The role of *HESXI* in septo-optic dysplasia and its variants.

A thesis submitted for the degree of Doctor of Philosophy

By

David Euan Graham McNay

Institute of Child Health.
Septo-Optic Dysplasia (SOD) is a highly variable developmental abnormality of the midline structures of the brain, classically resulting in hypoplasia of the optic nerves, absence of the septum pellucidum and dysgenesis of the pituitary gland. The homeobox gene Hesxl is a transcriptional repressor expressed at gastrulation within the anterior midline visceral endoderm, with subsequent expression in the prosencephalon and Rathke's pouch. Hesxl null mutant mice manifest a phenotype similar to SOD in man and this led to the successful characterisation of a recessive HESX1 mutation within two familial cases (Dattani et al., 1998). However, the cohort sample size was too small to draw general conclusion about the role of HESX1 in SOD.

This thesis investigated the role of HESX1 in SOD, particularly sporadic cases, and involved the collection of the largest cohort of SOD and SOD-like phenotypes to date (n=670). Mutation screening of HESX1 was carried out using both SSCP detection and d-HPLC heteroduplex detection. The mutation screen identified a number of heterozygous sequence variants of which two appear to be mutations (S170L and E149K).

An S170L mutation was identified in a sporadic case of growth hormone deficiency and an ectopic posterior pituitary. This mutation had been previously described in a sib pair and functional analysis carried out (Thomas et al., 2001 and Brickman et al., 2002).

An E149K variant was identified in a child demonstrating growth hormone deficiency, hypoplasia of the anterior pituitary and an ectopic posterior pituitary. The E149K variant lies at position 42 of the homeodomain, at the first position of the third helix. This change is highly conserved and has not been found in 140 control chromosomes. Functional studies were undertaken to investigate the function of this variant, including in-vitro DNA binding assays and in-vitro transfection studies.
1. Introduction.

1.1 Septo-optic-pituitary dysplasia: clinical spectrum.
1.1.1 Optic nerve hypoplasia.
1.1.2 Absence of the septum pellucidum.
1.1.3 Pituitary abnormalities.
1.1.4 Associated features.

1.2 Development.
1.2.1 Optic nerve development.
1.2.2 Development of the septum pellucidum.
1.2.3 Pituitary development.
1.2.4 Developmental origin of Septo-optic dysplasia.

1.3 Genetics.
1.3.1 HESXI.
1.3.2 Hesxl.
1.3.3 Xanf-1 and Xanf-2.
1.3.3.1 Xanf-1.
1.3.3.2 Xanf-2.
1.3.4 Other HESXI homologues.
1.3.5 The Anf class of homeobox genes.
1.3.6 Other genes implicated in SOD.

1.4 Aims.

2 HESXI mutation screening.

2.1 Introduction.

2.2 Mutation screening design and methodology.
2.2.1 Patient recruitment.
2.2.2 Mutation selection.

2.2.3 Regions to be screened.

2.2.4 Variance identification.

2.2.5 Causative mutation identification.

2.2.6 Overview.

2.3 Methods.

2.3.1 STS primer design.

2.3.2 Polymerase Chain Reaction.

2.3.2.1 Initial PCR conditions.

2.3.2.2 Optimisation of PCR conditions for phase 1 of mutation screening.

2.3.2.3 PCR conditions for phase 1 of mutation screening.

2.3.2.4 Optimisation of PCR conditions for phase 2 of mutation screening.

2.3.2.5 PCR conditions for phase 2 of mutation screening.

2.3.2.6 Summary.

2.3.3 SSCP detection.

2.3.3.1 The clean gel system.

2.3.3.2 Use of the clean gel system during phase 1 of mutation detection.

2.3.3.3 Use of the clean gel system during phase 2 of mutation detection.

2.3.4 dHPLC detection of heteroduplex using the WAVE system.

2.3.4.1 Heteroduplex detection.

2.3.4.2 The WAVE system.

2.3.4.3 Use of the WAVE system.

2.3.4.3.1 Putative *HESX1* alternative exon A.

2.3.4.3.2 Putative *HESX1* alternative exon B.

2.3.4.3.3 *HESX1* exon 1.

2.3.4.3.4 *HESX1* exon 2.

2.3.4.3.5 *HESX1* exon 3/4.

2.3.4.3.6 *HESX1* 5' LIM1 homology region.

2.3.5 DNA Sequencing.

2.3.5.1 Fluorescent DNA sequencing using BigDye and the ABI377.

2.3.5.2 Direct fluorescent cycle sequencing of PCR products.

2.3.5.3 Fluorescent cycle sequencing of plasmid DNA.

2.3.5.4 Sequence comparison.

2.3.6 cDNA studies of *HESX1* expression.

2.4 Results.
2.4.1 Patient recruitment. 83
2.4.2 Variance detection. 84
2.4.2.1 Putative exon alt A 86
2.4.2.2 Putative exon alt B 88
2.4.2.3 5' LIM1 homology region. 89
2.4.2.4 Exon 1. 91
2.4.2.5 Exon 2. 92
2.4.3.6 Exon 3 and 4. 95
2.4.4 DNA sequencing. 101
2.4.4.1 Putative exon A. 102
2.4.4.2 5' LIM homology region. 106
2.4.4.3 Exon 2. 106
2.4.4.4 Exon 3/4. 108
2.4.5 cDNA studies of HESX1 expression. 111
2.4.6 Bioinformatics. 112
2.4.6.1 Putative alternative exons. 112
2.4.6.2 Non-coding regions. 115
2.4.6.3 DHX3F sequence variation. 116
2.4.7 Patient phenotypes. 117
2.4.7.1 Genotype to phenotype correlations. 117
2.4.7.2 S170L. 118
2.4.7.3 E149K. 119

3 Functional investigations. 120

3.1 Sequence variants. 120
3.1.1 Putative alternative exon A. 120
3.1.2 5' LIM1 homology region. 120
3.1.3 Exons 1-4 121
3.1.3.1 Previous control data. 121
3.1.3.2 Silent variants. 122
3.1.3.3 Coding variants. 122
3.1.4 Further functional investigations. 123

3.2 E149K. 124
3.2.1 The effect of E149K on the HESX1 homeodomain. 124
3.2.2 Allelic tracking. 126
3.2.3 DNA binding studies. 128
3.2.3.1 Recombinant HESX1 production. 129
3.2.3.2 EMSA. 134
3.2.4 In-vitro transfection studies. 143
3.2.4.1 Construct production. 143
3.2.4.2 Transfections. 144
3.2.5 Western blotting. 146

4 Mutation screening of PTX1 and PTX2. 151

4.1 PTX1. 151
4.1.1 Introduction. 151
4.1.2 Methods. 153
4.1.2.1 dHPLC-heteroduplex using the WAVE system. 153
4.1.3 Results. 157
4.1.3.1 Exon 1. 157
4.1.3.2 Exon 2. 157
4.1.3.3 Exon 3. 158
4.1.3.4 Bioinformatics. 163

4.2 PTX2. 166
4.2.1 Introduction. 166
4.2.2 Methods. 167
4.2.2 Results. 168
4.2.3.1 DNA sequencing. 168
4.2.3.2 Bioinformatics. 168

4.3 Discussion. 177

5 Related work. 177

5.1 Mutation screening of candidate genes within the patient cohort. 177

5.2 Mutation screening of known pituitary disease genes. 177

5.3 Mutational studies into the role of HESX1 in SOD. 178
5.4 Functional studies of *HESX1* variants.  
5.4.1 R160C.  
5.4.2 S170L.  
5.4.3 N125S.

6 Discussion.  
6.1 *HESX1* mutation screening.  
6.1.1 Patient recruitment.  
6.1.2 cDNA studies of *HESX1* expression.  
6.1.3 Putative alternative exons.  
6.1.4 Non-coding regions.  
6.1.5 Variance detection.

6.2.1 *HESX1* variants.  
6.2.2 S170L.  
6.2.3 E149K.  
6.2.3.1 Further investigation of E149K.  
6.2.4 *HESX1* sequence variation.

6.3 The role of *HESX1* in the condition SOD.  
6.3.1 *HESX1* and growth hormone deficiency.

7 References.
Table of Figures.

| Figure 1.1.1a | The optic nerves. | 24 |
| Figure 1.1.2a | The location of the septum pellucidum. | 25 |
| Figure 1.1.3a | The pituitary gland. | 28 |
| Figure 1.1.3b | An ectopic posterior pituitary. | 30 |
| Figure 1.2.3a | Four stages of pituitary development. | 38 |
| Figure 1.2.3b | Pituitary development occurs via distinct stages. | 39 |
| Figure 2.1.6a | Sample screening protocol. | 61 |
| Figure 2.3.2.1a | Location of primers in relation to the HESX1 genomic structure. | 64 |
| Figure 2.3.2.1b | Initial PCR amplification using DHX1F-DHX1R. | 65 |
| Figure 2.3.2.1c | Initial PCR amplification using DHX2F-DHX2R. | 65 |
| Figure 2.3.2.1d | Initial PCR amplification using DHX3F-DHX3R. | 65 |
| Figure 2.3.2.1e | Initial PCR amplification using DHX4F-DHX4R. | 66 |
| Figure 2.3.2.2a | Example of PCR optimisation for phase 1 of mutation screening showing DHX2F-DHX2R at an annealing temperature of 52°C across a range of Mg²⁺ concentration from 0.5mM to 3.5mM. | 67 |
| Figure 2.3.2.4a | Example of PCR optimisation for phase 2 of mutation screening showing H5′F-H5′R across a temperature gradient of 50.0°C to 62.5°C. | 68 |
| Figure 2.3.2.4b | Location of oligonucleotides primers for putative alternative exons and 5′LIM1 homology region in relation to the genomic structure of HESX1. | 69 |
| Figure 2.3.2.5a | An example of PCR amplification from genomic DNA samples (n=19) under optimised conditions using DHX3F-DHX4R. | 70 |
The experimentally determined running temperature of
*HESX1* exon 1.

The experimentally determined running temperature of
*HESX1* exon 2.

The experimentally determined running temperature of
*HESX1* exon 3/4.

A *HESX1* exon A SSCP detection gel showing a sample with
a SCC change, an example of a false positive result.

An overlay of exon A dHPLC results showing three samples
4883, 4884, and 4885 with wild type.

An overlay of exon A dHPLC results showing four samples
4517, 4518, 4519, and 4521.

An overlay of exon A dHPLC results showing four samples
3881, 3882, 3883, and 389.

An overlay of exon B dHPLC results for two samples 4857
and 4867 with wild type.

An overlay of exon B dHPLC results for four samples 4869,
4800, 4814, and 4802.

An overlay of 5' LIM1 homology region dHPLC results at
61°C for 4 samples 4751, 4754, 4718, and 4714.

An overlay of 5' LIM1 homology region dHPLC results at
59°C for 4 samples 5057, 5071, 5072, and 4842.

An overlay of 5' LIM1 homology region dHPLC results at
59°C for 4 samples 4988, 4989, 4992, and 5001.

A *HESX1* exon 1 SSCP detection gel showing a sample with
a SSC change and samples with duplex changes.

An overlay of exon 1 dHPLC results showing four samples
4015, 4347, 4578, and 335.
An overlay of exon 1 dHPLC results showing four samples 4644, 4639, 4640, and 4647.

A HESX1 exon 2 SSCP detection gel.

An overlay of exon 2 dHPLC results showing three samples 4900, 4870, and 4871 with wild-type.

An overlay of exon 2 dHPLC results showing four samples 4820, 5103, 5120, and 5134.

An overlay of exon 2 dHPLC results showing four samples 4585, 4410, 4614, and 4624.

An overlay of exon 2 dHPLC results showing four samples 4203, 3711, 3936, and 4212.

An overlay of exon 2 dHPLC results showing four samples 4954, 4214, 4352, and 4489.

A HESX1 exon 3 SSCP detection gel showing a sample with a SCC change (arrow) and a number of failed samples.

A HESX1 exon 4 SSCP detection gel.

Exon 3 SSCP results for N125S samples.

Exon 4 SSCP results for the S170L sample.

Exon 4 SSCP results for 525g>a samples.

An overlay of exon 3/4 dHPLC results showing four samples 4748, 4750, 4751, and 4752.

An overlay of exon 3/4 dHPLC results showing four samples 3679, 4760, 4995, and 5050.

An overlay of exon 3/4 dHPLC results showing four samples 4766, 4767, 4768, and 4769.

An overlay of exon 3/4 dHPLC results showing three samples 4712, 4585, and 4721 with WT.
An overlay of exon 3/4 dHPLC results showing four samples 483, 4963, 4763, and 4701.

Exon 4 SSCP results for the E149K sample.

Location of HESX1 variants.

Alignment of the HESX1 open reading frame and amino acid sequence showing the location sequence variants.

Sequencing results of putative exon A for sample 4868 showing -50c>t in the heterozygous state.

Sequencing results of putative exon A for sample 3883 showing -4 t>g in the heterozygous state.

Sequencing results of putative exon A for sample 4904 showing the donor splice site IVS1A +5c>t variant in the heterozygous state.

Sequencing results of the 5' LIM1 homology region A for sample 4645 showing -277t>g in the heterozygous state.

Sequencing results of exon 2 for sample 4582 showing 183t>c in the heterozygous state.

Sequencing results of exon 2 for sample 4489 showing Y90H in the heterozygous state.

Sequencing results of exon 2 for sample 4212 showing 219c>t in the heterozygous state.

Sequencing results of exon 2 for sample 4405 showing the erroneous a161g variant.

Sequence results showing the N125S variant.

Sequence results showing the S170L variant.

Sequencing results of exon 3/4 for sample 4769 showing the E149K variant.

Sequencing results of exon 3/4 for sample 4508 showing
the 525g>a variant.

(Figure 2.4.5a) Results of *HESX1* transcript amplification of week 17 brain and week 8 cDNA from eye, ear, stomach, kidney, and heart.

(Figure 2.4.5b) Results of the second *HESX1* transcript amplification of week 17 brain and week 8 cDNA from eye, ear, stomach, kidney, and heart.

(Figure 2.4.6.1a) Partial sequence of contig 28 showing exons 1-4.

(Figure 2.4.6.1b) Pairwise BLAST of GenBank gi:1720505 and GenBank gi:1720505.

(Figure 2.4.6.1c) Partial sequence of contig 21, showing *HESX1* exons A and B, as well as, APPL exon 1.

(Figure 2.4.6.1d) The genomic structure of *HESX1*.

(Figure 2.4.6.2a) Partial sequence of contig 28 showing the 5’LIM1 homology region.

(Figure 2.4.6.3a) Variation within the sequence of oligonucleotide primer DHX3F.

(Figure 3.1.3.3a) Alignment of ANF homeodomains using ClustelW.

(Figure 3.2.1a) Prd Glu42 forms an inter-monomer salt-bridge with Arg3 and an intra-monomer salt-bridge with Arg31.

(Figure 3.2.1b) Prd Glu42 forms an inter-monomer salt-bridge with Arg3 and an intra-monomer salt-bridge with Arg31, a space fill view.

(Figure 3.2.1c) *HESX1* Glu42 forms an intra-monomer salt-bridge with Arg31.

(Figure 3.2.2a) E149K pedigree.

(Figure 3.2.2b) E149K allele inheritance.

(Figure 3.2.3.1a) Site-directed mutagenesis of *HESX1* cDNA.

(Figure 3.2.3.1b) Evaluation of E149K induction.

(Figure 3.2.3.1c) The progress of E149K purification.

(Figure 3.2.3.1d) Aliquots of purified *HESX1*.
Final SDS-PAGE results showing equal volumes of recombinant E194K and WT HESX1.

EMSA using a protein concentration gradient of 1nM - 1024Nm with a probe concentration of 200nM.

EMSA using a protein concentration gradient of 32nM - 192nM with a probe concentration of 2nM.

Proportion of each probe species from the EMSA using a protein concentration gradient of 32nM - 192nM with a probe concentration of 2nM.

The ratio of monomer to dimer from the EMSA using a protein concentration gradient of 32nM - 192nM with a probe concentration of 2nM.

EMSA using a protein concentration gradient of 128nM - 448nM with a probe concentration of 2nM.

Proportion of each probe species from the EMSA using a protein concentration gradient of 128nM - 448nM with a probe concentration of 2nM.

The ratio of monomer to dimer from the EMSA using a protein concentration gradient of 128nM - 448nM with a probe concentration of 2nM.

Cartoon of the GAL4 reporter construct (PG5SV40).

The mean dual Luciferase ratio of each transfection type as a percentage of pJBXi control with 95% confidence levels.

Antibody testing of murine monoclonal antibody and rabbit polyclonal antibody.

Results of Western blotting of HESX1 transfected COS-7 cells.
PCR results from cDNA derived from GAL4-HESX1 transfected cos-7 cells.

PCR results from cDNA derived from GAL4-HESX1 transfected cells.

PTXI genomic structure.

PTXI exon 3 melting profile.

PTXI exon 3a melting profile.

PTXI exon 3b melting profile.

An overlay of PTXI exon 1 dHPLC results showing three samples 3847, 4610, and 4244 with wild-type.

An overlay of PTXI exon 2 dHPLC results showing four samples 3679, 3681, and 3709 and 3810.

An overlay of PTXI exon 2 dHPLC results showing three samples 5004, 5025, and 5064 with wild-type.

An overlay of PTXI exon 3 dHPLC at 67°C results showing four samples 4701, 4712, 4731, and 4668.

An overlay of PTXI exon 3 dHPLC at 69°C results showing four samples 4668, 4831, 5004, and 5027.

An overlay of PTXI exon 3 dHPLC at 71°C results showing four samples 4701, 4712, 4731, and 4668.

An overlay of PTXI exon 3a dHPLC at 66°C results showing two samples 4668, and 5004 with wild-type.

An overlay of PTXI exon 3a dHPLC at 66°C results showing four samples 4701, 4712, 4714, and 4731.

An overlay of PTXI exon 3b dHPLC at 66°C results showing three samples 3679, 5027, and 4821 with wild-type.

An overlay of PTXI exon 3b dHPLC at 68°C results showing four samples 3679, 5027, 4821, and 4598.
(Figure 4.1.3.4a) Homo sapiens chromosome 5 clone CTC-276P9, gi:18449971.

(Figure 4.2.3.2a) Homo sapiens chromosome 4 clone RP11-380D23, gi:9247141.

(Figure 4.2.3.2b) \textit{PTX2} genomic structure.
Table of Tables.

(Table 1.1.3a) Hormones produced by the anterior pituitary. 29
(Table 1.1.3b) The clinical consequences of anterior pituitary hormone defects. 29
(Table 2.3.1a) Sequence of oligonucleotide primers identical to Dattani et al. (1998).
(Table 2.3.2b) Sequence of oligonucleotides created. 62
(Table 2.3.2.1a) Master mix constituents for each primer pair. 64
(Table 2.3.2.1b) Volume of master mix, DNA, and mQ H2O added to each tube. 64
(Table 2.3.2.1c) Annealing temperatures. 64
(Table 2.3.2.3a) Master mix constituents for each primer pair. 67
(Table 2.3.2.3b) Volume of master mix and DNA to each tube. 68
(Table 2.3.2.5a) Master mix constituents for each primer pair. 69
(Table 2.3.2.5b) Volume of master mix and DNA added to each well. 69
(Table 2.3.2.5c) Thermocycling profile. 70
(Table 2.4.1a) The categorisation of recruited patients. 84
(Table 2.4.2a) Stages of HESX1 mutation screening. 85
(Table 2.4.2b) Breakdown of mutation screening results. 86
(Table 2.4.4a) The total number of samples sequenced, sub-divided into familial and sporadic cases, alongside the number samples containing sequence variance. 102
(Table2.4.6.3a) Variation within the sequence of oligonucleotide primer DHX3F. 117
(Table 2.4.7.1a) HESX1 genotype to phenotype correlations. 118
(Table 3.1.3.3a) Allele specific primer digestion of N125S within Afro-Caribbean controls. 123
(Table 3.2.2a) Allele tracking data for E149K pedigree. 127
(Table 3.2.4.1a) Site-directed mutagenesis oligonucleotide primers. 129
In-vitro transfection experiments. 144

PTX1 oligonucleotide primers. 153

PTX1 column running temperatures. 156

PTX1 exon3 dHPLC- heteroduplex profiles. 163

PTX2 oligonucleotide primers. 167

Phenotypes of samples screened for mutations in PTX1 and PTX2. 175
Acknowledgements.

First and foremost I am indebted to my supervisors, Dr Mehul Dattani and Professor Mike Preece, without whose tireless effort this thesis would not have been possible. Likewise I would also like to acknowledge the patients and clinicians, without whom this thesis would be impossible.

I would also like to acknowledge the help and assistance of all the members of the Clinical and Molecular Lab at the Institute of Child Health. In particular, Professor Sue Malcolm for her inspiration, Dr Clare Beesely for her patience in that first year, Diana Blaydon for the provision of human cDNA and Paul Rutland for having an answer for every question. I am indebted to the DNA lab for processing all those DNA samples. I would like to acknowledge and thank the members of our group, Dr Kath Woods for her help and assistance during *HESX1* mutation screening and setting up the *in-vitro* transfections, Nele Haute for helping with *PTX1* mutation screening and sharing her *HESX1* mutation screening data, and James Turton for being an excellent research assistant. I would also like to acknowledge Dr Josh Brickman for providing the GAL4-*HESX1* clone and the GAL4 reporter construct, Dr Paul Thomas for sharing his *HESX1* mutation screening data, Luciani Carvalho for sharing her mutation screening data and Dr George Howat for his help with the DNA binding constants.

I would also like to thank Dr Kathy Mahon, for sharing her *Hesxl* null mutant data and the sequence of *Hesxl* regulatory regions, and Dr John Old for the Afro-Caribbean controls.

Finally, I would like to thank Rowan for everything.
The homeobox gene Hesx1/HESX1 has been implicated in the establishment of anterior pattern in the central nervous system (CNS) in a number of vertebrate species. Its role in pituitary development has been documented through loss-of-function studies in the mouse. A homozygous missense point mutation resulting in a single amino acid substitution, Arg160Cys (R160C), is associated with a heritable form of the human condition of septo-optic dysplasia (SOD). We have examined the phenotype of affected members in this pedigree in more detail and demonstrate for the first time a genetic basis for midline defects associated with an undescended or ectopic posterior pituitary. A similar structural pituitary abnormality was observed in a second patient heterozygous for another mutation in HESX1, Ser170Leu (S170L). Association of S170L with a pituitary phenotype may be a direct consequence of the HESX1 mutation since S170L is also associated with a dominant familial form of pituitary disease. However, a third mutation in HESX1, Asn125Ser (N125S), occurs at a high frequency in the Afro-Caribbean population and may therefore reflect a population-specific polymorphism. To investigate the molecular basis for these clinical phenotypes, we have examined the impact of these mutations on the regulatory functions of HESX1. We show that Hesx1 is a promoter-specific transcriptional repressor with a minimal 36 amino acid repression domain which can mediate promoter-specific repression by suppressing the activity of homeodomain-containing activator proteins. Mutations in HESX1 associated with pituitary disease appear to modulate the DNA-binding affinity of HESX1 rather than its transcriptional activity. Wild-type HESX1 binds a dimeric homeodomain site with high affinity (K(d) 31 nM) whilst HESX1(S170L) binds with a 5-fold lower activity (K(d) 150 nM) and HESX1(R160C) does not bind at all. Although HESX1(R160C) has only been shown to be associated with the SOD phenotype in children homozygous for the mutation, HESX1(R160C) can inhibit DNA binding by wild-type HESX1 both in vitro and in vivo in cell culture. This dominant negative activity of HESX1(R160C) is mediated by the Hesx1 repression domain, supporting the idea that the repression domain is implicated in interactions between homeodomain proteins. Our data suggest a possible molecular paradigm for the dominant inheritance observed in some pituitary disorders.

We have previously shown that familial septo-optic dysplasia (SOD), a syndromic form of congenital hypopituitarism involving optic nerve hypoplasia and agenesis of midline brain structures, is associated with homozygosity for an inactivating mutation in the homeobox gene HESX1/Hesx1 in man and mouse. However, as most SOD/congenital hypopituitarism occurs sporadically, the possible contribution of HESX1 mutations to the aetiology of these cases is presently unclear. Interestingly, a small proportion of mice heterozygous for the Hesx1 null allele show a milder SOD phenocopy, implying that heterozygous mutations in human HESX1 could underlie some cases of congenital pituitary hypoplasia with or without midline defects. Accordingly, we have now scanned for HESX1 mutations in 228 patients with a broad spectrum of congenital pituitary defects, ranging in severity from isolated growth hormone deficiency to SOD with panhypopituitarism. Three different heterozygous missense mutations were detected in individuals with relatively mild pituitary hypoplasia or SOD, which display incomplete penetrance and variable phenotype amongst heterozygous family members. Gel shift analysis of the HESX1-S170L mutant protein, which is encoded by the C509T mutated allele, indicated that a significant reduction in relative DNA binding activity results from this mutation. Segregation analysis of a haplotype spanning 6.1 cM, which contains the HESX1 locus, indicated that only one HESX1 mutation was present in the families containing the C509T and A541G mutations. These results demonstrate that some sporadic cases of the more common mild forms of pituitary hypoplasia have a genetic basis, resulting from heterozygous mutation of the HESX1 gene.
1. Introduction.

1.1 Septo-optic-pituitary dysplasia: clinical spectrum.

Septo-optic dysplasia (SOD) is a rare developmental disorder affecting the anterior midline in man. The diagnostic criteria for the condition are disputed, there is doubt as to whether the condition represents a single disorder, and the true incidence is unknown. However, it is universally accepted that there is an association between the classical triad of clinical features: absence of the septum pellucidum, hypoplasia of the optic nerves and pituitary dysfunction. The complete classical triad may occur within a single patient or a patient may only demonstrate two of the features of the classical triad in various combinations. Although cases of isolated features of the classical triad do not fulfil the diagnosis of SOD, it is unclear if they represent the mildest cases of the SOD spectrum.

In the limited number of autopsy investigations undertaken, defects can appear to be developmental malformations or destructive lesions, and both can be seen within the same patient (Roessmann, 1989). Occasionally, a number of secondary features are seen within individual patients, including additional defects relating to the anterior midline but also structures unrelated to the anterior midline.

Reeves first reported the association of optic nerve hypoplasia and the absence of the septum pellucidum in 1941. The affected child demonstrated congenital blindness with hypoplastic optic nerves and an absence of the septum pellucidum. However, this isolated report did not bring the syndrome to the attention of the medical community. De Morsier first used the term “la dysplasia septo-optique” (Septo-optic dysplasia) in 1956 to describe findings in a single post-mortem investigation. At autopsy, the 84-year-old female demonstrated a rudimentary septum pellucidum and a divided left optic nerve resulting from misorientation of the optic chiasma due to a dilated anterior ventricle. He then went on to review the literature for cases of defects of the septum pellucidum, and described a further 36 cases of which 9 showed optic nerve hypoplasia (de Morsier, 1962). In 1970, Hoyt et al. described nine patients with septo-optic dysplasia and pituitary dwarfism. In the same year, Kaplan et al described six patients with SOD and pituitary defects. They coined the term septo-optic-pituitary dysplasia to represent this new association. In 1978, Clark and Meyer emphasized the neonatal presentation of the syndrome, characterised by blindness, seizures, and hypoglycemia. They also noted variability in presentation. In 1989, Kaufman et al. described the association of pituitary stalk hypoplasia with the syndrome. A year later
Benner et al. (1990) described the first familial case in a brother and sister and six years later, Wales and Quarrell (1996) described another familial case within a highly in-bred family.

Due to the rarity of the disorder no single specialist medical field deals with septo-optic dysplasia. The techniques used to identify the triad of features are varied and due to the nature of these techniques, patient phenotypes are often poorly characterised. The term septo-optic dysplasia is used loosely to describe the association between the 3 classical features, although it does not explicitly refer to the pituitary; the more accurate term “septo-optic-pituitary dysplasia” is not commonly used. Thus, it can be unclear if usage of septo-optic dysplasia refers explicitly to defects of the septum pellucidum and optic nerves only, or encompasses the pituitary as well.

Therefore, within this thesis the classical triad will be defined as:

1) Unilateral or bilateral optic nerve hypoplasia, either as an ophthalmologic or radiological finding.

2) The radiological finding of either partial or complete absence of the septum pellucidum.

3) An endocrine abnormality of the pituitary.

In addition, the term septo-optic-pituitary dysplasia (SOPD) will be used to refer both to cases demonstrating the complete classical triad and cases demonstrating at least two of the three classical features. The use of a single term (SOPD) to describe the entire spectrum of the syndrome does not infer that there is a single syndrome. Neither does the terminology suggest that the full syndrome includes all three classical features, such that cases consisting of only two classical features are incomplete manifestations. Septo-optic dysplasia (SOD) will be limited to referring to cases in the literature, those clinically diagnosed as such, and the condition in general.

In practice, aplasia or hypoplasia of the corpus callosum may be included within the classical triad alongside absence of the septum pellucidum (Dattani, personal communication). It is unclear whether this represents a separate syndrome or an extension of septo-optic-pituitary dysplasia.

Several aetiologies have been postulated to account for the occurrence of SOD, such as viral infections, environmental teratogens, a vascularisation defect, or degenerative damage. Mendelian inheritance was ruled out as a cause of the disorder (Harris and Haas, 1972). However, the production of a null mutant of the murine Hesx1 gene displaying a phenotype...
with remarkably similar features to SOD led to the successful identification of a homozygous 
*HESX1* Arg160Cys mutation within the two siblings reported by Wales and Quarrell (1996) 
(Dattani *et al.*, 1998). However, no *HESX1* mutations were shown in another 18 patients. A 
heterozygous Ser170Leu mutation in two additional familial cases was identified and 
published at a later stage (Dattani, personal communication and Thomas *et al.*, 2001).

1.1.1 Optic nerve hypoplasia.

The optic nerve is the pathway for transduction of visual information from the retina to the 
brain and consists mainly of nerve fibres projecting from the retinal ganglion cells to the 
lateral geniculate body and pulvinar, and superior colliculus. These are of two types of nerve 
fibre, crossed and uncrossed. The crossed fibres cross the midline at the optic chiasma such 
that they project to the opposite side of the brain. On the other hand, uncrossed fibres do not 
cross the midline and project to the same side (Figure 1.1.1a) (Gray 1918).
Anatomically, the optic nerve originates from the confluence of ganglion cell axons as they traverse the scleral canal to exit the eye, and ends anatomically as these axons merge with the axons of the fellow optic nerve at the chiasm (Martin and Corbett, 2000).

Anatomic divisions of the optic nerve include intraocular, intraorbital, intracanalicular, and intracranial portions. The short intraocular course of the optic nerve is often referred to as the optic nerve head, and the portion that can be seen with the ophthalmoscope is called the optic disc. The intraorbital portion of the optic nerve is approximately 25 mm in length from the posterior aspect of the globe to the orbital apex. In the orbit, the optic nerve is surrounded by the optic nerve sheath, which is continuous with the intracranial dura through the optic canal posteriorly and bounded by the sclera anteriorly. After passing through the lamina cribrosa, the retinal ganglion cell axons acquire myelin sheathing, doubling the diameter of the optic nerve to greater than 3 mm. The myelin sheath is produced by oligodendrocytes, the same cell type as in the white matter tracts in the central nervous system. Peripheral nerves are
myelinated by Schwann cells. Thus, the optic "nerve" is histologically a white matter tract rather than a peripheral nerve. The intracanalicular portion of the optic nerve is about 10 mm long; beginning where the optic nerve enters the optic foramen in the lesser wing of the sphenoid, and ending at the point where the optic nerve exits the optic canal and enters the intracranial cavity. From the orbit, the optic canal moves medially and superiorly to enter the intracranial cavity. The optic canal is separated from the sphenoid sinus by very thin bone, and the course of the optic nerve can be seen as a convexity in the lateral wall of the sinus. In addition to the optic nerve, the optic canal also contains the ophthalmic artery. The intracranial portion of the optic nerve is approximately 15 mm long, extending from the nerve's entrance into the intracranial cavity to the chiasm, but this measurement may vary greatly depending on the relative location of the chiasm. The optic nerves angle superiorly at 45 degrees from the skull base and converge toward the midsagittal plane to form the chiasm. The anterior clinoid is superior and lateral to the optic nerve as the nerve emerges from the optic foramen. The frontal lobes and olfactory tracts are above the optic nerve (Martin and Corbett, 2000).

Optic nerve hypoplasia (ONH) is a relatively common cause of blindness and the loss of visual acuity may be partial or complete (Acers 1981). However, there is little correlation between the degree of optic nerve hypoplasia and visual acuity (Acers, 1981). Optic disc hypoplasia represents incomplete development of the optic disc, characterized by a small optic disc with a larger concentric variably pigmented ring. The condition may be diffuse or segmental, i.e. the entire optic nerve head may be hypoplastic, or only a segment of the optic nerve head may be hypoplastic. ONH may be unilateral or bilateral, with discordance common in bilateral cases (Acers, 1981). Acers (1981) produced a normal range of optic nerve head diameter. Ouvrier and Billson (1986) narrowed the definition of ONH to a developmental anomaly in which there is a reduction in the number of nerve fibres projecting from ganglion cells in the retina to the lateral geniculate body.

Optic nerve hypoplasia is not a specific disease but a morphological diagnosis (Burke et al., 1991 and Novakovic et al., 1988). The two stage development of the optic nerves (formation of the optic stalk followed by axon migration) results in confusion as to where the causative developmental defect underlying ONH occurs. It may be that the optic stalk either fails to form or forms abnormally, due to a defect within the optic stalk itself or an earlier defect of forebrain formation. Likewise, retinal ganglion axon migration may fail or occur abnormally, due to a central defect of the retinal ganglion cells or a defect in axon migration. Additionally, the forming optic nerve may be physically stretched and damaged by abnormal development of the forebrain leading to retrograde degeneration of the retinal ganglion cell axons (Acers and Warn, 1994).
ONH is associated with a number of syndromes, in particular abnormalities of the developing brain (Burke et al., 1991). Skarf and Hoyt (1984), found that developmental delay is the most common association (46%), while Margalith et al. (1984) identified neuropsychiatric handicaps in 73% of patients and Edwards and Layden (1970) found similar features in 50% of patients. These associations fail to throw light on the causative developmental defect underlying ONH, since all the proposed mechanisms could easily lead to, or result from, wider development defects.

ONH is usually detected on fundoscopy as a small optic nerve head of subnormal diameter with a yellowish rim surrounding the hypoplastic optic nerve head (Acers, 1981 and Ouvrier and Billson, 1986). In most cases, the retinal pigment epithelium overgrows the outer ring to terminate at the inner optic nerve resulting in the appearance of a “double ring” sign. In other cases, the retinal pigment epithelium does not overgrow the outer ring resulting in no double ring (Ouvrier and Billson, 1986). Using fundoscopy the diagnosis may be missed due to the outer yellowish rim being mistaken for the true optic nerve head (Ouvrier and Billson, 1986).

ONH can also be determined using radiological techniques (Acers, 1981). In this case, the ONH is defined as an optic nerve diameter less than two standard deviations below the mean control diameter (Acers, 1981). The identification of ONH by radiological means is difficult and standard techniques may underdiagnose the condition (Good et al., 2000).

1.1.2 Absence of the septum pellucidum.

The septum pellucidum is a double membrane separating the anterior horns of the lateral ventricles of the brain. It is situated in the median plane and is attached to the corpus callosum above and at the front, as well as, to the fornix below (Figure 1.1.2a). It is triangular in form, being broad at the front and narrow behind. Between the two membranes exists a cavity, the cavum septum pellucidum, which varies in size (Gray, 1918).
The location of the septum pellucidum (Gray 1918). Here the septum pellucidum is referred to as the septum lucidum.

The function of the septum pellucidum is unknown. Absence of the septum pellucidum in one patient has been implicated in specific learning difficulties (Griffiths and Hunt 1984). Absence of the septum pellucidum can be regarded as a mild form of holoprosencephaly and is seen in many patients with this syndrome. It may result from non-development of the septum pellucidum isolated from other defects. It is also possible that absence of the septum pellucidum may be due to degeneration secondary to a defect of the corpus callosum, which may or may not be readily apparent. The phenotype of the Hesx1 null mutant mouse suggests that absence of the septum pellucidum may arise from an early defect of forebrain formation (Dattani et al., 1998).

In the absence of a post-mortem investigation, the definition of an absent septum pellucidum is made radiologically, usually using magnetic resonance imaging.

1.1.3 Pituitary abnormalities.

The pituitary is a central regulator of physiology and is important for normal growth and development. There are two pituitary glands in man; the stellar pituitary commonly referred to as the pituitary gland and the lesser-known pharyngeal pituitary.
The stellar pituitary sits at the end of a stalk of connecting tissue, the pituitary stalk, which connects the gland to the hypothalamus. The stellar pituitary consists of two parts, the posterior lobe (also known as the posterior pituitary), and the endocrine lobe (also known as the anterior pituitary). The endocrine lobe can be subdivided into the anterior lobe furthest from the pituitary stalk and a small intermediate lobe lying close to the posterior lobe. However, usage of this subdivision is far from universal. Therefore, within this thesis the term anterior pituitary will be used to refer entire endocrine portion of the organ, while anterior lobe will be used to refer to the sub-region (Figure 1.1.3a).

The posterior lobe is the site of release of the hormones oxytocin and vasopressin, although the hormones are actually produced in the hypothalamus. The anterior pituitary secretes a number of hormones, which are regulated both by the hypothalamus and by feedback from the target tissues. The anterior pituitary contains a number of specialised hormone producing cell types, which secrete several different hormones (Table 1.1.3a). The actions of these hormones are diverse. The anterior hormones of greatest clinical importance are GH, PRL, TSH, FSH, LH, and ACTH. In general, defects in these cell types result in either an underproduction of a hormone (hypopituitarism) or an overproduction of the hormone (hyperpituitarism) (Table 1.1.3b). Subtle defects in the profile of hormone secretion may result in a clinical phenotype, such as acromegaly. All six hormones may be deficient, either singly or in any combination.
### Table 1.1.3a) Hormones produced by the anterior pituitary.

