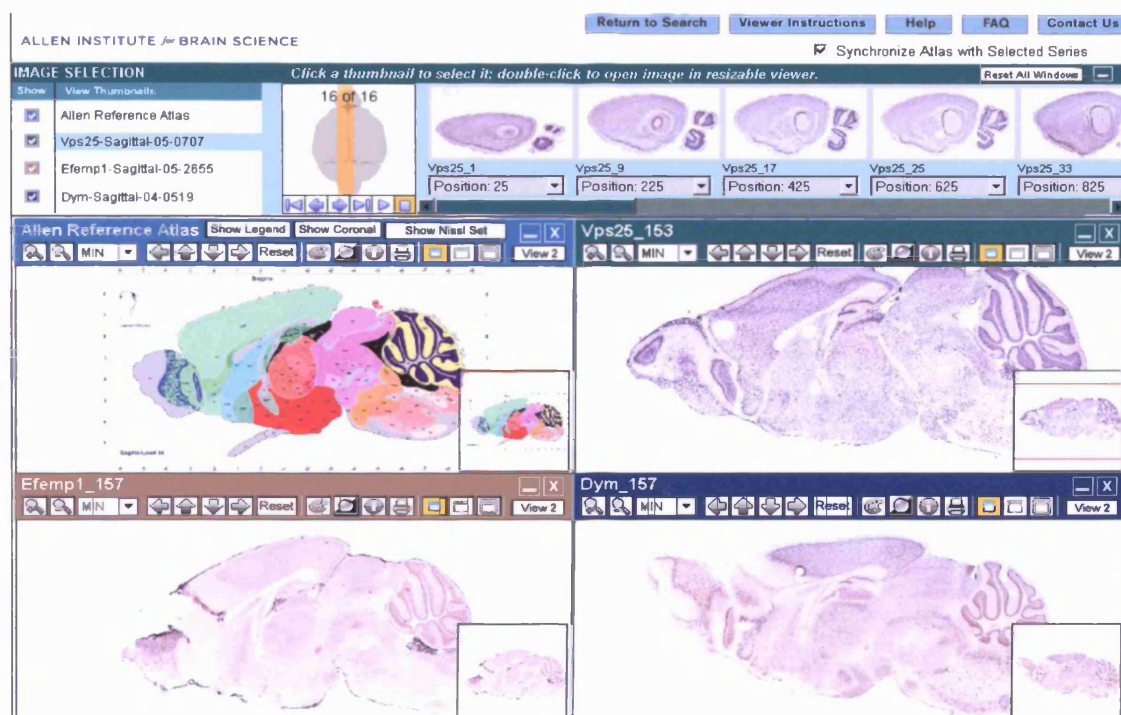


Figure D (i). Allen atlas mouse brain images, sagittal sections of in-situ hybridisation using *VPS25* (*EAP20*), *EFEMP1* and *DYM* probes.

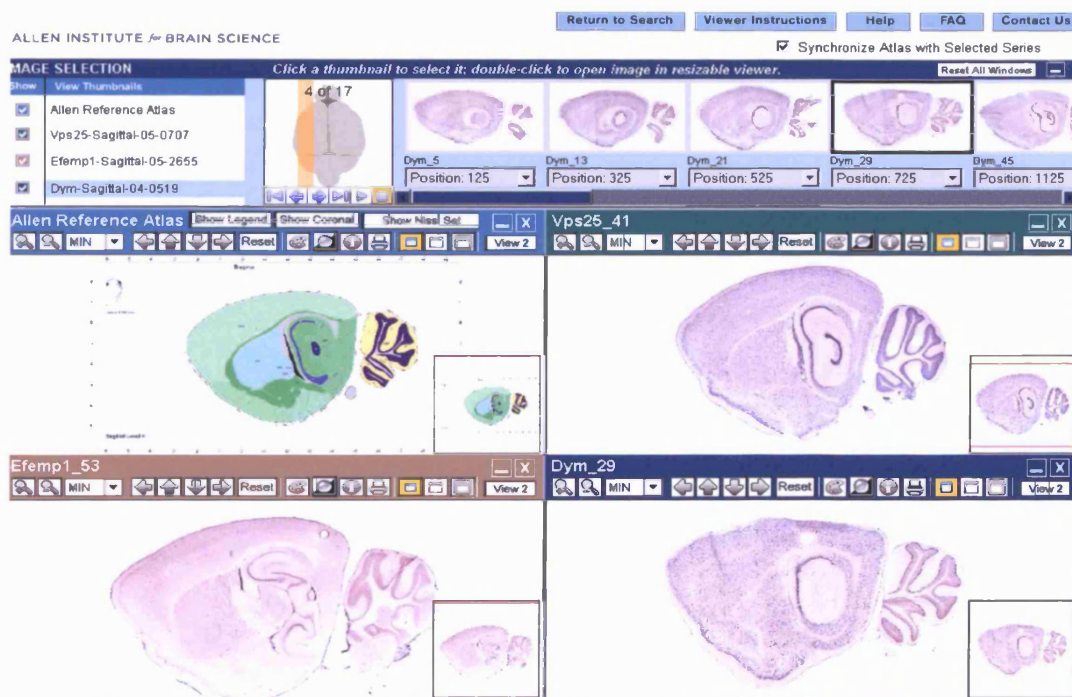
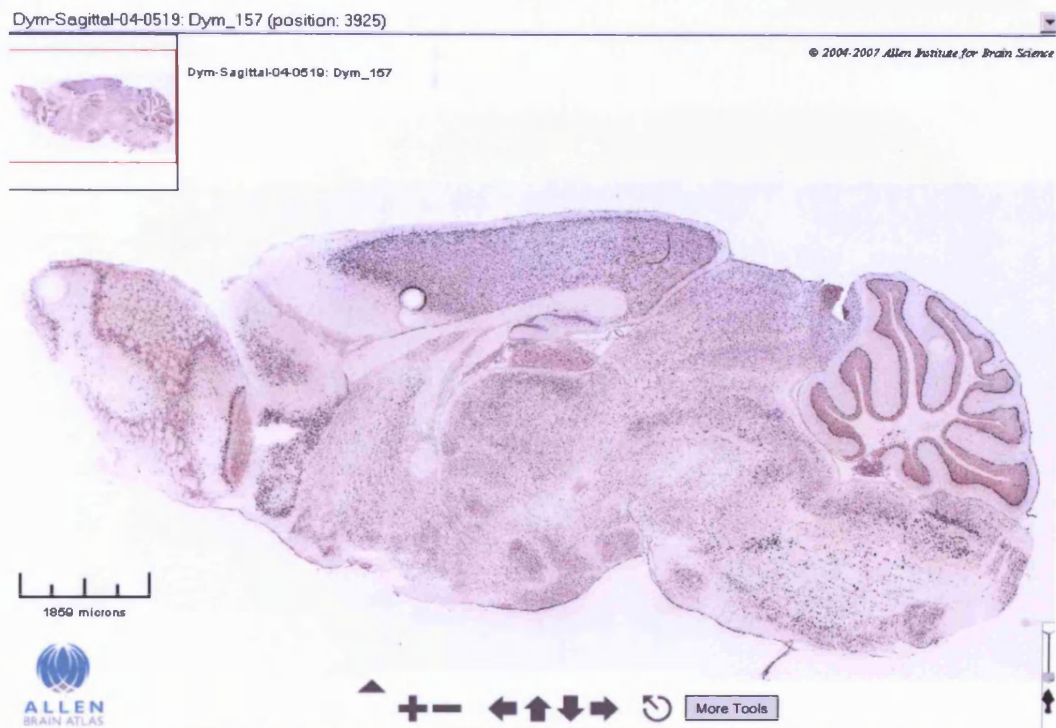


