European Journal of Human Genetics (2009) 17, 1423 – 1431

www.nature.com/ejhg

ARTICLE

Separation of the *PROX1* gene from upstream conserved elements in a complex inversion/translocation patient with hypoplastic left heart

Harinder K Gill[1,2], Sian R Parsons[3], Cosma Spalluto[3], Angela F Davies[4], Victoria J Knorz[3], Clare EG Burlinson[1], Bee Ling Ng[5], Nigel P Carter[5], Caroline Mackie Ogilvie[1,4], David I Wilson[3] and Roland G Roberts*[,1]

[1]*Division of Genetics and Molecular Medicine, Department of Medical and Molecular Genetics, King's College London, London, UK;* [2]*Department of Clinical Genetics, Guy's and St Thomas' NHS Foundation Trust, London, UK;* [3]*Centre for Human Development, Stem Cells and Regeneration, Human Genetics Division, University of Southampton, Southampton General Hospital, Southampton, UK;* [4]*Department of Cytogenetics, Guy's and St Thomas' NHS Foundation Trust, London, UK;* [5]*Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK*

Hypoplastic left heart (HLH) occurs in at least 1 in 10 000 live births but may be more common *in utero*. Its causes are poorly understood but a number of affected cases are associated with chromosomal abnormalities. We set out to localize the breakpoints in a patient with sporadic HLH and a *de novo* translocation. Initial studies showed that the apparently simple 1q41;3q27.1 translocation was actually combined with a 4-Mb inversion, also *de novo*, of material within 1q41. We therefore localized all four breakpoints and found that no known transcription units were disrupted. However we present a case, based on functional considerations, synteny and position of highly conserved non-coding sequence elements, and the heterozygous *Prox1*$^{+/-}$ mouse phenotype (ventricular hypoplasia), for the involvement of dysregulation of the *PROX1* gene in the aetiology of HLH in this case. Accordingly, we show that the spatial expression pattern of PROX1 in the developing human heart is consistent with a role in cardiac development. We suggest that dysregulation of *PROX1* gene expression due to separation from its conserved upstream elements is likely to have caused the heart defects observed in this patient, and that *PROX1* should be considered as a potential candidate gene for other cases of HLH. The relevance of another breakpoint separating the cardiac gene *ESRRG* from a conserved downstream element is also discussed.

European Journal of Human Genetics (2009) **17,** 1423–1431; doi:10.1038/ejhg.2009.91; published online 27 May 2009

Keywords: chromosome inversion; chromosome translocation; PROX1; hypoplastic left heart; position effect

Introduction

HLH represents a spectrum of heart defects characterized by severe underdevelopment of left heart structures. In classical HLH, the aortic and mitral valves are atretic and the left ventricle and ascending aorta are markedly small. HLH is a major cause of death in the first week of life, accounting for 25% of all neonatal deaths due to heart defects.[1] Without treatment, HLH is invariably lethal and

*Correspondence: Dr RG Roberts, Department of Medical and Molecular Genetics, King's College London, 8th Floor, Tower Wing, Guy's Hospital, London SE1 9RT, UK.
Tel: +44 207 188 3704; Fax: +44 207 188 2585;
E-mail: roli.roberts@genetics.kcl.ac.uk
Received 20 November 2008; revised 22 April 2009; accepted 23 April 2009; published online 27 May 2009

even with surgical managements, the 5-year survival is only 50–60%. The incidence of HLH is 0.1–0.25/1000 live births but higher in fetal life due to spontaneous intrauterine death.[2,3]

The evidence for a genetic aetiology in a proportion of cases is strong. A number of aneuploidies and microdeletion syndromes exist, in which HLH is seen more frequently than would be expected, such as trisomies 13 and 18 and monosomy X (Turner syndrome). There are reports of increased incidences of chromosomal abnormalities in both live-born and fetal populations with HLH (~5 and ~12%, respectively) and in particular monosomy 11q23-qter (Jacobsen syndrome).[1,4–6] The incidence of extracardiac abnormalities found in individuals with HLH on post mortem ranges from 3 to 28%, depending on the study.[6–8] A number of syndromes due to single gene mutations are associated with congenital heart defects that may include HLH (eg, Holt–Oram,[9] Ellis–van Crefeld[10] and Rubinstein–Taybi syndromes[11]), and several families have been reported in which non-syndromic HLH is inherited in an autosomal-recessive[12–14] or autosomal-dominant fashion,[15,16] suggesting a monogenic cause in these families. Additionally, the recurrence of left-sided heart defects in first degree relatives of a proband with HLH is thought to be about 10–15%, suggestive of a strong and simple genetic component.[17–19]