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Hormones produced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somatotropes</td>
<td>Growth hormone (GH)</td>
</tr>
<tr>
<td>Lactotropes</td>
<td>Prolactin (PRL)</td>
</tr>
<tr>
<td>Corticotropes</td>
<td>Adrenocorticotropic hormone (ACTH)</td>
</tr>
<tr>
<td></td>
<td>β-lipotropin</td>
</tr>
<tr>
<td>Thyrotropes</td>
<td>Thyroid-stimulating hormone (TSH)</td>
</tr>
<tr>
<td>Melanotropes</td>
<td>α-melanocyte stimulating hormone (α-MSH)</td>
</tr>
<tr>
<td></td>
<td>Corticotrophin-like intermediary peptide (CLIP)</td>
</tr>
<tr>
<td></td>
<td>γ-lipotropin</td>
</tr>
<tr>
<td></td>
<td>β-endorphin</td>
</tr>
<tr>
<td>Gonadotropes</td>
<td>Luteinizing hormone (LH)</td>
</tr>
<tr>
<td></td>
<td>Follicle-stimulating hormone (FSH)</td>
</tr>
</tbody>
</table>

### Table 1.1.3b) The clinical consequences of anterior pituitary hormone defects.

<table>
<thead>
<tr>
<th>Hormone</th>
<th>Hypopituitarism</th>
<th>Hyperpituitarism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth hormone</td>
<td>Dwarfism</td>
<td>Gigantism, Acromegaly</td>
</tr>
<tr>
<td>Prolactin</td>
<td>Failure of lactation</td>
<td>Galactorrhea</td>
</tr>
<tr>
<td>Adrenocorticotropic hormone</td>
<td>Hypoadrenalism</td>
<td>Cushing’s disease</td>
</tr>
<tr>
<td>Thyroid-stimulating hormone</td>
<td>Hypothyroidism</td>
<td>Hyperthyroidism</td>
</tr>
<tr>
<td>Gonadotrophins</td>
<td>Amenorrhea</td>
<td>Precocious Puberty</td>
</tr>
<tr>
<td></td>
<td>Delayed Puberty</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Infertility</td>
<td></td>
</tr>
</tbody>
</table>

The pharyngeal pituitary is an elusive tissue, which is only reliably detectable using immunohistochemistry (Fuller and Batsakis 1996). Lying under the sphenoid bone, it is smaller than the stellar pituitary, and may produce any combination of the anterior lobe hormones (Fuller and Batsakis 1996). It is unknown whether this tissue represents a redundant embryological tissue or a functional adult tissue.

The diagnosis of pituitary abnormalities is normally made by performing tests of anterior pituitary function. These are usually performed when a patient presents with clinical features suggestive of a pituitary abnormality, such as short stature with poor growth, hypoglycemia, lethargy and delayed puberty. However, a number of defects of the stellar pituitary may be seen radiologically. The anterior pituitary may be hypoplastic or aplastic and the pituitary stalk may be hypoplastic or apparently disrupted, with the posterior pituitary occupying an
ectopic position (Figure 1.1.3b). There appears to be no correlation between radiological findings and pituitary function. In particular, patients with an ectopic posterior pituitary may not manifest features of diabetes insipidus.

(Figure 1.1.3b) An ectopic posterior pituitary. A WT MRI scan is shown on the left while a MRI scan of a SOPD patient is shown on the right. Note that in addition to the ectopic posterior pituitary, the anterior pituitary is hypoplastic, the corpus callosum partially hypoplastic, while the pituitary stalk is not imaged. AP - anterior pituitary, CC - corpus callosum, OC - optic chiasm, PP - posterior pituitary, and PS - pituitary stalk.

1.1.4 Associated features

A number of associated features have been reported in SOD patients. These vary from defects that appear to be extensions of the SOD phenotype, through those affecting other anterior structures, to those affecting developmental systems quite separate from the anterior midline. In addition, a number of cases of SOD have been described in association with a second syndrome, as have a handful of cases of SOD arising as the result of exposure to a teratogen.

Atypical ophthalmological and endocrine features have been described in SOD patients. Three cases have been described with other ocular phenotypes. A case of SOD complicated with acute optic neuritis has been described (Nakagaki et al., 2001). MRI demonstrated agenesis of the septum pellucidum and hypoplasia of the corpus callosum without optic nerve hypoplasia. Gunduz et al. (1996) described an 8-month-old girl with SOD and bilateral complex microphthalmos. The child demonstrated bilateral microphthalmos, retinal dysplasia, anterior segment dysgenesis in the right eye and congenital cataract in the left eye. Finally, a patient with SOD and Duane syndrome has been characterised (Aguirre-Aquino et al., 2000).
A single case of a patient with SOD associated with presumed polyendocrine dysfunction has been described (Shammas et al., 1993). The patient had central diabetes insipidus, Hashimoto's thyroiditis, and gestational diabetes mellitus.

Defects of the corpus callosum have been described in a number of SOD patients. Likewise, olfactory nerve/bulb hypoplasia has also been described in a number of patients, suggesting that these two components may be part of a wider syndrome encompassing septo-optic dysplasia. Olfactory nerve hypoplasia/aplasia is a common autopsy finding, indeed de Morsier (1956) described this finding. Of the two autopsies described by Costello and Gluckman, one demonstrated aplasia of the olfactory bulbs and nerves and the other hypoplasia (Costello and Gluckman 1984). Coulter et al. described a fourth case as part of a spectrum of features (Coulter et al., 1993), whilst Levine et al. described a further case in 2001. Olfactory bulb hypoplasia is a feature in the Hesx1 null mutant mouse but has not been demonstrated in the index HESX1 mutant patients (Dattani et al., 1998). However, subclinical olfactory bulb hypoplasia cannot be ruled out.

Two syndromes resulting from defective neural development appear to be related to septo-optic dysplasia, namely holoprosencephaly (HPE) and Schizencephaly. Patients with HPE can manifest abnormalities of pituitary development as well as ONH, and agenesis of the septum pellucidum is often considered to be a mild form of HPE. Schizencephaly also appears to be related to SOD and has been demonstrated in a number of SOD patients (Denis et al., 2000 and De Smedt et al., 2000). Denis et al. (2000) reported 30 patients with schizencephaly of whom 13 lacked a septum pellucidum, 9 demonstrated dysgenesis of the corpus callosum, and 2 demonstrated septo-optic dysplasia.

Three cases of septo-optic dysplasia with associated digital abnormalities have been described. Pagon and Stephan (1984) reported two male SOD patients with digital abnormalities. Both children had similar anomalies of the hands and feet, including aplasia/hypoplasia of the distal and middle phalanges of fingers 2,3,and 4 and the great toes and constriction bands on a number of fingers. A further case has been reported by Faivre et al. (2002).

A number of single case reports of septo-optic dysplasia associated with a second syndrome have been described. For example, Sener (1996) described a case of SOD in conjunction with bilateral rolandic cortical dysplasia in a 3-year-old child as cortico-septo-optic dysplasia. Sener also described a case of SOD associated with total callosal absence, calloso-septo-optic dysplasia (Sener 1996). Rivkees (2001) described a child with arrhythmicity and SOD, the arrhythmicity being due to failure of the biological clock located within the hypothalamus. Median Cleft Face Syndrome has been associated with SOD in a single patient (Stewart et
al., 1983). Teng et al. (1989) reported a female infant demonstrating Apert syndrome associated with SOD. Finally, a case of Cornelia de Lange syndrome combined with SOD has also been described (Hayashi et al., 1996).

In addition to these associations, two cases of SOD associated with a second syndrome have been described where the causative mutation for the second syndrome is known. However, it is unclear whether the mutations in question are causative for SOD. Firstly, Carey et al. (1998) described a case of Waardenberg syndrome associated with SOD. The 6-month-old female infant from a non-consanguineous kindred, with no familial history of either disorder, carried a novel \textit{PAX3} mutation. Secondly, Schuelke et al. (2002) reported a case of SOD associated with exercise intolerance due to a mitochondrial mutation.

A case of Sotos syndrome associated with SOD has also been described (Buyukgebiz et al., 1996). 77% of Sotos patients carry a mutation in the NSD1 gene, and the lack of NSD mutations in 33% patients suggests the presence of at least one other causative locus (Kurotaki et al., 2002). In addition, the genomic location of the \textit{HESXI} gene (3p21) has been implicated in two patients. Schrander-Stumpel et al. (1990) described a 6 year-old boy with a t(3;6)(p21;p21) translocation, while Cole et al. (1992) described a 22 year-old woman with loss of heterozygosity at 3p21.

Three cases of SOD forming part of the spectrum of defects resulting from maternal teratogen consumption have been described. Phencyclidine was putatively implicated in the defects seen in a female newborn demonstrating a number of CNS malformations including hypoplasia of the optic nerves and aplasia of the septum pellucidum and posterior pituitary (Michaud et al., 1982). McMahon and Braddock (2001) described a typical case of valproic acid embryopathy complicated by SOD. Coulter (1993) reported an infant with fetal alcohol syndrome with numerous CNS malformations including SOD.

1.2 Development.

1.2.1 Optic nerve development.

Vertebrate eye development is a complex inductive process involving tissues of various embryonic origins: neural ectoderm, surface ectoderm, neural crest, and paraxial mesoderm in addition to blood vessels and autonomic nerves. The optic nerve consists almost entirely of neural ectoderm derived tissue, mostly retinal ganglionic axons. Development of the
vertebrate eye has been described morphologically in a large number of species ranging from the rhesus monkey, via the mouse and *Xenopus*, to the zebrafish. However, the mouse stands as the best model due to the ease of generating and selecting eye phenotypes combined with its position as a mammal. Optic nerve development occurs in three distinct stages: the formation of the optic vesicle during early forebrain formation, separation of the optic vesicle into the optic cup and optic stalk, and retinal ganglionic axon migration into and along the optic stalk forming the optic nerve proper.

In the mouse formation of the optic placode is the first stage of eye development. The optic placode is a circumscribed central region of the neuroepithelial surface of the cephalic neural folds that appears at around E7.5-8 (early Theiler stage [TS] 12, human Carnegie stage [CS] 9) when the folds start to elevate within the central part of the prospective forebrain region (Kaufman, *The atlas of mouse development* 1998). Shortly afterwards, by E8.5 (TS13, CS10) the optic sulcus, a shallow medially directed groove, connects the medial part of the placode to the diencephalic region of the developing forebrain, which will in time form the hypothalamus. During TS13, development of the optic pit begins as an indentation in the centre of the optic placode that deepens as the cephalic neural folds enlarge (Kaufman, *The atlas of mouse development* 1998). Between E8.5-9, the optic vesicles develop from the optic placodes as lateral invaginations of the forebrain. As the optic vesicles invaginate the optic stalks form, connecting the invaginations to the forebrain (Kaufman, *The atlas of mouse development* 1998).

It is not until E9.5 (TS 15, CS12) that the optic vesicles begin to expand while the optic stalks contract, morphologically distinguishing the two structures. Over this time period the optic vesicles collapse inward at the outer edge to form the optic cup (Kaufman, *The atlas of mouse development* 1998). At this stage, the optic cup comes in contact with the overlying ectoderm which is induced to form the optic pit, that then develops into the secondary optic cup and this goes on to form the lens vesicle (Kaufman, *The atlas of mouse development* 1998).

It is not until E11.5-12 (TS 19-20, CS16) that the formation of both the optic cup and optic stalks is complete. This is marked by the disappearance of the intra-retinal space and the lumen of the optic stalk. At this stage the optic stalks reach their narrowest point. The diameter of the optic stalk increases from this point as the axonal processes of the neural retinal cells grow along the stalks forming the optic nerves. Over this time period the optic stalk is not circular in cross-section, rather it has deep cleft on the ventral surface, the optic fissure, in which lies the hyaloid blood vessels (Kaufman, *The atlas of mouse development* 1998).
By E12 (TS20) axon projections from the retinal ganglion cells pass through the centre of the optic stalk and migrate into the narrow lumen of the optic stalk, occupying and thus obliterating the lumen. These unmyelinated axons migrate medially towards the optic chiasma to form the optic nerves and from there onto the lateral geniculate bodies to form the optic tracks. The optic stalk/nerve becomes covered by a layer of mesenchyme that becomes fibrous forming the dura mater, which is continuous with that of the brain and the fibrous coating of the eyeball.

Genetic dissection of vertebrate eye development has increased the understanding of optic nerve development and has confirmed the presence of the three separate phases of optic stalk development.

The development of the optic vesicle is the least understood and until recently genetic information was limited to expression data such as the bilateral expression of Six6 detectable at E8.25 in the region of the optic placode, representing the first genetic marker of the optic placode (Jean et al., 1999). Two mouse mutations do however provide valuable insights into this phase of development. Firstly, homozygous Pax6 mutant embryos show a complete absence of eyes by midgestation. These mice do however form “normal” optic vesicles but development does not proceed beyond this point (Grindley et al., 1995). Secondly, Hesx1 null mutant embryos show variable hypoplasia of the optic vesicles with phenotypic divergence between each side of the embryo. Recent experiments in the chick demonstrate a central role for the signalling molecules Shh, Fgf8, and Bmp4 in the coordinated development of the optic and telencephalic vesicles (Ohkubo et al., 2002).

By inserting BMP4, BMP2 or BMP7 soaked beads into the anterior neuropore Ohkubo et al. (2002) were able to induce variable defects of optic vesicle development. In the most severe cases (class E) the anterior prosencephalon is truncated such that optic vesicles are not morphologically identifiable. In less severe cases (class D) anterior prosencephalon structures are produced but are abnormal. The phenotype includes a complete absence of the optic stalk with the retinal tissue continuous with the forebrain. Less severe classes retain more forebrain-derived structures in concert with normalising FGF8 and SHH signal production. However the optic stalk is consistently affected with optic stalk hyperplasia without Pax2 expression in the mildest cases, demonstrating that, of the forebrain structures, the optic stalk is the most sensitive to Bmp4 beads. Interestingly the Bmp4 mediated inhibition of Shh signalling does not require protein synthesis whilst the reduction of Fgf8 does. In line with these findings Shh signalling has therefore been shown to be required for optic stalk formation and Pax2 expression.

34
The separation of the optic vesicle into the optic cup and optic stalk is well described and a central role for the genes Pax2 and Pax6 has been elucidated. Pax2 is expressed in the ventral half of the optic vesicle during early eye morphogenesis, being restricted thereafter to the optic stalk. Expression persists at the lips of the optic fissure along the entire optic stalk (Nornes et al., 1990 and Torres et al., 1996). In Pax2 deficient mice the optic nerves project only to the superior colliculus with pigmented retinal epithelium extending into the stalk/nerve (Torres et al., 1996). In man, PAX2 mutations result in optic nerve coloboma (Sanyanusin et al., 1995). On the other hand, Pax6 is expressed within the optic cup and the tissues derived from it. Mutations in Pax6 in the mouse underlie a number of small-eye phenotypes, while human PAX6 mutations result in aniridia (Nornes et al., 1990, Walther and Gruss, 1991, Hill et al., 1991, Grindley et al., 1995 and Torres et al., 1996). The evidence suggests that Pax2 defines the optic stalk; while Pax6 defines the optic cup with both genes capable of reciprocal transcriptional repression, leading to the definition of a sharp boundary. During optic stalk/optic cup definition and development a small region of overlap between the Pax2 and Pax6 expression domains occurs at the optic cup/stalk boundary, marking the future optic nerve head (Schwarz et al., 2000).

Retinal ganglionic axon migration into and along the optic stalk forming the optic nerve proper is well described both morphologically and genetically. Morphologically, axon migration occurs in a specific temporal pattern from different regions of the retina, with a specific and coordinated pattern of axon bundle formation. These axons then migrate along the optic stalks and cross the midline, forming the optic chiasma. A large number of cell adhesion molecules are expressed differentially within the optic stalk/nerve, forming part of an axonal migration pathway. A common phenotype of abnormal migration results from disturbance to this pathway (Plump et al., 2002). The most critical stage of axon migration in terms of optic nerve hypoplasia is the entrance of the retinal ganglion axons into the optic stalk lumen forming the optic nerve head, as a reduction at this point marks optic nerve hypoplasia. It is unclear if the future optic nerve head produces a long-range signal that attracts the axons towards it. When the retinal axons reach the region of the future optic nerve head they are attracted to the optic never head by the diffusible protein Netrin-1 (Deiner et al., 1997 and Hopker et al., 1999). When the retinal axons reach the optic nerve head, Laminin-1, which is expressed only on the retinal surface, converts the Netrin-1 attraction into repulsion, and this appears to result in the axons sharply turning into the optic nerve. The gene DCC1 encodes the main Netrin-1 receptor on the surface of retinal ganglion axons and plays the central role in mediating the transduction of the Netrin-1 signal into growth of the retinal ganglion axons (Stein et al., 2001, and Stein and Tessier-Lavigne, 2001). As the retinal ganglion axons migrate along the optic nerve, Netrin-1 signalling is
modified by Slit-1 and Slit-2 such that the axons migrate along their correct path, forming the optic chiasma at the correct location. Mice harbouring mutations in either DCC1 or Netrin-1 exhibit optic nerve hypoplasia due to a reduced number of retinal ganglion axons exiting the eye via the optic nerve head (Deiner et al., 1997).

Following the bilateral expression of Six6 at E8.25 in the region of the optic placode, expression of Six6 is limited to the ventral optic stalk and the ventral presumptive neural retina, corresponding to the boundary between neural and pigment retina. This expression persists in the retinal ganglion cells (Lopez-Rios et al., 1999 and Jean et al., 1999). In homozygous null mutant mice, loss of Six6 function in the retina precursors results in decreased cellular proliferation and thus retinal hypoplasia that often results in optic nerve hypoplasia.

1.2.2 Development of the septum pellucidum.

The processes underlying the development of the septum pellucidum are not well understood. The septum pellucidum consists of two membranes, connected to the corpus callosum above and the fornix below, which separate the lateral ventricles from each other and it originates from the rostral wall of the telencephalon, in particular the dorsal region of the lamina reuniens (figure 1.2.2a) (Rakic and Yakovlev, 1968). Its development is closely linked to that of the fornix and the corpus callosum (Rakic and Yakovlev, 1968).

In humans, between 25 and 50mm crown rump length (CRL) (8 to 11 weeks) the lamina reuniens is a densely cellular homogeneous cell mass lying across the midline. However, no nerve fibres cross the midline at this point.

During the early part of this period, (25-35mm CRL) nerve fibres grow medially from the ganglionic eminence, such that by 40mm CRL, the ends of the nerve fibres have met at the midline within the ventral lamina reuniens, forming the anlage of the anterior commissure, the area praecommissuralis (Rakic and Yakovlev, 1968). Around 40mm CRL two fibre bundles can be seen growing from the midline area praecommissuralis dorsally before converging to form a single median groove and by 45-55mm CRL nerve fibre exchange between the two bundles forms the primordium of the fornix (Rakic and Yakovlev, 1968). By 60mm CRL (12 weeks), the crossing fibre bundle has formed the anterior commissure proper (Rakic and Yakovlev, 1968).
Beginning around 23mm CRL the dorsal region of the lamina reuniens, between the hemispheric vesicles, folds into the median groove, such that by 40mm CRL the banks of the groove are juxtaposed (Rakic and Yakovlev, 1968). Cellular proliferation within the banks of the median groove leads to the fusion of the banks of the median groove forming the massa commissuralis (45mm CRL). Between 50 and 60mm CRL, the first fibres migrating from the isocortical region of the hemispheric vesicles penetrate the massa commissuralis, thus forming the corpus callosum (Rakic and Yakovlev, 1968). Over the next eight weeks of development the number of nerve fibres crossing the midline increases so that by 160mm CRL (20 weeks), the corpus callosum has reached its adult shape but not its full thickness (Rakic and Yakovlev, 1968).

The septum pellucidum itself arises from the floor of the massa commissuralis between the nerve fibres of the corpus callosum and the area praecommisuralis between 60-90mm CRL (Rakic and Yakovlev, 1968). The septum pellucidum first appears as a small triangular out-pocket, the apex of which points towards the area praecommisuralis. By 145-170 mm CRL the closure of the banks of the median groove have closed over the triangular out-pocket heading towards the area praecommisuralis (Rakic and Yakovlev, 1968). The sides of the triangular out-pocket form the septum pellucidum leaving a region between them, the cavum septi. The cavum septi progressively reduces in size until it forms a narrow cleft in the adult (Rakic and Yakovlev, 1968).

1.2.3 Pituitary development.

Of the three structures affected in septo-optic-pituitary dysplasia, the development of the pituitary is by far the best characterised. Development of the pituitary has been described not only in man but also across a wide array of model organisms. It is clear that the pituitary is derived from two distinct tissues; the posterior pituitary is derived from neuro-ectoderm whilst the anterior pituitary is derived from oral ectoderm. During pituitary development, these two tissues develop in unison into the adult gland. Although development of the pituitary has been described in a number of species, it is the development of the pituitary in the mouse that has become the model of human pituitary development and is well characterised both morphologically and genetically. However, development of the anterior pituitary has been investigated in far greater detail than the posterior pituitary.

Experiments in the mouse (reinforced by experiments in *Xenopus* and the chick) suggest that the anterior pituitary primordium is located within the midline of the anterior neural ridge at
E7.0 immediately anterior to the primordium of the neural plate which will give rise to the ventral diencephalon, including the posterior pituitary (Couly and Le Douarin 1988, Eagleson and Harris 1990, Fedtsova and Barabanov 1990 and Rubenstein et al., 1998). Although it is possible to trace the origin of the anterior pituitary to the anterior neural ridge, development of the anterior lobe as a separate tissue begins within the oral ectoderm shortly before rupture of the buccal or oral membrane. Development in the mouse occurs in four distinct stages: the placode stage, the Rathke’s pouch stage, the definitive pouch stage, and finally, the formation of the pituitary gland proper (Sheng and Westphal, 1999). The anterior pituitary begins as a thickening of the oral ectoderm, termed the pituitary placode, at E8 in the mouse. This invaginates to form Rathke’s pouch, by E9.5. Rathke’s pouch then closes to form the definitive pouch by E12 and finally, the pouch undergoes terminal differentiation to form the mature gland by E17 (Figure 1.2.3a).

(Figure 1.2.3a) Four stages of pituitary development (Sheng and Westphal, 1999). AL-anterior lobe, AN-anterior neural pore, DI-diencephalon, F-forebrain, H-heart, HB-hindbrain, I-infundibulum, MB-midbrain, N-notochord, NP-neural plate, O-oral cavity, OC-optic chiasma, OM-oral membrane, P-pontine flexure, PL-posterior lobe, PO-pons, RP-Rathke’s pouch and SC-sphenoid cartilage

The placode stage begins at E8, after the turning of the embryo has been accomplished at the anterior end. The stomodeum roof is flat and the beginnings of Rathke’s pouch are not apparent (Kerr, 1946). At this stage the oral ectoderm uniformly expresses the gene sonic hedgehog (Shh). The rapidly dividing paraxial mesoderm begins to separate the brain floor from the surface ectoderm such that by E8.5 the ectoderm becomes attenuated except at defined areas, one of which is the pituitary placode, which is becoming faintly defined due to its slight increase in thickness (Kerr, 1946). It is at this stage that Hesx1 is expressed, within the region of oral ectoderm immediately underlying and in contact with the diencephalon.
(Hermes et al., 1996). Thus, Hesx1 expression within the oral ectoderm defines the pituitary placode and is the first genetic marker of pituitary development. Shortly after Hesx1 expression begins, Shh expression is eliminated within the Hesx1 expression domain. Coinciding with this event the expression of the transcription factor Isl1 occurs within the oral ectoderm underlying the diencephalon.

Between E8.5 and E9.5, Rathke’s pouch becomes defined as a separate entity. A region of neural-ectodermal fusion is maintained while the two tissues are becoming separated in the surrounding area (Kerr, 1946). This causes the formation of a shallow dome shaped pocket, which is Rathke’s pouch (Figure 1.2.3b A). During this stage, Lhx3 is expressed within pouch tissue underlying the infundibulum. Lhx3 expression rapidly spreads throughout Rathke’s pouch. Interestingly, Lhx3 expression does not occur within the ventral domain of the posterior wall of Rathke’s pouch. Explant experiments have shown that Lhx3 expression is induced by FGF signalling from the ventral diencephalon, which in vivo expresses FGF8, FGF10, and FGF18 (Takuma et al., 1998). By E9, Hesx1 expression has been extinguished throughout the embryo except for Rathke’s pouch. At E9.5, although Rathke’s pouch has formed it has yet to begin folding towards closure (Kerr, 1946) (Figure 1.2.3b B). Progression beyond this stage requires Lhx3 (Sheng et al., 1996).

(Figure 1.2.3b) Pituitary development occurs via distinct stages (A-F) (Kerr 1946) A.t.-anterior thickening, B.v.-blood vessel, Hy.s.-hypophyseal stalk, Mes.-mesenchyme, Ne.u.-neural-ectoderm fusion, N.L.-neural lobe, O.c.-optic chiasma, Po.L.-postoptic lamina, Pl.n.-prechondral nuclei, Pm.l.-premammillary lamina, P.t.-posterior thickening, R.p.-Rathke’s pouch and Th-transitional border of cells.
Shortly after E9.5, Rathke’s pouch begins to fold over on itself. The pivot for this folding is the neural-ectodermal fusion region. Rathke’s pouch also bends halfway down the posterior wall, and this second bending occurs in an opposite direction to the main fold (Kerr, 1946) (Figure 1.2.3b B). By E10, expression of the pituitary specific transcription factor prop1 occurs and Rathke’s pouch has deepened and become more tubular, while maintaining a circular cross-section (Kerr, 1946). The two walls now lie parallel to each other. Both the posterior and anterior walls have thickened. The tissue posterior to the posterior wall bend has not undergone thickening and will not contribute to the adult gland (Kerr, 1946). By this stage, the paraxial mesoderm completely surrounds Rathke’s pouch (Figure 1.2.3b C). Over the next few hours, the two sides of Rathke’s pouch stay parallel to each other and the walls continue to thicken. By this time, the apex of Rathke’s pouch (fusion region) is flat and the pouch in sagittal section looks rectangular (Kerr, 1946) (Figure 1.2.3b D).

A consequence of this process is that the posterior and anterior walls of Rathke’s pouch are now equal in length (Figure 1.2.3b D). Under the action of both an internal folding mechanism and the pressure of the mesoderm, the lips of the mouth of Rathke’s pouch move closer together (Kerr, 1946) (Figure 1.2.3b E). Thus, when the mouth of Rathke’s pouch meets, it does so squarely (Kerr, 1946) (Figure 1.2.3b F). The closed pouch is now known as the definitive pouch. The progression of development from the invaginated pouch to the closed definitive pouch requires BMP signalling as deletion of BMP4 or the expression of BMP antagonist Noggin results in the failure of pituitary development to progress beyond E10. Likewise, Shh expression within the surrounding oral ectoderm is required for cellular proliferation after E10 in part due to Shh inducing BMP2 within the pouch and in part due to Shh acting in concert with FGF's to maintain Lhx3 expression (Treier et al., 2001). The closure of Rathke’s pouch also requires Lhx3/Lhx4 since in the Lhx3/Lhx4 double knockout mouse the pouch fails to close. In the Lhx3 null mutant, the pouch does close but development appears to arrest at this point, with only the corticotrope lineage forming (Sheng et al., 1996 and Sheng et al., 1997). Isl1 is also required for progression beyond this point since in the Isl1 null mutant, pouch closure does occur, but cellular differentiation does not.

The expression patterns of a number of genes change over this period. At E10, Isl1 expression is reduced in a wave from the dorsal to the ventral side, such that by E11, Isl1 is expressed only at the ventral side of the pouch (Ericson et al., 1998). At E11.5, the expression of Hesx1 begins to reduce. This reduction occurs from the ventral side to the dorsal side of the pituitary, such that by E15.5, expression is extinguished (Hermesz et al., 1996). The gene Prop1, which is first expressed at E10, reaches its maximum level of expression at E12 and begins to be down regulated, such that its expression is not reliably detected by E15. Over this period, a number of genes show this gradient of expression. Nkx-
3.1, Pax6, Six6, and Six-3 are expressed dorsally while Pitx1, Ptx2, Msx1, Brn-4, GATA-2, P-Frk and T-pit/Tbx19 are expressed ventrally.

As the brain floor and the roof of the mouth are pushed apart by the paraxial mesoderm, the definitive pouch is pulled away from the oral ectoderm. As the pouch moves away from the oral ectoderm, a stalk of tissue, the hypophyseal stalk, forms at E12. By E12.5, this stalk thins and breaks, leaving a small thickened area of the oral ectoderm, which eventually becomes indistinguishable from the surrounding tissue. The future of this tissue is unknown in the mouse; however in man it forms the pharyngeal pituitary.

Up to E12.5, the pouch consists of homogenous cells (Ericson et al., 1998, and Sheng and Westphal, 1999). However at E12.5 Rathke’s pouch reaches a crucial stage when cellular homogeny is lost when the first pituitary cell type, the rostral thyrotropes, differentiates as shown by expression of thyroid stimulating hormone (TSH) (Ericson et al., 1998, and Sheng and Westphal, 1999). These cells have no known function and disappear shortly after birth. It is unknown if rostral thyrotropes occur in man. Shortly after formation of the rostral thyrotropes, the corticotrope lineage differentiates just dorsal to the thyrotropes, demonstrated by the expression of adrenocorticotropic hormone (ACTH). Explant experiments carried out by Ericson et al. (1998) suggest that rostral thyrotropes differentiate from isl1⁺/lhx3⁺ cells whilst corticotropes differentiate from isl1⁻/lhx3⁻ cells. The T-box gene T-pit/Tbx19 appears to be critical for the balance between these two cell types as misexpression in vivo results in the misexpression of ACTH within the rostral tip suggesting conversion of rostral tip cells into corticotropes (Lamolet et al., 2001). T-pit/Tbx19 also appears necessary for the differentiation of the melanotrope lineage.

When hypophyseal stalk breakage commences, the definitive pouch, now free from the oral ectoderm, begins its reorientation into its position in the adult. The underlying mesoderm, which has begun to differentiate into cartilage in the surrounding areas, prevents enlargement ventrally. The neural lobe at this time does allow some dorsal enlargement and there is space for enlargement to proceed at both the anterior and posterior aspects of the definitive pouch (Kerr 1946). This growth pattern of the gland brings about a change in its relative position to the main axis of the hypophyseal cavity, as it is bent backwards and upwards by approximately 90 degrees. This movement brings the pouch and the neural lobe into the positions they occupy in the adult.

During the latter stages of pituitary development, large and irregularly shaped processes, the cell cords, are pushed out of the anterior pituitary (Kerr 1946 and Hashimoto et al., 1998). The expansion of the cell cords, both in the caudal direction and in the dorsal direction, results in the mesenchyme becoming located both between the cell cords and along the
anterior of the pouch. This mesenchyme then invades the pituitary, forming a mesenchyme network throughout the gland. Over this time frame, vascularisation of the pituitary also occurs. By E18, a capillary network exists throughout the anterior pituitary and connects to the posterior pituitary by a single vessel (Hashimoto et al., 1998).

During the final stages of pituitary development (E12.5-E17) the overlapping expression of transcription factors initiated between E10-E12.5 results in the division of the pituitary into distinct cell types in a specific temporal and spatial manner, such that an adult arrangement of cell types is achieved. Whilst formation of the rostral thyrotropes, corticotropes, and melanotropes is dependent on the early expression of Isl1, Lhx3, and T-pit/Tbx19, the central region of pituitary cells is dependent on a single gene, Prop-1, for differentiation. Within the mouse a hypomorphic allele of Prop-1 (the AMES dwarf mouse) leads to a reduction in numbers of somatotropes, lactotropes, and thyrotropes (Somson et al., 1996 and Gage et al., 1996). This region is subsequently divided into thyrotrope, gonadotrope, and a common somatotrope/lactotrope cell lineages by the opposed actions of two transcription factors, Pit-1 and GATA-2 (Dasen et al., 1999). These two genes are expressed in opposite ventral-dorsal gradients such that Pit-1 is expressed within the ventral and central regions, while GATA-2 is expressed within the dorsal and central regions. The entire region appears to be able to express Pit-1 while GATA-2 expression appears to require induction by a Bmp2 gradient produced by the dorsally located mesenchyme. High Gata-2 dorsal expression close to the BMP2 producing mesenchyme appears to inhibit Pit-1 expression within the dorsal regions whilst lower levels within the central region are unable to inhibit Pit-1 expression (Dasen et al., 1999). The ventral Pit-1 only region develops into somato-lactrotrope precursor cells under the action of Pit-1 protein, whilst the ventral Gata-2 only region develops into gonadotropes under the action of Gata-2 protein in conjunction with Sfl protein. The central region expressing both Pit-1 and Gata-2 develops into thyrotropes due to the actions of both Pit-1 and Gata-2 proteins. Pit-1 protein is produced in a large excess such that almost all Gata-2 protein is heterodimerised with Pit-1 within the region coexpressing the two proteins (Dasen et al., 1999).

1.2.4 Developmental origin of Septo-optic dysplasia.

The central issue in terms of the developmental origin of SOD is: How do defects in three separate systems arise? The most widely accepted solution is that SOD is a defect of the early anterior midline, perhaps at the anterior neural ridge stage, which later leads to the phenotypic defects. However, the central defect in the Hesxl null mutant mouse appears to
be within the early forebrain (Dattani et al., 1998 and Martinez-Barbera et al., 2000). A crucial difference between SOD and the phenotype of the Hesx1 null mutant mouse is that the classical triad is only seen in a subset of null mutant mice alongside generalised forebrain dysplasia. In man, the classical triad forms a distinct constellation of features separate from forebrain abnormalities are rare. In addition, optic nerve hypoplasia is not a feature of the Hesx1 null mutant mouse; rather generalised hypoplasia of the eyes is apparent.

A central defect of vascularisation to which optic nerve, pituitary and forebrain development is particularly susceptible has also been proposed, although a number of the assumptions upon which this hypothesis is based have been shown to be false (Brodsky, 1998). As well as a single defect, separate defects within each of the classical component of SOD are possible, such as a genetic factor independently expressed within each system. Indeed, it is unclear whether the defects observed in the Hesx1 mutant mice arise from a central defect of forebrain formation or whether there are separate defects of the eyes, pituitary and forebrain. Recent work by Dasen et al., (2001) suggests that the pituitary defect in the Hesx1 mutant mouse can stem from an independent source to that of the other forebrain defects. This reflects that a number of distinct Hesx1 expression zones exists. However, it is unclear if this concept can be applied to SOD.

1.3 Genetics.

1.3.1 HESX1

Kazanskaya et al. (1997) first cloned HESX1 cDNA, which they termed Hanf based on homology to the Xenopus gene Xanf-1. However, they carried out no investigation into the gene. Dattani et al. (1998) independently cloned the gene and they termed it HESX1 based on homology to the murine gene Hesx1. They cloned HESX1 cDNA using degenerate PCR and used this fragment to identify the genomic sequence.

Homology between Xanf-1 and Hesx1 places HESX1 in the Anf family of homeodomain genes. The HESX1 genomic structure is similar to that of Hesx-1 and Xanf-1. It is relatively small (1.7kb) and consists of 4 exons. Uniquely exons 2, 3, and 4 encode a homeodomain, which is of the paired-like class. Exon 1 encodes an amino-acid domain conserved across the Anf family, which shows identity to a region present in the engrailed family of homeodomain proteins (eh-1 domain), thought to play a pivotal role in transcriptional repression (Dattani et
al., 1998). Several lines of evidence suggest that Hesx1 acts as a transcriptional repressor and recently in-vitro transfection studies have indicated that HESX1 acts as a transcriptional repressor (Brickman et al., 2001) (see section 4.3).

Dattani et al. (1998) also referred to an expressed sequence tag (dbEST Id: 788895, GenBank gi: 1720505) generated by the IMAGE consortium (http://image.llnl.gov) from the Saoares_pregnant_uterus_NbHPU, which contains exons 2, 3, and 4 of HESX1 downstream of divergent sequence in the place of exon 1 which they believed to be a divergent exon. Using Southern blot analysis and PCR, Dattani et al., were unable to identify the origin of this sequence (Dattani personal communication). Using fluorescent in situ hybridisation HESX1 was located to 3p21.1-21.2 within a small region showing synteny to a region of murine chromosome 14, containing Hesx1.

Several studies that have been carried out into the PTX1 protein with respect to its ability to activate transcription of genes expressed within the pituitary have uncovered a number of PTX1 binding sites. Since PTX1 is a paired-like homeodomain protein and therefore related to HESX1, HESX1 should be able in principal to bind these sites. Furthermore, as Ptx1 is expressed during almost the entire period in which Hesx1 is expressed within the developing murine pituitary the two genes would be expected to functionally antagonise each other. Indeed, HESX1 has been shown to bind to the PTX1 binding sites within the αGSU and LHβ promoter (Quirk and Brown, 2002). Using in-vitro transfection experiments Quirk and Brown (2002) were able to demonstrate that HESX1 could directly repress these two genes, demonstrating that HESX1 may be able to directly repress the genes encoding various PTX1 activated pituitary hormones. Quirk and Brown (2002) went on to investigate these naturally occurring paired binding sites and demonstrated that HESX1 shows promoter specific transcriptional repression. Although HESX1 can repress both αGSU and LHβ in isolation, HESX1 is unable to reduce PTX1 activation of LHβ, whereas Hesx1 shows dosage dependent antagonism of the activation of αGSU by Ptx1. They demonstrated that HESX1 binds as a monomer to the PTX1 binding site within the LHβ promoter but binds as a homodimer at the αGSU promoter utilising both the known PTX1 binding site and a novel binding site 8bps downstream of the first binding site. Quirk and Brown (2002) conclude that HESX1 has two functions: 1) Repression of target genes as either a monomer or homodimer, and 2) blockage of specific transactivation sites as a homodimer.

Based upon the phenotype of the Hesx1 null mutant mouse, Dattani et al. (1998) screened 20 patients with SOD (including 2 familial cases), 26 patients with pituitary hypoplasia (including 2 familial cases), 9 patients with multiple pituitary hormone deficiency (including 5 familial cases), and 6 patients with holoprosencephaly. They used single strand
conformational polymorphism detection and southern-blot analysis and detected a single sequence change in two siblings with SOD. A c478t variant was identified due to the production of a single strand conformational polymorphism. No genomic rearrangements were discovered within 38 patients screened by Southern Blot analysis. Sequencing showed that both siblings were homozygous for the variant, while both parents were heterozygous for the variant. The siblings were the product of a consanguineous relationship. A number of relatives were shown to be heterozygous for the variant (n=8, including parents) but all were asymptomatic. A heterozygous relative suffering from persistent hyperinsulinemic hypoglycaemia of infancy was also identified (Dattani et al., 1998) within the pedigree. The affected siblings displayed agenesis of the corpus callosum, panhypopituitarism, optic nerve hypoplasia, and an absent septum pellucidum. These affected siblings were first reported by Wales et al. in 1996.

Dattani et al. (1998) investigated the function of the c478t variant and showed that it leads to an Arg160Cys amino acid substitution (Arg53Cys within the homeodomain). They demonstrated that this residue is conserved across the Anf family and almost completely across the entire family of homeodomain transcription factors. Crystal structure studies of the paired protein show that when the homeodomain is bound to DNA this residue makes contact with the DNA phosphate backbone. They generated recombinant Hisx6-tagged wild-type and Arg160Cys mutant protein using the pJBME21 vector and nickel affinity chromatography. Dattani et al. (1998) then investigated the ability of the Arg160Cys protein to bind to the consensus pIII oligonucleotide sequence using an in vitro EMSA assay. Unlike wild-type protein, Arg160Cys HESXI was unable to bind to the pIII probe. Therefore, Dattani et al. (1998) reasoned that in vivo the Arg160Cys protein would be unable to bind target DNA.

Dattani et al. (1998) also identified a heterozygous a374g (N125S) substitution within a child of Afro-Caribbean origin. N125S occurs at position 18 of the HESX1 homeodomain (Brickman et al., 2001) The variant was also present in the child’s mother and two siblings, suggesting that this variant represents either a polymorphism or an example of variable penetrance comparable to the Hesx1 null mutant mouse. During this study, Dattani et al., also identified a heterozygous c509t variant leading to a Ser170Leu amino acid substitution within two siblings displaying severe and evolving GH deficiency. In addition, the older sibling displayed optic nerve hypoplasia and mild cranio-facial dysplasia. (Thomas et al., 2001 and Dattani, personal communication).
1.3.2 Hesxl.