Figure D (ii). Allen atlas mouse brain images, parasagittal section of in-situ hybridisation using *VPS25* (*EAP20*), *EFEMP1* and *DYM* probes.

a)



b)

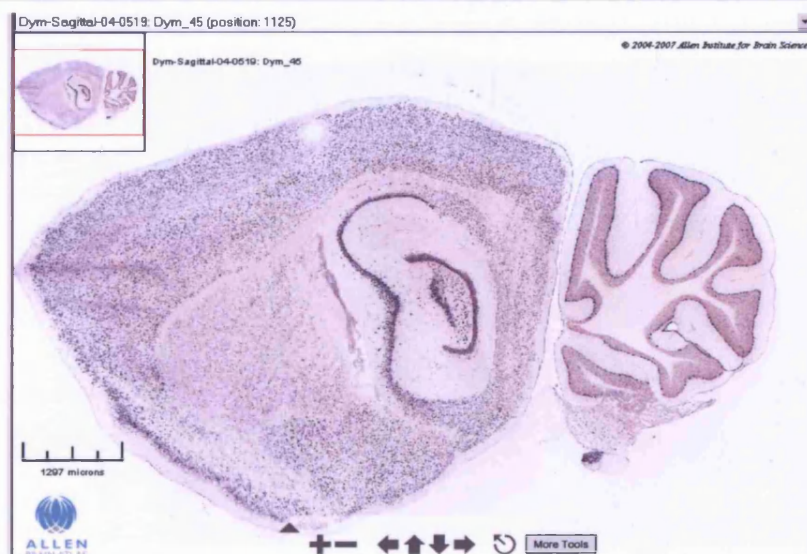
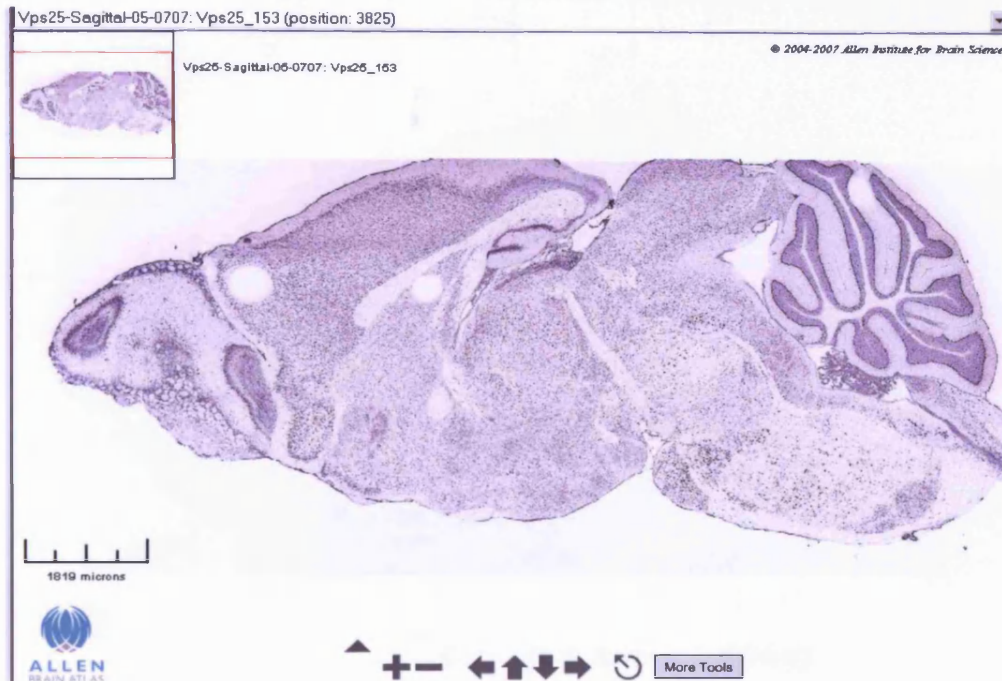


Figure D (iii). Allen atlas mouse brain images, sagittal (a) and parasagittal (b) sections of in-situ hybridisation using *DYM* probe.

a)



b)