De novo chromosomal aberrations (for example, microdeletions and balanced translocations) can provide opportunities to positionally map and clone specific genes involved in sporadic genetic defects. We ascertained a case with isolated HLH and an apparently balanced reciprocal *de novo* translocation between chromosomes 1 and 3. We hypothesized that the translocation might disrupt the normal function of a gene at one of the breakpoints and thereby cause the heart defect in the proband. Our high-resolution mapping of the breakpoints reveals a complex but grossly balanced rearrangement and enables us to identify the separation of the gene encoding the cardiac transcription factor PROX1 from conserved upstream elements as the likely cause of the defect in this child. We investigated PROX1 in fixed sections of developing human hearts and found that it localized to structures affected in HLH; this supports the role of PROX1 dysfunction in the pathogenesis of the heart defect in our case.

Materials and methods

Patient details

Through the fetal cardiology unit at Guy's Hospital, we antenatally ascertained a male fetus with isolated HLH, born at 38 weeks' gestation to non-consanguineous parents of Bangladeshi origin, with a birth weight of 2.63 kg (ninth centile). A diagnosis of classical HLH had been made at 20 weeks' gestation by ultrasound scan; this was subsequently monitored by fetal echocardiography. Postnatally no additional congenital malformations were found.

The patient has undergone two stages of the Norwood surgical correction, and is awaiting the final operation. His development was considerably delayed in a global manner at the 6-month examination, following which there has been significant improvement in all areas, suggesting that the delay was due to prolonged episodes as an in-patient following surgery. His growth parameters have been between the second and tenth centiles, which is appropriate for his family. He had some subtle facial dysmorphic features, with minimally posteriorly rotated ears and a long philtrum. There was no relevant family history. Ethical approval for this study was obtained from the Guy's and St Thomas' Local Research Ethics Committee (LREC).

Chromosome analysis

G-banded chromosome preparation from peripheral lymphocytes was carried out according to standard laboratory procedures. The resolution of preparations thus obtained was a minimum of 550 bands per haploid set.

Fluorescence *in situ* hybridization

Bacterial artificial chromosome (BAC) clones from the BACPAC resource centre and fosmid clones from the Sanger Institute were grown in Luria–Bertani (LB) liquid medium containing chloramphenicol, and DNA was prepared using a Qiagen Plasmid Mini Kit according to the manufacturers' instructions. The DNA was further purified by extraction with phenol/chloroform/isoamyl alcohol (25:24:1) and ethanol precipitation. Epstein–Barr virus-transformed lymphoblastoid cells from the proband were grown and metaphase chromosome spreads prepared according to standard procedures. The BAC and fosmid clones were used as fluorescence *in situ* hybridization (FISH) probes (prepared and analysed as described previously[20]) for hybridization to chromosome preparations from the proband and his parents.

Array comparative genome hybridization analysis

DNA from the proband was tested for genome imbalance using oligonucleotide arrays with 44 000 probes across the genome (Agilent). Hybridization was carried out according to the manufacturer's recommendation, and the commercial analysis software was used, incorporating a three-probe cutoff for imbalance flagging. The database of Genomic Variants (DGV; http://projects.tcag.ca/variation/)[21] was used to identify putatively benign copy number variants (CNVs).