Hesxl (originally referred to as HES-1) was originally identified in a study of homeobox genes expressed in embryonic stem (ES) cells (Thomas and Rathjen, 1992). Thomas et al. (1995) then described the genomic organisation, sequence, and expression of the gene. They found two species of Hesxl using Northern blotting, a 1.0kb and a 1.2kb transcript. Sequencing of cDNA clones from ES cells demonstrated that both the transcripts contained identical open reading frames and 5'UTR, but divergent 3'UTR. The Hesxl homeodomain shares 80% identity with the homeodomain of the Xenopus homeobox gene XANF-1. They identified the genomic arrangement and the relationship between the transcripts. Hesxl contains 5 exons; the first four exons encode the Hesxl open reading frame, while the fifth exon encodes the alternative 3'UTR contained in Hesxlb. Webb et al. (1993) mapped Hesxl to mouse chromosome 14, Bands A3-B.

The expression of Hesxl was originally described by Thomas et al. (1995) and independently investigated by Hermesz et al. (1996), who referred to the gene as Rathke’s pouch homeobox (Rpx). Thomas et al. (1995) detected Hesxl expression in total RNA from E16 embryos, but not from E10.5 or E12.5 embryos. On the other hand, Hermesz et al. (1996) detected expression in total RNA at E8.5 (the earliest time point analysed) until E9.5 but not afterwards. Thomas et al. (1995) described tissue specific expression in embryonic liver, viscera, amnion, and yolk sac at E14.5 whilst expression was only detectable within the liver at E16. No expression was observed in post-natal adult tissues. Hermesz et al. (1996) performed a more in-depth expression profile using in-situ hybridisation. At the earliest stage examined, they detected signal at E6.5 just as cells began to delaminate from the primitive ectoderm to form mesoderm. Slightly later during early gastrulation (E6.5-6.75), they detected Hesxl expression in “a very restricted region of the endoderm directly across from the T (Brachyury) expression domain”. They state that is possible that the Rpx expression was within the primitive visceral endoderm, a view supported by Dattani et al. (1998). At E7, Hermesz et al. (1996) detected expression in the anterior axial mesendoderm underlying the prospective cephalic neural plate. At E7.5, expression had spread to the juxtaposed prospective neural plate. They suggest that this may reflect induction of expression in the prospective neural plate by expression in the underlying endoderm/prechordal plate (mesendoderm). By early head fold stage, expression in the underlying mesendoderm cells was extinguished, with expression limited to the neurectoderm of the prospective prosencephalon by E8-8.5. At E8.5, Hermesz et al. (1996) detected expression in the oral ectoderm and a small region of the adjacent foregut. At E9, Hesxl expression was limited to a region of the oral ectoderm with no expression detectable in the neuroectoderm. This
region of oral ectoderm is the anlage of Rathke’s pouch, the primordium of the anterior pituitary (Hermesz et al., 1996). They reported uniform expression throughout the development of the pouch into the anterior pituitary until E12.5 when the first endocrine cell type, the rostral thyrotropes appear. The other endocrine cells appear over the next few days in a dorsal to ventral wave within the pouch. Hesx1 expression diminishes in a similar wave mirroring the differentiation of the pituitary cells such that by E13.5, expression is undetectable in the anterior pituitary, leading to only sparse expression close to the diencephalon at E14.5, and total loss of expression by E15.5. No reference was made to expression within the liver by Hermesz et al. (1996).

Two independent strains of mice carrying Hesx1 null alleles were then generated (Dattani et al., 1998 and Mahon, unpublished data). Both strains display similar phenotypes. At birth, expected Mendelian ratios were observed. However, 75% of homozygous mice die before weaning (Dattani et al., 1998). This group first detected defects in mutant mice at E8.5 when homozygous mice displayed a reduction in prospective forebrain tissue as well as a lack of optic vesicles. They reported that this phenotype was not only variable between mice, but also between each side of a single embryo. By E9.5, this variability was such that two classes of phenotype were identified. Class I mutants display a severe phenotype, characterised by the absence or substantial reduction of the telencephalic vesicles, eyes, olfactory placodes and frontonasal mass. Class II mutants displayed a milder phenotype. Left-right asymmetry in the development of anterior structures was observed in many of the embryos. No defects were apparent caudal to the forebrain region. Histological analysis of class I embryos at E12.5 demonstrated a reduced forebrain with no telencephalic vesicle or infundibulum and absent optic cups, optic lenses, olfactory placodes and Rathke’s pouch (Dattani et al., 1998). At this stage, class II mutant embryos displayed a general retardation of forebrain development and reduced telencephalic vesicles. The embryos also displayed defects in the hypothalamus and aberrant morphogenesis of Rathke’s pouch. Hypoplasia of the optic vesicles and the olfactory pits was also apparent and often asymmetrical (Dattani et al., 1998).

Mutant mice also demonstrated defects in the expression patterns of Six3 and Pax6 within the anterior neural plate at E8.5 (Dattani et al., 1998). Fgf8 expression was significantly reduced at this stage within the forebrain, however mid-hindbrain expression appeared normal. Expression of other factors within the midbrain and hindbrain also appeared normal (Dattani et al., 1998). Between E10.5 and E12.5, normal expression of Shh in the medial and lateral ganglionic eminences was absent in class I mutants but present in class II mutants.
At birth Class I mutants displayed a small head with a short nose and an absence of eyes or severe microphthalmia (Dattani et al., 1998). These mice died shortly after birth. The milder group displayed a less severe craniofacial defect and a milder eye defect. The eye when present was located more posteriorty than in WT littermates and often remained buried in the orbit (Dattani et al., 1998). Histological analysis demonstrated a variable phenotype of midline defects. The septum pellucidum was absent in 25% of mutant mice. The corpus callosum was absent in 3 out of 4 mice; failure of the nerve fibres to cross the midline was evident rather than an absence of the nerve fibres themselves. The anterior pituitary was hypoplastic and only loosely associated with the posterior pituitary, with partial bifurcations apparent. However, mice that survived weaning proved fertile and reproduced, demonstrating that the gonadal axis is intact. The mutant mice were also a similar size to WT mice, suggesting that the growth hormone axis was also intact (Dattani et al., 1998). Class I mutants appear to consist of entirely homozygous mice, whilst class II mutants consisted of homozygote and heterozygote mice. Approximately 1% of heterozygous mice displayed a phenotype (Dattani et al., 1998). The phenotype of the second strain of Hesx1 null mutant mice is similar to that of Dattani et al. (1998) (Mahon, unpublished data).

The temporal and spatial expression of Hesx1 does not allow the location of the primary defect affecting forebrain development on the basis of expression alone. Therefore, Martinez-Barbera et al. (2000 and 2001) investigated whether the location of the primary defect in Hesx1 mutant mice was located in the anterior visceral endoderm, the anterior axial mesendoderm, or the anterior neural ectoderm. Using chimeric mice with variable WT or mutant contributions to the epiblast, but not the anterior visceral endoderm, they were able to show that forebrain defects were rescued in epiblast containing a high proportion of WT cells within a mutant anterior visceral endoderm. In chimeric mice containing a high proportion of mutant cells within the epiblast, within a WT anterior visceral endoderm, forebrain defects were observed. Likewise, Martinez-Barbera et al. (2000 and 2001) were unable to find defects within the anterior visceral endoderm or the anterior axial mesendoderm in Hesx1 mutant mice. They were, however, able to show defects within the anterior neural ectoderm. The expression domains of Six3 and Rax/Rx were significantly reduced within the forebrain of mutant mice. Otx1 expression was found within the medial region of the anterior neural plate; thus, Otx1 expression occurred across the midline in mutant mice. Martinez-Barbera et al. (2000 and 2001) therefore concluded that the forebrain defects observed in Hesx1 mutant mice are due to a defect within the anterior neural ectoderm.

The strain of Hesx1 mutant mice generated by Dattani et al. (1998) was the subject of a study by Dasen et al. (2001) concentrating on the role of Hesx1 in pituitary development. They described a pituitary phenotype in most homozygous mutant mice with a complete lack of
pituitary tissue in ~5% of homozygous mutant mice; the pituitary placode stage occurred but
development did not proceed beyond it. They note that in these cases, the overlying
diencephalon appeared normal suggesting that the lack of pituitary tissue was due to a defect
within the developing pituitary. Targeted expression of Prop-1 within the oral ectoderm
under control of the Ptx1 promoter resulting in early Prop-1 expression leads to a complete
failure of anterior pituitary development with no diencephalic defect, similar to the phenotype
arising as a result of the lack of Hesxl expression. Dasen et al. (2001) were able to show that
in these cases ectopic Fgf8 expression occurred within the pituitary. They went on to show
that targeted misexpression of Fgf8 within the oral ectoderm results in an absence of Hesxl
expression, demonstrating that Fgf8 expression extinguishes Hesxl expression.

In the majority of Hesxl mutant mice, anterior pituitary development appeared to proceed
normally until E12.5 (Dasen et al., 2001). At this stage, some mutant mice were
characterised by multiple invaginations of oral ectoderm forming multiple pituitary pouches,
presumably similar to the bifurcations seen by Dattani et al. (1998). In these mutants, ectopic
expression of Lhx3 and Prop-1 was found within the oral ectoderm suggesting that the area of
oral ectoderm recruited to form the anterior pituitary had expanded. Dasen et al. (2001) were
able to show that expanded Lhx3 expression corresponded to an expanded domain of FGF8
and FGF10 expression. Therefore, Dasen et al. (2001) conclude that Hesxl is required for
defining FGF8 expression boundaries. These findings suggest that Hesxl and FGF8 exert
reciprocal feedback regulation during pituitary development (Dasen et al., 2001).

Dasen et al. (2001) noted that the two mutant Hesxl pituitary phenotypes are independent of
each other as some mutant mice demonstrated multiple pituitary invaginations but no over
proliferation, whilst others demonstrated a single pituitary invagination with over
proliferation.

The presence of the eh-1 domain suggested that Hesxl is a transcriptional repressor and Xu et
al. (1998) suggest that the homeodomain may play a role in repression. They went onto show
that Hesxl binds a corepressor complex containing N-CoR/SMRT, mSin3A/B, and histone
deacetylases and that repression by the eh-1 domain and the homeodomain require this
complex.

Dasen et al. (2001) followed on from this work by demonstrating that Hesxl interacts with
the TLE-Groucho family of corepressors, with the interaction domain mapped to the eh-1
region. Expression studies demonstrated that TLE1 is expressed within the pituitary in a
strikingly similar expression domain to Hesxl. Using in-vitro transfection studies, Dasen et
al. (2001) were able to show that without TLE1, Hesxl is able to reduce Prop-1 expression
but can completely block Prop-1 mediated transcriptional activation. They were able to show
in-vivo interaction between Hesx1 and TLE1. Homozygous Hesx1 mutant mice are characterised by a large degree of cellular over-proliferation within the pituitary between E13.5 and E15.5 suggesting that the final stage of Hesx1 expression within the anterior pituitary is required to limit cellular proliferation (Dasen et al., 2001). Indeed, persistent expression of Hesx1 within the developing pituitary longer than normal resulted in a small reduction of some pituitary cell types. However, coexpression of TLE1 resulted in the near-complete absence of pituitary cell types. The double misexpression of Hesx1 and TLE1 also revealed aberrant anterior pituitary morphogenesis in ~30% of mutants, very similar to that observed Prop-1 in mutant mice (Dasen et al., 2001). TLE1 misexpression alone resulted in no phenotype.

1.3.3 Xanf-1 and Xanf-2.

Three sequences homologous to HESX1 have been identified in Xenopus, XANF-1 (Zaraisky et al., 1992), XANF-2 (Mathers et al., 1995), and a presumed XANF-1 pseudoallele (Zaraisky et al., 1992). XANF-1 and XANF-2 appear to have diverged from a common ancestor relatively recently as they are 89% identical, with a single amino acid substitution within the homeodomain. Thisidentity is uniform across the sequence, whereas identity to other Anf genes is concentrated in specific regions (Mathers et al., 1995). It is unclear whether the function of the two genes has diverged.

1.3.3.1 Xanf-1.

Zaraisky et al. (1995) used a combined RT-PCR and whole-mount in situ hybridisation approach to investigate XANF-1 expression. They first detected XANF-1 expression at a low level in most of the cells of the animal half at the late blastula stage (stage 9). By gastrulation, expression is limited to the presumptive anterior mesoderm of the dorsal blastopore lip, where expression is increased. This tissue corresponds to the Spemann organiser (Zaraisky et al., 1995). By stage 10.5-11, expression is detected within the presumptive anterior neurectoderm and by stage 11.5-12, expression is limited to this tissue. Embryos dissected at stage 12.5-13.5 demonstrated that XANF-1 expression occurred within both of the neurectoderm layers but was absent within the underlying mesoderm (Zaraisky et al., 1995). By stage 15, expression is limited to a crescent shaped area within the anterior neural ridge with two intensive strips of expression at the midline. These strips will give rise to the telencephalon and anterior pituitary, with the remaining XANF-1 expressing tissue
fated to contribute to the diencephalon. Expression is then limited to anterior pituitary anlage by stage 20-22, where expression continues until the tail bud stage.

Zaraisky et al. (1995) investigated the early role of XANF-1 by microinjection of exogenous XANF-1 mRNA into various blastomeres, at the 32-cell stage. They were able to show that injection into blastomeres C4 resulted in the formation of an incomplete secondary axis with exogenous XANF-1 mRNA containing cells concentrating at the anterior end of the secondary axis within the anterior most mesenchyme. They were also able to demonstrate that exogenous XANF-1 mRNA has the ability to produce secondary axis in ventral marginal zone explants. Zaraisky et al. (1995) noted that exogenous XANF-1 mRNA positive cells formed the anterior portion of the secondary axis, while exogenous XANF-1 negative cells contributed to the posterior portion. Zaraisky et al. (1995) were able to demonstrate that exogenous XANF-1 mRNA injection results in migration of surface cells of the dorsal marginal region into the embryo forming a deep furrow similar to the blastopore lip. They found that large amounts of exogenous XANF-1 mRNA resulted in a large amount of migration of cells giving rise to total erosion of the dorsal blastopore lip and failure of the closure of the blastopore ring. However, they found that small amounts of exogenous XANF-1 mRNA resulted in furrow formation and cellular migration but no obvious phenotype by the tadpole stage, even though the exogenous XANF-1 mRNA had led to a considerable reorganisation of the developmental fate map.

Ermakova et al. (1999) investigated the result of exogenous XANF-1 mRNA injection in more detail using a hormone inducible approach where induction of the injected exogenous XANF-1 mRNA was achieved by dexamethasone (DEX). When exogenous XANF-1 mRNA was induced before the midgastrula stage, they noted that the neural plate was expanded in 58% of cases and that this was the result of conversion of cells, from prospective neural crest and epidermis fates to a neural fate. Ermakova et al. (1999) found that the neural plate expansions developed into abnormal outgrowths within the CNS. However, this was combined with inhibition of the terminal differentiation of the primary neurones. Their results suggested that the expansion and inhibition phenotype is a cell-autonomous effect (Ermakova et al. 1999). Ermakova et al. (1999) were also able to demonstrate that exogenous XANF-1 mRNA induction downregulated expression of Otx2, BF-1 and XBF-1 during gastrulation and neurulation as well as inhibiting expression of endogenous XANF-1 expression. In a small number of cases, a cyclopic phenotype with medial fusion of the eyes was observed. These results were replicated when the transcriptional repressor domain of the Drosophila engrailed protein was fused to the XANF-1 protein but not when the transcriptional activator domain of herpes simplex virus, VP16, was fused to the XANF-1
Recently, Eroshkin et al. (2002) investigated the control of XANF-1 expression. They found three regions of homology between XANF-1, HESX1, and Ganf within the 5' flanking region (up to -2200bp) (see Figure 2.4.6.2a). Luciferase reporter constructs containing serial deletions of the 5' flanking region, microinjected into Xenopus embryos, demonstrated that these regions correspond to functional transcriptional regulators. Deletion of Box 1 (-510 to -380) decreases expression 6-fold. As Box 1 includes a binding site for bicoid-type homeodomains, this suggests that Otx-2, which is co-expressed with XANF-1, within the anterior neuroectoderm, may upregulate the expression of XANF-1. The transcription factors Dlx1 and Dlx5, which are both coexpressed with XANF-1 within the anterior margin of the neural plate, may also regulate XANF-1 expression. Deletion of Box 2 (-189 to -167) decreases expression 4-fold. This region contains core motifs for SOX-type DNA binding proteins, of which Sox-2, Sox-3, and Sox-D are coexpressed with XANF-1 during Xenopus development. Indeed, Sox-D overexpression within animal caps can activate XANF-1 expression (Mizuseki et al., 1998). Deletion of Box 3 (-200 to -189) leads to a similar 4-fold reduction of expression to that of Box 2. Deletion of Box 3 also leads to an expansion of the XANF-1 expression domain within the posterior brain, suggesting that Box 3 may be responsible for inhibition of XANF-1 expression posterior to the midbrain-hindbrain boundary. This apparent paradox may be explained by the location of a palindromic target site for both HOX-type homeodomain factors and paired-like homeodomain factors. As both groups contain enhancers as well as repressors, the cellular environment within which Box 3 is located may determine whether it will act as an enhancer or a repressor. As XANF-1 is itself a paired-like homeodomain factor, this site may be the site of inhibition responsible for the ability of microinjected exogenous XANF-1 to inhibit endogenous XANF-1 (Zaraisky et al., 1995). This suggests that XANF-1 is auto-regulatory. Deletion of both Box 2 and Box 3 totally inhibits XANF-1 expression, demonstrating that Box 1 alone is insufficient for maintenance of XANF-1 expression.

1.3.3.2 Xanf-2

XANF-2 expression appears to be similar, but not identical, to that of XANF-1 (Mathers et al., 1995 and Zaraisky et al., 1995). Mathers et al. (1995), first detected expression at stage 10 with continuous expression detected until stage 35. Expression was first observed in inner dorsal ectoderm, with expression detected in the anterior dorsal ectoderm by stage 13. Expression appeared to be restricted to the inner anterior ectoderm during neurulation.
During tadpole stages, XANF-2 expression occurs within a field posterior to the cement gland in the upper sensory plate (Mathers et al., 1995). This region will eventually give rise to the stomodeum, anterior pituitary, optic vesicles, and olfactory placodes. XANF-2 expression then progressively becomes limited to the midline within this region, marking the anlage of the anterior pituitary.

XANF-2 expression can be induced within uncommitted ectoderm using NH₄Cl (Mathers et al., 1995). The inner ectodermal layer is more receptive to induction than outer ectoderm (which expresses the cement gland marker XCG 13 in response to NH₄Cl), suggesting that the ectoderm layers are predisposed to their normal fates early in development (Mathers et al., 1995). Induction by NH₄Cl requires protein synthesis as induction is blocked by cyclohexamide treatment, suggesting that NH₄Cl activates Xanf-2 indirectly by activating a Xanf-2 activator rather than activating Xanf-2 directly (Mathers et al., 1995). Induction is reduced but not blocked by retinoic acid, suggesting that retinoic acid may play a role in normal Xanf-2 expression. Mathers et al. (1995) were also able to show that XANF-2 expression can be induced by follistatin, hedgehog, and noggin. On the other hand, they demonstrated that basic FGF is unable to induce XANF-2 expression and indeed reduces the level of NH₄Cl induced expression, suggesting that FGF signalling acts to downregulate Xanf-2 expression. The induced XANF-2 expression does not undergo the typical restriction of expression to the anterior pituitary. However, expression of the first anterior pituitary hormone prologue, POMC, does occur (Holling et al., 2000).

1.3.4 Other HESXI homologues.

Kazanskaya et al. (1997) and Dhawan et al. (1997) both independently identified a chicken HESXI homologue (Ganf/chaf). Dhawan et al., (1997) identified the gene during a screen for homeobox transcripts expressed during eye development. However, they were unable to detect actual Ganf expression. Kazanskaya et al. (1997) also identified the chicken homologue during a screen for members of the Anf class of homeobox genes. They also identified sturgeon (Aanf), zebrafish (Danf), newt (Panf), and human (Hanf) homologues.

1.3.5 The Anf class of homeobox genes.

Eight members of the Anf class of homeobox genes have been identified from seven species, whilst a further three sequences have been mentioned in the literature. Kazanskaya et al.
(1997) investigated the relationships between these genes. They surmised that in all likelihood these genes represent orthologues but were unable to rule out additional Anf class genes existing within the genome of these species. The genes appear to be related to the paired class, with the homeodomains showing a maximum of 54% similarity. However, the Anf class does not contain a region of homology comparable to the paired domain of the paired class. They noted that the Anf class of genes contain a region of homology near the N-terminus (Ft/sld/eh/sILGL), which shows similarity to the paired class octapeptide (HSIAGILG). However, it appears that this region actually represents a region of homology to the repressor domain of engrailed (Dattani et al., 1998). Indeed this domain (engrailed-homology domain 1, eh1) has been shown to function as a repressor (see section 5.4).

Kazanskaya et al. (1997) remarked on the low level of identity across the class, especially outside the homeodomain. Indeed, they were only able to identify two regions of homology across the class out with the homeodomain and its flanking regions. In addition to the eh1 region, the class contains a second region of homology near the N-terminus (p/hh/yRPW). However, it is unclear if this small region is a functional entity. Kazanskaya et al. (1997) note that the majority of homeodomain classes appear to diverge by less than 5% across vertebrate species, with a maximal 10% divergence within the NOT2 class. The Anf class however show a maximum of 25% between any two members. This divergence is exclusively limited to regions presumed not to contact DNA. Kazanskaya et al. (1997) propose that this finding suggests either that protein-protein interactions are of less importance within the Anf class than other classes of homeodomains, or that the Anf class has coevolved with a co-factor protein. Xu et al. 1998, have shown that the homeodomain of the murine Anf homologue is capable of interacting with the nuclear receptor co-repressor (N-CoR). It is unknown if this interaction is maintained across the Anf class. As N-CoR interacts with a number of other proteins, co-evolution with N-CoR is unlikely.

Functionally, the Anf class of homeodomain proteins appear to be transcriptional repressors (Dattani et al., 1998, Brickman et al., 2001 and Dasen et al., 2001). They are expressed within the extreme anterior region of the embryonic body axis during gastrulation and neurulation. Expression begins during early gastrulation, and reaches a maximum around the transition between gastrulation and neurulation (Kazanskaya et al., 1997). Early expression occurs in ectoderm and mesendodermal cells, primarily within the prechordal plate, with later anterior neural plate expression within the expression domain of Otx2 homologues (Kazanskaya et al., 1997). Kazanskaya et al. (1997) also note that in some members of the class, later expression becomes intense in two regions, the first fated to develop into the dorsal telencephalon and the anterior part of the dorsal diencephalon, including the posterior pituitary, and the second fated to develop into the anterior pituitary.
1.3.6 Other genes implicated in SOD.

In addition to the findings of Dattani et al. (1998) demonstrating the role of \textit{HESX1} in SOD, two further genes, \textit{PAX3} and \textit{cytochrome b}, have been implicated in the disorder. A single mutation has been described in both genes resulting in a single case of SOD with an additional syndrome known to be the result of mutations within that gene.

Waardenburg syndrome is an autosomal dominant disorder characterised by pigmentary and facial anomalies, deafness and a variety of other clinical traits (Carey et al., 1998). Over 50 \textit{PAX3} mutations have been characterised in Waardenburg type one traits (Carey et al., 1998). Carey et al. (1998) identified a novel \textit{PAX3} mutation within a four-generation family with Waardenburg syndrome type one. The proband of this family exhibited SOD in addition to Waardenburg syndrome type one, whilst his affected father exhibited Waardenburg syndrome type one only (Carey et al., 1998). However, the nature of the SOD phenotype is not mentioned. The Q391H mutation is the first to be described within exon 7 of the \textit{PAX3} gene and Carey et al. (1998) note that in addition to changing the amino acid sequence of PAX3 this mutation is likely to affect splicing of \textit{PAX3} mRNA. The significance of the finding is unclear as in a study comparing \textit{PAX6} and \textit{PAX3} expression \textit{PAX6} was expressed in the structures affected in SOD, whilst PAX3 was not (Terzic et al., 1999).

The mitochondrially encoded gene \textit{cytochrome b} is a component of the respiratory chain complex III. Mutations within \textit{cytochrome b} have been shown to result in a number of disorders characterised by exercise intolerance with occasional muscle weakness and/or myoglobinuria, Leber hereditary optic neuropathy, histiocytoid cardiomyopathy and encephalopathy. In addition, multisystem manifestations have been described, including deafness, mental retardation, retinitis pigmentosa, cataract, growth retardation, and epilepsy (Keightley et al. 2000, Wibrand et al. 2001, Legros et al. 2001, and Andreu et al. 1999). All known mutations are heteroplasminic and appear to have arisen sporadically (Schuelke et al., 2002). Schuelke et al., (2002) report a single case of exercise intolerance combined with SOD. A non-conservative Ser35Pro mutation at a highly conserved residue was shown to result in an expected increased reactive oxygen species levels. Schuelke et al., (2002) note that other genetic defects shown to result in increased reactive oxygen species levels have been implicated in brain malformations, as have other mitochondriopathies. It is unclear whither the SOD phenotype reflects the normal multisystem manifestation of mitochondriopathies which in this case includes SOD or if the SOD phenotype represents a distinct phenotypic consequence of the mitochondriopathy. Likewise, Schuelke et al., (2002)
were unable to conclude whither the SOD phenotype was the direct result of Ser35Pro mutation or secondary to the increased reactive oxygen species levels.

Deletion of the genetic locus of SIX6 has been described in patients with anophthalmia, absence of the optic nerve and chiasma as well as pituitary defects (Gallardo et al., 1999). Recently, a Six6 null mutant mouse has been generated. This mouse manifests a phenotype characterised by pituitary hypoplasia and retinal hypoplasia often leading to optic nerve hypoplasia (Li et al., 2002). Although the pituitary is hypoplastic, it appears to function normally with all endocrine cell types apparent. The Six6 null mutant mouse shows a reduction in cellular proliferation of the affected structures. Although the phenotypes of the null mutant mouse and human hemizygous patients represent a more widespread phenotype than that observed in SOD, milder mutations could in theory give rise to some cases of optic nerve hypoplasia combined with pituitary defects. However, there is no evidence to date for a role of SIX6 in the development of the septum pellucidum and other forebrain structures.

1.4 Aims.

At the beginning of this project, a small-scale study had identified two siblings with the condition septo-optic dysplasia (SOD), both of whom harboured a single causative mutation in the gene HESXI. Additionally, a second pair of siblings were suspected of carrying a second heterozygous causative mutation within HESXI (Dattani et al., 1998). Therefore, the overall aim of this project was to further investigate the role HESXI in septo-optic dysplasia and related conditions, and in particular to investigate whether it is a major aetiological factor in sporadic SOD.

The individual aims of the project were:

1. To collect a large cohort of patients with septo-optic dysplasia and its milder variants.

2. To screen the cohort with a view to identifying HESXI mutations.

3. To investigate the functional significance of any mutations.
2 \textit{HESX1} mutation screening.

2.1 Introduction.

In order to elucidate the total contribution of the \textit{HESX1} locus to septo-optic dysplasia, the genetic and epigenetic status of the \textit{HESX1} locus must first be determined in all affected individuals. Although techniques are available to detect a wide variety of genetic and epigenetic mutations, an exhaustive cataloguing of \textit{HESX1} status was outside the scope of this study. Therefore, a sampling approach was undertaken. In essence, a sample group of patients from the population at large was obtained and a screen for a subset of DNA variants applied. The mutation screen was designed to balance the need to extract the maximum amount of data with the resources available. The mutation screen did not have the purpose of identifying all mutations within the affected group, neither was it designed to identify the level of a particular mutation within the group.

2.2 Mutation screening design and methodology.

The first step in design of this study was that of patient recruitment. Subsequently the type of mutation to be screened for was selected. The regions of DNA to be screened were then chosen. The methods used to detect sequence variance were decided. Likewise, the method of DNA sequencing was chosen. Finally, all of these had to be combined into a robust yet achievable and affordable mutation detection strategy.

2.2.1 Patient recruitment.

Dr M. Dattani, a Senior Lecturer in Paediatric Endocrinology, carried out all patient recruitment. Affected individuals were recruited from the affected population at large, and their genetic material was then subjected to a mutation screen of \textit{HESX1}. Given the phenotype of the index \textit{HESX1} SOPD family it was decided not only to recruit SOPD patients, but also to include patients with defects of the corpus callosum. Given that the full phenotype associated with mutations in \textit{HESX1} is unknown, patients with features outside the classical triad were not excluded. Likewise, patients who displayed only single features of the classical triad were also recruited. Thus, a broad based approach was undertaken. The
first group of patients recruited were those used by Dattani et al (1998) to discover the ARG160CYS mutation (referred to hence forth as the Dattani subset). Patients were recruited from within Great Ormond Street Children’s Hospital (GOSH) and the Middlesex Hospital forming the London Centre for Paediatric Endocrinology (LCPE), and from other national and international centres. Ethical committee approval was obtained from both GOSH and ICH. Centres were asked to contribute DNA samples, or 2x 5-10ml EDTA blood samples, from affected individuals and any family members available. The GOSH/ICH DNA laboratory processed blood samples and DNA was produced in 250 mg/ml aliquots of varying volumes. However, a number of samples arrived as DNA of variable concentration.

All available clinical data regarding each recruited patient was entered into a Microsoft Access database and in accordance to the data protection act (1998) each sample given a unique identifiable number by Dr Dattani and an anonymised working database was created. The database was used both to track the progress of HESX1 mutation screening and to calculate the number of patients in each clinical sub-group.

2.2.2 Mutation selection.

The second step in the design of this study was to focus on the type of mutation that would be screened for. In the case of HESX1 there is no knowledge or indeed suggestion of epigenetic effects. Therefore, the search for such effects was not considered. Similarly, long-range control elements dictating HESX1 expression have not been demonstrated as yet and therefore were not considered. Neither is there any evidence of genomic deletions of HESX1 or rearrangements of HESX1 as mutational mechanisms within SOD. Indeed Dattani et al. (1998) found no evidence of genomic reorganisation in 38 patients using Southern-blot analysis of restricted genomic DNA. However, it is clear that such mutations are not uncommon in genetic disorders and in all likelihood such genomic rearrangements will have a role in the contribution of HESX1 to SOD given the recessive loss-of-function nature of the ARG160CYS mutation (Dattani et al., 1998). Nevertheless, the search for such rearrangements was not undertaken on the grounds of feasibility, as the nature of samples was not expected to be suitable for techniques such as FISH and the expected amount of sample DNA too low for southern-blot studies. Thus, the study focused on point mutations and small genomic rearrangements.
2.2.3 Regions to be screened.

The third step in the design of the mutation screen was to determine the exact regions of the genome to be studied. Initially, the aims of the study were to examine only coding regions (exons 1-4) of the *HESX1* gene as undertaken by Dattani *et al.* (1998). As work progressed cis-acting control elements were identified in the murine *Hesx1* gene and the *Xenopus Xanf-1* gene. The sequence of these cis-acting control elements was compared with the genomic sequence of human *HESX1* and their homologues identified (see 2.4.6.2).

Furthermore, investigation into the identity of the "alternative exon 1" sequence found in EST GenBank gi:1720505 (Dattani *et al.*, 1998) discovered that the sequence was encoded by 2 putative exons found within genomic clone RP11-241K3 (GenBank gi:7798783) along with *HESX1* genomic sequences (see 2.4.6.1). These putative exons were included in the study.

2.2.4 Variance identification.

The next step was to decide upon the DNA sequence variance detection techniques to be used. The most robust technique is to elucidate the exact sequence of the DNA in question. This process is labour-intensive and expensive; thus only suitable for small numbers of samples and therefore only suitable for the familial cases. A number of techniques exist for DNA sequencing. As direct fluorescent cycle sequencing allows rapid sequencing of PCR products, this technique was chosen as the DNA sequencing method. In the main the Big Dye terminator sequencing kit (ABI) was used and the reactions were visualised using an ABI 377 DNA sequencer (Perkin-Elmer).

For the large number of sporadic samples, DNA sequencing would have been prohibitively expensive and therefore techniques which enable the identification of DNA fragments from individual samples that are likely to contain sequence variance, as well as, the identification of those unlikely to contain sequence variance were required. There are a number of such techniques. Dattani *et al.* (1998) had successfully used Single Strand Conformational Polymorphism detection (SSCP detection) to detect the index ARG160CYS mutation. It was therefore decided to use SSCP detection as the variance detection technique. Dattani *et al.* (1998) examined DNA from 20 patients with SOPD and 35 SOPD-PIT patients using the CleanGel system (APBiotech). As this system was available this was selected as the system for SSCP detection. SSCP detection does not detect all sequence variance; although a high sensitivity can be achieved by incorporating increased running temperatures and varying the
constituents of the polyacrylamide gel (Jaeckel et al., 1998). During the course of this study a new technology for variance detection became available, the dHPLC-heteroduplex detection based WAVE system (Transgenomics Inc). The WAVE system was incorporated into variance detection in order to increase the sensitivity of variance detection.

2.2.5 Causative mutation identification.

Once the DNA sequence of familial cases and suggestive sporadic cases was obtained it was compared with wild-type sequence and differences noted. Polymorphic status was given if the variant was present in control samples. Function was initially assessed based on the sequence of GenBank gi:3258660 (Dattani et al., 1998) and once available on the sequence of clone RP11-241K3 GenBank gi:7798783 (HuGO). In cases of variance giving rise to protein sequence variation, a homologue sequence comparison approach was undertaken. Those mutations suggestive of being causative where then investigated to identify whither they led to a variation in DNA binding, using EMSA, and/or repression, using in-vitro transfection assays.

2.2.6 Overview.

The overall mutation screening strategy adopted is shown in Figure 2.2.6a. Firstly, a DNA amplicon derived from the sequence of interest was generated using PCR with sequence-specific primers. Amplicons derived from sporadic cases were then subjected to either SSCP detection or dHPLC heteroduplex detection. Any amplicons suggestive of sequence variance were re-generated and sequenced. Amplicons not suggestive of sequence variance using the first technique were subjected to the second technique. Again, any amplicons suggestive of sequence variance were re-generated and sequenced. Samples which did not provide data suggestive of sequence variance were classified as clear of sequence variance.

Amplicons from familial cases were directly sequenced without being subjected to either technique. Only a single member of an affected family was sequenced.
2.3 Methods.

2.3.1 STS primer design.

Initially, identical oligonucleotides to those used by Dattani et al. (1998) were made to order by a commercial oligonucleotide manufacturer (MWG) for PCR amplification (Table 2.3.1a).

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence 5'-3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>DHX1F</td>
<td>5'−AGC TGT TGC TCT GTG CAG ACC ACG−3'</td>
</tr>
<tr>
<td>DHX1R</td>
<td>5'−ACA AAG AAT TGA AAC AAT TAA GCT GCA−3'</td>
</tr>
</tbody>
</table>
In the case of DHX3F, a sequence discrepancy between the Human Genome Project Organisation's published sequence of the HESX1 containing clone RP11-241K3 (GenBank gi: 7798783) and the sequence of the HESX1 genomic region published by Dattani et al. (1998) was noted after the oligonucleotide had been in use for some time. In response to this an investigation into the level of this variant in a control population was undertaken (see section 2.4.6.3).

As more regions of HESX1 were included in the analysis, oligonucleotides were designed for use in PCR amplification of these regions of HESX1. In all cases the sequence of clone RP11-241K3 (GenBank gi:7798783) published by the Human Genome Project Organisation was used as the basis of the oligonucleotide design. The Primer3 software located at http://www-genome.wi.mit.edu/cgi-bin/primer/primer3 www.cgi was utilised in the design of the oligonucleotide primers. The sequence of interest was processed using the default settings. If the output was not satisfactory the sequence was re-analysed using adjusted settings in a step-wise manner. In general, this process provided a number of primer pairs in rank order of suitability. The oligonucleotide sequences were then checked manually. Any that were not suitable were discarded and the next primer pair in rank order were manually checked for suitability. This process produced all subsequent oligonucleotides used in STS PCR (Table 2.3.2b).

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence 5'-3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCDNAIF</td>
<td>5' - CAA TTG AGA GAA TCT TAG GAC TGG A -3'</td>
</tr>
<tr>
<td>HCDNAIR</td>
<td>5' - TCC AGT CCT AAG ATT CTC TCA ATT G -3'</td>
</tr>
<tr>
<td>HCDNA3F</td>
<td>5' - TGC TAT CCT GGT ATC GAT ATT AGA GA -3'</td>
</tr>
<tr>
<td>HCDNA3R</td>
<td>5' - TCT CTA ATA TCG ATA CCA GGA TAG CA -3'</td>
</tr>
<tr>
<td>HESX1ALTX</td>
<td>5' - TGC ACC CCA GCT CTA AAG AA -3'</td>
</tr>
<tr>
<td>HESX1-5'UTR</td>
<td>5' - TCT AAA TCC AGC CCC CTT G -3'</td>
</tr>
<tr>
<td>HESX1-3'UTR</td>
<td>5' - CTG ATT CTT CAT GCT CTG CA -3'</td>
</tr>
<tr>
<td>H5'F</td>
<td>5' - GCC ACA TTT GTG CAT CAG TT -3'</td>
</tr>
<tr>
<td>H5'R</td>
<td>5' - CTC TGC CCC AGC TGT ATA -3'</td>
</tr>
<tr>
<td>VF16 5'</td>
<td>5' - TCG GAA TCC ACC ATG GTG TCT CCC AGC CTT CAG GAA -3'</td>
</tr>
<tr>
<td>VF16 3'</td>
<td>5' - CGC GGG ATC CTT CCA GCA GAT TTG TGT T -3'</td>
</tr>
</tbody>
</table>

(Table 2.3.1a). Sequence of oligonucleotide primers identical to Dattani et al. (1998).
2.3.2 Polymerase Chain Reaction.

The Polymerase Chain Reaction (PCR) is a technique allowing the amplification of DNA sequences without the need to clone such sequences. Details of the standard molecular biological technique are widespread and will not be discussed in full. In brief, PCR relies on the ability of a Thermostable DNA polymerase to polymerise dNTP's into DNA using a DNA template and a sequence specific oligonucleotide primer. PCR was carried out under two separate conditions. First, during phase 1 (before availability of the WAVE apparatus) PCR was carried out using both DMSO and oil, and second during phase 2 (after the availability of the WAVE apparatus) PCR was carried out without DMSO or oil.