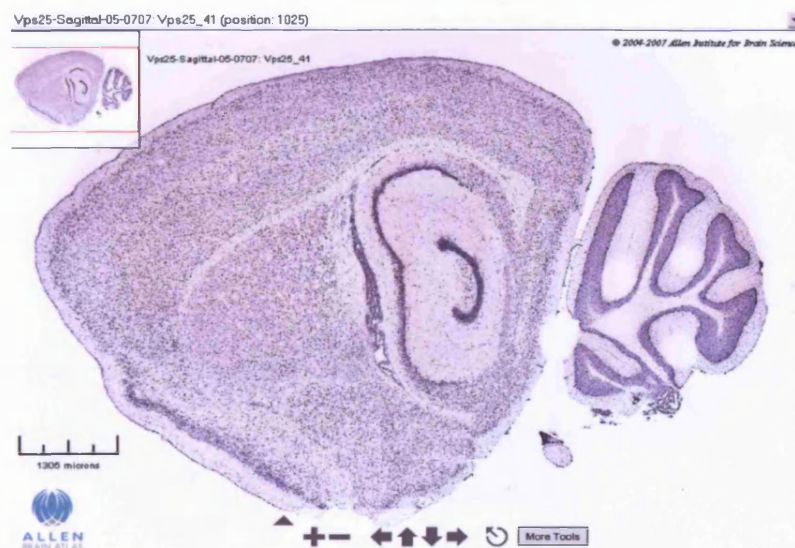
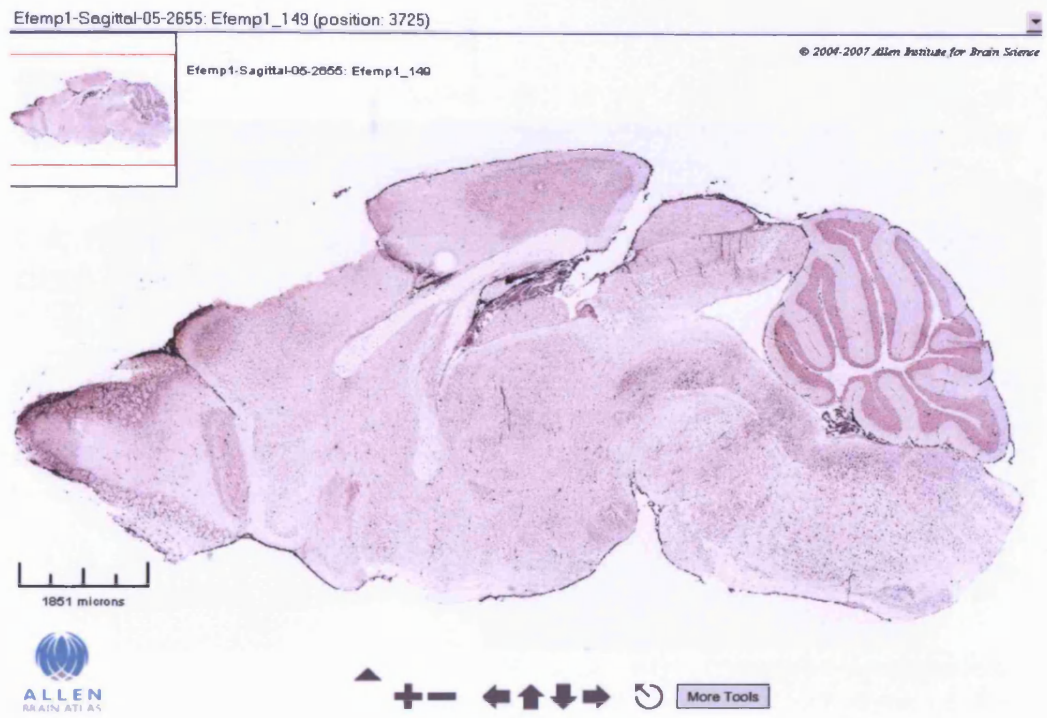


Figure D (iv). Allen atlas mouse brain images, sagittal (a) and parasagittal (b) sections of in-situ hybridisation using *VPS25* (*EAP20*) probe.

a)



b)

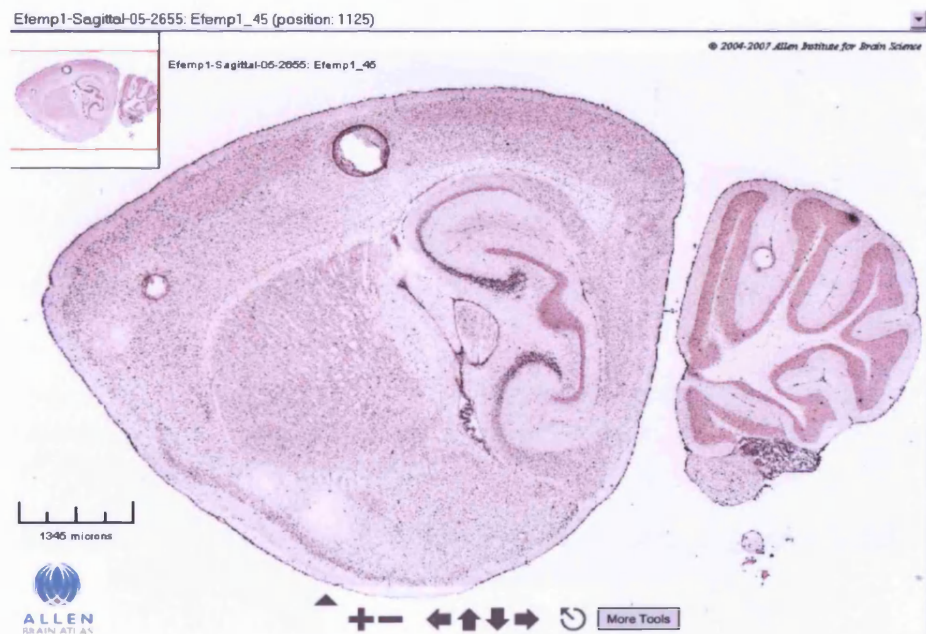


Figure D (v). Allen atlas mouse brain images, sagittal (a) and parasagittal (b) sections of in-situ hybridisation using *EFEMP1* probe.