Flow sorting of chromosomes

Purification of the translocated chromosomes, derivatives 1 and 3, was carried out using a flow cytometer (MoFlo; DAKO) as described previously.[22]

Polymerase chain reaction and sequence analysis

PCRs were carried out in a final volume of 25 μl using ~1–100 ng of genomic DNA as template. Each reaction contained 125 μM dNTPs (Amersham Biosciences), 0.5 μM each primer, 1 × PCR buffer (containing 1.5 mM $MgCl_2$) and 1 unit *Taq* DNA polymerase (Promega). All PCR primers were obtained from MWG Biotech (Milton Keynes, UK). The PCR reactions were overlaid with mineral oil and DNA was amplified in a PTC-100 thermal cycler (MJ Research Inc., Essex, UK), typically programmed to perform 30 cycles of 95°C for 1 min, 55–60°C for 30 s, 72°C for 1 min. Multiplex PCRs were optimized on whole genomic DNA before using on flow-sorted material. For amplification of DNA fragments greater than 1 kb, the Expand High Fidelity PCR system (Roche, Welwyn Garden City, UK) was used as directed by the manufacturer. Sequence analysis was performed using the ABI Prism Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Warrington, UK) according to manufacturers' instructions. Precipitated products were resuspended in 20 μl HiDi Formamide (Applied Biosystems) and electrophoresed in an ABI Prism 3100 Genetic Analyzer. Primer sequences are available on request.

Immunohistochemistry

We followed previously reported protocols for the use of human embryonic and fetal tissues and immunolocalization of proteins.[23,24] Briefly, tissues were obtained with informed consent and with permission from the Southampton and South West Hants joint LREC, staged according to the Carnegie classification or foot length, fixed in 4% paraformaldehyde in PBS, embedded in paraffin wax, and microtome sectioned at 5-μm thickness. Epitope retrieval was performed by immersion in boiling sodium citrate solution (0.01 M, pH6.0) for 40 min and then 60-s treatment with trypsin (1 mg/ml; Sigma-Aldrich). Immunochemistry was performed using primary antibodies that were raised to PROX1 (Abcam), ESRRG (Abcam), smooth muscle actin (Novocastra), CD34 (Novocastra) and D2-40 (Abcam). The secondary antibodies used were FITC-conjugated anti-rabbit Ig (Sigma) and Alexa 594-conjugated anti-mouse Ig (Molecular Probes). Dehydrated slides were mounted in Vectashield (Vector Laboratories) containing 4,6-diamidino-2-phenylindole (DAPI) nuclear counterstain. The slides were visualized and images captured with a Zeiss Axioplan fluorescence microscope and Axiovision software (Carl Zeiss, Welwyn Garden City, UK).

Results

Characterization of the translocation: revelation of an additional 4-Mb *de novo* inversion and definition of all four breakpoints

G-banded chromosome analysis of the proband revealed an apparently balanced reciprocal translocation between chromosomes 1 and 3 (46,XY,t(1;3)(q41;q27.1)dn). Both parents had normal karyotypes, and this rearrangement is therefore likely to have arisen *de novo*. Further testing of the patient using array comparative genome hybridisation (CGH)-detected imbalance only in regions previously identified as CNVs.

To identify gene(s) involved in the aetiology of the heart defect in this patient, we set out to characterize the 1;3 translocation breakpoints at the molecular level. The initial stage of this involved FISH using BACs approximately 1 Mb apart, followed by tiled BACs spanning the interval between those 1-Mb probes which flanked the breakpoint. Next, fosmid clones were used to identify ~40-kb regions, which spanned the breakpoint. The derivative chromosomes were then isolated by flow sorting (Supplementary Figure 2A), and breakpoints were further narrowed down by iterative multiplex PCR to intervals of a few kilobases (Supplementary Figure 2B). Finally long-distance PCR between primers either side of the translocation breakpoint enabled its characterization at the sequence level (Supplementary Figure 2C). The results of these studies are presented in detail in Supplementary Information, Section 1.