2.3.2.1 Initial PCR conditions.

Initially, reactions were performed as reported by Dattani et al., (1998) (Dattani, personal communication), including oligonucleotide primers (Figure 2.3.2.1a) but with the use of an Omni-E thermocycler and replacement of doubly distilled H₂O with mQ H₂O (Millipore). A master mix batch size of 11 was used (Table 2.3.2.1a). Master mix, DNA, and mQ H₂O was then aliquoted into each of 10 200μL Eppendorf tubes and overlain with 1 drop of mineral oil (Table 2.3.2.1b). Thermocycling was carried out as reported by Dattani et al. (1998). A single denaturation cycle of 94°C for 120 seconds was followed by 35 cycles of 94°C for 20 seconds, the specific annealing temperature (Table 2.3.2.1c) for 30 seconds and 72°C for 60 seconds. Finally, a single cycle of 72°C for 300 seconds was carried out. Products were separated by Agarose gel electrophoresis and visualised by EtBr staining.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>E149KRev</td>
<td>5'- CTG GAT TCT GTA TTT CTC TAG ATT CAA TTT -3'</td>
<td>-3'</td>
</tr>
<tr>
<td>E149KForw</td>
<td>5'- AAA TTG AAT CTA GAG AAA GAC AGA ATC CAG -3'</td>
<td>-3'</td>
</tr>
<tr>
<td>E149KTest</td>
<td>5'- CTA AAT TTG AAA ATT ACC AGG ATT CTG -3'</td>
<td>-3'</td>
</tr>
<tr>
<td>HINF1INS</td>
<td>5'- GGA GGA TAG CAG TTT ACT CTA AAG ACT -3'</td>
<td>-3'</td>
</tr>
<tr>
<td>HESX1HEX</td>
<td>5'- GAA GCC TTA TTA ATG AAA CAG TAG ACT ACC -3'</td>
<td>-3'</td>
</tr>
<tr>
<td>M13REV</td>
<td>5'- CAG GAA ACA GCT ATG AC -3'</td>
<td>-3'</td>
</tr>
<tr>
<td>NOVAT7T</td>
<td>5'- GCT ACT TAT TGC TCA GCG G -3'</td>
<td>-3'</td>
</tr>
<tr>
<td>NOVAT7P</td>
<td>5'- TAA TAC GAC TCA CTA TAG GG -3'</td>
<td>-3'</td>
</tr>
</tbody>
</table>

(Table 2.3.2b) Sequence of all oligonucleotides created.
(Figure 2.3.2.1a) Location of primers in relation to the HESXI genomic structure.

<table>
<thead>
<tr>
<th>Primer pair</th>
<th>DHX1F</th>
<th>DHX1R</th>
<th>DHX2F</th>
<th>DHX2R</th>
<th>DHX3F</th>
<th>DHX3R</th>
<th>DHX4F</th>
<th>DHX4R</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR I buffer (Promega)</td>
<td>0</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>0</td>
<td>22</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>PCR II buffer (Promega)</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>22</td>
<td>22</td>
<td>0</td>
<td>17.6</td>
<td>22</td>
</tr>
<tr>
<td>2mM dNTP's</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>25mM MgCl2</td>
<td>8.8</td>
<td>8.8</td>
<td>0</td>
<td>17.6</td>
<td>8.8</td>
<td>8.8</td>
<td>0</td>
<td>17.6</td>
</tr>
<tr>
<td>50ng/μL 1:1 mixture of primers</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>mQ H₂O</td>
<td>4.4</td>
<td>4.4</td>
<td>13.2</td>
<td>1.1</td>
<td>4.4</td>
<td>4.4</td>
<td>13.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Taq polymerase (Promega)</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
</tr>
</tbody>
</table>

(Table 2.3.2.1a) Master mix constituents for each primer pair. All volumes in μL.

<table>
<thead>
<tr>
<th>Primer pair</th>
<th>DHX1F</th>
<th>DHX1R</th>
<th>DHX2F</th>
<th>DHX2R</th>
<th>DHX3F</th>
<th>DHX3R</th>
<th>DHX4F</th>
<th>DHX4R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master mix</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mQ H₂O</td>
<td>12.5</td>
<td>12.5</td>
<td>12.5</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Table 2.3.2.1b) Volume of master mix, DNA, and mQ H₂O added to each tube. All volumes are in μL.

<table>
<thead>
<tr>
<th>Primer pair</th>
<th>DHX1F</th>
<th>DHX1R</th>
<th>DHX2F</th>
<th>DHX2R</th>
<th>DHX3F</th>
<th>DHX3R</th>
<th>DHX4F</th>
<th>DHX4R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annealing temperature</td>
<td>65</td>
<td>55</td>
<td>60</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Table 2.3.2.1c) Annealing temperatures. All temperatures are in degrees centigrade.

The results of the initial PCR amplification of sample DNA for mutation detection using the conditions of Dattani et al. (1998) are shown in figures 2.3.2.1b, 2.3.2.1c, 2.4.2.1d and 2.3.2.1e. Bands of expected size were produced by the initial PCR amplification, which were confirmed by direct cycle sequencing.
(Figure 2.3.2.1b) Initial PCR amplification using DHX1F-DHX1R. The expected product size is indicated by the arrow.

(Figure 2.3.2.1c) Initial PCR amplification using DHX2F-DHX2R. The expected product size is indicated by the arrow.

(Figure 2.3.2.1d) Initial PCR amplification using DHX3F-DHX3R. The expected product size is indicated by the arrow.
2.3.2.2 Optimisation of PCR conditions for phase 1 of mutation screening.

It was apparent that the uniformity of reactions across a batch of samples was poor, as was uniformity between batches. There are a number of causes of the poor uniformity. First, the use of 50ng/μL 1:1 mix of primers does not produce a 1:1 molar ratio of primers. Second, a high concentration of template DNA was used. Third, the use of a 1min extension time is greatly in excess of what is required for the amplification of fragments in this size range. Finally, it would be expected that the optimal conditions used by Dattani et al (1998) would be sub-optimal for use on a different thermocycler, namely the Omni-E. The conditions reported by Dattani et al., (1998) were therefore not used to amplify DNA for mutation screening. Therefore, optimisation of the above reactions was performed bearing the above limitations in mind. Primer usage was converted to a 1:1 molar ratio with each primer being added individually to a final concentration of 1μM. The volume of genomic template DNA was reduced to 0.3μL per reaction. In addition Biopro Taq polymerase (Bioline) was used. The reaction conditions were optimised for the Omni-E thermocycler.

For each primer pair an optimisation reaction was carried out using MgCl₂ concentrations ranging from 0.5mM to 3.5.mM at the annealing temperature reported by Dattani et al. (1998). The optimisation reactions were then repeated reducing the annealing temperature in 2°C steps until satisfactory results were obtained. Products were separated by Agarose gel electrophoresis and visualised by EtBr staining (Figure 2.3.2.2a).
Based upon the results of these reactions, an annealing temperature of 52°C and a MgCl₂ concentration of 3.0mM were used for DHX2F-DHX2R, an annealing temperature of 58°C and a MgCl₂ concentration of 1.5mM for DHX3F-DHX3R and an annealing temperature of 56°C and a MgCl₂ concentration of 1.5mM for DHX4F-DHX4R. Satisfactory results were not obtained for DHX1F-DHX1R. The DHX1F-DHX1R reaction was then repeated with the addition of the PCR additive DMSO. Satisfactory results were obtained at an annealing temperature of 55°C at a MgCl₂ concentration of 1.5mM.

### 2.3.2.3 PCR conditions for phase 1 of mutation screening.

The optimised reactions were then scaled to a batch size of 10. These optimised conditions were used during phase 1 of mutation screening (Tables 2.3.2.3a and 2.3.2.3b). Each reaction was overlain with 1 drop of mineral oil and the thermocycling carried out within an Omni-E thermocycler using thermocycling conditions identical to those used in the optimisation.

<table>
<thead>
<tr>
<th>Primer pair</th>
<th>DHX1F DHX1R</th>
<th>DHX2F DHX2R</th>
<th>DHX3F DHX3R</th>
<th>DHX4F DHX4R</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x PCR buffer (Bioline)</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>2mM dNTP’s</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>25mM MgCl₂ (Bioline)</td>
<td>13.2</td>
<td>26.4</td>
<td>13.2</td>
<td>13.2</td>
</tr>
<tr>
<td>10µM forward primer</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>10µM reverse primer</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>mQ H₂O</td>
<td>112.2</td>
<td>100.1</td>
<td>113.2</td>
<td>113.2</td>
</tr>
<tr>
<td>DMSO</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BioPro Taq polymerase (Bioline)</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
</tr>
</tbody>
</table>

(Table 2.3.2.3a) Master mix constituents for each primer pair. All volumes in µL.
2.3.2.4 Optimisation of PCR conditions for phase 2 of mutation screening.

The use of dHPLC heteroduplex analysis using the WAVE system has a number of advantages, but has the major disadvantage that the WAVE system is intolerant of a wide array of chemicals normally associated with PCR (Transgenomics Inc). Therefore, the PCR conditions used during phase 1 are unsuitable for use with the WAVE system as the reaction buffer supplied with the polymerase enzyme contains detergents that damage the column (Transgenomics Inc.). A new set of PCR conditions for amplification of the regions of \textit{HESXI} to be screened for mutation had to be devised. In addition, the use of oil as a vapour seal during thermocycling is also incompatible with the WAVE machine. Combined with the risk of experimental failure caused by air entering the WAVE machine due to low reaction volume resulting from excessive vapour loss during thermocycling, the need for a thermocycler fitted with a hot lid used in conjunction with consumables designed to limit vapour loss was paramount. This led to the use of a new thermocycler in phase 2, the Eppendorf Master gradient thermocycler. This switch in PCR conditions coincided with a rapid increase in the rate of sample acquisition with a consequential need for a larger throughput where required.

A similar optimisation strategy was devised to that used in phase 1, adapted for use with the HotStar Taq polymerase (Qiagen) following manufacturer’s guidelines. A single optimisation strategy was sufficient for all amplicons. Each reaction was carried out either with or without Q solution (Qiagen) (replacing DMSO which reduces column life (Transgenomics)) and utilised the gradient ability of the Eppendorf mastercycler gradient thermocycler (Figure 2.3.2.4a).

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
Primer pair & DHX1F & DHX1R & DHX2F & DHX2R \\
\hline
Master mix & 19.7 & 19.7 & 19.7 & 19.7 \\
DNA & 0.3 & 0.3 & 0.3 & 0.3 \\
\hline
\end{tabular}
\caption{Volume of master mix and DNA to each tube. All volumes are in µL.}
\end{table}

(Figure 2.3.2.4a) Example of PCR optimisation for phase 2 of mutation screening showing H5’F-H5’R across a temperature gradient of 50.0°C to 62.5°C. The expected product size is indicated by the arrow.
Additionally, the two putative alternative exons were included in the mutation screen as was the 5’ LIM1 homology domain, using the oligonucleotide primer pairs, HAXAF-HAXAR, HAXBF-HAXBR, and H5’F-H5’R (Figure 2.3.2.4b).

(Figure 2.3.2.4b) Location of oligonucleotides primers for putative alternative exons and 5’ LIM1 homology region in relation to the genomic structure of HESX1.

2.3.2.5 PCR conditions for phase 2 of mutation screening.

The optimised reactions were scaled to a batch size of 60. These optimised conditions were used during phase 2 of mutation screening (Tables 2.3.2.5a and 2.3.2.5b). Thermocycling was carried out within an Eppendorf mastercycler gradient using a single thermocycling profile (Table 2.3.2.5c). A representative example of PCR amplification using the least robust oligonucleotide pair of DHX3F-DHX4R is shown in Figure 2.3.2.5a.

<table>
<thead>
<tr>
<th>Primer pair</th>
<th>HAXAF</th>
<th>HAXBF</th>
<th>H5’F</th>
<th>H5’R</th>
<th>DHX1F</th>
<th>DHX1R</th>
<th>DHX2F</th>
<th>DHX2R</th>
<th>DHX3F</th>
<th>DHX3R</th>
<th>DHX4F</th>
<th>DHX4R</th>
<th>DHX3F</th>
<th>DHX4R</th>
</tr>
</thead>
<tbody>
<tr>
<td>HotStar buffer (Qiagen)</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
</tr>
<tr>
<td>2mM dNTP’s</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
</tr>
<tr>
<td>25mM MgCl₂ (Qiagen)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10μM forward</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
</tr>
<tr>
<td>10μM reverse</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
</tr>
<tr>
<td>mQ H₂O</td>
<td>494</td>
<td>494</td>
<td>754</td>
<td>754</td>
<td>754</td>
<td>754</td>
<td>754</td>
<td>494</td>
<td>494</td>
<td>468</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q solution (Qiagen)</td>
<td>260</td>
<td>260</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>260</td>
<td>260</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HotStar Taq polymerase (Qiagen)</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Table 2.3.2.5a) Master mix constituents for each primer pair. All volumes in μL.

<table>
<thead>
<tr>
<th>Primer pair</th>
<th>HAXAF</th>
<th>HAXBF</th>
<th>H5’F</th>
<th>H5’R</th>
<th>DHX1F</th>
<th>DHX1R</th>
<th>DHX2F</th>
<th>DHX2R</th>
<th>DHX3F</th>
<th>DHX3R</th>
<th>DHX4F</th>
<th>DHX4R</th>
<th>DHX3F</th>
<th>DHX4R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master mix</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
</tr>
<tr>
<td>DNA</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Table 2.3.2.5b) Volume of master mix and DNA added to each well. All volumes are in μL.
<table>
<thead>
<tr>
<th>Stage</th>
<th>Temperature</th>
<th>Time</th>
<th>Number of cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denaturation.</td>
<td>95°C</td>
<td>900s</td>
<td>1</td>
</tr>
<tr>
<td>Denaturation.</td>
<td>94°C</td>
<td>30s</td>
<td></td>
</tr>
<tr>
<td>Annealing.</td>
<td>56°C</td>
<td>30s</td>
<td>35</td>
</tr>
<tr>
<td>Extension.</td>
<td>72°C</td>
<td>30s</td>
<td></td>
</tr>
<tr>
<td>Extension.</td>
<td>72°C</td>
<td>600s</td>
<td>1</td>
</tr>
</tbody>
</table>

(Table 2.3.2.5c) Thermocycling profile.

(Figure 2.3.2.5a) An example of PCR amplification from genomic DNA samples (n=19) under optimised conditions using DHX3F-DHX4R. The expected product size is indicated by the arrow.

2.3.2.6 Summary.

The PCR amplification of DNA sequences was performed in two ways. Firstly, each exon of each sample was amplified in batches of 10. The reactions were carried out in 200uL Eppendorf tubes and thermocycling performed using an Omni-E thermocycler equipped with a hot lid. An attempt to use the conditions used by Dattani et al., (1998) was unsuccessful and the conditions for PCR amplification had to be re-optimised. However, the rate of patient recruitment increased and it became apparent that this strategy was not capable of dealing with the rate at which patients were recruited. PCR conditions also required re-optimisation when dHPLC heteroduplex analysis using the WAVE system (Transgenomics) became available, this being unable to function in the presence of many common PCR components. Therefore, an Eppendorf Master Gradient thermocycler equipped with a hot lid was obtained and used. The batch size was increased to 60 and the reactions carried out in a 96 well PCR plate. The reaction components used in PCR amplification had to be changed to remove any detergents, resulting in the use of the HotStar Taq polymerase. Both these conditions for PCR amplification required optimisation so that efficient mutation screening could be achieved.
2.3.3 SSCP detection.

Under non-denaturing conditions an individual strand of DNA will adopt a three dimensional conformation, a single strand conformation, determined by the primary sequence of the DNA. Thus, different primary sequences will adopt different single strand conformations. A single strand conformation is a biological trait and therefore when variant single strand conformations are evident within a biological population they are known as a single strand conformational polymorphisms (SSCPs). SSCP s are therefore due to primary sequence polymorphisms (Orita et al., 1989).

The detection of an SSCP should lead to the identification of a primary sequence polymorphism. However, SSCP detection can be variable due to two distinct problems (Jordanova et al., 1997). Firstly, the SSC adopted by a variant DNA strand may not sufficiently differ from that of a WT DNA strand to be detected. This can be overcome to a great extent by varying the temperature at which SSCP detection is carried out. Nonetheless, a number of SSCP s will remain undetectable. The second problem is that of duplex formation, lowering the signal to noise ratio. This can be overcome by denaturation of the DNA followed by rapid cooling, favouring intra-strand hydrogen bond formation over inter-strand hydrogen bonding. These problems are compounded by a significant false positive rate in SSCP detection.

There are a number of methods available for the detection of SSCP s. Previously, slab based gel electrophoresis had been used to detect HESXI SSCP's (Dattani et al., 1998). The system used for performing this method of detection was the CleanGel system (APBiotech). As this system had been used successfully and the equipment was available, this system was therefore used for SSCP detection.

2.3.3.1 The CleanGel system.

The clean gel system is a slab based polyacrylamide electrophoresis (PAGE) system using pre-cast gels on a plastic backing. It is designed for use with the Multiphor II Electrophoresis Unit (APBiotech) and the corresponding One Plus Silver Staining Kit (APBiotech). To maximise throughput, the 48 well pre-cast gel format was used. The pre-cast gel slabs combined with the single kit system have the benefit of speed and reliability over more labour intensive systems. However, the pre-cast nature of the gel prohibits the adjustment of gel constituents to improve detection efficiency. The system also requires a particular electrophoresis and cooling set-up, and this was available within the research group.
2.3.3.2 Use of the clean gel system during phase 1 of mutation detection.

Prior to beginning SSCP detection, the denaturation buffer (F reagent) was prepared using a 7:5 ratio of SSCP solutions A and B. F reagent was stored at -18°C for a maximum of 12 months.

SSCP solution A:

<table>
<thead>
<tr>
<th>Volume</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>5ml</td>
<td>Formamide</td>
</tr>
<tr>
<td>100µL</td>
<td>500mM EDTA</td>
</tr>
<tr>
<td>250µL</td>
<td>1% Bromophenol blue</td>
</tr>
<tr>
<td>250µL</td>
<td>1% Xylene cyanol</td>
</tr>
</tbody>
</table>

SSCP solution B:

<table>
<thead>
<tr>
<th>Volume</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>50µL</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>100µL</td>
<td>500mM EDTA</td>
</tr>
<tr>
<td>4.3mL</td>
<td>mQ H₂O</td>
</tr>
</tbody>
</table>

PCR amplified DNA samples from a number PCR reactions, using a given primer pair, were pooled into groups of 47. To each batch of 47 samples, a wild type control was added to produce a batch size of 48, for use in SSCP detection. 1-3µL of each sample was aliquoted into 500µL Eppendorf tubes, depending on signal strength, and stored for analysis at -18°C.

The use of the clean gel system was in line with manufacturer's instructions (APBiotech). Firstly, the pre-cast gel was rehydrated. 25ml gel buffer was added to the chamber of the “GelPool” and the pre-cast gel lain upon it. The “GelPool” was then placed upon a rotating shaker for 60mins.

During the 60mins of pre-cast gel rehydration the experimental apparatus was set up and the samples prepared. First, the Multiphor II electrophoresis unit was connected to the MultiTemp thermostatic circulator and the temperature set to 4°C and left to attain working temperature. Secondly, F reagent was added to all samples to a final volume of 7µL.

Once rehydration was complete, 1ml of light mineral oil was pipetted onto the Multiphor II electrophoresis unit and the gel was placed upon it, so as to form an even layer of oil under the gel. Two electrode strips were then placed in the “PaperPool” and 22ml of electrode...
buffer added to each. The electrode strips were then added to the Multiphor II electrophoresis unit and the electrodes lowered into position.

All samples were then denatured at 98°C for 2mins before snap cooling in ice and then loaded into the gel. Electrophoresis was carried out at a voltage limit of 200V (23mA, 5W) for 10mins followed by a limit of 600V (30mA, 18W) until the Bromophenol blue reached the anode.

Silver staining was carried out using the One Plus Silver Staining Kit as per manufacturer’s instructions (APBiotech). First, the gel was fixed in 100mL 24% Ethanol + 25mL 5x Fixing solution for 30mins. Secondly, the gel was stained using 100mL mQ H₂O + 25mL Staining solution for 30mins, followed by a brief wash in 100ml mQ H₂O. The gel was then developed using 100ml mQ H₂O + 25ml Sodium Carbonate + 125μL Thiosulphate + 125μL Formaldehyde. Finally, the gel was added to 100mL mQ H₂O + 25mL Stop and Preserve solution for at least 1 hour.

The results of the SSCP detection experiment were then viewed and recorded. The gel was then placed on a sheet of A4 paper, gel side down. The gel was then placed plastic side down and a number of paper towels placed on top of the paper and a weight applied. The gel was then allowed to dry and the paper-backed gel was stored.

2.3.3.3 Use of the clean gel system during phase 2 of mutation detection.

The clean gel SSCP detection system was used exactly as above except that a 96-well PCR plate format was adopted; with one half of each 96-well PCR plate being used at a time.

2.3.4 dHPLC detection of heteroduplex using the WAVE system.

2.3.4.1 Heteroduplex detection.

In simple terms, DNA duplex consists of two complementary strands of DNA, held together by hydrogen bonds forming a double helix. However, DNA duplex can also be formed from two strands of DNA sharing less than 100% complementary sequence. In this case the region of sequence divergence will not form a double helix, but instead, a "bubble" of non-
complementary DNA will be formed. This arrangement is known as heteroduplex, as opposed to 100% complementary homoduplex. In reality, this view is simplistic. In-vitro, DNA duplex can form a number of basic stable formations, including A form, B form, and Z form. Most duplex appears to be a hybrid of these forms. On a smaller scale, particular runs of base pairs can adopt particular formations, such as A-tract DNA. Additionally, conditions such as tautomeration and base-pair rotation lead to major changes at the base pair level. In reality, DNA duplex of any particular sequence exists as a distribution of possible structures, dictated by local sequence and environmental conditions. However, the simplified view of DNA duplex structure and formation is adequate for the purpose of this study.

PCR amplicons from heterozygote DNA will produce two distinct DNA duplex species, each representing a single allele. These two duplexes can be denatured and annealed to form 4 duplex species, two homoduplex and two heteroduplex. Assuming insignificant differential amplification, the four species will be produced in equal amounts. As heteroduplex can be detected in a number of ways, this phenomenon can be used as the basis of a mutation detection screen.

2.3.4.2 The WAVE system.

The WAVE system (Transgenomics) is a denaturing High Performance Liquid Chromatography (dHPLC) based DNA fractionation device which can be used to perform temperature modulated heteroduplex analysis. The WAVE system is based around a hydrophobic solid phase column, which can bind DNA via a bridging molecule, combined with a UV-light based DNA detection system. The WAVE system collects and stores the detection data on an attached PC.

DNA binds to the WAVE column matrix (polystyrene-divinylbenzene copolymers) through the interaction of a bridging molecule, triethylammonium acetate (TEAA). The ammonium ion can bind to DNA phosphate ions, while the alkyl groups bind to the surface of the matrix. The amount of TEAA binding is directly proportional to DNA fragment length. The interaction between TEAA and the matrix is disrupted by acetonitrile. Thus across an acetonitrile gradient at non-denaturing temperatures, elution time will be directly proportional to DNA fragment length. At denaturing temperatures however differential elution of DNA fragments of equal length will occur based on the proportion of single stranded DNA. The elution time will be inversely proportional to the level of single stranded DNA. At a temperature where the region around a heteroduplex mismatch begins to denature, the heteroduplex DNA will elute before homoduplex DNA.
For accurate differential elution of heteroduplex DNA both an acetonitrile gradient profile
where homoduplex DNA elutes at the end of the gradient must be combined with a running
temperature where homoduplex DNA is 75% denatured. The WAVE system contains
software (WAVE-Maker) to predict both the acetonitrile gradient profile and the running
temperature. A further method of determining running temperature is available. If WAVE-
Maker derived gradient information suggests the fragment in question consists of a single
melting domain, the information can be used to run wild-type DNA sequentially over a range
of column temperatures. Data derived from this enables the temperature at which 75%
denaturation takes to be experimentally estimated. These two methods can be supplemented
by running known DNA variants within the fragment of interest to check that all are
recognised as heteroduplex. If any are undetected, the column temperature can be adjusted in
1°C steps until all are detected. Ideally, a single oven temperature is sufficient to cover all
base pairs. However, most DNA fragments contain multiple melting domains and would
require multiple running temperatures. It may be required therefore to design PCR to
produce fragments containing one or two melting domains (Transgenomics Inc).

In practice, a DNA sample is injected into the column within a constant flow of buffer. The
initial buffer has a low acetonitrile concentration, below the threshold required to elute the
DNA. The concentration of acetonitrile within the buffer is then increased in a steady
gradient, slightly beyond that required to elute the DNA. The flow of buffer exiting the
column is passed through a UV-spectrometer, which measures the absorbance at 260nM.
Finally, a high acetonitrile concentration wash buffer is passed over the column to remove
any DNA or contaminants. To produce the changing acetonitrile concentration the WAVE
system varies the ratio of two buffers (A and B), which differ only in their acetonitrile
concentration. The acetonitrile wash is achieved by use of a third buffer (C) containing a
high acetonitrile concentration. After each injection the injection syringe is washed in
syringe solution.

**Buffer A:** 50ml TEAA (Transgenomics Inc), 250μL dHPLC-grade acetonitrile, made to 1
litre with mQ H2O.

**Buffer B:** 50ml TEAA (Transgenomics Inc), 250ml dHPLC-grade acetonitrile, made to 1
litre with mQ H2O.

**Buffer C:** 750ml dHPLC-grade acetonitrile (Transgenomics Inc), made to 1 litre with mQ
H2O.

**Syringe solution:** 80ml acetonitrile, made to 1 litre with mQ H2O.
2.3.4.3 Use of the WAVE system.

The running conditions of all DNA regions were calculated before the regions were amplified from sample DNA. Only the helical fraction across base pairs of interest was used to calculate the column temperature.

All DNA samples were treated identically. First, the region of interest was amplified by PCR using a 96 well PCR plate format in 60x batches. Second, a 5μL aliquot of the reaction was separated using agarose gel electrophoresis and the quantity and quality of each amplicon checked. Failed reactions were repeated. Third, reactions from a number of batches were pooled to produce a single 96-well PCR plate. Fourth, the 96-well PCR plate was placed within an Eppendorf mastercycler gradient thermocycler and reactions were denatured and slowly cooled allowing heteroduplex formation. Finally, reactions were placed within the WAVE system and the run order constructed, forming the well order from which the WAVE machine will inject DNA into the column and the conditions to be applied. If more than one temperature was to be used within the run, the cooler run was performed first. The run was then started and left until complete. In general 5μL of each reaction was injected. An increased volume of reactions of poor quantity was injected.

Results were visualised using the software within the WAVE package. Firstly, sample results were viewed singly and any samples suggestive of variance noted. Then three sample results were overlain together alongside the wild-type result and any samples suggestive of variance noted. Finally, four samples (the maximum) were overlain without wild-type and the figure printed for later reference.

2.3.4.3.1 Putative HESX1 alternative exon A.

Primary sequence of the amplified region derived from the Human Genome Project was entered into the WAVE-Maker program. Both gradient information and melting profile estimation were obtained. Melting profile estimation suggested that a single melting domain (60°C) covered the entire region of interest; an oven temperature of 60°C was therefore chosen.

2.3.4.3.2 Putative HESX1 alternative exon B.

Primary sequence of the amplified region derived from the Human Genome Project was entered into the WAVE-Maker program. Both gradient information and melting profile
estimation were obtained. Melting profile estimation suggested that a single melting domain (68°C) covered the entire region of interest; an oven temperature of 68°C was therefore chosen.

2.3.4.3.3 HESX1 exon 1.

Primary sequence of the amplicon derived from the Human Genome Project was entered into the WAVE-Maker program. Melting profile estimation was obtained. It was apparent that HESX1 exon 1 amplicon contained a single melting domain. Therefore, a standard melting curve was produced as per manufacturer’s guidelines. Briefly, wild-type DNA was run through the column at various oven temperatures ranging from 50°C to 74°C. The retention time of the wild-type DNA was then plotted against oven temperature (°C). The resulting graph was then used to calculate the oven temperature corresponding to 75% retention time, which equals 75% denaturation (Transgenomics Inc). This predicted 62°C as opposed to the predicted 61°C (Figure 2.3.4.3.3a). A standard heteroduplex detection gradient profile (Transgenomics Inc) was then adapted to suit requirements as per manufacturer’s instructions.

(Figure 2.3.4.3.3a) The experimentally determined running temperature of HESX1 exon 1.

2.3.4.3.4 HESX1 exon 2.

Primary sequence of the amplicon derived from the Human Genome Project was entered into the WAVE-Maker program. Melting profile estimation was obtained. It was apparent that
HESX1 exon 2 amplicon contained a single melting domain. Therefore, a standard melting curve was produced as per manufacturer's guidelines. An oven temperature of 58°C was calculated as opposed to the predicted 57°C (Figure 2.3.4.3.4a). A standard heteroduplex detection gradient profile (Transgenomics Inc) was then adapted to suit requirements as per manufacturer's instructions.

\[
\begin{align*}
&\text{(Figure 2.3.4.3.4a) The experimentally determined running temperature of HESX1 exon 2.}
\end{align*}
\]

2.3.4.3.5 HESX1 exon 3/4.

The small size of HESX1 intron 3 allows the production of a fragment containing exons 3 and 4 which falls within the optimal size range for heteroduplex detection using the WAVE system (Transgenomics Inc). Primary sequence of the amplicon derived from the Human Genome Project was entered into the WAVE-Maker program. Melting profile estimation was obtained. It was apparent that the HESX1 exons 3 and 4 amplicon contained a single melting domain. Therefore, a standard melting curve was produced as per manufacturer's guidelines. An oven temperature of 55°C was calculated agreeing with the predicted 55°C (Figure 2.3.4.3.5a). A standard heteroduplex detection gradient profile (Transgenomics Inc) was then adapted to suit requirements as per manufacturer's instructions.

78
As three sequence variants within this fragment were known at the time of optimisation, these were used to confirm sequence variance detection. All three variants (525g>a, S170L and N125S) were detected.

2.3.4.3.6 HESXI 5' LIM1 homology region.

Primary sequence of the amplified region derived from the Human Genome Project was entered into the WAVE-Maker program. Both gradient information and melting profile estimation were obtained. Melting profile estimation suggested that two melting domains (59°C and 61°C) covered the entire region of interest; two oven temperatures of 59°C and 61°C were therefore used sequentially.

2.3.5 DNA Sequencing.

2.3.5.1 Fluorescent DNA sequencing using BigDye and the ABI377.

Sequencing of DNA was carried using the dideoxy method (Sanger et al., 1977), using fluorescent labelled dideoxy terminators. The BigDye kit (Perkin-Elmer) was used to generate the fluorescent-labelled dideoxy fragments. This kit comprises a modified Thermus aquaticus DNA polymerase with reduced discrimination against dideoxy nucleotides and
greatly reduced 5'→3' nuclease activity, combined with dideoxy nucleotides modified to contain a donor dye linked to an acceptor dye (Perkin-Elmer). The labelled dideoxy fragments were separated and visualised using an ABI 377 DNA Sequencer (Perkin-Elmer).

2.3.5.2 Direct fluorescent cycle sequencing of PCR products.

The PCR products were treated as per manufacturer’s instructions except that one half reaction volume was used. The reaction therefore consisted of 4μL BigDye reaction mix + 3.2μL 1μM primer + 1.8μL DNA. The DNA concentration was not quantified before the reaction due to the large throughput required. All sequencing was carried out in the forward and reverse directions and where needed using internal primers. During phase 1 of mutation detection, reactions were placed in 500μL Eppendorf tubes, overlain with mineral oil and the cycling carried out using a Techne GeneE thermocycler. During phase 2, reactions were carried out in 96 well PCR plates using an Eppendorf master gradient thermocycler.

The aqueous phase of reactions containing oil was separated due to differential adhesion to a sheet of Parafilm (American National Can). All reactions were purified using isopropanol as per manufacturer’s instructions. A poly-acyrlamide gel was poured between two glass plates for use in the ABI 377. Either a 32-well plastic comb or a 48-well paper comb was used. All reactions were resuspended in loading dye (Perkin-Elmer) and denatured at 95°C for 5mins before snap cooling in ice. 1.5μL of sequencing reaction was added to each well used. Electrophoresis was carried out in 1x TBE.

Sequence reactions providing poor data were repeated. Any reaction providing positive sequence variance data was repeated using a new PCR reaction derived, if possible, from a separate aliquot of genomic DNA.

2.3.5.3 Fluorescent cycle sequencing of plasmid DNA.

Fluorescent cycle sequencing of plasmid DNA was carried out as above, except that the reaction was adjusted so that it consist of 4μL BigDye reaction mix + 1.6μL 1μM primer + 2μL DNA + 2.4μL mQ H₂O.
2.3.5.4 Sequence comparison.

The ultimate goal of variance detection methods is to identify sequence variance within a test population. Therefore, careful analysis of sequence data is of prime importance.

All sequence data was viewed in the raw gel data format and where necessary manual tracking and contrast adjustments were undertaken. The raw data was then extracted and analysed using Sequencing Analysis 4.3.1 (ABI Prism). All good quality sequencing data was printed using a colour printer at a resolution such that individual peaks were clearly visible. The test sequence electropherogram was then compared against that of a wild-type sample.

Each individual printout was compared with a wild-type printout in three sequential stages. After each stage the reverse sequence was analysed, as was any overlapping sequence. If any sequence variance was identified the subsequent stages were not undertaken. First, each individual base peak was visually checked off against wild type and any dual peaks noted on a text printout of the wild-type sequence. Second, the level of each base peak to the surrounding base peaks was compared and noted as above. Third, any small peaks under the normal base peaks were compared with wild type and checked for a corresponding peak in the reverse sequence.

Samples providing poor data were rerun using the ABI 377 if the data looked favourable. Otherwise, new samples were generated using PCR and resequenced.

Sequence variants were allocated a standard designation. Samples within exons 1-4 and the 5' LIM1 homology region were numbered with reference to the open reading frame generated by initiation codon within exon 1 (Dattani et al., 1998). Intronic variants were numbered with reference to the nearest splice site. Variants within the putative alternative exons were numbered with reference to the open reading frame generated by an initiation codon within the putative exon alt A (i.e. alt -4a>g).

2.3.6 cDNA studies of HESXI expression.

The HESXI/Rpx/Xanf family are thought to be anterior restricted factors (see section 1.3). This produced a paradox during the recruitment of patients, as 37% of patients demonstrated secondary features, many of which are not limited to the anterior of the embryo. As little is known about the expression of HESXI in Homo sapiens a preliminary expression study was
carried out to provide data upon which a decision to exclude patients with non-anterior secondary features could be made.

cDNA derived from various tissues of a week 8 human embryo was readily available (gift from Miss D Blaydon). Two oligonucleotide primers (*HESX1* 5'UTR and *HESX1* 3'UTR) sited within the untranslated regions of *HESX1* were designed to allow the amplification of the entire *HESX1* exonic sequence.

Primers *HESX1* 5'UTR and *HESX1* 3'UTR were used to amplify the complete coding sequence of the *HESX1* gene from wild type genomic DNA in an effort to optimise the annealing temperature of the primers (both in the presence of Q solution and without Q solution). The resulting products were separated using agarose gel electrophoresis and visualised using Ethidium bromide staining. Only the -Q optimisation amplification of full-length transcript produced any signal, the expected 1.8kb fragment, and a 1.1kb fragment. Yield of the 1.1kb fragment decreased as annealing temperature rose, while yield of the 1.8kb fragment increased. The 1.1kb fragment was directly sequenced using fluorescent cycle sequencing and shown to be an artefact derived from a single primer.

In order to produce a suitable positive control amplified cDNA derived from complete head cDNA of human foetuses of weeks 6, 7, 8, 9, 10, and 12 was obtained (a kind from Dr K. Woods). This consisted of total embryological brain RNA of all stages (obtained from Virogen) treated with a cDNA synthesis kit to produce cDNA followed by blunt-end ligation of oligonucleotide adapters allowing the cDNA + adapters to be amplified by PCR using a proof reading polymerase enzyme.

This cDNA was then used as the template for PCR amplification using the primers *HESX1* 5'UTR and *HESX1* 3'UTR. This produced weak signal from weeks 6 – 8. In order to increase the production of smaller PCR products the 2 min extension time (used for the 1.8kb genomic fragment) was reduced to 45s and the reaction repeated. This increased the production of the expected 732bp fragment but also produced a smaller band in all lanes including the H_{2}O control (presumably a primer artefact). In addition the background noise level was increased significantly. The 732bp fragment was extracted, purified, directly sequenced and found to be derived from *HESX1* cDNA. The week 6 cDNA was used as a positive cDNA control.

Amplification from week 8 cDNA was carried out alongside a week 6 head cDNA positive control and a negative H_{2}O control. In addition cDNA from week 17 brain was included.
49µL of each master mix was aliquoted into two rows of a 96 well PCR plate (Abgene Thermo-Fast 96, Skirted). 1µL of each cDNA or 1µL of mQ H₂O was added to each well. The resulting products from each 50µL reaction were combined and concentrated using Microcon-100 columns. The concentrated products were then separated using agarose gel electrophoresis and visualised using Ethidium bromide staining.

**Primer pair:** 5’ primer (*HESXI* 5'UTR) - *HESXI* 3'UTR

**Reaction master mix:**

- 100µL: 10x HotStar Taq buffer (Qiagen)
- 100µL: 2mM dNTP’s (concentration of each dNTP)
- 100µL: 10µM 5’ primer
- 100µL: 10µM *HESXI* 3'UTR
- 5µL: HotStar Taq (Qiagen)
- 575µL: mQ H₂O

**Cycling conditions:**

1 cycle:
95°C - 5mins

35 cycles:
95°C - 30s
61°C - 30s
72°C - 45s

1 cycle:
72°C - 5mins

2.4 Results.

2.4.1 Patient recruitment.

It became readily apparent that patients fell into two broad categories. Some patient samples were sent with extensive clinical details, whilst others arrived with minimal information, often consisting of only a diagnosis. Clinical phenotypes allowed the categorisation of
patients (see section 1.1). The number of patients recruited within each category is shown in Table 2.4.1a, as is the number of familial cases, samples not screened, samples partially screened, and those with additional features including a breakdown of cases with ectopic pituitary dysplasia and defects of the corpus callosum.

<table>
<thead>
<tr>
<th>Group</th>
<th>Total</th>
<th>Familial</th>
<th>Not done</th>
<th>Not completed</th>
<th>Additional features</th>
<th>EPP</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOPD</td>
<td>20</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>16</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Optic nerve hypoplasia only</td>
<td>22</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Pituitary defect only</td>
<td>347</td>
<td>26</td>
<td>54</td>
<td>5</td>
<td>162</td>
<td>97</td>
<td>9</td>
</tr>
<tr>
<td>Septum Pellucidum defect only</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SOPD-PIT</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOPD-ONH</td>
<td>12</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>SOPD-SP</td>
<td>56</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>24</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>SOD</td>
<td>113</td>
<td>7</td>
<td>8</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>87</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>N/a</td>
<td>N/a</td>
<td>N/a</td>
</tr>
<tr>
<td>Corpus Callosum defect only</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2*</td>
<td>0</td>
<td>N/a</td>
</tr>
<tr>
<td>Total</td>
<td>670</td>
<td>42</td>
<td>80</td>
<td>9</td>
<td>251†</td>
<td>111</td>
<td>31‡</td>
</tr>
</tbody>
</table>

(Table 2.4.1a) The categorisation of recruited patients. SOD — diagnosed as such by referring clinician but without sufficient clinical information, EPP - ectopic posterior pituitary, CC - defect of the corpus callosum.