References

- Agatep, R., Kirkpatrick, R., Parchchaliuk, D., *et al.* 1998. Transformation of *Saccharomyces cerevisiae* by the lithium acetate/single stranded carrier DNA/polyethylene glycol protocol. *Technical Tips Online*, **1**(pp. P01525).
- Ailion, M. & Thomas, J. H. 2003. Isolation and characterization of high-temperature-induced Dauer formation mutants in *Caenorhabditis elegans*. *Genetics*, **165**(1), pp. 127-44.
- Akiyama, H., Chaboissier, M. C., Martin, J. F., *et al.* 2002. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. *Genes Dev*, **16**(21), pp. 2813-28.
- Argraves, W. S., Greene, L. M., Cooley, M. A., *et al.* 2003. Fibulins: physiological and disease perspectives. *EMBO Rep*, **4**(12), pp. 1127-31.
- Arikawa-Hirasawa, E., Wilcox, W. R., Le, A. H., *et al.* 2001. Dyssegmental dysplasia, Silverman-Handmaker type, is caused by functional null mutations of the perlecan gene. *Nat Genet*, **27**(4), pp. 431-4.
- Baker, K. E. & Parker, R. 2004. Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr Opin Cell Biol*, **16**(3), pp. 293-9.
- Ballock, R. T. & O'Keefe, R. J. 2003. Physiology and pathophysiology of the growth plate. *Birth Defects Res C Embryo Today*, **69**(2), pp. 123-43.
- Beighton, P. 1990. Dyggve-Melchior-Clausen syndrome. *J Med Genet*, **27**(8), pp. 512-5.
- Beighton, P., Giedion, Z. A., Gorlin, R., *et al.* 1992. International classification of osteochondrodysplasias. International Working Group on Constitutional Diseases of Bone. *Am J Med Genet*, **44**(2), pp. 223-9.
- Bonafede, R. P. & Beighton, P. 1978. The Dyggve-Melchior-Clausen syndrome in adult siblings. *Clin Genet*, **14**(1), pp. 24-30.
- Bond, J. & Woods, C. G. 2006. Cytoskeletal genes regulating brain size. *Curr Opin Cell Biol*, **18**(1), pp. 95-101.
- Bonifacino, J. S. & Traub, L. M. 2003. Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annu Rev Biochem*, **72**(pp. 395-447).
- Boutin, J. A. 1997. Myristoylation. *Cell Signal*, **9**(1), pp. 15-35.
- Boyadjiev, S. A., Fromme, J. C., Ben, J., *et al.* 2006. Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. *Nat Genet*, **38**(10), pp. 1192-7.

- Brent, R. & Ptashne, M. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell*, **43**(3 Pt 2), pp. 729-36.
- Chandrasekhar, S., Kleinman, H. K., Hassell, J. R., *et al.* 1984. Regulation of type I collagen fibril assembly by link protein and proteoglycans. *Coll Relat Res*, **4**(5), pp. 323-37.
- Cheung, V. G., Nowak, N., Jang, W., *et al.* 2001. Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature*, **409**(6822), pp. 953-8.
- Chiang, A. P., Beck, J. S., Yen, H. J., *et al.* 2006. Homozygosity mapping with SNP arrays identifies TRIM32, an E3 ubiquitin ligase, as a Bardet-Biedl syndrome gene (BBS11). *Proc Natl Acad Sci U S A*, **103**(16), pp. 6287-92.
- Chien, C. T., Bartel, P. L., Sternglanz, R., *et al.* 1991. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc Natl Acad Sci U S A*, **88**(21), pp. 9578-82.
- Cohen, M. M., Jr. 2006. The new bone biology: pathologic, molecular, and clinical correlates. *Am J Med Genet A*, **140**(23), pp. 2646-706.
- Cohn, D. H., Ehteshami, N., Krakow, D., *et al.* 2003. Mental retardation and abnormal skeletal development (Dyggve-Melchior-Clausen dysplasia) due to mutations in a novel, evolutionarily conserved gene. *Am J Hum Genet*, **72**(2), pp. 419-28.
- Colnot, C., Lu, C., Hu, D., *et al.* 2004. Distinguishing the contributions of the perichondrium, cartilage, and vascular endothelium to skeletal development. *Dev Biol*, **269**(1), pp. 55-69.
- Cooper, R. R., Pedrini-Mille, A. & Ponseti, I. V. 1973. Metaphyseal dysostosis: a rough surfaced endoplasmic reticulum storage defect. *Lab Invest*, **28**(1), pp. 119-25.
- Cosson, P. & Letourneur, F. 1997. Coatamer (COPI)-coated vesicles: role in intracellular transport and protein sorting. *Curr Opin Cell Biol*, **9**(4), pp. 484-7.
- Cotterill, S. L., Jackson, G. C., Leighton, M. P., *et al.* 2005. Multiple epiphyseal dysplasia mutations in MATN3 cause misfolding of the A-domain and prevent secretion of mutant matrilin-3. *Hum Mutat*, **26**(6), pp. 557-65.
- Debeer, P., Schoenmakers, E. F., Twal, W. O., *et al.* 2002. The fibulin-1 gene (FBLN1) is disrupted in a t(12;22) associated with a complex type of synpolydactyly. *J Med Genet*, **39**(2), pp. 98-104.
- Deininger, P. L. & Batzer, M. A. 1999. Alu repeats and human disease. *Mol Genet Metab*, **67**(3), pp. 183-93.
- DeLise, A. M., Fischer, L. & Tuan, R. S. 2000. Cellular interactions and signaling in cartilage development. *Osteoarthritis Cartilage*, **8**(5), pp. 309-34.
- Delot, E., Brodie, S. G., King, L. M., *et al.* 1998. Physiological and pathological secretion of cartilage oligomeric matrix protein by cells in culture. *J Biol Chem*, **273**(41), pp. 26692-7.