An unexpected result of this analysis was the revelation that the chromosome 1 breakpoint of the 1;3 translocation lies in the middle of a ~4-Mb inversion of chromosome 1 material (Figure 1a; Supplementary Figure 1). We found that this inversion is not present in either parent, and that it has therefore, like the translocation, arisen *de novo* in the proband. As the inversion breakpoints are as likely to be involved in the aetiology of the patient's phenotype as the translocation breakpoints, we set out to characterize these too. Our analysis confined the proximal and distal inversion breakpoints to a 7.1-kb region (chr1: 212145790 to 212152902) and a ~26-kb region (chr1: 216193862 to 216220000), respectively, with no detectable imbalance (see Supplementary Information). The implication of this is that the proximal half of band 1q41 (from ~212150000 to ~216210000) had undergone inversion at some point before (or possibly in concert with) the translocation event (Figure 1a). We were able to define the translocation breakpoints to base-pair resolution by sequencing a PCR product from der(3) (accession number FJ377539); this showed that translocation had occurred immediately distal to bases chr1: 214715823 and chr3: 185320520 (Figure 1a and c; Supplementary Figure 2C).

Bioinformatic analysis of breakpoint regions

Of the four *de novo* breakpoints in this HLH patient, none interrupts either an annotated transcription unit or a region across which there is any EST-based evidence of transcription. We therefore assumed that any causative link between any of the breakpoints and the patient's phenotype is likely to stem from a *cis*-acting effect on the transcription of neighbouring gene(s). To assess candidature we first identified the nearest and next-nearest genes