†Not including corpus callosum defect only patients, ‡Including corpus callosum defect only patients.

### 2.4.2 Variance detection.

Due to the nature of the techniques involved, the variance detection strategy was not applied in a serial manner; that is, all amplicons were not all generated and screened from a given sample before proceeding to the next sample. Rather a parallel processing approach was employed whereby amplicons derived from a single region of *HESXI* were generated from a number of samples and screened as a batch. Similarly, due to the nature of patient recruitment and the WAVE system only becoming available during the course of study and the ongoing investigation into *HESXI* genomic arrangement, the two variance detection techniques were not applied serially. Rather SSCP detection was used on exons 1-4 from all samples recruited until the WAVE system became available. Then dHPLC-heteroduplex detection was used on exons 1-4 from samples previously screened by SSCP detection and those recruited in the meantime. When the appropriate genomic sequence became available, dHPLC-heteroduplex detection was applied to putative exons alt A and alt B from all samples.
recruited at that time, followed by dHPLC-heteroduplex detection being applied to the 5’LIM1 homology region. SSCP detection was then performed on exons 1-4 of those samples recruited since the WAVE system became available. Finally, all amplicons outstanding were screened. Exons alt A, 1, 2, 3, and 4 were screened using both dHPLC-heteroduplex detection and SSCP detection. Exon alt1b and the 5’ LIM1 homology region were screened by dHPLC-heteroduplex detection only. Finally, a small number of amplicons could not be amplified and thus a number of patients were not completely screened (Table 2.4.2a).

<table>
<thead>
<tr>
<th>Stages</th>
<th>Regions</th>
<th>Method</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-4</td>
<td>SSCP</td>
<td>Samples recruited until WAVE became available.</td>
</tr>
<tr>
<td>2</td>
<td>1-4</td>
<td>WAVE</td>
<td>Samples screened by SSCP in stage 1.</td>
</tr>
<tr>
<td>3</td>
<td>1-4</td>
<td>WAVE</td>
<td>Samples recruited since beginning of stage 2.</td>
</tr>
<tr>
<td>4</td>
<td>1-4</td>
<td>SSCP</td>
<td>Samples recruited since beginning of stage 2.</td>
</tr>
<tr>
<td>5</td>
<td>alt A - alt B</td>
<td>WAVE</td>
<td>All samples recruited at time point.</td>
</tr>
<tr>
<td>6</td>
<td>5’ LIM1 homology region</td>
<td>WAVE</td>
<td>Subset of all samples recruited at time point.</td>
</tr>
<tr>
<td>7</td>
<td>alt A, 1, 2,3,4,</td>
<td>SSCP</td>
<td>All samples recruited since beginning of stage 5.</td>
</tr>
<tr>
<td>8</td>
<td>All</td>
<td>Both</td>
<td>Any samples outstanding</td>
</tr>
</tbody>
</table>

(Table 2.4.2a) Stages of HESX1 mutation screening.

In total 5546 DNA amplifications, consisting of 2673 SSCP detection reactions and 2873 dHPLC-heteroduplex detection reactions, were performed from sporadic cases and screened for mutations within HESX1. This figure excludes samples from which poor results were obtained due to either poor amplification, or poor variance detection, and thus amplification and screening were repeated. Based upon these results of this variance detection protocol, 188 amplicons were regenerated and directly sequenced for changes within HESX1. This identified 14 sequence variants. SSCP detection also produced 4 false negative results as SSCP detection failed to detect 4 sequence variants that were later identified using dHPLC-heteroduplex detection. No false negative results were identified within the dHPLC-heteroduplex negative samples (Table 2.4.2b).
Table 2.4.2b) Breakdown of mutation screening results. *known false negatives, †excludes duplex artefacts, ‡excludes 4489.

<table>
<thead>
<tr>
<th>Region</th>
<th>SSCP false positive†</th>
<th>SSCP false negative*</th>
<th>SSCP positive</th>
<th>dHPLC heteroduplex false positive</th>
<th>dHPLC heteroduplex false negative*</th>
<th>dHPLC heteroduplex positive</th>
<th>Total screened amplifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>alt A</td>
<td>523 13</td>
<td>0</td>
<td>554 13</td>
<td>0</td>
<td>1</td>
<td>1077</td>
<td></td>
</tr>
<tr>
<td>alt B</td>
<td>n/a n/a</td>
<td>n/a n/a</td>
<td>554 0</td>
<td>0</td>
<td>0</td>
<td>554</td>
<td></td>
</tr>
<tr>
<td>Liml</td>
<td>n/a n/a</td>
<td>n/a n/a</td>
<td>183 15</td>
<td>0</td>
<td>4</td>
<td>183</td>
<td></td>
</tr>
<tr>
<td>Exon 1</td>
<td>536 16</td>
<td>0</td>
<td>537 33</td>
<td>0</td>
<td>0</td>
<td>1073</td>
<td></td>
</tr>
<tr>
<td>Exon 2</td>
<td>541 10</td>
<td>0</td>
<td>530 11</td>
<td>0</td>
<td>2†</td>
<td>1071</td>
<td></td>
</tr>
<tr>
<td>Exon 3</td>
<td>537 20</td>
<td>1</td>
<td>n/a 1</td>
<td>n/a n/a</td>
<td>n/a n/a</td>
<td>537</td>
<td></td>
</tr>
<tr>
<td>Exon 4</td>
<td>536 28</td>
<td>1</td>
<td>2</td>
<td>n/a n/a</td>
<td>n/a n/a</td>
<td>536</td>
<td></td>
</tr>
<tr>
<td>Exon 3/4</td>
<td>n/a n/a</td>
<td>n/a n/a</td>
<td>515 25</td>
<td>0</td>
<td>2†</td>
<td>515</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2673 87</td>
<td>4</td>
<td>2873 97</td>
<td>0</td>
<td>9†</td>
<td>5546</td>
<td></td>
</tr>
</tbody>
</table>

2.4.2.1 Putative exon alt A

A total of 554 samples were screened by dHPLC-heteroduplex detection, while SSCP detection was carried out on a total of 523 samples. dHPLC-heteroduplex detection identified 13 positive peak-shifts, 1 of which was due to sequence variance, while SSCP detection identified 13 positive band-shifts none of which were due to sequence variance. Representative putative exon alt A SSCP detection results are shown in Figure 2.4.2.1a while representative putative exon alt A dHPLC detection results are shown in figures 2.4.2.1b, and 2.4.2.1c. A result later shown to correspond to a sequence change is shown in figure 2.4.2.1d.

Figure 2.4.2.1a. A HESX1 exon A SSCP detection gel showing a sample with a SCC change (arrow), this is an example of a false positive result. Wild type is far right.
Figure 2.4.2.b) An overlay of exon A dHPLC results showing three samples 4883 (blue), 4884 (green), and 4885 (red), with wild-type (black).

Figure 2.4.2.c) An overlay of exon A dHPLC results showing four samples 4517 (blue), 4518 (green), 4519 (red), and 4521 (black). Note that the peak height of 4518 (green) has been equalised to that of the other 3 samples resulting in a lack of vertical scale. 4518 (green) is an example of a false positive result.
2.4.2.2 Putative exon alt B

A total of 554 samples were screened by dHPLC-heteroduplex detection. Due to time constraints and in the knowledge that dHPLC-heteroduplex detection had produced no false positive results SSCP detection was not carried out. Representative alternative exon B dHPLC results are shown in figures 2.4.2.2a and 2.4.2.2b. dHPLC detection identified no changes suggestive of sequence variance.
2.4.2.3 5’ LIM1 homology region.

A total of 183 samples were screened by dHPLC-heteroduplex. Due to time constraints and in the knowledge that dHPLC-heteroduplex detection had produced no false negative results SSCP detection was not carried out. dHPLC detection identified 15 changes suggestive of sequence variance, 4 of which were due to sequence variance. Representative 5’ LIM1 domain dHPLC detection results are shown in figures 2.4.2.3a, and 2.4.2.3b. A result later shown to correspond to sequence change is shown in figure 2.4.2.3c.
(Figure 2.4.2.3a) An overlay of 5' LIM1 homology region dHPLC results at 61°C for four samples 4751 (blue), 4754 (green), 4718 (red), and 4714 (black).

(Figure 2.4.2.3b) An overlay of 5' LIM1 homology region dHPLC results at 59°C for four samples 5057 (blue), 5071 (green), 5072 (red), and 4842 (black).
(Figure 2.4.2.3c) An overlay of 5' LIM1 homology region dHPLC results at 59°C for four samples 4988 (blue), 4989 (green), 4992 (red), and 5001 (black). 4988 (blue) results from the sequence variant -277t>g

2.4.2.4 Exon 1

A total of 537 samples were screened by dHPLC-heteroduplex detection, while SSCP detection was carried out on a total 536 samples. dHPLC-heteroduplex detection identified 33 positive peak-shifts, none of which was due to sequence variance, while SSCP detection identified 16 positive band-shifts, none of which was due to sequence variance.

Representative exon 1 SSCP detection results are shown in figures 2.4.2.4a, while representative exon 1 dHPLC detection results are shown in figures 2.4.2.4b and 2.4.2.4c.

(Figure 2.4.2.4a) A HESX1 exon 1 SSCP detection gel showing a sample with a SSC change (arrow) and samples with duplex changes (brackets). Wild type is far right.
2.4.2.5 Exon 2

A total of 530 samples were screened by dHPLC-heteroduplex detection, while SSCP detection was carried out on a total 541 samples. dHPLC-heteroduplex detection identified 11 positive peak-shifts, 2 of which were due to sequence variance, whilst SSCP detection identified 10 positive band-shifts, none of which was due to sequence variance. SSCP detection also failed to detect the 2 samples containing sequence variance that were later
identified by dHPLC-heteroduplex detection. Representative exon 2 SSCP detection results are shown in figure 2.4.2.5a, while representative exon 2 dHPLC detection results are shown in figures 2.4.2.5b and 2.4.2.5c. Results later shown to correspond to a sequence change are shown in figures 2.4.2.5i (183t>c), 2.4.2.5j (219c>t), and 2.4.2.5k (Y90H).

(Figure 2.4.2.5a) A HESX1 exon 2 SSCP detection gel. Wild type is far left.

(Figure 2.4.2.5h) An overlay of exon 2 dHPLC results showing three samples 4900 (blue), 4870 (red), and 4871 (black), with wild-type (green).
(Figure 2.4.2.5c) An overlay of exon 2 dHPLC results showing four samples 4820 (blue), 5103 (green), 5120 (red), and 5134 (black). 5134 (black) is an example of a false positive result; note that the primer peak and the main peak are retained longer than the other three samples and that the wash peak is absent.

(Figure 2.4.2.5d) An overlay of exon 2 dHPLC results showing four samples 4585 (blue), 4410 (green), 4614 (red), and 4624 (black). 4614 (red) results from the sequence variant 183t>c. Note the failure of 4624 (black).
2.4.2.5e An overlay of exon 2 dHPLC results showing four samples 4203 (blue), 3711 (green), 3936 (red), and 4212 (black). 4212 (black) results from the sequence variant 219c>t.

(Figure 2.4.2.5f) An overlay of exon 2 dHPLC results showing four samples 4954 (blue), 4214 (green), 4352 (red), and 4489 (black). 4489 (black) results from the sequence variant Y90H.

2.4.2.6 Exon 3 and 4

A total of 537 samples were screened by SSCP detection for variance within exon 3, while 536 samples were screened by SSCP detection for variance within exon 4. The small size of intron 3 allowed both exon 3 and exon 4 to be screened together by HPLC-heteroduplex detection, thus 515 samples were screened for variance within exon3-intron3-exon4. Exon 3
SSCP detection identified 20 positive band-shifts, 3 of which were due to sequence variance, while Exon 4 SSCP detection identified 28 positive band-shifts, 2 of which were due to sequence variance. Exon 3 SSCP detection failed to detect 1 sample with known sequence variance, as did Exon 4 SSCP detection. dHPLC-heteroduplex detection identified 25 positive peak-shifts, 2 of which were due to sequence variance.

Representative exon 3 SSCP detection results are shown in figure 2.4.2.6a, while representative exon 4 SSCP detection results are shown in figure 2.4.2.6b. Results later shown to correspond to a sequence change are shown in figures 2.4.2.6c (N125S), 2.4.2.6d (S170L), and 2.4.2.6e (525g>a). Representative dHPLC detection results are shown in figures 2.4.2.6f and 2.4.2.6g. Results later shown to correspond to a sequence change are shown in figures 2.4.2.6h (N125S), 2.4.2.6i (S170L), and 2.4.2.6j (525g>a). SSCP detection failed to detect 2 samples with sequence variance shown in figures 2.4.2.6m (E149K) and 2.4.2.6e (525g>a).
(Figure 2.4.2.6b) A HESX1 exon 4 SSCP detection gel. Wild type is slightly left of centre.

(Figure 2.4.2.6c) Exon 3 SSCP results for N125S samples (arrow).
(Figure 2.4.2.6d) Exon 4 SSCP results for the S170L sample (arrow).

(Figure 2.4.2.6e) Exon 4 SSCP results for 525g>a samples (arrow). The sample on the left was identified as containing a SSC change, while the sample on the right was not identified as containing a SSC change.
(Figure 2.4.2.6f) An overlay of exon 3/4 dHPLC results showing four samples 4748 (blue), 4750 (green), 4751 (red), and 4752 (black).

(Figure 2.4.2.6g) An overlay of exon 3/4 dHPLC results showing four samples 3679 (blue), 4760 (green), 4995 (red), and 5050 (black).
(Figure 2.4.2.6h) An overlay of exon 3/4 dHPLC results showing four samples 4766 (blue), 4767 (green), 4768 (red), and 4769 (black). 4769 (black) results from the sequence variant E149K.

(Figure 2.4.2.6i) An overlay of exon 3/4 dHPLC results showing three samples 4712 (blue), 4585 (green), and 4721 (red) alongside WT (black). 4585 (green) results from the sequence variant 525g>a.
An overlay of exon 3/4 dHPLC results showing four samples 483 (blue), 4963 (green), 4763 (red), and 4701 (black). 4701 (black) results from the sequence variant N125S.

Exon 4 SSCP results for the E149K sample (arrow). This sample was not identified as containing a SSCP change.
2.4.4 DNA sequencing.

In total 330 HESXI amplicons were sequenced. This consisted of 129 amplicons from familial cases and 201 amplicons from sporadic cases. 25 amplicons were found to contain sequence variants (Table 2.4.4a). The location of each variant is shown in figure 2.4.4a.

<table>
<thead>
<tr>
<th>Region</th>
<th>Total</th>
<th>Familial</th>
<th>Sporadic</th>
<th>Sequence variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>alt A</td>
<td>47</td>
<td>20</td>
<td>27</td>
<td>3</td>
</tr>
<tr>
<td>alt B</td>
<td>20</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lim1</td>
<td>39</td>
<td>20</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>Exon 1</td>
<td>72</td>
<td>23</td>
<td>49</td>
<td>0</td>
</tr>
<tr>
<td>Exon 2</td>
<td>48</td>
<td>23</td>
<td>25</td>
<td>6</td>
</tr>
<tr>
<td>Exon 3/4</td>
<td>104</td>
<td>23</td>
<td>81</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>330</td>
<td>129</td>
<td>201</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 2.4.4a: The total number of samples sequenced, sub-divided into familial and sporadic cases, alongside the number samples containing sequence variance.

(Figure 2.4.4a) Location of HESXI variants.
2.4.4.1 Putative exon A.

Three sequence variants were found during sequencing of putative exon A, alt -50c>t (Figure 2.4.4.1a), alt -4 t>g (Figure 2.4.4.1b) and IVS A +5c>t (Figure 2.4.4.1c). Each variant was found in a single patient. All samples were regenerated and resequenced, which produced identical results.

(Figure 2.4.4.1a) Sequencing results of putative exon A for sample 4868 showing alt -50c>t in the heterozygous state. The alt -50c>t reverse sequence is shown on the left and the wild-type reverse on the right. The site is too close to the primer site to produce good quality forward sequence.
(Figure 2.4.4.1b) Sequencing results of putative exon A for sample 3883 showing -4 t>g in the heterozygous state. The forward variant sequence is shown on the top left and the variant reverse on the bottom left. The forward WT sequence is shown on the top right and the WT reverse on the bottom right.

(Figure 2.4.4.1c) Sequencing results of putative exon A for sample 4904 showing the donor splice site IVS1A +5c>t variant in the heterozygous state. The forward variant sequence is shown on the top left and the variant reverse on the bottom left. The forward WT sequence is shown on the top right and the WT reverse on the bottom right.
2.4.4.2 5’ LIM homology region.

A single sequence variant was found during sequencing of the 5’ LIM1 domain, -277t>g (Figure 2.4.4.2a). The variant was found in four patients, 4645, 4821, 4975, and 4988.

(Figure 2.4.4.2a) Sequencing results of the 5’ LIM1 region for sample 4645 showing -277t>g in the heterozygous state. The forward variant sequence is shown on the top left and the variant reverse on the bottom left. The forward WT sequence is shown on the top right and the WT reverse on the bottom right.

2.4.4.3 Exon 2.

Three sequence variants were found during sequencing of exon 2, 183t>c (Figure 2.4.4.3a), 219c>t (Figure 2.4.4.3b), and Y90H (Figure 2.4.4.3c). The 219c>t variant was found in a single patient, 4212, while the 183t>c variant was found in three patients, 4582, 4614, and 5113. The Y90H variant was identified in a parental sample, which had been included in the mutation screen in error. All samples were regenerated and resequenced, which produced identical results. A fourth sequence variant, 161a>g (Figure 2.4.4.3d), was identified in a single patient (4405). However, the variant was only seen in the forward direction and when regenerated and resequenced the variant was not apparent. On closer examination of the ABI 377 gel file, it was readily apparent that the “sequence variant” was in fact a file error.
(Figure 2.4.4.3a) Sequencing results of exon 2 for sample 4382 showing 183t>c in the heterozygous state. The forward variant sequence is shown on the top left and the variant reverse on the bottom left. The forward WT sequence is shown on the top right and the WT reverse on the bottom right.

(Figure 2.4.4.3b) Sequencing results of exon 2 for sample 4489 showing Y90H in the heterozygous state. The forward variant sequence is shown on the top left and the variant reverse on the bottom left. The forward WT sequence is shown on the top right and the WT reverse on the bottom right.
2.4.4 Exon 3/4

Four sequence variants were found during sequencing of exon 3/4; N125S (Figure 2.4.4a), S170L (Figure 2.4.4b), E149K (Figure 2.4.4c) and 525g>a (Figure 2.4.4c). The N125S variant was found in 8 patients, 4610, 4661, 4701, 4762, 5006, 5103, 5125, and 5164. The N125S variant was found in the heterozygous state in 7 patients and in the homozygous state in a single patient, 5103. The S170L variant was found in a single patient, 4691. The E149K variant was found in a single patient, 4769. The 525g>a variant was found in two patients, 4508 and 4585. All samples were regenerated and resequenced, which produced identical results.
(Figure 2.4.4.4a) Sequence results showing the N125S variant. Homozygous variant sequence is shown on the left, heterozygous sequence is shown in the centre while homozygous wild-type sequence is shown on the right.

(Figure 2.4.4.4b) Sequence results showing the S170L variant. The forward variant sequence is shown on the top left and the variant reverse on the bottom left. The forward WT sequence is shown on the top right and the WT reverse on the bottom right.
(Figure 2.4.4.4c) Sequencing results of exon 3/4 for sample 4769 showing the E149K variant. The variant sequence is shown on the left and the WT on the right.

(Figure 2.4.4.4d) Sequencing results of exon 3/4 for sample 4508 showing the S25g>a variant. The variant sequence is shown on the left and the WT on the right.
2.4.5 cDNA studies of HESXI expression.

Signal of the expected size was detectable in the stomach and kidney, with faint signal detectable in the ear (Figure 2.4.5a). Again, additional bands of a smaller size were detectable in all lanes including the H2O control (presumably primer artefacts). The signal produced from pre-amplified week 6 cDNA was of a poorer quality than the week 8 cDNA.

(Figure 2.4.5a) Results of HESXI transcript amplification of week 17 brain (B) and week 8 cDNA from eye (E), ear (Ea), stomach (S), kidney (K), and heart (H).

The experiment was repeated with the addition of a pre-amplified week 6 cDNA positive control (Figure 2.4.5b). The results were similar with signal detected in the stomach and kidney lanes with faint signal in the ear lane. In addition signal could be detected in the eye lane.

(Figure 2.4.5b) Results of the second HESXI transcript amplification of week 17 brain (B) and week 8 cDNA from eye (E), ear (Ea), stomach (S), kidney (K), and heart (H). Two positive controls were used week 6 pre-amplified cDNA (6+) and week 8 pre-amplified cDNA (8+).
2.4.6.1 Putative alternative exons.

During the course of this study, the Human Genome Organisation produced partial genomic sequence by Dattani et al. (1998) within sequence contig 28 (Figure 2.4.6.1a).
As clone RP11-241K3 contains ~160kb of genomic sequence, the 5' divergent HESXI sequence from GenBank gi:1720505 was compared with it using pairwise BLAST (http://www.ncbi.nlm.nih.gov) (Figure 2.4.6.1b).

\[
\begin{align*}
\text{Score} &= 81.4 \text{ bits (42)}, \text{ Expect} = 1e-13 \\
\text{Identities} &= 42/42 (100\%), \text{ Positives} 42/42 (100\%)
\end{align*}
\]

\[
\begin{array}{ll}
\text{Query} & 44 \quad \text{aacaccatgcaccaccagctcttaagagaagagactttcaagg 85}
\end{array}
\]

\[
\begin{array}{ll}
\text{Sbjct} & 72483 \quad \text{aacaccatgcaccaccagctcttaagagaagagactttcaagg 73524}
\end{array}
\]

\[
\begin{align*}
\text{Score} &= 66.4 \text{ bits (34)}, \text{ Expect} = 6e-09 \\
\text{Identities} &= 42/44 (95\%), \text{ Positives} 42/44 (95\%), \text{ Gaps} = 1/44 (2\%)
\end{align*}
\]

\[
\begin{array}{ll}
\text{Query} & 1 \quad \text{ggcagtctgcn-tgggatggccttccctgcacgcgcgtt} 43
\end{array}
\]

\[
\begin{array}{ll}
\text{Sbjct} & 72241 \quad \text{ggcagtctgcn-tgggatggccttccctgcacgcgcgtt 72284}
\end{array}
\]

(Figure 2.4.6.1b) Pairwise BLAST of GenBank gi:1720505 (Query) and GenBank gi:1720505 (Sbjct)

This identified the complete divergent sequence within the sequence contig 21 (Figure 2.4.6.1c) of clone RP11-241K3 (GenBank gi:1720505). The divergent sequence was encoded within two regions, EST nucleotides 1-43 between clone nucleotides 72241-72284 (referred hence forth as putative exon alt A) and EST nucleotides 44-84 between clone nucleotides 73483-73523 (referred hence forth as putative exon alt B). The genomic data contains a C dinucleotide at nucleotide 13, which within the EST sequence is undefined.

Clone RP11-241K3 GenBank gi:7798783 also contains the genomic sequence of the gene APPL and the first exon of APPL is located close to HESXI putative exon alt A, at either side of a CpG island (GI:1021667). Although only partial 5'UTR sequence for both transcripts is known in all likelihood the 5’UTR’s of both transcripts will overlap (Figure 2.4.6.1c).
Sequence contig 21 contains 4052 bases downstream of putative exon alt B, while sequence contig 28 contains 6043 upstream of exon 2, therefore the minimum distance between putative exon alt B and exon 2 is ~10kb. Excluding contigs 21 and 28, as well as, contigs containing APPL sequence the maximum distance between putative exon alt B and exon 2 is ~98kb.

Based upon these findings and the results of Dattani et al. (1998) the putative genomic structure of HESXI including the two putative exons is shown in figure 2.4.6.1d.

(Figure 2.4.6.1c) Partial sequence of contig 21, showing HESXI exons A and B, as well as, APPL exon 1. Known exonic sequence is shown in bold with both initiation codon’s are highlighted in grey. Sequence variants are indicated with a * and shown in bold capitals.

Sequence contig 21 contains 4052 bases downstream of putative exon alt B, while sequence contig 28 contains 6043 upstream of exon 2, therefore the minimum distance between putative exon alt B and exon 2 is ~10kb. Excluding contigs 21 and 28, as well as, contigs containing APPL sequence the maximum distance between putative exon alt B and exon 2 is ~98kb.

Based upon these findings and the results of Dattani et al. (1998) the putative genomic structure of HESXI including the two putative exons is shown in figure 2.4.6.1d.

(Figure 2.4.6.1d) The genomic structure of HESXI. The homeodomain is shown in red and the engrailed-like repressor domain is shown in green. Intron size is shown in base pairs and exonic coding sequence size is shown in nucleotides. Untranslated regions are not included.
2.4.6.2 Non-coding regions.

Murine *Hesx1* genomic sequence representing three transcriptional regulation regions was obtained (a kind gift of Dr. K. Mahon):

Region 1. A forebrain specific *Lim1* response element within the 5'UTR (5' LIM1 domain)

Region 2. A pituitary specific enhancer within the 3' UTR.

Region 3. An enhancer within intron 1.

The murine sequence was compared to *HESX1* genomic sequence contained within clone RP11-241K3 GenBank gi:7798783. Identity within regions 1 and 2 was readily identifiable (Figure 2.4.6.2a). No identity to region 3 could be detected within intron 1. Of the three recently identified *Xanf-1* cis-regulatory regions, box 3 and box 2 are contained with the 5' LIM1 homology region (Eroshkin et al., 2002).
During the bioinformatic investigation into the genomic structure, a single nucleotide difference was observed between the published HESXI genomic sequence (Dattani, et al., 1998) and that of clone RP11-241K3, GenBank gi:7798783 (HUGO). This variation (a-g) occurs in intron 2, within the sequence from which the oligonucleotide primer DHX3F was derived (Figure 2.4.6.3a). This position corresponds to the third nucleotide from the 3' end of the oligonucleotide primer DHX3F (see Table 2.4.6.3a).

(Figure 2.4.6.2a) Partial sequence of contig 28 showing the 5' LIMI homology region highlighted in grey. The translated region of exon 1 is shown in bold while cis-regulatory regions, box 1-3, are underlined (Eroshkin et al., 2002). The -277t>g variant is indicated with a * and show in bold capitals. Nucleotides within the 3'UTR homologous between man and mouse are shown underlined in bold.

2.4.6.3 DHX3F sequence variation.

(Figure 2.4.6.3a) Variation within the sequence of oligonucleotide primer DHX3F. Exonic sequence is underlined in bold, oligonucleotide primer DHX3F is shown in bold, while the oligonucleotide primer HESX1HEX is highlighted in grey. The variation is indicated by an * and the HUGO allele is shown.
If we assume that this variation represents a SNP, then PCR amplification from patient alleles carrying the HUGO allele (g) may fail. Therefore, the validity of this assumption was tested.

The HUGO allele forms part of a Dde I restriction site while the sequence published by Dattani et al. (1998) does not. Therefore, a 5' HEX labelled oligonucleotide primer upstream from DHX3F was obtained commercially to enable a restriction digest based test for the HUGO allele was undertaken (Figure 2.4.6.3a).

<table>
<thead>
<tr>
<th>Dattani DHX3F</th>
<th>5'-AGC TCA TTT TTG AGA CAT ACT GAA TA -3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>HUGO DHX3F</td>
<td>5'-AGC TCA TTT TTG AGA CAT ACT GAG TA -3'</td>
</tr>
</tbody>
</table>

(Table 2.4.6.3a) Variation within the sequence of oligonucleotide primer DHX3F (shown in bold).

PCR amplification was performed using control DNA with the primer pair HESX1HEX – DHX3R. The product was then digested by Dde I according to manufacturer’s guidelines (NEB) and purified by gel extraction. The products were then separated and visualised using an ABI 377 (ABI Prism) and the GeneScan software (ABI Prism).

The allele derived from the sequence published by Dattani et al. (1998) produces an expected product size of 163nt, while the HUGO allele produces an expected product size of 141nt. All control alleles (n=94) produced a product of 163nt.

Therefore, the HUGO allele may represent a rare SNP or a sequencing error. If the HUGO allele does represent a rare SNP, PCR amplification using oligonucleotide primer DHX3F may fail.

2.4.7 Patient phenotypes.

2.4.7.1 Genotype to phenotype correlations.

The clinical phenotype associated with each variant genotype is shown in Table 2.4.7.1a.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Genotype</th>
<th>State</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>4868</td>
<td>alt -50c&gt;t</td>
<td>+/-</td>
<td>GH deficiency</td>
</tr>
<tr>
<td>3883</td>
<td>alt -4 t&gt;g</td>
<td>+/-</td>
<td>Panhypopituitarism and pituitary hypoplasia</td>
</tr>
<tr>
<td>4904</td>
<td>IVS A +5c&gt;t</td>
<td>+/-</td>
<td>GH deficiency</td>
</tr>
<tr>
<td>4645</td>
<td>-277t&gt;g</td>
<td>+/-</td>
<td>Absent Septum pellucidum and hypoplastic optic chiasma</td>
</tr>
<tr>
<td>4821</td>
<td>-277t&gt;g</td>
<td>+/-</td>
<td>GH deficiency, ACTH deficiency, TSH deficiency, palatal cleft and atypical optic nerves and chiasma</td>
</tr>
<tr>
<td>Chromosome</td>
<td>Mutation</td>
<td>SG2T</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>----------</td>
<td>------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>4975</td>
<td>-277t&gt;g</td>
<td>+/-</td>
<td>SOD</td>
</tr>
<tr>
<td>4988</td>
<td>-277t&gt;g</td>
<td>+/-</td>
<td>?SOD, Diabetes Insipidus, nasal hypoplasia and disturbed thermoregulation</td>
</tr>
<tr>
<td>4582</td>
<td>183t&gt;c</td>
<td>+/-</td>
<td>SOD: GH deficiency, cortisol deficiency, central hypothyroidism, bilateral optic nerve hypoplasia and hypoglycemia</td>
</tr>
<tr>
<td>4614</td>
<td>183t&gt;c</td>
<td>+/-</td>
<td>SOD: GH deficiency and unilateral optic nerve hypoplasia</td>
</tr>
<tr>
<td>5113</td>
<td>183t&gt;c</td>
<td>+/-</td>
<td>GH deficiency</td>
</tr>
<tr>
<td>4212</td>
<td>219c&gt;t</td>
<td>+/-</td>
<td>?SOD: GH deficiency, anophthalmia, optic nerve hypoplasia, small optic chiasma, white matter lesions in cerebral hemispheres, and bilateral undescended testes</td>
</tr>
<tr>
<td>n/a</td>
<td>Y90H</td>
<td>+/-</td>
<td>Unaffected parent.</td>
</tr>
<tr>
<td>4610</td>
<td>N125S</td>
<td>+/-</td>
<td>GH deficiency and unilateral optic nerve hypoplasia</td>
</tr>
<tr>
<td>4661</td>
<td>N125S</td>
<td>+/-</td>
<td>GH deficiency</td>
</tr>
<tr>
<td>4701</td>
<td>N125S</td>
<td>+/-</td>
<td>Neuro-secretory GH deficiency</td>
</tr>
<tr>
<td>4762</td>
<td>N125S</td>
<td>+/-</td>
<td>SOD: Absent corpus callosum, thin optic nerves, hypoplastic pituitary, GH deficiency and TSH deficiency</td>
</tr>
<tr>
<td>5006</td>
<td>N125S</td>
<td>+/-</td>
<td>SOD: optic nerve hypoplasia, small optic chiasma, absent septum pellucidum and pituitary hypoplasia</td>
</tr>
<tr>
<td>5103</td>
<td>N125S</td>
<td>+/-</td>
<td>Pituitary hypoplasia, ectopic posterior pituitary, cortisol deficiency and TSH deficiency</td>
</tr>
<tr>
<td>5125</td>
<td>N125S</td>
<td>+/-</td>
<td>SOD</td>
</tr>
<tr>
<td>5164</td>
<td>N125S</td>
<td>+/-</td>
<td>Panhypopituitarism, wrinkly skin, bone disorder, GH deficiency, cortisol deficiency and thyroxin deficiency</td>
</tr>
<tr>
<td>4769</td>
<td>E149K</td>
<td>+/-</td>
<td>GH deficiency, pituitary hypoplasia, ectopic posterior pituitary and supernumerary digits</td>
</tr>
<tr>
<td>4691</td>
<td>S170L</td>
<td>+/-</td>
<td>GH deficiency, pituitary hypoplasia and ectopic posterior pituitary</td>
</tr>
<tr>
<td>4508</td>
<td>525g&gt;a</td>
<td>+/-</td>
<td>SOD</td>
</tr>
<tr>
<td>4585</td>
<td>525g&gt;a</td>
<td>+/-</td>
<td>SOD: agenesis of the corpus callosum and hypothyroidism</td>
</tr>
</tbody>
</table>

(Table 2.4.7.1a) HESXI genotype to phenotype correlations.

2.4.7.2 S170L

S170L was identified in the male patient 4691 (DNA supplied by the London Centre for Paediatric Endocrinology). Patient 4691 was born at full-term with a birth weight of 7lb 11oz. The neonatal period was complicated by jaundice. A diagnosis of growth hormone deficiency was made in 1990 at age 6yrs. Patient 4691 was treated with growth hormone replacement through childhood until 1998, when it was stopped. The peak GH in response to
insulin induced hypoglycaemia was 1.1mU/L. A basal FSH level of 2.8U/L and a basal LH level of 4.4U/L was recorded. TRH provocation showed a basal prolactin level of 252mU/L that rose in response to TRH to 799mU/L. His basal TSH level was 2.7mU/L with a peak of 8.2mU/L. His basal cortisol level was 252nmol/L with a peak of 920nmol/L.

Neuro-imaging was performed using MRI scanning and showed the presence of a small anterior pituitary with an ectopic posterior pituitary. In conclusion, anterior pituitary hypoplasia with an ectopic posterior pituitary tissue was noted, with no optic nerve hypoplasia and no midline defects.

2.4.7.3 E149K.

E149K was identified in the male patient 4769 (DNA supplied by Prof. S. Shalet, Christie hospital). Patient 4769 was born at full-term by a normal vaginal delivery, with a birth weight of 2.8kg. The neonatal period was complicated by jaundice, hypoglycaemia, and umbilical cord infection. Shortly after birth, small supernumerary digits were excised from both hands. During infancy and early childhood, his growth was poor (height well below 3rd centile) and a diagnosis of growth hormone deficiency was made in July 1980 at age 6yrs. At diagnosis, the height of patient 4769 was below the 3rd centile and he was diagnosed as having growth hormone deficiency. Patient 4769 was treated with growth hormone replacement through childhood until November 1990. The most recent peak GH provocation using the insulin tolerance test (9/9/96) gave a peak of 1.9mU/L with a peak response to arginine (10/9/96) of 1.3mU/L. At presentation he was considered to have hypoplasia of the scrotum, with a small penile size noted at 12.5 years. LHRH provocation (3/6/91) showed a basal FSH level of 2 U/L which did not rise in response to LHRH. His basal LH level was 4U/L with a peak of 8U/L. He received testosterone replacement therapy until 17 years old. This was stopped and he subsequently proved to be fertile with the birth of a child in 1999. His son appears to be an unaffected carrier of the E149K mutation.

Neuro-imaging by MRI scanning was performed in 2001 and the findings showed the presence of a small stella turcica with a small and thin infundibulum. The anterior pituitary was very small. The total depth of the pituitary gland was only 3 to 4 mm. Additionally, an enhancing nodule was observed in the region of the tuber cinereum, consistent with an ectopic posterior pituitary. In conclusion, pituitary gland hypoplasia with possible ectopic posterior pituitary tissue was noted, with no optic nerve hypoplasia and no midline abnormalities.
3 Functional investigations.

3.1 Sequence variants.

Of the sequence variants identified, putative exon A -50c>t (n=1), putative exon A -4 t>g (n=1), IVS1A +5c>t (n=1), exon 1 -277t>g (n=4), 183t>c (n=3), 219c>t (n=1), Y90H (n=1), 525g>a (n=2) and E149K (n=1) are novel changes, while N125S (n=8) and S170L (n=1) have been previously identified (Dattani et al., 1998 and Thomas et al., 2001).

3.1.1 Putative alternative exon A.

The three DNA sequence variants identified within putative exon A were all identified within single patient samples. Given the lack of data regarding the function of this putative exon and the likely insignificance of a minor population study, no control data regarding to these alleles were obtained. The function of the above alleles could in principle be assessed genetically by comparison to a large control population; however this would be a difficult undertaking requiring the recruitment of a large control cohort and, given the difficulty in classification of both patient and control samples, it would likely be a futile exercise. In the absence of a large genetic study the variants cannot be formally classified as being associated with SOD or not. IVS1A +5c>t is the most likely to be a mutation as the putative exon A donor site shows very poor homology to the consensus sequence and may in fact be a site of regulation of alternative HESXI expression. The function of all three variants could be investigated by standard methods assaying their affect on both transcription and splicing; however these studies could not be carried out due to time constraints.

3.1.2 5’ LIM1 homology region.

The -277t>g allele was identified in 4 out of 183 patient samples screened by dHPLC-heteroduplex detection (2.2%) and although a minor population study into this allele would be unlikely to give significant results, a small population study was carried out given the potential for functional investigation into this allele in the mouse.
Given the level of -277t\(\rightarrow\)g within the patient cohort, the variant would be expected to occur in 1 control allele out of 125 control alleles. Therefore, a control sample of mixed ethnic origin were screened by dHPLC-heteroduplex detection using the WAVE system. The control cohort consisted of 56 Caucasian and 32 Afro-Caribbean individuals (kind gift of Dr J. Old, Oxford). Results were achieved from 87/88 control samples. One Caucasian control sample gave a dHPLC-heteroduplex result identical to that of the -277t\(\rightarrow\)g allele (1.1%) and DNA sequencing confirmed that control sample SC60 is a -277t\(\rightarrow\)g heterozygote. One Afro-Caribbean control sample gave a dHPLC-heteroduplex result suggestive of variance. DNA sequencing identified that the sample is a t-85g heterozygote. This variant is out with the region of homology between man and mouse.

These results suggest that -277t\(\rightarrow\)g represents a polymorphism. Although -277t\(\rightarrow\)g has been identified in a single control chromosome, it could represent a causative mutation or a susceptibility allele, as the variant would disrupt a potential binding site for the GATA family of transcription factors (GATAA – GAGAA). This is supported by evidence that the 5’ LIM1 homology region in the mouse contains a Gata-2 response element (K. Mahon, personal communication). In collaboration with Dr. K. Mahon (National Institutes of Health, Bethesda, USA) transgenic mice carrying -277t\(\rightarrow\)g are being generated in an attempt to investigate the function of -277t\(\rightarrow\)g.