- Dhami, P., Coffey, A. J., Abbs, S., *et al.* 2005. Exon array CGH: detection of copy-number changes at the resolution of individual exons in the human genome. *Am J Hum Genet*, **76**(5), pp. 750-62.
- Duden, R. 2003. ER-to-Golgi transport: COP I and COP II function (Review). *Mol Membr Biol*, **20**(3), pp. 197-207.
- Dyggve, H. V., Melchior, J. C. & Clausen, J. 1962. Morquio-Ullrich's disease: an inborn error of metabolism? . *Archives of Disease in Childhood* **37**(pp. 525-534).
- Ehlermann, J., Weber, S., Pfisterer, P., *et al.* 2003. Cloning, expression and characterization of the murine Efemp1, a gene mutated in Doyme-Honeycomb retinal dystrophy. *Gene Expr Patterns*, **3**(4), pp. 441-7.
- Ehtesham, N., Cantor, R. M., King, L. M., *et al.* 2002. Evidence that Smith-McCort dysplasia and Dyggve-Melchior-Clausen dysplasia are allelic disorders that result from mutations in a gene on chromosome 18q12. *Am J Hum Genet*, **71**(4), pp. 947-51.
- El Ghouzzi, V., Dagoneau, N., Kinning, E., *et al.* 2003. Mutations in a novel gene Dymeclin (FLJ20071) are responsible for Dyggve-Melchior-Clausen syndrome. *Hum Mol Genet*, **12**(3), pp. 357-64.
- Engfeldt, B., Bui, T. H., Eklof, O., *et al.* 1983. Dyggve-Melchior-Clausen dysplasia. Morphological and biochemical findings in cartilage growth zones. *Acta Paediatr Scand*, **72**(2), pp. 269-74.
- Enomoto, H., Enomoto-Iwamoto, M., Iwamoto, M., *et al.* 2000. Cbfa1 is a positive regulatory factor in chondrocyte maturation. *J Biol Chem*, **275**(12), pp. 8695-702.
- Fields, S. 2005. High-throughput two-hybrid analysis. The promise and the peril. *Febs J*, **272**(21), pp. 5391-9.
- Fields, S. & Song, O. 1989. A novel genetic system to detect protein-protein interactions. *Nature*, **340**(6230), pp. 245-6.
- Forlino, A., Piazza, R., Tiveron, C., *et al.* 2005. A diastrophic dysplasia sulfate transporter (SLC26A2) mutant mouse: morphological and biochemical characterization of the resulting chondrodysplasia phenotype. *Hum Mol Genet*, **14**(6), pp. 859-71.
- Frantz, S. A., Thiara, A. S., Lodwick, D., *et al.* 1999. Exon repetition in mRNA. *Proc Natl Acad Sci U S A*, **96**(10), pp. 5400-5.
- Fredman, D., White, S. J., Potter, S., *et al.* 2004. Complex SNP-related sequence variation in segmental genome duplications. *Nat Genet*, **36**(8), pp. 861-6.
- Fu, L., Garland, D., Yang, Z., *et al.* 2007. The R345W mutation in EFEMP1 is pathogenic and causes AMD-like deposits in mice. *Hum Mol Genet*, **16**(20), pp. 2411-22.

- Gedeon, A. K., Colley, A., Jamieson, R., *et al.* 1999. Identification of the gene (SEDL) causing X-linked spondyloepiphyseal dysplasia tarda. *Nat Genet*, **22**(4), pp. 400-4.
- Gerber, H. P., Vu, T. H., Ryan, A. M., *et al.* 1999. VEGF couples hypertrophic cartilage remodeling, ossification and angiogenesis during endochondral bone formation. *Nat Med*, **5**(6), pp. 623-8.
- Gietz, R. D. 2006. Yeast two-hybrid system screening. *Methods Mol Biol*, **313**(pp. 345-71.
- Girisha, K. M., Cormier-Daire, V., Heuertz, S., *et al.* 2007. Novel mutation and atlantoaxial dislocation in two siblings from India with Dyggve-Melchior-Clausen syndrome. *Eur J Med Genet.* epub.
- Gissen, P., Johnson, C. A., Gentle, D., *et al.* 2005. Comparative evolutionary analysis of VPS33 homologues: genetic and functional insights. *Hum Mol Genet*, **14**(10), pp. 1261-70.
- Gleghorn, L., Ramesar, R., Beighton, P., *et al.* 2005. A mutation in the variable repeat region of the aggrecan gene (AGC1) causes a form of spondyloepiphyseal dysplasia associated with severe, premature osteoarthritis. *Am J Hum Genet*, **77**(3), pp. 484-90.
- Gorski, J. L., Estrada, L., Hu, C., *et al.* 2000. Skeletal-specific expression of Fgd1 during bone formation and skeletal defects in faciogenital dysplasia (FGDY; Aarskog syndrome). *Dev Dyn*, **218**(4), pp. 573-86.
- Griffith, E., Walker, S., Martin, C. A., *et al.* 2008. Mutations in pericentrin cause Seckel syndrome with defective ATR-dependent DNA damage signaling. *Nat Genet*, **40**(2), pp. 232-6.
- Haque, F., Lloyd, D. J., Smallwood, D. T., *et al.* 2006. SUN1 interacts with nuclear lamin A and cytoplasmic nesprins to provide a physical connection between the nuclear lamina and the cytoskeleton. *Mol Cell Biol*, **26**(10), pp. 3738-51.
- Hecht, J. T., Makitie, O., Hayes, E., *et al.* 2004. Chondrocyte cell death and intracellular distribution of COMP and type IX collagen in the pseudoachondroplasia growth plate. *J Orthop Res*, **22**(4), pp. 759-67.
- Hecht, J. T., Nelson, L. D., Crowder, E., *et al.* 1995. Mutations in exon 17B of cartilage oligomeric matrix protein (COMP) cause pseudoachondroplasia. *Nat Genet*, **10**(3), pp. 325-9.
- Hierro, A., Sun, J., Rusnak, A. S., *et al.* 2004. Structure of the ESCRT-II endosomal trafficking complex. *Nature*, **431**(7005), pp. 221-5.
- Horton, W. A. & Scott, C. I. 1982. Dyggve-Melchior-Clausen syndrome. A histochemical study of the growth plate. *J Bone Joint Surg Am*, **64**(3), pp. 408-15.
- Hou, P., Estrada, L., Kinley, A. W., *et al.* 2003. Fgd1, the Cdc42 GEF responsible for Faciogenital Dysplasia, directly interacts with cortactin and mAbp1 to modulate cell shape. *Hum Mol Genet*, **12**(16), pp. 1981-93.