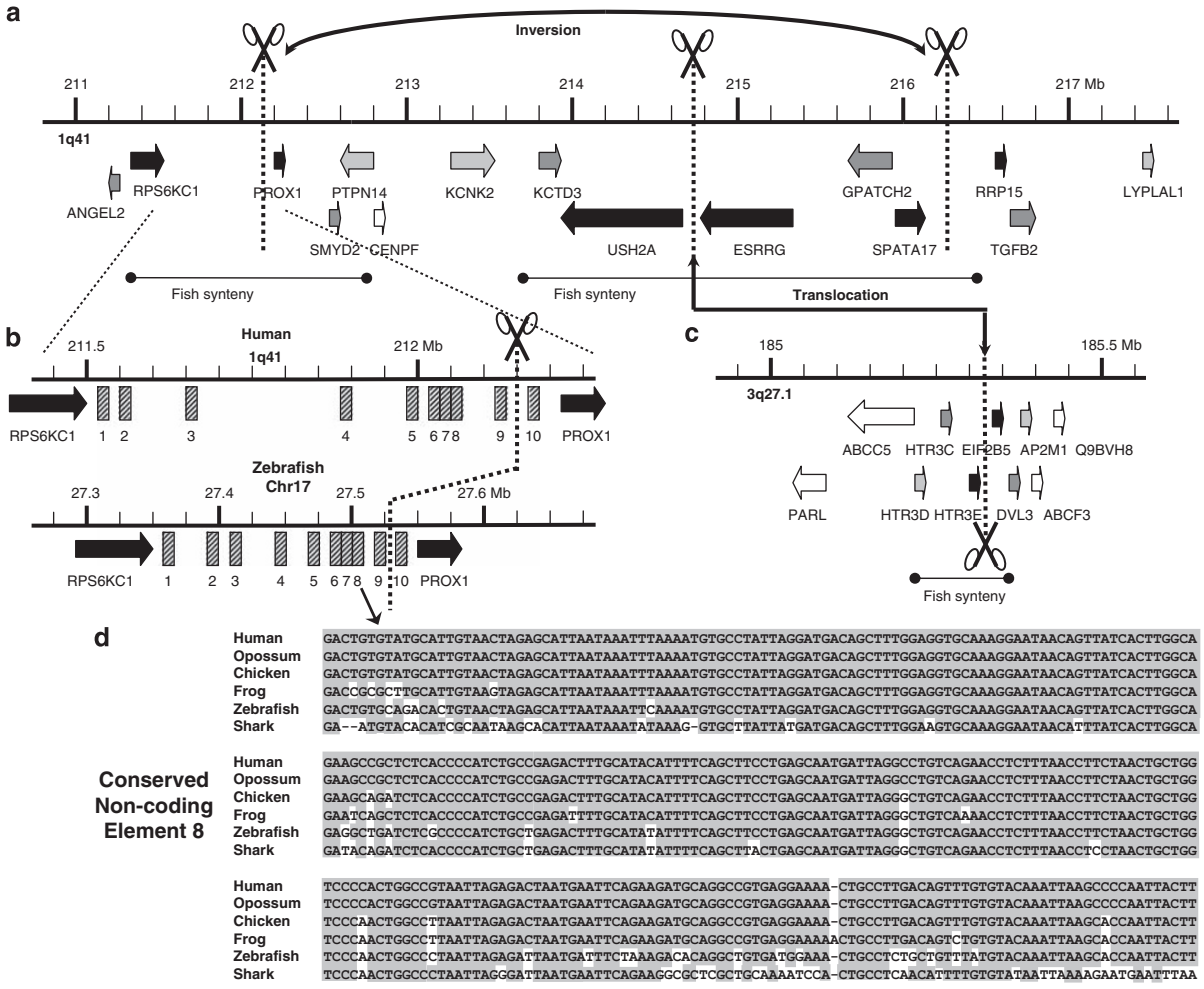

Figure 1 Detailed gene context of rearrangement breakpoint regions. (**a**) Scale diagram of the ~6-Mb region of 1q41 showing the positions of the inversion and translocation breakpoints with respect to genes. Genes are shaded black, dark grey, light grey and white according to their decreasing proximity to the breakpoints. The regions of synteny conserved in all vertebrate genomes examined are shown below ('fish synteny'). Positions of breakpoints are shown by bold dotted lines and scissor symbols. (**b**) Fourfold expansion of the environs of the proximal inversion breakpoint (comprising the *RPS6KC1*/*PROX1* intergenic region) together with the corresponding region of the zebrafish genome. Note that differing scales have been used for clarity; the intergenic regions are ~710 and ~190 kb in the respective species. CNEs 1–10 are shown as hatched boxes; note conserved order and juxtaposition. (**c**) Scale diagram of the region of 3q27.1 containing the translocation breakpoint. Note that a fourfold difference in scale between representations of the two chromosomes has been used for clarity. Annotation and shading as in (**a**). (**d**) Sequence alignment of a 300-bp section of CNE8. Human shows 100, 98, 95, 90 and 81% identity to opossum, chicken, frog, zebrafish and shark, respectively. Grey shading indicates identity to human.

on either side of each breakpoint, and then examined the literature for evidence of relevant function, expression and haploinsufficient phenotype (Supplementary Information, Section 2). Then, on the premise that long-distance *cis*-acting influences are likely to be mechanistically complex and therefore evolutionarily conserved, we assessed the conservation of synteny between neighbouring genes, the presence of conserved non-coding elements (CNEs) in the disrupted intergenic region, and the site of the breakpoint itself (Supplementary Information, Section 3). We assumed that a good candidate for the causative lesion would separate a gene of aetiologically relevant function from syntenically and structurally conserved elements.

We make a detailed case in Supplementary Information (summarized in Table 1) for the elimination of all but one of the candidate genes, with only *PROX1* fulfilling all of these criteria. Thus a breakpoint ~78 kb upstream of the *PROX1* promoter separates the gene from nine intergenic CNEs conserved in all vertebrates examined; heterozygosity for *Prox1* null mutations in mouse leads to a relevant phenotype, namely severe ventricular hypoplasia. *ESRRG* was also investigated because it partially fulfilled these