3.1.3 Exons 1-4

3.1.3.1 Previous control data.

The homeodomain encoding exons of HESXI (exons 2-4) have been sequenced in 70 control samples (140 chromosomes) of mixed Caucasian and Asian origin (Brickman et al., 2002). No DNA sequence variants were identified (Dr M. Dattani, personal communication). In addition, exon 1 has been sequenced in 50 Australian control samples (100 chromosomes) (Thomas et al., 2001). No DNA sequence variants were identified (Dr P. Thomas, personal communication). Furthermore, exons 1-4 have been sequenced in 20 German control samples (N. Haute, unpublished data). A single control sample contained a DNA sequence variant, 525g\(\rightarrow\)a in the heterozygous state (N. Haute, personal communication). Overall, no DNA sequence variants have been found within exon 1 from 140 control chromosomes and, with the exception of the 525g\(\rightarrow\)a allele, no DNA sequence variants have been found within exons 2-4 from 180 control chromosomes.
3.1.3.2 Silent variants.

Of the silent DNA sequence variants identified within *HESXI* exons 1-4 (183t>c, 219c>t, and 525g>a) only 525g>a has been identified within a control sample (1/180 alleles). Control data suggests that these alleles are rare polymorphisms. 183t>c appears to be a non-functional polymorphism, given the occurrence of 183t>c within a patient who was later shown to carry a causative GH deletion (J. Turton, personal communication). 219c>t does not result in an amino-acid sequence change and has not been found in control chromosomes. Although no functional defect is apparent, 219c>t cannot be ruled out as a causative mutation at this stage. 525g>a appears to be a polymorphism as this variant has been identified in a single control sample.

3.1.3.3 Coding variants.

The occurrence of Y90H within a parent of a patient suggests that this change represents a rare polymorphism. However a non-septo-optic dysplasia phenotype cannot be formally ruled out as the referring endocrine centre did not provide this information.

Except for N125S, the three DNA sequence variants within *HESXI* exons 1-4 which change the amino acid sequence were identified within single samples. Neither of these variants has been identified in 180 control chromosomes, which suggests that these variants are rare polymorphisms. N125S on the other hand was identified in a number of patient samples (n=8), all of which are of Afro-Caribbean extraction. No *HESXI* control data derived from the Afro-Caribbean population were available and given the frequency of the allele within the patient cohort a study of N125S within the Afro-Caribbean population was carried out.

Genomic DNA from Afro-Caribbean individuals was obtained (a kind gift of Dr J. Old, Oxford). This consisted of sickle-cell anaemia patients from the Commonwealth, of predominantly Nigerian and Jamaican extraction. Allele specific primer digestion was performed to detect N125S. Briefly, an oligonucleotide primer complementary to the sequence immediately adjacent to the N125S site and containing a partial *HinfI* restriction endonuclease restriction was obtained. The oligonucleotide was designed such that PCR amplified DNA from N125S DNA would contain a complete *HinfI* endonuclease restriction site, while PCR amplified DNA from WT DNA would contain a single mismatch. 42 control samples were amplified (84 chromosomes), digested with *HinfI* and the restriction digest products visualised by agarose gel electrophoresis. Based on this data, a N125 level of 0.64
and a S125 level of 0.36 was calculated and the null hypothesis of Hardy-Weinberg equilibrium could not be rejected; \( \chi^2 = 0.06 \) and \( p = 0.97 \) (Table 3.1.3.3a).

<table>
<thead>
<tr>
<th>N125/N125</th>
<th>N125/S125</th>
<th>S125/S125</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>20</td>
<td>5</td>
</tr>
</tbody>
</table>

(Table 3.1.3.3a) Allele specific primer digestion of N125S within Afro-Caribbean controls.

Thus, N125S appears to be a non-selective polymorphism, at least within the Afro-Caribbean population. This is in agreement with the results of Brickman et al. (2002) who were unable to demonstrate that N125S results in a defect of HESX1 function.

Of the sequence variants identified within SOD patients three lead to changes in HESX1 amino-acid sequence; N125S (position 18 of the homeodomain), E149K (position 42 of the homeodomain), and S170L (conserved motif C-terminal to homeodomain). E42 is completely conserved across the ANF class suggesting that E149K represents a mutation, whilst serine occurs at position 18 of the ANF homeodomain in *P. walti, A. baeri* and *D. rerio* suggesting that N125S represents a polymorphism (Figure 3.1.3a). S170 lies within a 6 amino acid motif (RESQFL) immediately C terminal to the homeodomain. The motif is identical in all ANF class protein and is unique to the ANF class, suggesting that S170L represents a mutation (Thomas et al., 2001). It is of note that 18 of the 60 residues within the homeodomain are not conserved across the class, a sequence divergence level of 30%.

<table>
<thead>
<tr>
<th>Species</th>
<th>Conserved Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>O. cuniculus</em></td>
<td>GRRPRFA---Q---LE---F---Y---ID-RE-la-KL-L---EDRIQFWQRRAK---KRS---</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>GRRPRFA---Q---LE---F---Y---ID-RE-la-KL-L---EDRIQFWQRRAK---KRS---</td>
</tr>
<tr>
<td><em>M. musculus</em></td>
<td>GRRPRFA---Q---LE---F---Y---ID-RE-la-KL-L---EDRIQFWQRRAK---KRS---</td>
</tr>
<tr>
<td><em>G. gallus</em></td>
<td>GRRPRFA---Q---LE---F---Y---ID-RE-la-KL-L---EDRIQFWQRRAK---KRS---</td>
</tr>
<tr>
<td><em>P. walti</em></td>
<td>GRRPRFA---Q---LE---F---Y---ID-RE-la-KL-L---EDRIQFWQRRAK---KRS---</td>
</tr>
<tr>
<td><em>A. baeri</em></td>
<td>GRRPRFA---Q---LE---F---Y---ID-RE-la-KL-L---EDRIQFWQRRAK---KRS---</td>
</tr>
<tr>
<td><em>D. rerio</em></td>
<td>GRRPRFA---Q---LE---F---Y---ID-RE-la-KL-L---EDRIQFWQRRAK---KRS---</td>
</tr>
<tr>
<td><em>X. Leavis</em></td>
<td>GRRPRFA---Q---LE---F---Y---ID-RE-la-KL-L---EDRIQFWQRRAK---KRS---</td>
</tr>
</tbody>
</table>

(Figure 3.1.3.3a) Alignment of ANF homeodomains using ClustelW. Positions 18 (N125S) and 42 (E149K) are highlighted in grey. *Xanf-l.

3.1.4 Further functional investigations.

Based upon the occurrence of N125S within the Afro-Caribbean population at Hardy-Weinberg equilibrium, N125S appears to be a benign polymorphism. The finding of N18 and S18 within other ANF class proteins strengthens this view. Therefore further functional
characterisation of this variant was not undertaken, especially in the light of the findings of Brickman et al. (2002). The functional effect of S170L had previously been investigated (see section 4.3), and thus was not investigated further. Although the -227t>g allele is found in both patient and control samples, the mutation effects a potential binding site for the GATA family of transcription families and GATA-2 has been placed upstream of Hesx1 in the mouse (Dr K. Mahon, personal communication). One possible investigation path would involve the generation of the -277t>g allele in the mouse Hesx1 gene and assessing the function of this change. This would be a major undertaking out with the scope of this study. Therefore, all data regarding the -277t>g allele was donated to Dr K. Mahon (National Institutes of Health, Bethesda, USA) who has undertaken this investigation. E149K affects a completely conserved amino acid residue and has not been found in 180 control chromosomes. Additionally, two mutations at this residue (E42A and E42K) within a second paired-like transcription factor (CRX) have been shown to be causative in dominant cone-rod retinal dystrophy (Freund et al., 1997 and Sankila et al., 2000). Therefore, only the functional significance of E149K was investigated further.

3.2 E149K

3.2.1 The effect of E149K on the HESX1 homeodomain.

Glu at homeodomain position 42 (E149) is conserved between HESX1 (Paired-like) and Paired, as well as across the ANF class. As the structure of the Paired (Prd) protein bound to target DNA is known (Wilson et al., 1995 - pdb code IFJL) it was used as a model to investigate the consequences of E149K with the assumption that HESX1 bound to DNA as a dimer is likely to occupy a conformation similar to that of Prd. Glu42 (helix III/IV) does not contact DNA, although it does form an intra-monomer salt-bridge with Arg31 (helix II). In the Prd bound dimer structure, Glu42 forms an additional inter-monomer salt-bridge with Arg3 (flexible N-terminal arm) (Figures 3.2.1a and 3.2.1b).
(Figure 3.2.1a) Prd Glu42 (red) forms an inter-monomer salt-bridge with Arg3 (purple) and an intra-monomer salt-bridge with Arg31 (yellow).

(Figure 3.2.1b) Prd Glu42 (centre) forms an inter-monomer salt-bridge with Arg3 (right) and an intra-monomer salt-bridge with Arg31 (left), a space fill view.
Recently, the structure of the HESX1 homeodomain bound to DNA as a monomer has been determined by NMR (de la Mata et al., 2002). The solved monomeric structure clearly shows the interaction between Glu42 and Arg31 (Figure 3.2.1c). As a monomeric structure, no inter-dimer interaction between Glu42 and Arg3 exists. Although Arg3 is resolved in the structure, it is disorganised (de la Mata et al., 2002 personal communication). The N-terminal arm (residues 1-6) occupies a position quite different from that observed in the bound paired dimer, in that it has a bend which points away from the location occupied by a second monomer in the paired dimer.

(Figure 3.2.1c) HESX1 Glu42 (red) forms an intra-monomer salt-bridge with Arg31 (yellow). Arg3 is shown in purple.

3.2.2 Allelic tracking.

DNA was obtained from all available family members of patient 4769 and the E149K status of each individual determined (Figure 3.2.2a). Allelic tracking studies were carried out as per Thomas et al. (2001).
Four polymorphic markers flanking \textit{(HESX1 D3S1606, D3S3616, D3S1295, and D3S1300)} (HuGO) were used to tract the mutation in all family members. Oligonucleotide primers for each marker were commercially obtained using the sequence given by Dr Jean Weissenbach’s protocol (http://www.gdb.org). Each forward oligonucleotide was labelled with HEX. PCR amplification was carried out as per Dr Jean Weissenbach’s protocol (http://www.gdb.org). The products of each reaction were visualised using an ABI 377 (ABI Prism), using the GeneScan software (ABI Prism). The product size from each marker in each individual was recorded and each size given an arbitrary label (Table 3.2.2a). The E149K carrying allele was then determined and the inheritance of the allele through the pedigree tracked (Figure 3.2.2b). The results do not exclude compound heterozygosity.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
 & D3S1606 & D3S3616 & D3S1295 & D3S1300 \\
\hline
I.1 & 206 (1) & 209 (1), 213 (3) & 135 (2) & 236 (5) \\
\hline
I.2 & 206 (1) & 217 (5) & 125 (1), 135 (2) & 218 (1), 226 (2) \\
\hline
II.1 & 206 (1) & 209 (1), 217 (5) & 135 (2) & 226 (2), 236 (5) \\
\hline
II.2 & 206 (1) & 213 (3), 217 (5) & 135 (2) & 226 (2), 236 (5) \\
\hline
II.3 & 206 (1) & 213 (3), 217 (5) & 125 (1), 135 (2) & 218 (1), 236 (5) \\
\hline
III.1 & 206 (1) & 209 (1), 215 (4) & 125 (1), 135 (2) & 226 (2), 228 (3) \\
\hline
III.2 & 206 (1) & 213 (3), 217 (5) & 125 (1), 135 (2) & 226 (2), 230 (4) \\
\hline
III.3 & 206 (1) & 211 (2), 213 (3) & 125 (1), 135 (2) & 236 (5), 238 (6) \\
\hline
\end{tabular}
\caption{Allele tracking data for E149K pedigree.}
\end{table}
3.2.3 DNA binding studies.

To investigate the function of E149K on the DNA binding ability of HESX1, in-vitro electrophoretic-mobility shift assays (EMSA) using recombinant HESX1 protein and artificial probe DNA was carried out as per Dattani et al. (1998). Unlike Dattani et al. (1998), recombinant HESX1 was produced using the pET30a vector (Novagen) and the BL21(DE3) (Novagen) strain of *E. coli*, as opposed to the pJBME21 vector. The pseudo-binding affinity determination procedure used by Dattani et al. (1998) where both monomeric and dimeric species are combined into a single bound species was not carried out as the possibility of a dimerisation defect resulting from disturbance of the interaction between Glu42 and Arg3 invalidates one of the assumptions upon which procedure is based, namely that the monomeric-dimeric dissociation constant is equal between the test protein and WT.
3.2.3.1 Recombinant HESX1 production.

The use of the pJBME21 vector by Dattani et al. (1998) to produce recombinant HESX1 resulted in two distinct protein species (Dattani, unpublished data). This raises the possibility that putative monomer and dimer bands seen by Dattani et al. (1998) do not represent HESX1 monomers and homodimers, but in fact represent the two recombinant protein species. To eliminate this potential source of error, the pET30a vector (Novagen) and the BL21(DE3) (Novagen) strain of *E. coli* were used to produce recombinant HESX1 protein.

A four-way site-directed mutagenesis strategy was undertaken using *HESX1* cDNA template combined with one set of oligonucleotide primers containing complementary sequence to that surrounding the E149K genomic sequence as well as the E149K variant sequence (E149Kforw and E149Krev) and a second set of primers containing complementary sequence to the extreme 5' and 3' regions of translated *HESX1* cDNA in addition to adapter sequence allowing in-frame insertion into the pET30a vector (Table 3.2.3.1a and Figure 3.2.3.1a).

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP16 5'</td>
<td>5'-TCG GAA TTC ACC ATG GTG TCT CCC AGC CTT CAG GAA -3'</td>
</tr>
<tr>
<td>VP16 3'</td>
<td>5'-CGC GGG ATC CTT CCA GCA GAT TTG TGT T -3'</td>
</tr>
<tr>
<td>E149KForw</td>
<td>5'-AAA TTG AAT CTA GAG AAA GAC AGA ATC CAG -3'</td>
</tr>
<tr>
<td>E149Krev</td>
<td>5'-CTG GAT TCT GTC TTT CTC TAG ATT CAA TTT -3'</td>
</tr>
</tbody>
</table>

(Table 3.2.3.1a) Site-directed mutagenesis oligonucleotide primers.
ttcacccatgg tgtctccccag ccttcaggaa ->
a tgtctccccag ccttcaggaa ggcgctcagc tcggggaaaa
caaacccctca acttgctcct ttcaatgga gagaatctta ggactggacc
agaagaaaga ctgtgttcca ttaatgaaac cccacagccc tgggcagac
acctgcagct catcagggaa agatggtaac ttatgtctac atgtccccaa
ctgccgagc gggatctccat tcctctaggt ggtggatcag ccaatgccag
aagaaagagc ttcgaatatg gaaaaattact ttccccacctc agaaagactg
tcctggaaaa gagaatgttga tgtgtataga ggcgaagac caagaactgc
ttttactca aaccagattg aagtgttaga aaatgtctttt agagtaaact
  aa attgaatcta E14KForw
gctatcctgt tatcctgatatt agagaagact tagctcaaaa atttgaatcta
  <-tt taacttagat
  gagaaagaca gaatccag ->
gaggaagaca gaatccagat ttgttttcaaa aatgcgcgtg cacaactgaa
cctctctgtcttaggtc E14KRev
  *
aagttcccat agagaatcac agtttctaat ggcgaaaaa aatttcaaca
  <-ttgt VP16 3'
caaatctgtcg gaaa
gtttagagca cttctaggg cgc

(Figure 3.2.3.1a) Site-directed mutagenesis of HESX1 cDNA.

130
The site-directed mutagenesis was carried out in three steps using a standard PCR protocol utilizing a proofreading Taq polymerase. Firstly, PCR using oligonucleotide primers E149Kforw and VP16 3’ with a HESX1 cDNA template was carried out. Secondly, PCR using oligonucleotide primers E149KRev and VP16 5’ with a HESX1 cDNA template was carried out. Finally, the products of these two reactions were purified by agarose gel electrophoresis, combined, and used as the template for a third PCR using oligonucleotide primers VP16 5’ and VP16 3’.

The mutant PCR product was then cleaved with BamHI and NcoRI and treated with calf intestinal phosphatase (CIP) as per manufacturer’s instructions (NEB). The digested insert was then ligated to similarly digested pET30a by incubation with T4 DNA ligase at 16°C for 8 hours. The ligated plasmid was then transformed into XL1-Blue competent cells (Invitrogen) as per manufacturer’s instructions. Transformed E. coli were then grown on LB Agar plates containing kanamicin as a selective agent. 6 colonies were then picked and a mini-prep of each produced using a mini-prep kit (Qiagen) as per manufacturer’s instructions. The sequence of both the insert and the vector surrounding the insertion site was elucidated using vector sequencing oligonucleotide primers as per manufacturer’s instructions (SP6 and T7T). A maxi-prep of a single plasmid [pET-30a(E149K)] carrying the E149K mutation was carried out using the maxi-filter kit (Qiagen) as per manufacturer’s instructions.

Wild-type recombinant HESX1 insert was produced as above by PCR using oligonucleotide primers VP16 5’ and VP16 3’ with wild type cDNA as the template. The insert was then treated as above. A maxi-prep of a single plasmid [pET-30a(WT)] carrying the WT sequence was performed using the maxi-filter kit (Qiagen) as per manufacturer’s instructions.

Production and purification of recombinant protein was performed as per Dattani et al. (1998), adjusted for the pET30a vector. Plasmid DNA was transferred into the BL21(DE3) strain of E. coli (Novagen). A 3ml starter culture was grown in selective LB media to an OD600 of ~0.5 and then added to 100ml of selective LB media. The 100ml culture was grown to an OD600 of 0.75 and then 100mM IPTG was added to a final concentration of 1mM. Before induction 1ml of uninduced culture was taken and stored at 4°C. After induction for 2.5hrs, a 1ml aliquot of induced culture was removed. Both the induced and uninduced samples were centrifuged and pellet resuspended in 1x SDS buffer. The samples were then denatured at 90°C and separated by acrylamide gel electrophoresis and induction evaluated (Figure 3.2.3.1b).
Following the demonstration of successful induction, large scale production of recombinant protein was undertaken. Two 1L cultures of selective LB media were grown to a OD<sub>600</sub> of 0.3-0.5, when 100mM IPTG was added to a final concentration of 1mM. Prior to induction, a 1ml aliquot of uninduced culture was removed. The induced cultures were grown for 2.5 hours at 30°C when a 1ml aliquot of induced culture was removed and the rest of the culture centrifuged and the pellets stored at -80°C.

The pellet was then resuspended in bugbuster plus benzonase (Calbiochem) + 1mM PMSF (Sigma) and rocked for 20 minutes at room temperature. A sample aliquot was removed and stored prior to centrifugation at 10,000rpm for 30 minutes at 4°C. The supernatant was removed and a sample aliquot stored. The pellet was re-dissolved in buffer A1 using a bouncer and was stirred for 1 hour at room temperature before a sample aliquot was removed. The insoluble material was removed by centrifugation at 10,000rpm for 20 minutes at room temperature and an aliquot of supernatant stored. A small amount of each sample was resuspended in 1x SDS buffer then denatured at 90°C before separation by acrylamide gel electrophoresis and the progress of purification assayed (Figure 3.2.3.1c).
Meanwhile a Ni-NTA His-Bind resin column (Novagen) was prepared and equilibrated with buffer A1. The supernatant was then bound to the column and washed with buffer A1. The bound protein was then eluted using buffer A2 at a flow rate of 1ml per minute. 30 x3ml samples were serially collected. A small amount of each sample was resuspended in 1x SDS buffer then denatured at 90°C before separation by acrylamide gel electrophoresis and aliquots containing high concentrations of recombinant protein noted (Figure 2.3.4.1d).

These aliquots were then dialysed against buffer B1, 2/3rds of which were replaced by buffer B2 every 4 hours until an approximate Urea concentration of 0.3M was achieved when the protein solution was aliquoted against buffer B2 for 4 hours. The protein solutions were then removed and flash frozen in 1ml aliquots using liquid nitrogen.

The concentration of the recombinant protein was determined using the Bio-Rad protein assay kit (Bio-Rad) and a BSA standard, as per manufacturer’s instructions. The
concentration of recombinant protein was low and therefore the recombinant protein was concentrated using YM-10 Microcon columns as per the manufacturer’s instructions (Microcon). The final concentration of the E149K protein was determined to be 2.47μM while the concentration of the wild-type protein was determined to be 4.16μM. The final quality of purified proteins was assessed by SDS-PAGE (Figure 2.3.3.1e).

(Figure 2.3.3.1e) Final SDS-PAGE results showing equal volumes of recombinant E194K and WT HESX1. MW = molecular weight standards.

### 3.2.3.2 EMSA

EMSAs were carried out using a slab electrophoresis chamber (ATTO) holding a 6% polyacrylamide gel. A 14-well comb was used for all EMSAs. With the exception of the electrophoresis equipment, each EMSA was prepared and carried out as per Dattani et al. (1998).

First, a 6% polyacrylamide gel was poured (6ml acrylamide, 3ml 5x TBE, 3ml 10% glycerol, 300μL 10% APS, 30μL Temed and 17.7ml mQH₂O). While the gel set dilution buffer (10mM Tris pH8, 50mM KCl, 10% glycerol, 1mg/ml bovine serum albumin, 1mM EDTA and 1mMDTT) and reaction buffer (10mM Tris pH8, 50mM KCl, 10% glycerol, 1mg/ml bovine serum albumin, 1mM EDTA and 1mMDTT) were prepared. Then the set polyacrylamide gel was pre-run at 100V in 0.5xTBE buffer with 1% glycerol at 4°C for a minimum of 1hr. While the gel was prerunning, recombinant protein was diluted to the required concentrations using dilution buffer, probe DNA was labelled with ³²p-CTP (Amersham Pharmacia) using the ready-to-go DNA labelling beads (-dCTP) as per manufacturer’s instructions (Amersham Pharmacia). Unincorporated dCTP was removed using Micropin G-25 columns (Amersham Pharmacia). The required amount of probe was then added to the reaction buffer and 10μL aliquoted into each of 14 0.5ml tubes. 10μL of
diluted proteins (2x final concentration) was added to the appropriate tubes and the reaction incubated for 30mins at room temperature. 2μL of loading dye was then added to each tube and the entire contents loaded into the appropriate well of the pre-run gel. The gel was then run at 180V at 4°C for 45mins, followed by 250V at 4°C until the dye front had migrated the required distance down the gel.

After electrophoresis, the gel was fixed and dried using a gel drying film system as per manufacturer’s instructions (Promega). The results were then visualised by auto-radiography and the film developed using an X-OGRAPH automatic film developer (Imaging systems). If after development film exposure time was too short or too long, exposure was repeated for an adjusted duration.

A GS-800 calibrated densitometer system (Bio-Rad) was used to quantify the results. The developed film was scanned and the data extracted from the resulting image. A standard sized box was placed over each probe species within each lane as well as a region of the lane without any apparent signal as a background control. The signal density of each box was then calculated and the data extracted.

The species density data was then entered into a Microsoft Excel worksheet. The background density was then subtracted from each species density and the adjusted density transformed into a percentage of total species density. The fastest migrating species was assumed to represent free probe, while the slowest migrating species was assumed to represent dimeric bound probe with the median species assumed to represent monomeric bound probe.

Results of an initial EMSA using a probe concentration of 200nM over a large protein concentration gradient (1nM – 1024nM) shows that the binding profile of E149K is dissimilar to that of WT (Figure 3.2.3.2a) protein. At 16nM of E149K retarded probe can be seen while no retarded probe can be seen in the WT lane. Similar results can be seen for 64nM where E149K retarded probe can be seen whereas WT retarded probe cannot be seen. At 256nM and 1024nM retarded probe can be seen in both WT and E149K lanes. Across the gradient of E149K two distinct retarded probe species can be seen whereas two retarded probe species are difficult to resolve in WT lanes. Retarded probe occurs at lower concentration of E149K compared to WT. However, at 1024nm the amount of retarded probe is clearly greater in the WT lane than the E149K lane, suggesting the presence of more bound DNA.
Following these results, an EMSA using a probe concentration of 2nM over a small protein concentration gradient (32nM – 192nM) was carried out (Figure 3.2.3.2b). The density of each probe species was determined and an intra-lane background reading subtracted. A plot of the % of the total probe signal of each species was plotted across the protein gradient with the assumptions that the slowest migrating species is dimeric bound probe, the fastest is free probe and the medial species is monomeric bound probe (Figure 3.2.3.2c). Finally, the ratio of monomer to dimer was determined by dividing the %monomer by the %dimer and a plot of the ratio across the protein gradient was produced, (Figure 3.2.3.2d).
(Figure 3.2.3.2.b) EMSA using a protein concentration gradient of 32nM - 192nM with a probe concentration of 2nM. From left to right: lane 1 = 5μL Dattani WT positive control, lane 2 = 192nM WT, lane 3 = 160nM WT, lane 4 = 128nM WT, lane 5 = 96nM WT, lane 6 = 64nM WT, lane 7 = 32nM WT, lane 8 = 192nM E149K, lane 9 = 160nM E149K, lane 10 = 128nM E149K, lane 11 = 96nM E149K, lane 12 = 64nM E149K, lane 13 = 32nM E149K, and lane 14 = probe only.
Figure 3.2.3.2c  Proportion of each probe species from the EMSA using a protein concentration gradient of 32nM - 192nM with a probe concentration of 2nM.
32nm-192nm ratio of monomer to dimer.

(Figure 3.2.3.2.d) The ratio of monomer to dimer from the EMSA using a protein concentration gradient of 32nM - 192nM with a probe concentration of 2nM.

From visual inspection, two clear and distinct retarded probe species can be seen across the E149K lanes while only a single species can be seen across the WT lanes. The signal density of each visually identifiable species was measured, and where a probe species could not be visually seen the signal density was measured in the region were the probe species would be expected to occur (i.e. WT assumed monomer). The level of the slower migrating retarded probe species (assumed dimer) rises across the protein concentration gradient and the level of E149K dimer is consistently higher than that of WT dimer. As a consequence the level of un-bound probe falls as each protein concentration rises with the level of WT un-bound probe consistently higher than that of E149K. The ratio of monomer to dimer of E149K across the gradient is non-zero and is of an approximately similar value, while the ratio of monomer to dimer of WT is non-zero only at 160nM and 192nM. The ratio of monomer to dimer is higher for E149K than WT across the gradient.
Following this EMSA, a third EMSA using a probe concentration of 2nM over a higher protein concentration gradient (128nm – 448nm) was carried out (Figure 3.2.3.2e). The density of each probe species was determined and an intra-lane background reading subtracted. A plot of the % of the total probe signal of each species was plotted across the protein gradient (Figure 3.2.3.2f), again with the assumption that the slowest migrating species is dimeric bound probe, the fastest is unbound probe and the medial species is monomeric bound probe. Finally, the ratio of monomer to dimer was determined by dividing the %monomer by the %dimer and a plot of the ratio across the protein gradient was produced (Figure 3.2.3.2g).

(Figure 3.2.3.2e) EMSA using a protein concentration gradient of 128nM - 448nM with a probe concentration of 2nM. From left to right: lane 1 = probe only, lane 2 = 448nM WT, lane 3 = 384nM WT, lane 4 = 320nM WT, lane 5 = 256nM WT, lane 6 = 192nM WT, lane 7 = 128nM WT, lane 8 = 448nM E149K, lane 9 = 384nM E149K, lane 10 = 320nM E149K, lane 11 = 256nM E149K, lane 12 = 192nM E149K, and lane 13 = 128nM E149K.
Figure 3.2.3.2f. Proportion of each probe species from the EMSA using a protein concentration gradient of 128nM - 448nM with a probe concentration of 2nM.
From visual inspection, two clear and distinct retarded probe species can be observed across the E149K lanes whereas the distinction between the two species across the WT lanes is less clear. By accurate measurement of the signal density of each probe species for each lane, a low level of a faster migrating (assumed monomer) retarded probe species can be detected in all WT protein concentration lanes except for the lowest concentration (128nM). Monomeric species can be seen in all E149K lanes. The level of the WT monomer would appear to be consistently below that of E149K. The amount of the slower migrating retarded probe species (assumed dimer) rises across the protein concentration gradient and the amount of E149K dimer is consistently higher than that of WT dimer. The ratio of E149K monomer to E149K dimer across the gradient is non-zero, while the ratio of WT monomer to WT dimer of is non-zero except at 128nM. Again the ratio of monomer to dimer is higher from E149K than WT across the gradient.
3.2.4 *In-vitro* transfection studies.

*In-vitro* transfection studies were carried out as per Brickman *et al.* (2001).

### 3.2.4.1 Construct production.

A modified pJBXi plasmid (Emami and Carey, 1992) encoding a wild-type GAL4-HESX1 fusion protein, pJBME[WT], was obtained (a kind gift of Dr. J. Brickman). E149K *HESX1* insert was produced by PCR across the insert of pET-30a[E149K], using a proof-reading *Taq* polymerase. The PCR product was purified by Agarose gel electrophoresis. The purified PCR product was then digested with EcoR1 and BamH1 as per manufacturer’s instructions (NEB) to produce digested E149K insert.

Meanwhile, a maxi-prep of pJBME[WT] was carried out using the QIAfilter plasmid maxi kit (Qiagen) as per manufacturer’s instructions. An aliquot of pJBME[WT] was then digested with EcoR1 and BamH1 as per manufacturer’s instructions (NEB) to remove the WT HESX1 insert and to produce digested pJBME vector. The digested E149K insert was then ligated into the digested pJBME vector by incubation with T4 DNA ligase at 16°C for 8 hours. The ligated plasmid was then transformed into XL1-Blue subcloning grade competent cells (Stratagene) as per manufacturer’s instructions. Transformed *E.coli* were then grown on LB Agar plates containing Ampicillin as a selective agent. 6 colonies were then picked and a mini-prep of each produced using a mini-prep kit (Qiagen) as per manufacturer’s instructions. The sequence of both the insert and the vector surrounding the insertion site was elucidated using oligonucleotides HCDNA1F, HCDNA1R, HCDNA3F, HCDNA3R and compared to that of pJBME[WT]. Finally, a maxi-prep of a single plasmid (pJBME[E149K]) carrying the E149K mutation was carried out using a kit (Qiagen) as per manufacturer’s instructions.

A GAL4 reporter construct (pG5SV40) was obtained (a kind gift of Dr. J. Brickman) (Figure 3.2.4.1a).

![5 x 17mers SV40 Luc](image)

*Figure 3.2.4.1a* Cartoon of the GAL4 reporter construct (pG5SV40).
### 3.2.4.2 Transfections.

COS-7 cells were cultured in DMEM supplemented with 10% fetal calf serum, penicillin plus streptomycin, and L-gluconamide. DNA was transfected by lipofection using Lipofectamine (Gibco BRL), as per manufacturer’s instructions.

In total, 1.2μg of DNA was transfected. This consisted of 100ng of pG5SV40, 75ng of pRenilla (Promega) and 1025ng of vector DNA. Vector DNA consisted of the required amount of pJBME[WT], pJBME[E149K], or pJBXi (empty vector) supplemented with pBSV to 1025ng total DNA. The Renilla Luciferase Assay System was used to measure Luciferase activity as per manufacturer’s instructions (Promega).

Three *in-vitro* transfection experiments were independently performed, each in triplicate (Table 3.2.4.2a).

<table>
<thead>
<tr>
<th></th>
<th>pBSV</th>
<th>pBSV</th>
<th>pBSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100ng pJBXi</td>
<td>100ng pJBXi</td>
<td>100ng pJBXi</td>
</tr>
<tr>
<td>2</td>
<td>200ng WT</td>
<td>100ng WT</td>
<td>100ng WT</td>
</tr>
<tr>
<td>3</td>
<td>400ng WT</td>
<td>200ng WT</td>
<td>200ng WT</td>
</tr>
<tr>
<td>4</td>
<td>600ng WT</td>
<td>400ng WT</td>
<td>400ng WT</td>
</tr>
<tr>
<td>5</td>
<td>800ng WT</td>
<td>600ng WT</td>
<td>600ng WT</td>
</tr>
<tr>
<td>6</td>
<td>100ng E149K</td>
<td>100ng E149K</td>
<td>100ng E149K</td>
</tr>
<tr>
<td>7</td>
<td>200ng E149K</td>
<td>200ng E149K</td>
<td>200ng E149K</td>
</tr>
<tr>
<td>8</td>
<td>400ng E149K</td>
<td>400ng E149K</td>
<td>400ng E149K</td>
</tr>
<tr>
<td>9</td>
<td>600ng E149K</td>
<td>600ng E149K</td>
<td>600ng E149K</td>
</tr>
<tr>
<td>10</td>
<td>800ng E149K</td>
<td>800ng E149K</td>
<td>800ng E149K</td>
</tr>
</tbody>
</table>

*(Table 3.2.4.2a) In-vitro transfection experiments. 12 separate transfections were performed in triplicate within each experiment and three separate experiment were carried out.*

For each separate experiment, the dual Luciferase data from each well was entered into a Microsoft Excel worksheet and the dual Luciferase ratio calculated for each well. The mean dual Luciferase ratio for the 100ng pJBXi (empty vector) transfections was calculated and all other dual Luciferase ratios were converted to a percentage of the mean pJBXi dual Luciferase ratio. The mean of each *in-vitro* transfection type (pooling the independent experiments) was calculated. Similarly, the standard deviation of each *in-vitro* transfection
type (pooling the independent experiments) was calculated. Based upon these figures, the mean percentage of empty vector control (pJBXi) for each *in-vitro* transfection type was plotted with 95% confidence levels (Figure 3.2.4.2a).

(Figure 3.2.4.2a) The mean dual Luciferase ratio of each transfection type as a percentage of pJBXi control with 95% confidence levels.

GAL4-HESX1 [WT] transfections show a dose dependent reduction in Luciferase activity, suggesting that GAL4-HESX1 [WT] is able to repress transcription from the SV40 promoter. GAL4-HESX1 [E149K] transfections do not show a dose dependent reduction in Luciferase activity. Paradoxically, at a concentration of 800ng GAL4-HESX1 [E149K] gives rise to a significantly higher Luciferase activity, suggesting that at this level GAL4-HESX1 [E149K] activates transcription at the SV40 promoter.
3.2.5 Western blotting.

*In-vitro* transfections were carried out as above, scaled up for the use of a 6-well format, with each transfection done in triplicate. Four transfections were carried out: cells only, pJBXi only, pJBME [WT], and pJBME [E149K]. Transfected cells were harvested by scraping in 500µL ice-cold RIPA buffer (50mM Tris-Cl (pH 7.6), 150mM NaCl, 1% NP-40, 0.5% DOC, 0.1% SDS, 1mM DTT*, 1mM PMSF*, *added fresh before use) followed by collection in an Eppendorf tube. An addition 200µL ice-cold RIPA buffer was washed over the cells to remove any residue and then combined with the harvested cells. The harvested cells were left for 30mins on ice to lyse followed by passing the solution through a 21g needle to shear genomic DNA. The lysate was then centrifuged at 15,000g for 20mins at 4°C. The supernatant was then taken, aliquoted and stored at -80°C.

Cellular proteins were separated by a 12.5% gel SDS-PAGE gel using the Bio-Rad Mini Protean II (Bio-Rad). 2x Laemml gel loading buffer was added to 20µL of each transfection lysate, followed by the samples being heated to 99°C for 5mins. The samples were then loaded onto the gel alongside a lane consisting of prestained size standards (Prestained SDS-PAGE Standards, Low Range, BioRad). Electrophoresis was carried out at 30mA until the dye front approached the bottom of the gel.

Two Gal-4 binding domain antibodies were available, murine monoclonal GAL4 BD mAb (Clontech) and rabbit polyclonal GAL4 DBD (Santa Cruz). Both were tested by comparing pJBME [WT] protein to a positive control (a kind gift of Dr. K. Woods). The rabbit polyclonal antibody produced a better signal than the murine monoclonal antibody (Figure 3.2.5a). The pJBME [WT] protein was less concentrated than positive control.

*(Figure 3.2.5a)* Antibody testing of murine monoclonal antibody (left) and rabbit polyclonal antibody (right). KD = kilo Dalton, MW = molecular weight standards, WT = pJBME [WT] and + = positive control.
The protein extracts of each transfection were separated by SDS-PAGE alongside the positive control and were detected using the rabbit polyclonal antibody (Figure 3.2.5b).

\[ \text{MW} \quad C \quad V \quad \text{WT} \quad E149K \quad + \]

(Figure 3.2.5b) Results of Western blotting of HESX1 transfected COS-7 cells. KD = kilo Dalton, MW = molecular weight standards, C = cells only, V = vector only, E149K = pJBME [E149K], WT = pJBME [WT] and + = positive control.

GAL-4 binding domain positive protein of the expected size could be detected in the WT lane but not in the E149K lane. A smaller protein was detected in the E149K lane but its significance is unclear. The cells only and vector only lanes were not resolved well. These results suggest that no GAL4-HESX1[E149K] protein occurs within the pJBME [E149K] transfected cells. However, the results of this Western are inconclusive due to a low protein concentration. Thus the experiment was repeated using a higher number of transfected cells (4 wells of a 6 well plate) (Figure 3.2.6c).

\[ \text{MW} \quad C \quad V \quad E149K \quad \text{WT} \quad + \]

(Figure 3.2.5c) Results of Western blotting of HESX1 transfected COS-7 cells. KD = kilo Dalton, MW = molecular weight standards, C = cells only, V = vector only, E149K = pJBME [E149K], WT = pJBME [WT] and + = positive control.

The results are similar in that GAL-4 binding domain positive protein of the expected size could be detected in the WT lane but not in the E149K lane. A slight signal could be detected
in the E149K lane, however a similar band could be detected in the vector only lane. To confirm this result the experiment was repeated using the murine monoclonal antibody (Figure 3.2.5d).

(Figure 3.2.5d) Results of Western blotting of HESX1 transfected COS-7 cells. KD = kilo Dalton, MW = molecular weight standards, C = cells only, V = vector only, E149K = pJBME [E149K], WT = pJBME [WT] and + = positive control.

The results were identical using the murine monoclonal antibody. In order to confirm these findings the experiment was repeated with total cellular protein extracted over a number of time points (6 hours, 12 hours, 24 hours and 36 hours), with the transfection media replaced with media + serum after the 24 hour time point (Figure 3.2.5e).

(Figure 3.2.5e) Results of Western blotting of HESX1 transfected COS-7 cells over a 36 hour time period. KD = kilo Dalton, MW = molecular weight standards, E6/E12/E24/E36 = pJBME [E149K] at 6 hours/12 hours/24 hours/36 hours respectively, V36= vector only at 36 hours and W6/W12/W24/W36 = pJBME [WT] at 6 hours/12 hours/24 hours/36 hours respectively.
Although the insert of the pJBME [E149K] construct was sequenced and found to be as expected, the pJBME [E149K] plasmid may contain a mutation that inhibits the production of GAL4-HESX1[E149K] protein. On the other hand, the E149K variant would be expected to interfere with folding of the HESX1 homeodomain which may result in degradation of the GAL4-HESX1[E149K] protein. Therefore, total RNA was extracted from transfected cells (cells only, vector only, GAL4-HESX1[E149K] and GAL4-HESX1[WT], 3 wells of a 6 well plate each) using TRIzol reagent (Gibco-BRL). Special care was taken not to disturb the DNA phase so as to minimise DNA contamination. The total RNA was treated with the cDNA Synthesis System (Roche) as per manufacturer’s instructions. This cDNA was then used as the template for PCR using HESX1 cDNA specific primers (HCDNA1F and HCDNA3R). PCR optimisation was performed as described in section 2.3.2.4 using GAL4-HESX1[WT] transfected cDNA as a template. The optimised PCR reaction was used to test for HESX1 mRNA sequences within the cDNA (Figure 3.2.5f).