- Hurley, J. H. & Emr, S. D. 2006. The ESCRT complexes: structure and mechanism of a membrane-trafficking network. *Annu Rev Biophys Biomol Struct*, **35**(pp. 277-98.
- Hussain, S., Witt, E., Huber, P. A., *et al.* 2003. Direct interaction of the Fanconi anaemia protein FANCG with BRCA2/FANCD1. *Human Molecular Genetics*, **12**(19), pp. 2503-10.
- Huynh, D. P., Scoles, D. R., Nguyen, D., *et al.* 2003. The autosomal recessive juvenile Parkinson disease gene product, parkin, interacts with and ubiquitinates synaptotagmin XI. *Human Molecular Genetics*, **12**(20), pp. 2587-97.
- Iafrate, A. J., Feuk, L., Rivera, M. N., *et al.* 2004. Detection of large-scale variation in the human genome. *Nat Genet*, **36**(9), pp. 949-51.
- James, P., Halladay, J. & Craig, E. A. 1996. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics*, **144**(4), pp. 1425-36.
- Kandziora, F., Neumann, L., Schnake, K. J., *et al.* 2002. Atlantoaxial instability in Dyggve-Melchior-Clausen syndrome. Case report and review of the literature. *J Neurosurg*, **96**(1 Suppl), pp. 112-7.
- Kinning, E., Tufarelli, C., Winship, W. S., *et al.* 2005. Genomic duplication in Dyggve Melchior Clausen syndrome, a novel mutation mechanism in an autosomal recessive disorder. *J Med Genet*, **42**(12), pp. e70.
- Kirkpatrick, T. J., Au, K. S., Mastrobattista, J. M., *et al.* 2003. Identification of a mutation in the Indian Hedgehog (IHH) gene causing brachydactyly type A1 and evidence for a third locus. *J Med Genet*, **40**(1), pp. 42-4.
- Korf, U. & Wiemann, S. 2005. Protein microarrays as a discovery tool for studying protein-protein interactions. *Expert Rev Proteomics*, **2**(1), pp. 13-26.
- Kornak, U. & Mundlos, S. 2003. Genetic disorders of the skeleton: a developmental approach. *Am J Hum Genet*, **73**(3), pp. 447-74.
- Kuijjer, R., van de Stadt, R. J., de Koning, M. H., *et al.* 1988. Influence of cartilage proteoglycans on type II collagen fibrillogenesis. *Connect Tissue Res*, **17**(2), pp. 83-97.
- Lander, E. S. & Botstein, D. 1987. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science*, **236**(4808), pp. 1567-70.
- Lang, M. R., Lapierre, L. A., Frotscher, M., *et al.* 2006. Secretory COPII coat component Sec23a is essential for craniofacial chondrocyte maturation. *Nat Genet*, **38**(10), pp. 1198-203.
- Laurier, V., Stoetzel, C., Muller, J., *et al.* 2006. Pitfalls of homozygosity mapping: an extended consanguineous Bardet-Biedl syndrome family with two mutant genes (BBS2, BBS10), three mutations, but no triallelism. *Eur J Hum Genet*, **14**(11), pp. 1195-203.

- Le, A., Graham, K. S. & Sifers, R. N. 1990. Intracellular degradation of the transport-impaired human PiZ alpha 1-antitrypsin variant. Biochemical mapping of the degradative event among compartments of the secretory pathway. *J Biol Chem*, **265**(23), pp. 14001-7.
- Leighton, M.P., Nundlall, S., Starborg, T., *et al.* 2007. Decreased chondrocyte proliferation and dysregulated apoptosis in the cartilage growth plate are key features of a murine model of epiphyseal dysplasia caused by a *matn3* mutation *Hum Mol Genet*, **15**;16(14):1728-41
- Lim, J., Hao, T., Shaw, C., *et al.* 2006. A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. *Cell*, **125**(4), pp. 801-14.
- Lindmo, K. & Stenmark, H. 2006. Regulation of membrane traffic by phosphoinositide 3-kinases. *J Cell Sci*, **119**(Pt 4), pp. 605-14.
- Lloyd, D. J., Trembath, R. C. & Shackleton, S. 2002. A novel interaction between lamin A and SREBP1: implications for partial lipodystrophy and other laminopathies. *Hum Mol Genet*, **11**(7), pp. 769-77.
- Lupski, J. R. 1998. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet*, **14**(10), pp. 417-22.
- Maddox, B. K., Keene, D. R., Sakai, L. Y., *et al.* 1997. The fate of cartilage oligomeric matrix protein is determined by the cell type in the case of a novel mutation in pseudoachondroplasia. *J Biol Chem*, **272**(49), pp. 30993-7.
- Mari-Beffa, M., Santamaria, J. A., Murciano, C., *et al.* 2007. Zebrafish fins as a model system for skeletal human studies. *ScientificWorldJournal*, **7**(pp. 1114-27.
- Marmorstein, L. Y., McLaughlin, P. J., Peachey, N. S., *et al.* 2007. Formation and progression of sub-retinal pigment epithelium deposits in *Efemp1* mutation knock-in mice: a model for the early pathogenic course of macular degeneration. *Hum Mol Genet*, **16**(20), pp. 2423-32.
- Marmorstein, L. Y., Munier, F. L., Arsenijevic, Y., *et al.* 2002. Aberrant accumulation of EFEMP1 underlies drusen formation in Malattia Leventinese and age-related macular degeneration. *Proc Natl Acad Sci U S A*, **99**(20), pp. 13067-72.
- McLaughlin, P. J., Bakall, B., Choi, J., *et al.* 2007. Lack of fibulin-3 causes early aging and herniation, but not macular degeneration in mice. *Hum Mol Genet*, **16**(24), pp. 3059-70.
- Michaelides, M., Jenkins, S. A., Brantley, M. A., Jr., *et al.* 2006. Maculopathy due to the R345W substitution in fibulin-3: distinct clinical features, disease variability, and extent of retinal dysfunction. *Invest Ophthalmol Vis Sci*, **47**(7), pp. 3085-97.
- Minina, E., Kreschel, C., Naski, M. C., *et al.* 2002. Interaction of FGF, *Ihh*/*Pthlh*, and BMP signaling integrates chondrocyte proliferation and hypertrophic differentiation. *Dev Cell*, **3**(3), pp. 439-49.