Table 1 Properties of candidate genes

		Breakpoint relationship								
	Gene	*1*[a]	*2*[b]	*3*[c]	*4*[d]	*5*[e]	*Homozygous phenotype*[f]	*Heterozygous phenotype*[f]	*Function*	*Expression*
Chr1	*VASH2*	3′	2	NA	917	959	Slightly reduced angiogenesis (M)	ND	Angiogenesis regulator	Endothelial cells
	ANGEL2	5′	1	NA	892	892	ND	ND	Phosphatase related	Ubiquitous
	RPS6KC1	3′	0	1	633	859	ND	ND	Endosome trafficking (kinase)	Ubiquitous
→										
	PROX1	5′	0	9	78	78	Lymph defects, 50% vent. hypo. (M)	Neonatal death, 30% vent. hypo. (M)	Transcription factor	Wide, including heart
	SMYD2	5′	1	NA	371	371	ND	ND	Lysine methyltransferase	Heart
	PTPN14	3′	2	NA	447	644	ND	ND	Protein tyrosine phosphatase	Wide
	KCNK2	5′	2	NA	1097	1097	Resistance to 'depression' state (M)	None (M)	Potassium channel	Neurons
	KCTD3	3′	1	NA	854	908	ND	ND	Channel-like multimerisation	Ubiquitous
	USH2A	5′	0	0	51	51	Deafness, blindness (H)	None (H)	Extracellular cilia link	Ear, eye
→										
	ESRRG	3′	0	1	26	663	9% vent. hypo. (M)	7% vent. hypo. (M)	Orphan nuclear receptor	Heart
	GPATCH2	5′	1	NA	335	335	ND	ND	Unknown	Ubiquitous
	SPATA17	3′	0	6	94	337	ND	ND	Calmodulin-binding	Gonads
→										
	RRP15	5′	0	0	317	317	ND	ND	Ribosome maturation	Ubiquitous
	TGFB2	5′	1	NA	378	378	Neonatal death, heart/skeletal defects (M)	None (M)	Growth factor	Heart
	LYPLAL1	5′	2	NA	1206	1206	ND	ND	Lysophospholipase-like	Wide
Chr3	*HTR3D*	3′	2	NA	81	88	ND	No cardiac defects (H)	5HT-gated ion channel	Gut/nervous
	HTR3C	3′	1	NA	58	68	ND	No cardiac defects (H)	5HT-gated ion channel	Nervous system
	HTR3E	3′	0	0	12	21	ND	No cardiac defects (H)	5HT-gated ion channel	Gut
→										
	EIF2B5	5′	0	0	14	14	Central demyelination (H)	None (H)	Translation initiation factor	Ubiquitous
	DVL3	5′	1	NA	34	34	Cardiac outflow defects (M)	None (M), no cardiac defects (H)	Wnt signalling pathway	Wide
	AP2M1	5′	2	NA	55	55	ND	No cardiac defects (H)	Endocytosis	Ubiquitous

Arrows indicate positions of breakpoints in this patient.
[a]Whether nearest breakpoint is 5′ or 3′ of gene.
[b]Number of genes between gene and nearest breakpoint.
[c]Number of syntenic conserved non-coding elements (CNEs) separated from gene by nearest breakpoint. NA, not applicable.
[d]Distance from nearest breakpoint to gene (in kilobases).
[e]Distance from nearest breakpoint to promoter (in kilobases).
[f]Reported homozygous and heterozygous loss-of-function phenotypes in human (H) and/or mouse (M). vent. hypo., ventricular hypoplasia; ND, not determined. References in Supplementary Information.

criteria (a breakpoint ~26 kb downstream of the *ESRRG* polyadenylation site, and therefore ~660 kb from the promoter, separates the gene from a single intergenic CNE; heterozygosity for *ESRRG* null mutations in mouse leads to mild ventricular hypoplasia).

PROX1 is a homeodomain-containing transcription factor, related to the *prospero* gene in *Drosophila*. Although PROX1 research has largely focussed on its role in development of the lens and lymphatic vessels, expression is seen in the developing heart in both mouse[25] and chicken.[26] Mice null for *Prox1* (*Prox1*$^{-/-}$) suffer from essentially complete failure of lymphatic development, dying *in utero* by day E15 with severe oedema;[27] they also have a hollow lens.[28] Moreover, 100% of *Prox1*$^{+/-}$

heterozygotes on most genetic backgrounds died within a few days of birth (only one strain showed an ~3% survival rate), showing that mice are highly sensitive to haploinsufficiency for *Prox1*.[27,28] This is consistent with problematic consequences of neonatal closure of the ductus arteriosus, which occurs in mice at about 3 h after birth.[29] Death of heterozygotes was accompanied with cyanosis and respiratory distress,[30] consistent with cardiac insufficiency. Critically, work published during revision of this article shows that neonatal *Prox1*$^{+/-}$ mice do indeed have severe ventricular hypoplasia, with hearts ~30% smaller than those of wild-type mice.[31] We believe that this makes *PROX1* a prime functional candidate for HLH in humans.