![Image](image_url)

(Figure 3.2.5f) PCR results from cDNA derived from GAL4-HESX1 transfected cells.

No signal was detected from cells only cDNA or vector only cDNA. Signal of the expected size was detected from both GAL4-HESX1[E149K] cDNA and GAL4-HESX1[WT] cDNA. The experiment was repeated with identical results. To test for plasmid DNA contamination GAL4-HESX1[E149K] cDNA, GAL4-HESX1[WT] cDNA and GAL4-HESX1[WT] plasmid DNA was used to transform XL1-Blue subcloning grade competent cells (Stratagene) as per manufacturer’s instructions. The E.coli were then grown on LB Agar plates containing Ampicillin as a selective agent. Neither GAL4-HESX1[E149K] cDNA or GAL4-
HESX1[WT] cDNA produced any transformed colonies, while control GAL4-HESX1[WT] plasmid DNA produced a number of transformed colonies.
4 Mutation screening of PTX1 and PTX2.

As few HESX1 mutations were identified other candidate genes were considered which could account for the SOD phenotype. However, this produced a large number of potential SOD candidate genes. Therefore the portion of the patient cohort with additional features was examined for phenotypes strongly suggestive of particular candidate genes. Two such phenotypes were readily identifiable. The first consisted of 13 patients with cleft lip and/or palate in addition to components of SOPD, in particular pituitary defects. This group were screened for mutations in the gene PTX1. The second consisted of 8 patients with components of SOPD, in particular pituitary defects, in addition to defects of left-right asymmetry. This group was sequenced for mutations in PTX2.

4.1 PTX1.

4.1.1 Introduction.

Ptx1 (PITX1, POTX1, BFT) is a bicoid-like homeodomain transcription factor. Originally cloned in the mouse as a gene expressed in the pituitary that is able to activate the POMC gene (Lamonerie et al., 1996), it is the founder member of the pituitary OTX-related factors. In addition to its ability to activate POMC, it has been shown to have a role in limb development, craniofacial development, and in the expression of interferon α genes.

The murine Ptx1 gene is first expressed during gastrulation at E6-7 within the newly formed mesoderm at the posterior margin of the embryonic endoderm, where the primitive streak will initially appear. Expression is maintained in derivatives of this tissue, the allantois and posterior lateral mesoderm (E7.5). By E8, expression is observable in the somatic and splanchnic mesoderm. Later, by E9-9.5, expression is additionally detectable in the umbilical vessels. Expression within the posterior is limited to the lateral plate mesoderm. Following this pattern of expression, Ptx1 is expressed within the hind limb mesenchyme by E11.5 (Lanctot et al., 1997). Ptx1 is only expressed in the latter stages of fore limb development (Szeto et al., 1999). At E8, an additional zone of expression within the anterior occurs within the oral ectoderm, oral membrane, and rostral foregut, extending into a v-shaped segment within the facial ectoderm from underneath the future lens placode to the first branchial arch. At E9, Ptx1 expression appears in the first branchial arch mesenchyme. Expression is maintained until at least E10 in a number of tissues derived from these regions.
(Lanctot et al., 1997). Crucially, expression is maintained throughout anterior pituitary development and Ptxl is expressed in the adult post-natal pituitary. Expression of Ptxl initially occurs throughout the pituitary, but after E16 expression is limited to the POMC expressing cells (Lamonerie et al., 1996). Expression of the chick gene, cPtxl, is very similar to that seen in the mouse (Lamonerie et al., 1996).

At the transcription level, Ptxl has been shown to interact with a number of other transcription factors, SF-1 and Pit1 (Tremblay et al., 1998), Tpit (Lamolet et al., 2001), NeuroD1/β2 (Poulin et al., 1997), and Hesxl (Quirk and Brown 2002). The role of Ptxl in expression of αGSU and βLH is well characterised, and two target binding sites 5'−CTTA−3' (αGSU) and 5'−ATTA (βLH) have been identified (Tremblay et al. 1998 and Quirk et al., 2001). Hesxl is also able to use these two sites to repress transcription. However, Ptxl is able to bypass Hesxl monomer mediated repression at the αGSU site, whilst Hesxl cooperative dimerisation is able to block Ptxl access to the βLH site (Quirk and Brown 2002). The significance of these data remains unclear, given that HESX1 expression is extinguished before gonadotrope differentiation occurs. It might be that HESX1 acts to limit PTX1 activation of these Gonadotrope specific genes during early pituitary development.

Ptxl -/- mice show a striking phenotype where the distal hind limbs resemble the fore limbs (Szetao et al., 1999). In addition, defects of the anterior pituitary are also apparent. βLH, βFSH, βTSH and αGSU levels are diminished in Ptxl -/- mice indicating that the number of gonadotropes and thyrotropes is reduced and/or defective. There appears to be no reduction in POMC transcripts, indeed ACTH levels are increased. It is a variable phenotype and Ptxl -/- mice can also manifest midline abnormalities, such as cleft palate.

PTXI maps to 5q31 (Crawford et al., 1997) and based upon the expression and function of Ptxl, as well as, the phenotype of Ptxl -/- mice PTXI is considered to be a candidate gene for anterior pituitary defects associated with facial clefting.

A sub-population of the patient cohort (n=31) was screened for mutations in PTXI. 13 patients with cleft lip and/or palate in addition to components of SOPD and 1 patient with facial dysplasia in addition to SOPD were screened for PTXI mutations given the craniofacial phenotype of the Ptxl null mutant mouse. Similarly, 3 patients with multiple midline abnormalities were screened. As Ptxl plays an important role in pituitary development 7 patients with isolated pituitary disease were screened, as were 7 patients with SOPD. The sporadic cases (n=28) were screened by dHPLC-heteroduplex detection with all samples "spiked" with WT DNA, whilst familial cases (n=3) were directly sequenced.
4.1.2 Methods.

Oligonucleotides were derived from genomic contig gi:7708834 using identical procedures as in section 2.3.1 (Table 4.1.2a).

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ptxlx1xF</td>
<td>5'-ctc ttg ttc ccg cag gtc -3'</td>
</tr>
<tr>
<td>Ptxlx1xR</td>
<td>5'-tcg aga acg gga aaa aga aag -3'</td>
</tr>
<tr>
<td>Ptxlx2xF</td>
<td>5'-ctt tcc agt ggg gac agt tc -3'</td>
</tr>
<tr>
<td>Ptxlx2xR</td>
<td>5'-tgt ggg tct aag ctt tgg ag -3'</td>
</tr>
<tr>
<td>Ptxlx3xF</td>
<td>5'-ggt cga agg aga gaa acg ac -3'</td>
</tr>
<tr>
<td>Ptxlx3xR</td>
<td>5'-gag tgg gaa gag gga gga -3'</td>
</tr>
<tr>
<td>Ptxlx3bxF</td>
<td>5'-agc tcc atc cct cca tga cc -3'</td>
</tr>
<tr>
<td>Ptxlx3x3aR</td>
<td>5'-cgg tga ggt tgt tga tgt tg -3'</td>
</tr>
</tbody>
</table>

(Table 4.1.2a) PTX1 oligonucleotide primers.

PCR optimisation was performed as described in section 2.3.2. dHPLC-heteroduplex detection was carried out using the WAVE system using identical procedures as given in section 2.3.3. DNA sequencing was accomplished via direct cycle sequencing using the BigDye terminator kit (Perkin-Elmer), utilising identical procedures as in section 2.3.5.

4.1.2.1 dHPLC-heteroduplex using the WAVE system.

Primary sequence of the amplified regions (derived from Homo sapiens chromosome 5 clone CTC-276P9, gi:18449971) was entered into the WAVE-Maker program (Figure 4.1.2.1a).

Both gradient information and melting profile estimation were obtained, from which column running temperatures were calculated (Table 4.1.2.1a). The exon 3 melting profile indicated
that 3 column running temperatures were required for this exon (Figure 4.1.2.1b), however results of WAVE using the entire exon were poor. Therefore the exon was divided in two and run separately (Figure 4.1.2.1c and 4.1.2.1d).

(Figure 4.1.2.1b) PTX1 exon 3 melting profile.
(Figure 4.1.2.1c) PTX1 exon 3α melting profile.
(Figure 4.1.2.1d) PTX1 exon 3b melting profile.

<table>
<thead>
<tr>
<th>Amplicon</th>
<th>Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 1</td>
<td>69</td>
</tr>
<tr>
<td>Exon 2</td>
<td>66</td>
</tr>
<tr>
<td>Exon 3</td>
<td>67, 69, 71</td>
</tr>
<tr>
<td>Exon 3a</td>
<td>66</td>
</tr>
<tr>
<td>Exon 3b</td>
<td>66, 68</td>
</tr>
</tbody>
</table>

(Table 4.1.2.1a) PTX1 column running temperatures.
4.1.3 Results.

4.1.3.1 Exon 1.

dHPLC-heteroduplex detection did not identify any samples suggestive of sequence variance. A sample dHPLC-heteroduplex result is shown in figure 4.1.3.1a. Direct sequencing of familial samples did not identify any sequence variance.

(Figure 4.1.3.1a) An overlay of PTX1 exon 1 dHPLC results showing three samples 3847 (blue), 4610 (green), and 4244 (red), with wild-type (black).

4.1.3.2 Exon 2.

dHPLC-heteroduplex detection identified 1 sample suggestive of sequence variance (Figure 4.1.3.2a). A sample dHPLC-heteroduplex result is shown in figure 4.1.3.2b. Direct sequencing of the suggestive sporadic and familial samples did not identify any sequence variance.
An overlay of PTX1 exon 2 dHPLC results showing four samples 3679 (blue), 3681 (green), and 3709 (red), and 3810 (black). 3679 (blue) is an example of a false positive result.

(Figure 4.1.3.2a) An overlay of PTX1 exon 2 dHPLC results showing three samples 5004 (blue), 5025 (green), and 5064 (red), with wild-type (black).

(Figure 4.1.3.2b) An overlay of PTX1 exon 2 dHPLC results showing three samples 5004 (blue), 5025 (green), and 5064 (red), with wild-type (black).

4.1.3.3 Exon 3.

dHPLC-heteroduplex detection of exon 3 produced results which were difficult to interpret (Figure 4.1.3.3a, 4.1.3.3b, 4.1.3.3c).
(Figure 4.1.3.3a) An overlay of PTX1 exon 3 dHPLC at 67°C results showing four samples 4701 (blue), 4712 (green), 4731 (red), and 4668 (black).

(Figure 4.1.3.3b) An overlay of PTX1 exon 3 dHPLC at 69°C results showing four samples 4668 (blue), 4831 (green), 5004 (red), and 5027 (black).
To improve the signal quality, exon 3 was divided into two parts: exon 3a (Ptxlx3F-Ptxlx3aR) and exon 3b (Ptxlx3bF-Ptxlx3R). dHPLC-heteroduplex detection of exon 3a identified 10 samples suggestive of sequence variance. 3 patterns of dHPLC-heteroduplex profile were seen: Profile 3aA n=1 (Figure 4.1.3.3d), Profile 3aB n=7 (Figure 4.1.3.3d and 4.1.3.3e) and Profile 3aC n=2 (Figure 4.1.3.3e).
dHPLC-heteroduplex detection using exon 3b at both 66°C and 68°C identified 11 samples suggestive of sequence variance. 10/11 samples correspond to the samples identified during exon 3a dHPLC-heteroduplex detection. All 11 samples showed similar profiles at 66°C (Figure 4.1.3.3f). At 68°C, three separate profiles were identifiable, profile 3bA n=8, profile 3bB n=1 and profile 3bC n=2 (Figure 4.1.3.3g).
In total, exon 3 from 13 samples (11 sporadic and 2 familial) was sequenced. This identified a number of sequence variants. Within exon 3, two sequence variants were identified, dbSNP474853 and an additional single nucleotide variant, IVS2 -73g>a was found in both a heterozygous state (n= 2) and a homozygous state (n= 1). Variants within these regions were not expected to give rise to detectable heteroduplex using the WAVE system and expectedly these two variants do not correspond to specific dHPLC-heteroduplex profiles. Two novel sequence variants; A299G (896c>g) and (gggcgc)n were identified in all 11 sporadic samples, as well as, a single familial sample. 3 samples were found to contain a heterozygous R140R (417c>a) variant. Two known SNP’s, dbSNP561552 and dbSNP479632, were not found to be heterozygous within the sequenced samples.

The exon 3a and 3b profiles were found to correspond to specific variant combinations in all but 2 cases (Table 4.1.3.3a). Both the profiles 3aA and 3aB appear to be due to A299G and (gggcgc)n=2. The difference between the two is likely due to the two variants being found either in cis- or in trans-.

(Figure 4.1.3.3g) An overlay of PTX1 exon 3b dHPLC at 68°C results showing four samples 3679 (blue), 5027 (green), 4821 (red), and 4598 (black). Sample 3679 (blue) shows profile 3bA while sample 4598 (black) shows profile 3bB and samples 5027 (green) and 4821 (red) show profile 3bC.
<table>
<thead>
<tr>
<th>Profile</th>
<th>Variant combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>3aA</td>
<td>A299G and ((gggcgc)_{m=2})</td>
</tr>
<tr>
<td>3aB</td>
<td>R140R and ((gggcgc)_{m=2})</td>
</tr>
<tr>
<td>3aC</td>
<td>R140R, A299G and ((gggcgc)_{m=2})</td>
</tr>
<tr>
<td>3bA</td>
<td>A299G and ((gggcgc)_{m=2})</td>
</tr>
<tr>
<td>3bB</td>
<td>A299G and ((gggcgc)_{m=2})</td>
</tr>
<tr>
<td>3bC</td>
<td>((gggcgc)_{m=2})</td>
</tr>
</tbody>
</table>

*(Table 4.1.3.3a) PTXI exon3 dHPLC- heteroduplex profiles.*

### 4.1.3.4 Bioinformatics.

During the course of this study the sequence of contig gi:7708834 was progressively updated as the sequence contigs derived from the clone were overlapped, until a final complete sequence gi:18449971 was produced. *PTXI* is located within this genomic clone (Figure 4.1.3.4a and 4.1.3.4b).
Three SNP's are known within the PTXI gene: dbSNP474853, dbSNP561552 and dbSNP479632. On comparison, an amino acid difference between published PTXI cDNA sequence (gi:20547128) and Homo sapiens chromosome 5 clone CTC-276P9 (gi:18449971), A299G is apparent, suggesting that A299G is a polymorphism. Additionally, this comparison suggests that the (gggccc)n repeat is polymorphic as n=2 in the PTXI cDNA sequence, and n=3 in PTXI genomic DNA sequence.
4.2 **PTX2.**

4.2.1 Introduction.

PTX2 (Pitx2, Rieg, Arpl, Brxl) is a member of the pituitary OTX-related factors and, like Pitx1, is a bicoid-like homeodomain transcription factor. The gene was first identified in man as a gene involved in Rieger syndrome (Semina et al., 1996) and, independently in the mouse, as a pituitary homeodomain transcription factor (Gage and Camper, 1997). The homeodomains of PTX1 and PTX2 are 97% identical with identity extending to the C-terminus (67%). There is no significant identity between the N-terminus of the two proteins (Gage and Camper, 1997). Four Ptx2 isoforms have been identified, Ptx2a, Ptx2b, Ptx2c and Ptx2d, all of which differ at the N-terminus (Gage and Camper 1997, Schweickert et al., 1999 and Cox et al., 2002). Additionally, these isoforms show promoter specific activity and synergism with Ptx2a, Ptx2b and Ptx2c acting as transcriptional enhancers, whilst Ptx2d acts to repress the function of the other three isoforms (Cox et al., 2002).

RT-PCR shows that Ptx2 is expressed in the pituitary at E13.5 as well as in several adult rodent cell-lines (Gage and Camper, 1997). In addition to the pituitary, expression is evident in the adult brain, eye, kidney, lung, testis, and tongue. During development, Ptx2 is expressed in embryonic head tissue from E8.5 (Gage and Camper, 1997). Additionally Semina et al. (1996), identified Ptx2 expression within maxillary and mandibular epithelia, and the umbilical cord. Ptx2 is also expressed within the left lateral plate mesoderm (l-LPM) at E8.5 (Schweickert et al., 1999).

Pitx2-deficient mice with the lacZ gene inserted in the homeobox-containing exon of Pitx2 have been generated (Kitamura et al., 1999). In these mutant mice, Pitx2a and Pitx2b showed symmetrical expression, whilst Pitx2c showed asymmetrical expression. The mutant mice also showed an increase in mesodermal cells in the distal end of the left lateral body wall and an amnion continuous with the lateral body wall thickened in its mesodermal layer, resulting in a failure of ventral body wall closure. Defects of laterality involving the heart and lungs were also noted, as were defects of the ocular system resulting in enophthalmos. Schweickert et al., (1999), were able to demonstrate that in the chick and Xenopus, only the isoform Ptx2c is asymmetrically expressed within the l-LPM, heart and gut. They were able to show that during the formation of laterality Ptx2c expression is induced by nodal with subsequent auto-regulation of expression. A number of authors have expanded these findings to show that Ptx2c is a critical part of the signalling cascade that induces laterality in vertebrates (Lin et al., 1999, Yu et al., 2001, Faucort et al., 2001, and Linask et al., 2002).
Ptx2 is an important regulator of pituitary development, being necessary for initiating expansion of Rathke's pouch and maintaining expression of Hesx1 and Prop1 and for specification and expansion of the gonadotrope lineage and the Pit1 dependant lineage (Suh et al., 2002). Ptx1 shows considerable overlap with Ptx2, whilst expression of Pit1, Gata2, Egr1, and Sfl appears to be critically sensitive to Ptx2 dosage, suggesting that pituitary development is dependent upon concentration of the Ptx family (Suh et al., 2002).

Rieger syndrome is an autosomal-dominant disorder characterised by ocular anterior chamber anomalies, dental hypoplasia, and umbilical stump abnormalities (Semina et al., 1996). Mutations within human Ptx2, also known as RIEG, have been shown to be responsible for a number of cases of Rieger syndrome, resulting from either haploinsufficiency or a dominant negative effect (Semina et al., 1996 and Saadi et al., 2001). Additionally, defects of the pituitary are occasionally seen in Rieger syndrome patients. However, no RIEG mutations have been associated with a pituitary phenotype to date, although identical mutations have clear in-vitro effects on the function of Ptx2 on pituitary hormone promoters (Quentien et al., 2002). RIEG mutations are limited to the homeodomain and intron 5, with no mutations described in the Ptx2c specific exon 4.

Given the role of Ptx2 in both pituitary development and left-right axis formation a sub-group (n=8) of patients with components of SOPD, in particular pituitary defect, in addition to defects of left-right asymmetry were screened for mutations in PTX2. All samples were directly sequenced.

4.2.2 Methods.

Exons 2-6 of PTX2, which contain translated sequence, were directly sequenced from all samples. Oligonucleotides were derived from genomic contig gi:7708834 using identical procedures as in section 2.3.1 (Table 4.2.2a).

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PTX2x2F</td>
<td>5' - AAA AAC ACG CCT GAA GCC TA -3'</td>
<td></td>
</tr>
<tr>
<td>PTX2x2R</td>
<td>5' - GGA TGG CGT ACT CTT TCT GC -3'</td>
<td></td>
</tr>
<tr>
<td>PTX2x3F</td>
<td>5' - TGG GTC TTT GCT CTT TGT CC -3'</td>
<td></td>
</tr>
<tr>
<td>PTX2x3R</td>
<td>5' - CCA GAG GCG GAG TGT CTA AG -3'</td>
<td></td>
</tr>
<tr>
<td>PTX2x4F</td>
<td>5' - CAG CTT GGC TTG AGA ACT CG -3'</td>
<td></td>
</tr>
<tr>
<td>PTX2x4R</td>
<td>5' - GAG CGG GAA TGT CTG CAG GCC -3'</td>
<td></td>
</tr>
</tbody>
</table>
PCR optimisation was carried as given in section 2.3.2.4. DNA sequencing was accomplished via direct cycle sequencing with the BigDye terminator kit (Perkin-Elmer), using identical procedures as in section 2.3.3.

4.2.3 Results.

4.2.3.1 DNA sequencing.

No sequence variants were discovered during DNA sequencing.

4.2.3.2 Bioinformatics.

During the course of this study, the sequence of the human genomic clone RP11-380D23 was progressively updated as the sequence contigs derived from the clone were overlapped. For bioinformatic purposes, sequence contig version gi9247141 was used. PTX2 is located within this genomic clone. The published mRNA sequence Arpla (gi:3708006), Arplb (gi:3708008) and Arplc (gi:3708010) was used to elucidate the genomic structure of PTX2 (figs. 4.2.3.2a and 4.2.3.2b). PTX2d had not been identified when this investigation was undertaken.

<table>
<thead>
<tr>
<th>Primer (Table 4.2.2a)</th>
<th>Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTX2x5F</td>
<td>5'-CAG CTC TTC CAC GGC TTC T -3'</td>
<td></td>
</tr>
<tr>
<td>PTX2x5R</td>
<td>5'-GGC CTG TAC CTC CAC AAC AT -3'</td>
<td></td>
</tr>
<tr>
<td>PTX2x6F</td>
<td>5'-AAA AAC ACG CCT GAA GCC TA -3'</td>
<td></td>
</tr>
<tr>
<td>PTX2x6R</td>
<td>5'-GGG G TGTA GG T GG A G A A GAT -3'</td>
<td></td>
</tr>
<tr>
<td>PTX2x6aF</td>
<td>5'-AAAG AGC TTC CCC TTC TTC AAC TCT A -3'</td>
<td></td>
</tr>
<tr>
<td>PTX2x6aR</td>
<td>5'-AGG GGA AAA CAT GCT CTG TGA -3'</td>
<td></td>
</tr>
<tr>
<td>PTX2xNewR</td>
<td>5'-AAAT TCA GCG ACG GGC TAC T -3'</td>
<td></td>
</tr>
</tbody>
</table>

(PTX2 oligonucleotide primers.)
7020  agtggacttt ggtggaatct ctctgtgaagt caagtcacac cccacacgga gtggagcag
7026  agggagagaa gaggagtgag agggagggag aggagagaga gtcggagacc gaggagaaaa
7032  ggggccgcca gacagcagct tcctctcgtg aaagccgctc cccacacgga gtggagcag
7038  cctgggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7044  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7050  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7056  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7062  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7068  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7074  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7080  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7086  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7092  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7098  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7104  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7110  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7116  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7122  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7128  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7134  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7140  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7146  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7152  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7158  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7164  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7170  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7176  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7182  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7188  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7194  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7200  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7206  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7212  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7218  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7224  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7230  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7236  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7242  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7248  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7254  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7260  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7266  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7272  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7278  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7284  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7290  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7296  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7302  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7308  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7314  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7320  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
7326  gctggagagg agcagaaaga aactgccagt ggcggctaga tttcggaggc cccagtgcac
169
36001  agagccgacg ggctcactg cccgcgctac ccgcgatcct acgagctggc ctgcatgggc
36061  aatccctgcc acagagaatt ggtgttctgg cttgcagttg ccagagcaga gctaataaaa
36121  tccctaccag gccaagagcc gcgaacaggc tccaacctgt gagctcctaa caagggaaac
36181  ccgccagaga cacggaagag ttggccctcc ctgggaaacc tttgtcccgg ccctggccca
36241  gctttttccc tcctgggctc gcgcttctta caccttcttt acggttgttt cggccattca
36301  ggtctctccc acacacccta tttcctagtt ttgtgatctc cgggagcaaa gttttaatac
36361  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
36421  aagagggagt cgtgtctggc gagccgacaag cggcagattc gctttaaaac ccctagctct
36481  ggtttggttc aagaatcgtc aagggggtgc ggtttgtttg tcctgggctc gcgcttctta
36541  ccgccgctac ccacccacag aatgctactt ccctctctac ctggtgagag aagggggtgc
36601  ataaagcccc cggttagagc agagcaaaca aaaagaagaa aacaagacta aaagaaaga
36661  aagagggagt cgtgtctggc gagccgacaag cggcagattc gctttaaaac ccctagctct
36721  gccctggtat
36781  gacaccaaaat gccaaaccct cccagacttt ccgaactctg gccttttggg aagggggtgc
36841  gccctggtat
36901  tcctgggctc gcgcttctta caccttcttt acggttgttt cggccattca
36961  ggtctctccc acacacccta tttcctagtt ttgtgatctc cgggagcaaa gttttaatac
37021  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
37081  aagagggagt cgtgtctggc gagccgacaag cggcagattc gctttaaaac ccctagctct
37141  tctctgggtg gcatttgtgc aagggggtgc ggtttgtttg tcctgggctc gcgcttctta
37201  ccgccgctac ccacccacag aatgctactt ccctctctac ctggtgagag aagggggtgc
37261  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
37321  ggtctctccc acacacccta tttcctagtt ttgtgatctc cgggagcaaa gttttaatac
37381  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
37441  aagagggagt cgtgtctggc gagccgacaag cggcagattc gctttaaaac ccctagctct
37501  aagagggagt cgtgtctggc gagccgacaag cggcagattc gctttaaaac ccctagctct
37561  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
37621  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
37681  cagagggagt cgtgtctggc gagccgacaag cggcagattc gctttaaaac ccctagctct
37741  ggtctctccc acacacccta tttcctagtt ttgtgatctc cgggagcaaa gttttaatac
37801  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
37861  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
37921  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
37981  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
38041  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
38101  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
38161  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
38221  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
38281  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
38341  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
38401  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
38461  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
38521  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
38581  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
38641  acaactacta gtcctcttag aaggagaaag aaaaagagaa gaaagacctt ttggtgctgt
4.3 Discussion.

In total 31 patients were screened for mutation in PTX1 while 8 samples were screened for mutations in PTX2 (Table 4.3a).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Phenotype</th>
<th>Number of Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTX1</td>
<td>Cleft lip/palate with components of SOPD</td>
<td>13</td>
</tr>
<tr>
<td>PTX1</td>
<td>Facial dysplasia with components of SOPD</td>
<td>1</td>
</tr>
<tr>
<td>PTX1</td>
<td>Multiple midline abnormalities including components of SOPD</td>
<td>3</td>
</tr>
<tr>
<td>PTX1</td>
<td>SOPD</td>
<td>7</td>
</tr>
<tr>
<td>PTX1</td>
<td>Isolated pituitary dysfunction</td>
<td>7</td>
</tr>
<tr>
<td>PTX2</td>
<td>Defects of left-right axis formation with components of SOPD</td>
<td>8</td>
</tr>
</tbody>
</table>

(Table 4.3a) Phenotypes of samples screened for mutations in PTX1 and PTX2.
No mutations were identified in either *PTXI* or *PTX2*. The two novel *PTXI* variants g1932a and c417a (Arg140Arg) appear to be benign polymorphisms, although an effect on *PTXI* function, perhaps at the splicing level, cannot be ruled out. It is therefore theoretically possible that these variants represent susceptibility alleles. Although non-coding mutations cannot be ruled out, neither *PTXI* nor *PTX2* appear to play a role in the aetiology of variant forms of SOD. It would be unlikely that mutations in these genes would produce classical SOD without secondary features.
5  Related work.

5.1  Mutation screening of candidate genes within the patient cohort.

A SOPD candidate gene mutation screen is being carried out on subgroups of the patient cohort (Dr. K. Woods, unpublished data and Mr. J. Turton, unpublished data). As yet this screen has not identified any coding region mutations in other candidate genes. Thus no other genes have been implicated in the aetiology of SOD within the patient cohort.

5.2  Mutation screening of known pituitary disease genes.

A subset of patients from the patient cohort was screened for mutations in the genes PROP1, PIT1, LHX3, and GH1 following HESX1 mutation screening.

Overall, 28 of the isolated pituitary defect patients within the HESX1 mutation screening cohort potentially harbour mutations within genes known to give rise to anterior pituitary defects. This represents 8% of the pituitary only patients. This finding suggests that known pituitary disease genes do not contribute significantly to the low HESX1 mutation rate within the pituitary only patient sub-group.

5.3  Mutational studies into the role of HESX1 in SOD.

In collaboration with this study, HESX1 from 20 Australian SOD patients was directly sequenced by Dr. Paul Thomas (Thomas et al., 2001). This identified 2 mutations, Q6H and T181A. These mutations were not found in 100 control chromosomes and have now been shown by our group to have functional consequences at the protein level.

A second study, in which HESX1 from 20 German SOD patients was directly sequenced, did not identify any sequence variants (Miss. N Haute, unpublished data). However, HESX1
from 20 control samples were also sequenced and this identified the 525g>a variant in an apparently un-affected adult male.

In a third study, *HESX1* from 20 Brazilian SOD patients was screened by SCCP followed by direct sequencing, and this identified a homozygous I26T mutation. This variant was found in a heterozygous state in a single control sample. Both the patients and control originate from the same region of Brazil and this may represent a founder mutation (Miss. L Carvalho. unpublished data). This mutation occurs within the N-terminal repressor domain.

A fourth study using SSCP detection did not detect any *HESX1* mutations within a cohort of 74 patients consisting of 48 isolated GHD and 26 combined pituitary hormone deficiency patients (Pfaffie et al., 2001).

A fifth study identified an *Alu* insertion within exon 3 in a consanguineous family (Netchine et al., 2002). Two siblings were both found to be homozygous for the insertion whilst both parents and a third sibling were found to be heterozygous. Both siblings demonstrated hypoplasia/aplasia of the anterior pituitary, whilst the older sibling was found to have an optic nerve coloboma. The phenotype in the siblings again underscores the variability of the phenotypic consequences of *HESX1* mutations and also demonstrates that the optic phenotype is not limited to optic nerve hypoplasia and thus SOD. The heterozygous relatives provide evidence that *HESX1* haploinsufficiency does not play a part in SOD or its variants.

Finally, two siblings carrying a heterozygous S170L mutation were identified during the initial *HESX1* mutation screen (Dattani et al., 1998 and Thomas et al., 2001). The two brothers exhibit GH deficiency whilst only the older brother exhibits unilateral optic nerve hypoplasia. Haplotype analysis demonstrated that the mutant allele was of paternal origin and that the two siblings inherited different maternal alleles. This excluded *HESX1* compound heterozygosity and indicates that S170L is either a dominant gain-of-function mutation or a recessive mutation resulting in haploinsufficiency. The fact that both brothers are affected when no R160C heterozygotes are affected provides strong evidence against S170L resulting in haploinsufficiency.
5.4 Functional studies of *HESX1* variants.

5.4.1 R160C.

In addition to the work carried out by Dattani *et al.* (1998) into the function of R160C two further studies have been carried out into the function of this allele. The first was carried out by Brickman *et al.*, (2001) in collaboration with this study. Brickman *et al.*, (2001) used *in-vitro* transfection experiment to investigate the function of R160C within COS cells.

A construct containing R160C HESX1 fused to the DNA binding domain of the GAL4 protein (GAL4-HESX1[R160C]) was co-transfected with a reporter in which Luciferase expression was driven by a GAL4 response element (identical to that used in section 3.2.5). They were able to show that R160C does not affect the ability of HESX1 to repress transcription. In addition, Brickman *et al.* (2001) were able to map the repressor activity of HESX1 in this context to the putative N-terminal repression domain, indicating that this domain does indeed function as a transcriptional repressor.

A second construct was made containing R160C HESX1 fused to the DNA binding domain of the GAL4 protein and the transcriptional activation domain of VP16:GAL4 – HESX1 – λVP4. The stability of these constructs was tested using a reporter in which Luciferase expression was driven by a GAL4 response element (same as above), where all constructs proved to be potent transcriptional activators. The tripartite fusions were then tested against each other using a reporter in which Luciferase expression was driven by a pIII containing response element, to which HESX1 can bind. As expected the R160C containing tripartite fusion was unable to enhance reporter expression. Brickman *et al.* (2001) were also able to show that deletion of the N-terminal repression domain does not affect the function of HESX1 in this context.

Brickman *et al.* (2001) showed that R160C is able to function as a dominant negative using EMSA, in which R160C inhibits WT binding. The results suggested that HESX1 is able to form dimers in solution, and that R160C leads to the formation of inactive heterodimers in solution. Hence, HESX1(R160C) has a potent dominant negative effect *in vitro*. Brickman *et al.* (2001) then produced a fusion protein containing R160C within the N-terminal deletion construct. This construct was unable to act as a dominant negative, demonstrating that the repressor domain is required for the dominant negative activity of R160C and that the repressor domain contains an additional HESX1 dimerisation domain. It is unclear whither the N-terminal dependent dimerisation requires homeodomain dimerisation.
De la Mata et al. (2002) investigated the consequences of R160C on the structure of HESX1 bound to DNA. They show that oxidised R160C (R53C) produces only a slight modification of HESX1 structure. However, they show that within the 3D structure of HESX1 arg53 lies close to cys24 and that cys53 can form a disulfide bond with cys24, which results in significant structural perturbation. They conclude that the disulfide bond formation is the likely cause of the loss of function associated with R160C.

5.4.2 S170L.

S170L is located within an RESQFL motif, which is unique to the Anf class and completely conserved across the class. This motif occurs at the C-terminal tail of the homeodomain. The exact function of this motif is unknown and may either add to DNA binding specificity or interact with HESX1 protein partners. In line with genetic evidence, Thomas et al. (2001) were able to show that S170L results in a 10-fold reduction in DNA binding. Brickman et al., (2001) further investigated the binding of S170L to DNA in greater detail and were able to refine the level of DNA binding defect to a 5-fold reduction.

As disrupted DNA binding does not appear to be the central defect of the S170L mutation, Brickman et al., (2001) investigated the ability of S170L to repress transcription using identical techniques to that of R160C. They were unable to demonstrate any defect of repression using an in-vitro transfection assay. It would therefore appear that S170L alters the interactions between HESX1 and its partners.

5.4.3 N125S.

The effects of the N125S polymorphism on both DNA binding and repression were investigated by Brickman et al. (2002). They demonstrated by EMSA that N125S-HESX1 binds to both the pIII site and the monomeric GBS site with a slightly higher affinity than WT-HESX1. Interestingly, N125S-HESX1 migrates differently from WT-HESX1 during EMSA suggesting a change in charge or conformation. This was correlated by studies showing that N125S-HESX1 does not interact with an antibody against the Hisx6 tail suggesting that the conformation of the protein is altered. They investigated the ability of N125S to repress transcription using identical techniques to that of R160C and were unable to show any difference from WT. Although the functional data suggest that in-vitro N125S may
produce an altered HESX1 conformation, within a cellular environment N125S does not appear to be deleterious.
6 Discussion.

6.1 HESX1 mutation screening.

6.1.1 Patient recruitment.

In total, 670 patients were recruited for this study into the role of HESX1 in SOD and its variants. Recruitment was ongoing throughout the duration of this project, and indeed is still continuing. Patients with established phenotypes were recruited, as were newly diagnosed patients. The study could be criticised in that the patient cohort is biased towards an endocrine phenotype, with 52% of patients showing a pituitary only phenotype. This is unlikely to be representative of the population as a whole since recruitment was naturally biased towards those centres known to Dr. Dattani and other members of the endocrine department at the London Centre for Paediatric Endocrinology. 31% of patients fell within the broad category of septo-optic dysplasia, but in only 45% of these patients was sufficient phenotypic detail available to classify the disorder accurately within the spectrum of septo-optic-pituitary dysplasia. 87 patients were classified as miscellaneous. These were samples accompanied with either little information beyond a statement that the sample was for the HESX1 study or patients with anterior phenotypes similar to variant SOD.

80 patients were recruited after mutation screening was completed. These patients and those subsequently recruited will undergo routine mutation screening within our laboratory. The sporadic recruitment led to problems with mutation screening, since samples had to be batched for PCR amplification and variance detection.

The expression of the HESX1/Rpx/Xanf family of transcription factors is believed to be restricted to the anterior of the embryo (see section 1.3). This produced a problem during the recruitment of patients, as 37% of patients demonstrated secondary features, many of which are not limited to the anterior of the embryo. As there is little known about the expression of HESX1 in Homo sapiens a preliminary expression study was carried to provide data upon which a decision to exclude patients with non-anterior secondary features could be made (see section 2.4.5). The results suggest that HESX1 is expressed out with the anterior and therefore patients with non-anterior secondary features could not be excluded from the analysis.
6.1.2 cDNA studies of *HESX1* expression.

Although, this study was of a preliminary nature, the data suggest that *HESX1* expression is not limited to the anterior of the embryo. Further testing of cDNA derived from an independent source would need to be carried out. Additionally, *HESX1* expression in tissues would need to be directly tested using *in-situ* hybridisation. Further experiments were not carried out due to time constraints. It was decided to include patients with secondary non-anterior features in the mutation screen as any *HESX1* mutations detected within this groups would of themselves provide valuable data on potential non-anterior functions of *HESX1*. Indeed a *HESX1* mutation was identified in a patient with a non-anterior secondary feature (supernumerary digits). However it is unclear if this feature is related to the *HESX1* mutation or is a chance occurrence.

6.1.3 Putative alternative exons.

Both the putative alternative exons were included in the mutation screen, as any *HESX1* mutations detected within these putative exons would of themselves provide valuable data on the role of these sequences. Three sequence variants were identified within the genomic sequence surrounding the putative exons (alt -50c>t, alt -4 t>g, and IVS1A+5 c>t); however no sequence variants were identified within the putative coding sequence. Each variant occurred in a single sample. Although the lack of putative amino acid variance would suggest a function for the alternative exon, the number of non-coding variants is too low to conclude that this is a significant finding. Although the evidence suggests that these sequences may give rise to a splice variant of *HESX1*, this remains to be proved.

6.1.4 Non-coding regions.

Towards the end of this study the genomic sequence of the murine *Hesx1* gene became available (a kind gift of Dr. K. Mahon) and was used to identify a region of homology between man and mouse upstream of the main *HESX1* transcription start-site, which in the mouse contains binding sites for both Gata-2 and Lim-1 (although the exact sites have not been mapped). This 5' LIM1 homology region was included in the mutation screen. Two other regulatory regions were identified in the mouse. No homologous sequence to a pituitary specific enhancer within intron 1 of the murine *Hesx1* could be identified in the
human *HESX1* gene, indeed a large portion of intron 1 in the human *HESX1* gene consists of an *Alu* element. This may reflect the fact that as yet, no alternative *Hesx1* transcript has been identified in the mouse. The third region of homology found in the mouse within the 3'UTR was also identified in human genomic *HESX1* sequence. However, this region consists of a number of small regions of homology, which would require individual screening as part of a separate study. Therefore due to time constraints, the 3' homology regions were not screened as part of this study. Again towards the end of this study, three regions known to play a role in the control of *Anf* expression were identified in the Xenopus *Xanf-1* gene. Fortunately, homology to two of these regions is included in the previously identified 5' LIM1 homology region. The third *Xanf-1* regulatory region lies out with the 5' LIM1 homology region and could not be included in the mutation screen due to time constraints.