- Miosge, N., Gotz, W., Sasaki, T., *et al.* 1996. The extracellular matrix proteins fibulin-1 and fibulin-2 in the early human embryo. *Histochem J*, **28**(2), pp. 109-16.
- Mueller, R. F. & Bishop, D. T. 1993. Autozygosity mapping, complex consanguinity, and autosomal recessive disorders. *J Med Genet*, **30**(9), pp. 798-9.
- Mundlos, S., Otto, F., Mundlos, C., *et al.* 1997. Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. *Cell*, **89**(5), pp. 773-9.
- Muragaki, Y., Mundlos, S., Upton, J., *et al.* 1996. Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13. *Science*, **272**(5261), pp. 548-51.
- Naffah, J. 1976. The Dyggve-Melchior-Clausen syndrome. *Am J Hum Genet*, **28**(6), pp. 607-14.
- Nakamura, K., Kurokawa, T., Nagano, A., *et al.* 1997. Dyggve-Melchior-Clausen syndrome without mental retardation (Smith-McCort dysplasia): morphological findings in the growth plate of the iliac crest. *Am J Med Genet*, **72**(1), pp. 11-7.
- Neumann, L. M., El Ghouzzi, V., Paupe, V., *et al.* 2006. Dyggve-Melchior-Clausen syndrome and Smith-McCort dysplasia: clinical and molecular findings in three families supporting genetic heterogeneity in Smith-McCort dysplasia. *Am J Med Genet A*, **140**(5), pp. 421-6.
- Ng, L. J., Wheatley, S., Muscat, G. E., *et al.* 1997. SOX9 binds DNA, activates transcription, and coexpresses with type II collagen during chondrogenesis in the mouse. *Dev Biol*, **183**(1), pp. 108-21.
- Nguyen, D. Q., Webber, C. & Ponting, C. P. 2006. Bias of selection on human copy-number variants. *PLoS Genet*, **2**(2), pp. e20.
- Nichols, W. C., Seligsohn, U., Zivelin, A., *et al.* 1998. Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. *Cell*, **93**(1), pp. 61-70.
- Oka, T. & Krieger, M. 2005. Multi-component protein complexes and Golgi membrane trafficking. *J Biochem*, **137**(2), pp. 109-14.
- Paupe, V., Gilbert, T., Le Merrer, M., *et al.* 2004. Recent advances in Dyggve-Melchior-Clausen syndrome. *Mol Genet Metab*, **83**(1-2), pp. 51-9.
- Pope, F. M., Nicholls, A. C., McPheat, J., *et al.* 1985. Collagen genes and proteins in osteogenesis imperfecta. *J Med Genet*, **22**(6), pp. 466-78.
- Pirog-Garcia, K. A., Meadows, R. S., Knowles, L., *et al.* 2007. Reduced cell proliferation and increased apoptosis are significant pathological mechanisms in a murine model of mild pseudoachondroplasia resulting from a mutation in the C-terminal domain of COMP. *1;16(17):2072-88*
- Provot, S. & Schipani, E. 2005. Molecular mechanisms of endochondral bone development. *Biochem Biophys Res Commun*, **328**(3), pp. 658-65.

- Rastogi, S. C., Clausen, J., Melchior, J. C., *et al.* 1977. The Dyggve-Melchior-Clausen syndrome. *Clin Chim Acta*, **78**(1), pp. 55-69.
- Rastogi, S. C., Clausen, J., Melchior, J. C., *et al.* 1978. Biochemical abnormalities in Dyggve-Melchior-Clausen syndrome. *Clin Chim Acta*, **84**(1-2), pp. 173-8.
- Rastogi, S. C., Clausen, J., Melchior, J. C., *et al.* 1980. Abnormal serum alpha 2-macroglobulin in Dyggve-Melchior-Clausen syndrome. *J Clin Chem Clin Biochem*, **18**(1), pp. 67-8.
- Rauch, A., Thiel, C. T., Schindler, D., *et al.* 2008. Mutations in the pericentrin (PCNT) gene cause primordial dwarfism. *Science*, **319**(5864), pp. 816-9.
- Reiter, L. T., Hastings, P. J., Nelis, E., *et al.* 1998. Human meiotic recombination products revealed by sequencing a hotspot for homologous strand exchange in multiple HNPP deletion patients. *Am J Hum Genet*, **62**(5), pp. 1023-33.
- Roesel, R. A., Carroll, J. E., Rizzo, W. B., *et al.* 1991. Dyggve-Melchior-Clausen syndrome with increased pipelicolic acid in plasma and urine. *J Inherit Metab Dis*, **14**(6), pp. 876-80.
- Rodgers, K. D., Sasaki, T., Aszodi, A., *et al.* 2006. Reduced perlecan results in skeletal dysplasia: A new mouse model. *Hum Mol Genet*, **16**(5), pp. 515-28.
- Rutishauser, J. & Spiess, M. 2002. Endoplasmic reticulum storage diseases. *Swiss Med Wkly*, **132**(17-18), pp. 211-22.
- Saiki, R. K., Scharf, S., Faloona, F., *et al.* 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**(4732), pp. 1350-4.
- Schipani, E., Kruse, K. & Juppner, H. 1995. A constitutively active mutant PTH-PTHrP receptor in Jansen-type metaphyseal chondrodysplasia. *Science*, **268**(5207), pp. 98-100.
- Schmitt-John, T., Drepper, C., Mussmann, A., *et al.* 2005. Mutation of Vps54 causes motor neuron disease and defective spermiogenesis in the wobbler mouse. *Nat Genet*, **37**(11), pp. 1213-5.
- Sen, S. K., Han, K., Wang, J., *et al.* 2006. Human genomic deletions mediated by recombination between Alu elements. *Am J Hum Genet*, **79**(1), pp. 41-53.
- Shaw, C. J. & Lupski, J. R. 2004. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet*, **13 Spec No 1**(pp. R57-64.
- Shaw, C. J., Withers, M. A. & Lupski, J. R. 2004. Uncommon deletions of the Smith-Magenis syndrome region can be recurrent when alternate low-copy repeats act as homologous recombination substrates. *Am J Hum Genet*, **75**(1), pp. 75-81.
- Shiang, R., Thompson, L. M., Zhu, Y. Z., *et al.* 1994. Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. *Cell*, **78**(2), pp. 335-42.

- Shimozawa, N., Suzuki, Y., Zhang, Z., *et al.* 1999. A novel nonsense mutation of the PEX7 gene in a patient with rhizomelic chondrodysplasia punctata. *J Hum Genet*, **44**(2), pp. 123-5.
- Shinomura, T., Kimata, K., Oike, Y., *et al.* 1984. Appearance of distinct types of proteoglycan in a well-defined temporal and spatial pattern during early cartilage formation in the chick limb. *Dev Biol*, **103**(1), pp. 211-20.
- Skibinski, G., Parkinson, N. J., Brown, J. M., *et al.* 2005. Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. *Nat Genet*, **37**(8), pp. 806-8.
- Smith, R. & Mc, C. J. 1958. Osteochondrodystrophy (Morquio-Brailsford type); occurrence in three siblings. *Calif Med*, **88**(1), pp. 55-9.
- Spranger, J., Bierbaum, B. & Herrmann, J. 1976. Heterogeneity of Dyggve-Melchior-Clausen dwarfism. *Hum Genet*, **33**(3), pp. 279-87.
- Spranger, J., Maroteaux, P. & Der Kaloustian, V. M. 1975. The Dyggve-Melchior-Clausen syndrome. *Radiology*, **114**(2), pp. 415-21.
- Stoll, C., Dott, B., Roth, M. P., *et al.* 1989. Birth prevalence rates of skeletal dysplasias. *Clin Genet*, **35**(2), pp. 88-92.
- Stone, E. M., Lotery, A. J., Munier, F. L., *et al.* 1999. A single EFEMP1 mutation associated with both Malattia Leventinese and Doyne honeycomb retinal dystrophy. *Nat Genet*, **22**(2), pp. 199-202.
- Superti-Furga, A., Bonafe, L. & Rimoin, D. L. 2001. Molecular-pathogenetic classification of genetic disorders of the skeleton. *Am J Med Genet*, **106**(4), pp. 282-93.
- Superti-Furga, A. & Unger, S. 2007. Nosology and classification of genetic skeletal disorders: 2006 revision. *Am J Med Genet A*, **143**(1), pp. 1-18.
- Tagariello, A., Schlaubitz, S., Hankeln, T., *et al.* 2005. Expression profiling of human fetal growth plate cartilage by EST sequencing. *Matrix Biol*, **24**(8), pp. 530-8.
- Thauvin-Robinet, C., El Ghouzzi, V., Chemaitilly, W., *et al.* 2002. Homozygosity mapping of a Dyggve-Melchior-Clausen syndrome gene to chromosome 18q21.1. *J Med Genet*, **39**(10), pp. 714-7.
- Thiel, C. T., Horn, D., Zabel, B., *et al.* 2005. Severely incapacitating mutations in patients with extreme short stature identify RNA-processing endoribonuclease RMRP as an essential cell growth regulator. *Am J Hum Genet*, **77**(5), pp. 795-806.
- Tickle, C. 2002. Molecular basis of vertebrate limb patterning. *Am J Med Genet*, **112**(3), pp. 250-5.
- Van Criekinge, W. & Beyaert, R. 1999. Yeast Two-Hybrid: State of the Art. *Biol Proced Online*, **2**(pp. 1-38.

- van Ommen, G. J. 2005. Frequency of new copy number variation in humans. *Nat Genet*, **37**(4), pp. 333-4.
- Voeltz, G. K., Rolls, M. M. & Rapoport, T. A. 2002. Structural organization of the endoplasmic reticulum. *EMBO Rep*, **3**(10), pp. 944-50.
- Vogel, K. G. & Trotter, J. A. 1987. The effect of proteoglycans on the morphology of collagen fibrils formed in vitro. *Coll Relat Res*, **7**(2), pp. 105-14.
- Vortkamp, A., Lee, K., Lanske, B., *et al.* 1996. Regulation of rate of cartilage differentiation by Indian hedgehog and PTH-related protein. *Science*, **273**(5275), pp. 613-22.
- Ward, C. L., Omura, S. & Kopito, R. R. 1995. Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell*, **83**(1), pp. 121-7.
- White, D. R., Ganesh, A., Nishimura, D., *et al.* 2007. Autozygosity mapping of Bardet-Biedl syndrome to 12q21.2 and confirmation of FLJ23560 as BBS10. *Eur J Hum Genet*, **15**(2), pp. 173-8.
- Xia, Y., Yu, H., Jansen, R., *et al.* 2004. Analyzing cellular biochemistry in terms of molecular networks. *Annu Rev Biochem*, **73**(pp. 1051-87.
- Yoon, B. S. & Lyons, K. M. 2004. Multiple functions of BMPs in chondrogenesis. *J Cell Biochem*, **93**(1), pp. 93-103.