6.1.5 Variance detection.

Two techniques were used to detect variance; single strand conformation polymorphism detection and heteroduplex detection. Although SSCP is known to have less than 100% detection rate (Jordanova et al., 1997) there are several methods of increasing the detection rate. The clean gel system (Amersham Pharmacia) uses pre-poured gels and therefore modification of the gel was unachievable. To increase the detection rate, running temperatures could be adjusted such that each sample could be run at two or more temperatures. However, this would only lead to a modest increase in the detection rate. In order to increase the detection rate, a second detection method was used. dHPLC-heteroduplex detection using the WAVE system is capable of detection rates approaching 100% (Transgenomics Inc). However, this is only achievable for fragments with a uniform melting profile. *HESX1* melting profiles are reasonably uniform and detection rates would be expected to be high. Indeed, a number of variants that had previously been missed by SSCP detection were detected by this method.

The WAVE system is unable to detect homozygous substitutions and small deletions since the WAVE profile derived from these variants will not differ from that of WT. To detect these changes, the test DNA has to be “spiked” with wild type DNA and then heteroduplex formation facilitated. This procedure was not carried out for a number of reasons. Firstly, only a small portion of homozygous changes will not be detected by SSCP detection. Secondly, the S170L mutation detected in two siblings demonstrates that *HESX1* mutations are not exclusively homozygous (Thomas et al., 2001). Finally, spiking test DNA with wild type DNA would double the variation detection time as well as double the cost. Thus,
spiking with wild type DNA was not performed so as to allow more samples to be screened with the available resources, in the knowledge that the false negative rate for homozygous mutations would be small.

Nevertheless, a single undetected homozygous mutation occurring at a relatively high level could theoretically be present within the cohort. This situation is unlikely given that this theoretical mutation has not been detected in the heterozygous state either within the cohort or in the carrier state within the control population. Only the -277t>g allele is found at a high enough level (2.2%) such that if this were indeed a mutation then the undetectable homozygous state would be a significant cause of \textit{HESX1} mutations within SOD. Although significant, the number of -277t>g heterozygotes suggest that the number of undetected -277t>g homozygotes would be small. Indeed, -277t>g was not detected within 16 sporadic SOPD, which were sequenced, nor in 42 familial cases that were completely sequenced. Therefore, a common undetected causative \textit{HESX1} mutation in the homozygous state is highly unlikely. In addition, the false negative rate was divergent between the two techniques with four variant samples being detected by dHPLC-heteroduplex detection that were not detected by SSCP detection. This represents just under half of heterozygous changes (5/9) and by extrapolation suggests that 50% of homozygous mutations should be detected by SSCP. As no homozygous mutations were detected by SSCP detection this would suggest that undetected homozygous mutations are indeed rare.

The major flaw of any mutation screen is that it cannot detect mutations out with the regions screened. These undetectable mutations would range from single nucleotide substitutions within transcription factor binding sites to large-scale genomic rearrangements. In addition the mutation screen may have failed to detect small deletions.

Deletions ranging from single exons to the entire \textit{HESX1} gene are undetectable by the mutation screen when they occur in the heterozygous state, due to the nature of PCR amplification. Quite simply, PCR amplification using a large number of cycles to increase both the quantity and uniformity of the signal produces product irrespective of the copy number of the region amplified. Only when the region is not present at all does PCR amplification fail and thus the technique can give some evidence of homozygous deletions, but not heterozygous deletions. There were 9 samples in which screening was undertaken but not completed. The non-completion of screening in 5 samples was due to a lack of DNA (due to other studies being carried out on the samples), whilst 4 samples were not completed due to PCR amplification failure of the DHX3F-DHX4R fragment. Without exception however, both the DHX3F-DHX3R and the DHX4F-DHX4R fragments were amplified from these 4 samples excluding homozygous deletions. The DHX3F-DHX4R fragment was the
most difficult to amplify in terms of both the number of samples from which amplification failed as well as the uniformity of signal strength, which was highly variable. The 4 DNA samples from which PCR amplification of the DHX3F-DHX4R fragment failed also appeared to be of poor quality. Given the lack of evidence pointing to homozygous deletions, one would suppose that heterozygous deletions of HESX1 are unlikely to contribute significantly to SOD. Both the R160C mutation (Dattani et al., 1998) and the Alu insertion mutation (Netchine et al., 2002) occur in a heterozygous state in a number of unaffected individuals. As these two mutations appear to represent a complete loss of function, other heterozygous loss of function deletions are unlikely to be causative in terms of SOD. Thus, the number of undetected causative heterozygous deletions is likely to be very low.

The false positive rate was relatively high at 3.3%. However, the classification of samples into “suggestive of variance” rather than “unsuggestive of variance” was biased towards “suggestive of variance” whenever there was any doubt.

6.2 HESX1 variants.

6.2.1 S170L.

The HESX1 mutation S170L was identified in a single sample, sample 4691. Patient 4691 was diagnosed with GH deficiency during childhood. Additionally, hypoplasia of the anterior pituitary with an ectopic posterior pituitary was uncovered during MRI. No defects of any other midline structure were noted, and optic nerve hypoplasia was not observed on either MRI or fundoscopy. Hence, this case represents isolated pituitary dysfunction rather than septo-optic-pituitary dysplasia. This S170L phenotype is in line with the two other cases of S170L, which show milder phenotypes than full SOPD with unilateral optic nerve hypoplasia only shown in one case.

S170L is located within an RESQFL motif (at the C-terminal tail of the homeodomain), which is unique to the Anf class and completely conserved across the class. The exact function of this motif is unknown and may either add to DNA binding specificity or interact with HESX1 protein partners.

Two siblings carrying heterozygous S170L mutations have been previously described as demonstrating variant SOD (Thomas et al., 2001). Genetic evidence ruled out compound heterozygosity. In line with genetic evidence, Thomas et al. (2001) were able to show that
S170L results in a 10-fold reduction in DNA binding. Brickman et al., (2001) further investigated the binding of S170L to DNA in greater detail and were able to refine the level of DNA binding defect to a 5-fold reduction.

This reduction in DNA does not explain the phenotype associated with S170L. The previously described R160C mutation represents a complete loss of function mutation that is unable to bind DNA and follows a recessive mode of inheritance (Dattani et al., 1998). The two R160C homozygous siblings exist within a large multi-generational kindred containing a number of unaffected R160C heterozygotes. Although the result of allelic tracking of R160C through the kindred is consistent with a recessive mode of inheritance, the data is insufficient to exclude a semi-dominant mode of inheritance with reduced penetrance. In addition a homozygous Alu insertion within exon 3 has been identified in two siblings with pituitary dysfunction (Netchine et al., 2002). Both parents and a third sibling were found to be heterozygous for the Alu insertion providing evidence that HESX1 haploinsufficiency does not play a part in SOD or its variants. Again the data is insufficient to exclude a semi-dominant mode of inheritance with reduced penetrance. Although it is unlikely, a HESX1 loss of function mutation may rarely give rise to a clinical phenotype comparable with the Hesxl null mutant mouse that results in a heterozygous phenotype in 1% of cases. In the case of the S170L siblings, the probability of this rare phenotypic expression of a HESX1 loss of function mutation occurring in both siblings is extremely low especially given that S170L does not result in a complete loss of DNA binding.

As disrupted DNA binding does not appear to be the central defect of the S170L mutation, Brickman et al., (2001) investigated the ability of S170L to repress transcription using identical techniques to that of R160C. They were unable to demonstrate any defect of repression using an in-vitro transfection assay. They conclude that S170L may alter the interactions between HESX1 and its partners. Further investigation into the S170L mutation is currently being carried out (Dattani, personal communication).

Given that S170L had previously been shown to be a mutation following a dominant mode of inheritance with reduced penetrance, further investigations into the mutation were not carried out as part of this investigation.

6.2.2 E149K.

The HESX1 mutation E149K was identified in a single sample, sample 4769, in the heterozygous state. Patient 4769 was diagnosed with GH and gonadotrophin deficiency
during childhood. Recently the diagnosis of gonadotrophin deficiency has been questioned, given that he has a child, who was conceived without assistance. MRI scanning revealed a hypoplastic anterior pituitary with an ectopic posterior pituitary. No defects of any other midline structure were noted, and optic nerve hypoplasia was not observed on either MRI or fundoscopy. Hence, this case represents isolated pituitary dysfunction rather than septo-optic-pituitary dysplasia. The E149K phenotype is in line with other heterozygous HESX1 mutations that show milder phenotypes rather than full SOPD. Intriguingly, supernumerary digits were removed from both hands shortly after birth. This is an uncommon but not rare occurrence. Since, digital abnormalities have been demonstrated in three other SOD patients this may represent a role for HESX1 out with the anterior of the embryo or a second abnormality unrelated to HESX1.

E149 is a highly conserved residue, not only across the ANF class but also across the paired-like family and homeodomains in general. The E149K variant has not been found in 180 sequenced control chromosomes (Dattani, personal communication and Haute, personal communication). E149K represents a change from a glutamic acid (Glu) residue at position 42 of the homeodomain to lysine (Lys) (E42K). This occurs at the first residue of the recognition helix and is a highly conserved residue, with a glutamic acid occurring in 226 of 346 homeodomains examined by Douboule (1994), and the homeodomain consensus sequence contains Glu at position 42 (Freund et al., 1997). Within the paired-like homeodomain family Glu42 is absolutely conserved (Freund et al., 1997). As shown in the 3-D structure of the Prd homeodomain, Glu42 forms an electrostatic interaction with Arg31 and indeed Arg31 is highly conserved and invariably coupled with Glu42 in paired-like homeodomains (Freund et al., 1997). Within paired-like homeodomains this interaction is involved in the folding of paired-like homeodomains, in particular the interaction between helix 2 and helix 3/4 and disruption of this interaction would likely give rise to a misfolded HESX1 protein (Wilson, 1995). Within the Prd homeodomain, when bound as a dimer, Glu42 also forms a salt-bridge in-trans with Arg3, as well as, forming a water mediated hydrogen bond with Arg44. These reactions are critical to dimerisation forming a major interaction face between the two Prd homeodomains and disruption of this interaction would likely give rise to a defect in dimerisation (Wilson et al., 1995).

If the major defect resulting from E149K were a defect in HESX1 protein folding, E149K would be expected to represent a loss of function mutation. In order for E149K to represent a loss of function allele that results in the phenotype of patient 4769, it must either produce haploinsufficiency of the HESX1 gene product or occur along side a second mutation, either within HESX1 (compound heterozygosity) or within a second locus (bi-allelic inheritance). For identical reasons given above in relation to S170L, haploinsufficiency of HESX1 does not
appear to have phenotypic consequences. However in the case of E149K a rare clinical phenotype associated with haploinsufficiency cannot be ruled out. Although all known functional regions of HESX1 were sequenced without a second mutation being identified, a second mutation out of these regions cannot be ruled out. On the other hand, if the major defect resulting from E149K were a defect in HESX1 dimerisation, E149K would be expected to represent a dominant negative mutation, which would show a dominant mode of inheritance. As a number of first-degree relatives of patient 4769 carry E149K without clinical pituitary dysfunction the potential dominant mode of inheritance is associated with reduced penetrance. Unfortunately, allelic tracking data was not sufficient to distinguish between a compound recessive mode of inheritance and a dominant mode of inheritance with reduced penetrance.

Other functional possibilities for the E149K allele exist; it may lead to an increased expression range of HESX1, an entirely novel function, or effect HESX1 mediated repression. HESX1 expression normally occurs in somatotrope precursor cells and HESX1 overexpression could in theory give rise to growth hormone deficiency, although a mechanism through which E149K could result in HESX1 overexpression is unclear. A gain of function mediated through a novel function of E149K HESX1 cannot be excluded, but is unlikely. A defect of repression is unlikely as the HESX1 homeodomain does not appear to function in repression (Brickman et al., 2001 and personal communication), although recruitment of co-repressors by the Hesxl homeodomain has been suggested in the mouse (Xu et al., 2000).

Two mutations of E42 within Paired-like homeodomains are known to result in human disease, E42A and E42K (Freund et al., 1997 and Sankila et al., 2000). Both of these mutations occur within the paired-like gene CRX and result in dominant Cone-Rod Dystrophy. Unfortunately, no functional investigations of these two mutations have been carried out to date, presumably due to the strength of the genetic data that demonstrated that the variants are mutations following a dominant mode of inheritance. Freund et al. (1997) hypothesized that on the balance of probability, the E42A allele most likely represented a haploinsufficient loss of function, with a dominant dimerisation defect a possibility and a loss of function due to a defect in the interaction between CRX and partner proteins excluded.

In summary, based upon the structure of the HESX1 homeodomain a loss of function due to homeodomain misfolding is a likely consequence of E149K, as is a defect in HESX1 dimerisation. However the likelihood of a HESX1 loss of function mutation resulting a phenotype is unlikely, but preselection during patient recruitment does make this a possibility. Both a defect in HESX1 mediated repression and a novel function due to E149K
are unlikely, but possible. To distinguish between these possibilities an investigation into the
function of the HESX1 E149K protein was performed to investigate DNA binding,
dimerisation and HESX1 mediated repression.

An initial EMSA examining the DNA binding properties of both WT and E149K proteins
over a wide protein concentration range (1nM-1024nM) clearly shows that recombinant
E149K is capable of binding DNA as two distinct species. This suggests that E149K is
capable of binding DNA both as a monomer (P.DNA) and as a dimer (P₂.DNA). Thus, an
inability of E149K to bind DNA under EMSA conditions is clearly ruled out. A mild defect
in DNA binding remains a possibility but would not explain the apparent dominant nature of
E149K. The results of this EMSA also demonstrate an increased concentration of a faster
moving, presumably monomeric, species within E149K lanes as compared with WT. This
would indicate that E149K may well result in a dimerisation defect.

This apparent dimerisation defect in line with the modelling data invalidates the method used
by Dattani et al. (1998) to estimate the HESX1 dissociation constant for WT, R160C and
S170L. This method combines the monomer and dimer species into one bound species and
calculates the dissociation constant between bound and unbound. This produces a pseudo-
dissociation constant that is related to both the monomer formation dissociation constant and
the dimer formation dissociation constant

\[
\text{pseudo-Kd} = \frac{[\text{P.DNA}][\text{DNA}][\text{P}^2][\text{DNA}]}{[\text{P}][\text{DNA}]} \cdot \frac{[\text{P}][\text{DNA}]}{[\text{P}][\text{DNA}][\text{P}][\text{DNA}][\text{P}][\text{DNA}]}
\]

[p represents the free
protein concentration]. To use this pseudo-dissociation constant to compare the binding of
mutant and WT HESX1 protein requires that the ratio of monomeric to dimeric bound
HESX1 is identical between the test protein and the WT protein, i.e. dimerisation is not
affected by the test mutation. The result of this first EMSA clearly invalidates this
assumption and thus the technique. However, the dissociation constant for formation of
monomer species (Kₐ = [P.DNA]/[P][DNA]) can be calculated using probe DNA containing a
single HESX1 binding site, such as GBS (Dattani et al., 1998). However, the evidence of
HESX1 dimerisation in solution provided by the R160C mutation (Brickman et al., 2001)
invalidates this technique as it requires that both test and WT proteins occur in solution as a
monomer, or that the dissociation constants for dimer formation in solution for both test and
WT proteins are known (Kₐ = [P₂]/[P]²). This technique can be used to compare the bound-
monomer dissociation constant of a test protein against WT protein given an assumption that
the dissociation constant for dimer formation in solution for the test protein is identical to that
of the WT protein. However, the above EMSA result invalidates this assumption, as it cannot
be shown where the potential difference in dimerisation occurs i.e. dimerisation in solution or
cooperative dimerisation bound to DNA.
Cooperative dimerisation can be calculated as a cooperativity factor (τ) defined as $K_{d1}/K_{d2}$ where $K_{d1}$ represents the dissociation constant for monomer formation (P+DNA→P.DNA) while $K_{d2}$ represents the dissociation constant for dimer formation (P+P.DNA→P$_2$.DNA) and τ equals 4[P$_2$.DNA][DNA]/[P.DNA]$^2$ (Wilson et al., 1993). The method used to determine τ compares the concentration of protein required such that [P$_2$.DNA] = [DNA] using a DNA probe amenable to cooperative dimerisation compared to that using a DNA probe which is not amenable to cooperative dimerisation, where the ratio of protein concentrations required to shift a noncooperative (pV) versus a cooperative (pIII) DNA probe to the point where [P$_2$.DNA] = [DNA] is equal to the square root of τ (Wilson et al., 1993). This technique again requires the assumption that the dissociation constant for dimer formation in solution for the test protein is identical to that of the WT protein to be valid. Classically, Paired-like homeodomains are not capable of cooperative dimerisation on monomer sites separated by greater than 4bp and therefore DNA probes containing monomer sites separated by 5bp (pV) can only be bound by Paired-like homeodomains in a noncooperative fashion (Wilson et al., 1993). However, Quirk and Brown (2002) have recently described a natural HESX1 binding site which consists of two monomer binding sites separated by 8bp (pVIII), on which HESX1 is capable of cooperative dimerisation. This finding also invalidates the technique used by Wilson et al. (1993) as it is unclear to which binding sites HESX1 can bind cooperatively and to which HESX1 binds noncooperatively.

A qualitative method for identifying cooperative dimerisation exists using EMSA. When a small amount of probe is shifted, noncooperative dimerisation can be detected when [P.DNA]/[P$_2$.DNA] is high. Indeed in noncooperative dimerisation the P$_2$.DNA species is not observed until [P.DNA] approaches saturation. On the other hand, cooperative dimerisation can be detected when [P.DNA]/[P$_2$.DNA] is low (Wilson et al., 1993). Where dimerisation occurs in solution and the dimer has an increased affinity for DNA, such as is suspected with HESX1 (Brickman et al., 2001), this method is not capable of discriminating between this situation and cooperative dimerisation. This technique will identify a qualitative difference in dimerisation whether or not dimerisation occurs in solution or bound to DNA.

EMSA was carried out to compare [P.DNA], [P$_2$.DNA] and [DNA] between E149K and WT protein over a total protein concentration of 128nM-448nM. Using the data from this assay, the percentage of total probe found as a monomer (fastest migrating retarded species) divided by the percentage of total probe found as a dimer (slower migrating retarded species) was plotted against total protein concentration. This graph shows that the highest monomer/dimer ratio occurs at a total protein concentration of 128nM for E149K while the highest WT monomer/dimer ratio occurs at a total protein concentration of 448nM. WT protein produces a result expected of a protein capable of cooperative dimerisation (Wilson et al., 1993), such
that the monomer/dimer ratio is lowest at low concentrations (the 128nM result represents lowest amount of probe detectable). On the other hand, E149K does not produce a result expected of a protein capable of cooperative dimerisation. The monomer/dimer ratio of E149K is consistently higher than WT over the concentration gradient, suggesting a defect of dimerisation.

To further investigate the relative monomer/dimer ratio of E149K compared with WT, EMSA was carried out over a lower concentration range of 32nM-192nM. Using the data from this assay, the percentage of total probe found as a monomer (fastest migrating retarded species) divided by the percentage of total probe found as a dimer (slower migrating retarded species) was plotted against total protein concentration. This graph shows that the highest monomer/dimer ratio occurs at a total protein concentration of 32nM for E149K while the highest WT monomer/dimer ratio occurs at a total protein concentration of 192nM. WT protein again produces a result expected of a protein capable of cooperative dimerisation, such that no monomer species is detected at low concentrations with the monomer species first detected at 160nM (1% of total probe). The E149K profile again is not typical of that of a protein that is capable of cooperative dimerisation.

The three EMSA experiments provide reproducible evidence that the E149K mutation results in impaired dimerisation. However, it is unclear if this is a defect of cooperative dimerisation or dimerisation in solution. The EMSA experiments also show that E149K is capable of binding DNA. The first EMSA (1nM-1024nM) shows shifted probe at lower concentrations of E149K than WT; similarly both subsequent EMSA’s show a greater amount of E149K bound (monomeric + dimeric) probe than WT bound probe across a wide total protein concentration of 32nM-448nM. This situation could be due to either an increased DNA binding affinity of E149K compared to WT or a reduction in the dimerisation (in solution) dissociation constant resulting in a higher E149K free monomer concentration.

In summary, E149K appears to lead to impaired dimerisation, although this is not a complete impairment as E149K does appear to be able to bind as a dimer. However, it cannot be determined if this apparent dimer species represents cooperative dimerisation or binding of pre-existing HESX1 dimers formed in solution. As E149K occurs at position 42 of the homeodomain, impaired cooperative dimerisation is more likely than impaired dimerisation in solution. However formal investigation into the exact nature of the dimerisation defect is beyond the scope of this thesis.

Xu et al. (1999) demonstrated that the murine Hesx1 homeodomain is able to recruit the co-repressor Ncor1. However transfection studies performed in our laboratory have failed to show that the Hesx1/HESX1 homeodomain is capable of repression in the absence of the eh-
1 repression domain (Dattani and Brickman personal communication). Indeed, three HESX1 variants within the homeodomain and its carboxy terminus (N125S, R160C, and S170L) have no effect on repression. Hence, it is highly unlikely that E149K could directly lead to impairment of repression. Nevertheless, in order to test the ability of E149K to repress transcription an identical approach to that described by Brickman et al. (2001) was undertaken; namely a construct containing the GAL4 DNA binding domain fused to wild type HESX1 was mutated to E149K and compared with the WT construct.

In-vitro transfection experiments were repeated a total of three times each in triplicate. It is clear that, within this assay, GAL4-HESX1[WT] is capable of repressing Luciferase activity in a dose dependent manner as shown by Brickman et al. (2001). On the other hand, GAL4-HESX1[E149K] did not result in significant repression of Luciferase activity. Indeed, the highest transfected quantity of GAL4-HESX1[E149K] (800ng) appears to increase the level of Luciferase activity significantly above that of the pJBXi baseline.

These in-vitro transfection data suggest that either GAL4-HESX1[E149K] has an effect on the ability of HESX1 to repress transcription or that GAL4-HESX1[E149K] protein is either degraded due to protein misfolding or is never made. For GAL4-HESX1[E149K] to have an effect on the ability of HESX1 to repress transcription, either the mutation could directly interfere with recruitment of co-repressors, or it could recruit a co-activator. It is unlikely that the mutation has an effect on the folding of the N-terminal repression domain, since three other HESX1 homeodomain variants do not have this effect. However, this cannot be formally ruled out on the strength of these data. Although GAL4-HESX1[E149K] appears to fold normally in-vitro (given the ability of the recombinant protein to bind DNA) GAL4-HESX1[E149K] would be expected to misfold in a cellular environment given the disrupted Glu42-Arg31 interaction.

To address this, a Western blot of GAL4-HESX1[E149K] and GAL4-HESX1[WT] from in-vitro transfected COS-7 cells was performed. The Western blot demonstrates that GAL4-HESX1[E149K] protein is undetectable within the protein extract of GAL4-HESX1[E149K] in-vitro transfected COS-7 cells. This finding demonstrates that the lack of GAL4-HESX1[E149K] transcriptional repression is due to a lack of GAL4-HESX1[E149K] protein. However, the data do not distinguish between non-production of GAL4-HESX1[E149K] protein and GAL4-HESX1[E149K] protein degradation. Although the HESX1 insert and the region surrounding were shown by DNA sequencing to be as expected, a mutation within the plasmid out with this region cannot be excluded. To address this total RNA was extracted from in-vitro transfected COS-7 cells and poly [A]+ mRNA was transcribed into cDNA which was then used a template for the PCR amplification using HESX1 cDNA specific
oligonucleotide primers. This experiment indicates that \textit{HESX1} containing mRNA occurs within both GAL4-HESX1[E149K] transfected cells and GAL4-HESX1[WT] transfected cells but not within non transfected cells or cells transfected with vector only. This strengthens the argument for GAL4-HESX1[E149K] protein degradation but does not formally rule out non production of GAL4-HESX1[E149K] protein, although no mutations are present within the insert.

In conclusion, the evidence indicates that E149K is a causative mutation, which results in the phenotype of patient 4769. It is unclear if this mutation represents a dominant mutation resulting in a defect in HESX1 dimerisation or is a recessive mutation resulting in misfolding of mutant HESX1.

\textbf{6.2.2.1 Further investigation of E149K.}

Formal demonstration of a defect of dimerisation in solution could be achieved by direct examination of the relative concentrations of E149K monomer and dimer in solution, in the absence of DNA. Techniques such as electrospray mass spectrometry and cross-linking studies may be able to formally identify or exclude a defect of dimerisation in solution. However, these techniques are technically difficult and time consuming and depend on the formation of a stable dimer. Such an investigation is out with the scope of this thesis. An alternative approach could be undertaken to further investigate dimerisation utilising the dominant negative properties of R160C mutation. Transfection experiments using R160C-HESX1 fused to a transcriptional activation domain derived from the Herpes Simplex Virus, VP16, have indicated that the R160C dominant negative phenotype results from R160C forming inactive heterodimers in solution and that this dimerisation is mediated by the eh-1 repression domain (Brickman \textit{et al.}, 2001). However it is unclear how to data produced by a study should be interpreted as such a study would address dimerisation mediated by the eh-1 domain and not by the homeodomain. In addition, it would be difficult to use this technique to investigate mild defects in dimerisation and the EMSA data suggests that EI49K does not lead to a complete loss of dimerisation.

Homeodomains of the \textit{paired-like} class have been shown to bind to a core TAAT consensus site and are able to cooperatively dimerise to two core sites separated by 2 or 3 bp (P2 and P3 respectively), but can only bind to two core sites separated by an increased amount (i.e. 5bp, pV) as two separate monomers (Wilson \textit{et al.}, 1995). Quirk and Brown (2002) have uncovered two naturally occurring HESX1 binding sites, a dimeric binding site at the LHβ promoter and a monomeric binding site at the αGSU promoter. They have demonstrated
apparent cooperative binding to the dimeric site, which consists of two core consensus sites separated by 8bp (P8). This finding is not compatible with the traditional view of homeodomain dimerisation i.e. HESX1 should not be able to cooperatively dimerise to this P8 sequence. However, this paradox can be solved if the apparent cooperative dimerisation at the LHβ promoter is mediated not by homeodomain dimerisation, but by repressor domain mediated dimerisation. Repressor domain mediated dimerisation could potentially allow dimerisation to two core sites separated by more than 4bp. If, as the work of Quirk and Brown (2002) suggests, HESX1 can dimerise to two core consensus sites separated by greater than 3bp, this would allow HESX1 to out compete other proteins containing paired-like homeodomains, which are believed to bind to such sites as monomers. However, the majority of research in this field has used isolated homeodomains rather than complete proteins and therefore this hypothesis requires further investigation in which E149K may play an important role.

6.2.3 HESX1 sequence variation.

With the exception of a recent duplication in Xenopus, only a single HESX1 homologue has been described in vertebrate species, and thus the HESX1 homologues are the only members of the Anf family of homeodomain transcription factors. Gene families are derived from a common ancestor that over time has undergone duplication with subsequent functional redundancy leading to diversification of sequence and function. Indeed there are many families of homeodomain transcription factors, such as the Hox genes, which are believed to have formed in this manner. This theory would predict that large gene families with multiple members within any given species would show greater sequence diversity across the family than a family of genes with a single member present in any given species. However, HESX1 shows the greatest sequence diversity of any homeodomain family (Kazanskaya et al. 1997).

HESX1 contains a paired-like homeodomain encoded by 3 exons, an engrailed-like repressor domain, and a six amino acid motif C-terminal to the homeodomain of no known function. It may be that the HESX1 family of proteins have only a single cellular function, such as binding to paired-like consensus sequences with a high affinity so as to compete with paired-like activators, and hence repress the transcription of target genes. This would suggest that amino acid residues, both within the homeodomain and out-with the homeodomain, which do not play a role in this single function, are not subject to stringent selection, allowing mutations within the coding sequence of these amino acids to rapidly accumulate. If this were indeed the case it would be expected that numerous polymorphisms would be found...
within the coding sequence of $HESX1$ which would lead to amino acid substitutions out-with the conserved regions. Yet only two such polymorphisms were found. The most likely explanation is that within the human species, $HESX1$ is under a high degree of selective pressure, perhaps due to co-evolution between $HESX1$ and some other gene or due to a novel $HESX1$ function within the human lineage, such as recruitment into a large co-repressor complex. If this were indeed the case then $HESX1$ mutations would be expected to be fully penetrant, otherwise they would accumulate within the population. Yet this is not the case and a large number of unaffected carriers have been described within 4 families (R160C, S170L, E149K, and the Alu insertion family described by Netchine et al. (2002). To date, functional work in man has been unable to demonstrate any $HESX1$ function out-with the three regions of homology. Detailed research into the function of $HESX1$ combined with population studies within other species may provide evidence that will solve this apparent paradox and answer the enigmatic finding of very little variation in the $HESX1$ amino acid sequence.

6.3 The role of $HESX1$ in SOD and its variants.

SOD is a developmental abnormality of the midline structures of the brain, in particular optic nerve hypoplasia, absence of the septum pellucidum and pituitary abnormalities. This study involved the collection of the largest cohort of SOD patients described to date (n=206). This study has been unable to detect $HESX1$ coding region mutations in any classical case of SOD. Indeed, the two $HESX1$ mutations found during this study result in pituitary defects alone. Of 590 patients screened, only 2 could be shown to harbour causative $HESX1$ mutations (E149K and S170L), although a further 14 patients were found to contain sequence variants that do not appear to be mutations (183t>c, Y90H, 525g>a and N125S), and 8 patients were found to contain sequence variants that may possibly be mutations (alt -50c>t, alt -4 t>g, IVS1A c+5t, exon1 -277t>g and 219c>t). Thus, $HESX1$ does not appear to be a major causative factor underlying septo-optic dysplasia within this patient cohort.

This low mutation detection rate may suggest that there remains a large group of undetected $HESX1$ mutations within the cohort, or that $HESX1$ is in itself not a major causative factor underlying septo-optic dysplasia within the general populace. Although there may be a very small number of homozygous mutations that were not detected by SSCP, this study suggests that mutations within the four coding regions, the 5' LIM1 homology domain, and the two putative alternative exons of $HESX1$ are not a common cause of SOD. There is little evidence for a role of $HESX1$ mutations in the aetiology of SOD out with the regions
screened, being limited to a screen of 38 patients using Southern blot analysis (Dattani et al., 1998). The time-consuming nature of mutation screening by SSCP and dHPLC-heteroduplex analysis did not allow time for screening of the patient cohort by techniques such as Southern blot analysis and FISH, which are able to detect large genomic rearrangements. Additionally, these techniques require large quantities of genomic DNA or cell lines, neither of which were readily available. Thus, the evidence suggests that HESX1 is not a major causative factor in septo-optic dysplasia within the general populace.

HESX1 mutations also appear to be only a rare cause of pituitary dysfunction, having been identified in only 2/347 patients with pituitary dysfunction. Although no HESX1 mutations were identified in cases of isolated optic nerve hypoplasia and isolated midline forebrain abnormalities such as agenesis of the septum pellucidum or hypoplasia of the corpus callosum, the number of these patients screened is too small to draw general conclusions concerning the population at large.

37% of SOD patients recruited were described as having additional features, which brings into question the definition of SOD as a single disorder. These additional features are not limited to anterior structures, and as HESX1 expression was previously described as being restricted to the anterior of the embryo, the ability of HESX1 to account for these cases of “variant” SOD was brought into question. In an attempt to address this question, a limited expression study of HESX1 within the human embryo was carried out. The data are insufficient to exclude HESX1 as a candidate gene for these cases of SOD with non-anterior features. Indeed the data suggests that HESX1 may not be limited to the anterior. As the mutation screening study would uncover causative mutations within HESX1 in these cases, the mutation screening study itself could produce evidence towards a role for HESX1 within SOD with non-anterior features. This view could possibly be supported by the identification of a patient with non-anterior features associated with a causative HESX1 mutation during the HESX1 mutation screening study (E149K).

Prior to the research carried out by Dattani et al. (1998), evidence of a genetic contribution to the aetiology of SOD was extremely poor. When the study is combined with other known studies into the role of HESX1 a total of 9 cases of SOD, or milder SOD phenotypes, have been shown to be due to causative HESX1 mutations representing approximately 1% of all known cases. If HESX1 does play a major role in SOD it would most likely be as a susceptibility locus. In particular the N125S polymorphism is a good candidate for a SOD susceptibility allele, although the contribution would likely to be low given Hardy-Weinberg equilibrium within the Afro-Caribbean population. However, this would be extremely
difficult to address given the crucial role of *HESX1* in early anterior patterning, the lack of a definitive SOD phenotype, and the rarity of the condition.

Although *HESX1* is not a major cause of SOD within this cohort, and by inference not a major cause of SOD within the general population, causative mutations at other genetic loci may play a significant role in SOD. Mutation screening of candidate genes within the recruited patient cohort was carried out in conjunction with this study (Dr K. Woods and Dr M. Dattani, personal communication). This ongoing study investigated the role of candidate genes mainly encoding transcription factors expressed during anterior patterning, which could give rise to the full SOD phenotype. No mutations were identified within the candidate genes screened, although a number of polymorphisms have been described. Although, transcription factors expressed during early anterior patterning are excellent candidate genes for SOD, given that their expression occurs within tissue that gives rise to all three structures involved in SOD, this finding is not unexpected as mutations within these genes are likely to give rise to severe phenotypes resulting in embryonic lethality. On the other hand, genes expressed exclusively within the three structures involved in SOD would be excellent candidate genes for the condition. However, no genes are known to fall into this category as yet.

The major cohort of patients recruited encompassed those with isolated pituitary defects. Of the three structures implicated in SOD, the development of the pituitary is by far the best characterised. It is therefore an excellent starting point to investigate the role of genes that may be involved in the aetiology of SOD. To this end the role of *PTXI* and *PTX2* in variant SOD was investigated in two subgroups with phenotype suggestive of *PTXI* and *PTX2* involvement. However, no mutations within these genes could be identified. Although both genes are excellent candidate genes for pituitary phenotypes they are very closely related and would be expected to have overlapping function. In addition, two other related genes *OTXI* and *OTX2* are also expressed within the developing pituitary and may compensate for defects within *PTXI* or *PTX2*.

Although a number of candidate transcription factors have been investigated with respect to an aetiological role in SOD, little evidence has been found to implicate any locus other than *HESX1* in the aetiology of the condition. Thus further work on the role of *HESX1* in SOD should be aimed at investigating the pathway through which *HESX1* functions and identifying other candidate genes within the *HESX1* pathway, both upstream and downstream of *HESX1*. This research would encompass the identification of signalling pathways, *HESX1* partner proteins, and genes regulated by *HESX1*.  

198
Signalling pathways including \textit{shh} and the FGF and BMP families may in fact be the best candidates for potential SOD candidates. In the mouse, \textit{Hesxl} has been shown to have a role in the regulation of an FGF signalling centre within the diencephalon as well as suppressing FGF signal production within Rathke's pouch. As FGF signalling is involved in a number of developmental stages within the brain, including the formation of the optic stalks, mutations with the FGF signalling pathway may well cause SOD. On the other hand, \textit{shh} is expressed within the murine pituitary placode and the first known event following \textit{Hesxl} expression in the pituitary placode is the repression of \textit{shh} expression. It is likely that \textit{HESXl} is directly or indirectly involved in this repression. Additionally, SOD shows considerable phenotypic similarity to holoprosencephaly, a condition associated with mutations within the \textit{shh} signalling pathway.

A number of studies have provided valuable information regarding the function of HESX1, from which the search for HESX1 partner proteins can begin. HESX1 has been shown to bind to the TLE family of co-repressors, with TLE1 being the likely \textit{HESXl} co-repressor within the pituitary (Dasen \textit{et al.}, 2001). Thus the TLE family of transcriptional co-repressors may be potential candidate genes for SOD. However, this family of proteins is widely expressed, suggesting that only partial mutations would be likely to give rise to SOD. Further investigations into the interactions of HESX1 with partner protein can be carried out using techniques such as coimmuno-precipitation and the yeast two-hybrid system. This work would hopefully provide additional candidate genes for SOD.

Research into genes regulated by \textit{HESXl} is progressing and recently the first HESX1 binding site has been identified. Quirk and Brown (2002) have uncovered two naturally occurring HESX1 binding sites, a dimeric binding site at the LH\textbeta{} promoter and a monomeric binding site at the \textalpha{}GSU promoter. Although the sequence of the core binding sites conform to the known \textit{paired-like} binding consensus, the separation of the dimeric core sites at the LH\textbeta{} promoter (8bp) represents a novel arrangement. Clearly, hormone genes are not candidates for SOD, but this novel core binding site arrangement may allow the elucidation of other genes regulated by HESX1. Before this novel arrangement can be used to search for genes downstream of \textit{HESXl} it must be confirmed \textit{in-vivo} that this site functions as HESX1 binding site during development.

In addition to candidate genes linked to HESX1, mutations in genes such as \textit{Netrin-1} or \textit{DCC-1}, which are involved in axon projection, may be implicated in SOD as all three structures affected in SOD involve axon projection. However mutation screening of \textit{Netrin-1} or \textit{DCC-1} within the patient cohort did not identify any mutations within these genes (Dr K. Woods, personal communication).
6.3.1 \textit{HESX1} and growth hormone deficiency.

All \textit{HESX1} mutations thus far identified have been associated with growth hormone deficiency, and most cases are associated with an ectopic posterior pituitary. Although growth hormone deficiency and subsequent growth failure is perhaps the most common of the classical SOD features, the likelihood that this association is due to chance is low, as a large number of patients with a non-growth hormone deficiency phenotype have been screened. Rather, it would appear that the somatotropes are the most susceptible tissue to \textit{HESX1} mutations and indeed the current data regarding \textit{HESX1} function suggests a plausible mechanism through which the pituitary may be the most susceptible tissue to \textit{HESX1} mutations.

Three independent \textit{HESX1} dependant stages occur during murine pituitary development. First expression of \textit{HESX1} within the anterior neural fold is required for proper anterior patterning and forebrain formation (Martinez-Barbera and Beddington, 2001.). Second, \textit{HESX1} limits a FGF8 expression domain within the developing diencephalon, which is required for pituitary development (Dasen \textit{et al.}, 2001). Third, \textit{HESX1} is required within the pituitary itself for correct development and function (Dasen \textit{et al.}, 2001). If an assumption that human pituitary development is similar to murine pituitary development is valid, then human pituitary development would similarly pass through at least three \textit{HESX1} dependant steps. This would suggest that the pituitary may be more susceptible to \textit{HESX1} mutations than the optic nerves or septum pellucidum which appear to be dependent only on \textit{HESX1} expression within the anterior neural fold. The case for the somatotropes being the most susceptible endocrine cell type to \textit{HESX1} mutations is not clear. The somatotropes are the last endocrine cell type within the pituitary to develop and this may result in this particular association.
7 References.


novel class of vertebrate homeobox genes expressed at the anterior end of the main embryonic axis. Gene 200, 25-34.


Lanctot, C., Lamolet, B., and Drouin, J. (1997). The bicoid-related homeoprotein Ptx1 defines the most anterior domain of the embryo and differentiates posterior from anterior lateral mesoderm. Development 124, 2807-2817.