In addition to this functional candidacy, we show that the *RPS6KC1*/*PROX1* intergenic region contains an array of 11 intergenic CNEs (see Figure 1b) shared by all vertebrates examined (see Supplementary Information, Section 3). In Figure 1d we show one particularly striking example in which 90% identity is shared by a 300-bp segment of CNE8 between human and fish. Such degree of conservation between these species would be unusual for a coding region, is exceptional for an intergenic region, and is highly suggestive of a critical biological function for which conserved synteny is important (such as *cis*-regulation of transcription). Of particular relevance to our HLH patient, the proximal inversion breakpoint separates 9 of the 11 CNEs (CNEs 1–9) from the *PROX1* gene. By contrast, only one intergenic CNE is separated from the 3′ end of the *ESRRG* gene, and this is more likely to be involved in regulating the adjacent *USH2A* gene; there are also six intronic CNEs within the *USH2A* gene (Supplementary Information, Section 3).

Expression of PROX1 in the developing heart

The detailed expression of Prox1 has been examined in certain tissues of presumed biological activity, including the lens of the eye,[32] the lymphatic system[27] and the CNS.[33] However Prox1 is also highly expressed in the developing mouse heart from E10.5 to E16.5,[25] along with other organs such as skeletal muscle, CNS, pancreas and eye; similar expression is seen in the developing chicken,[26] where it is confined to the cardiomyocytes and to endothelial cells on the concave side of the valves.[34,35]

We set out to examine the localization of PROX1 in developing human embryonic heart. Embryonic tissue was available from ~50–70 days after conception; during this period valvular development is completed, which may be critical in the pathogenesis of HLH. PROX1 was detected in all four valves (Figure 2) and double immunofluorescent microscopy showed cellular colocalization within endothelial cells (CD34-positive) and cells positive for smooth muscle actin (SMA). There was also significant expression in the walls of the aorta and pulmonary artery in addition to the atria and ventricles (left and right) – data not shown. The highest levels of expression were detected in nuclei of CD34-positive endothelial cells lining coronary arteries and veins, and D2-40-positive endothelial cells lining lymphatic vessels (data not shown). These data indicate that PROX1 has a role in cardiac development and is expressed in structures principally malformed in HLH (an expression pattern also seen in chicken[35]); this supports the proposal that disruption of PROX1 function plays a role in our patient's defect.

As one of the more likely, and testable, outcomes of the translocation upstream of the *PROX1* gene would be reduced or increased expression of the *PROX1* allele, which is in *cis* with the translocation, we looked for heterozygous sequence variants in the *PROX1* exons of the patient. This would enable us to test a small amount of heart tissue, made available after one of the patients' operations, for unbalanced expression from the two *PROX1* alleles. Unfortunately no heterozygosity was seen so we were unable to distinguish between the two alleles at the transcript level (data not shown).

We also examined the localization of ESRRG in developing human embryonic heart (see Supplementary Information, Section 4). ESRRG was detected in all four valves, walls of the aorta and pulmonary artery and cardiomyocytes in the ventricles and atria. Thus expression pattern alone cannot differentiate between the candidacy of PROX1 and ESRRG in this patient.

Discussion

We have described a patient whose HLH is associated with a complex *de novo* balanced chromosome rearrangement involving four separate breakpoints; one in 3q27.1 and three in 1q41. All four breakpoints lie in intergenic regions, suggesting that any causative link between the complex rearrangement and the heart defect probably arises as a result of dominantly acting disruption of expression of a neighbouring gene. We make the following case for the involvement of *PROX1* in this patient: (1) The *de novo* rearrangement disrupts a spatial relationship between the *PROX1* gene and numerous upstream elements that has been conserved for half a billion years of evolution; (2) Heterozygosity for a null mutation of *Prox1* in the mouse causes neonatal death in ~100% of pups, almost certainly related to their substantial ventricular hypoplasia; (3) PROX1 is found to be expressed in the developing mitral and aortic valves, proximal aorta and left ventricle, locations in which aberrant development might be expected to cause hypoplastic left heart; (4) Genes adjacent to the three remaining breakpoints are either poor functional candidates for HLH or are not separated from conserved elements by the rearrangement.

The complex rearrangement that we see in this patient is in itself interesting. Both the inversion and the

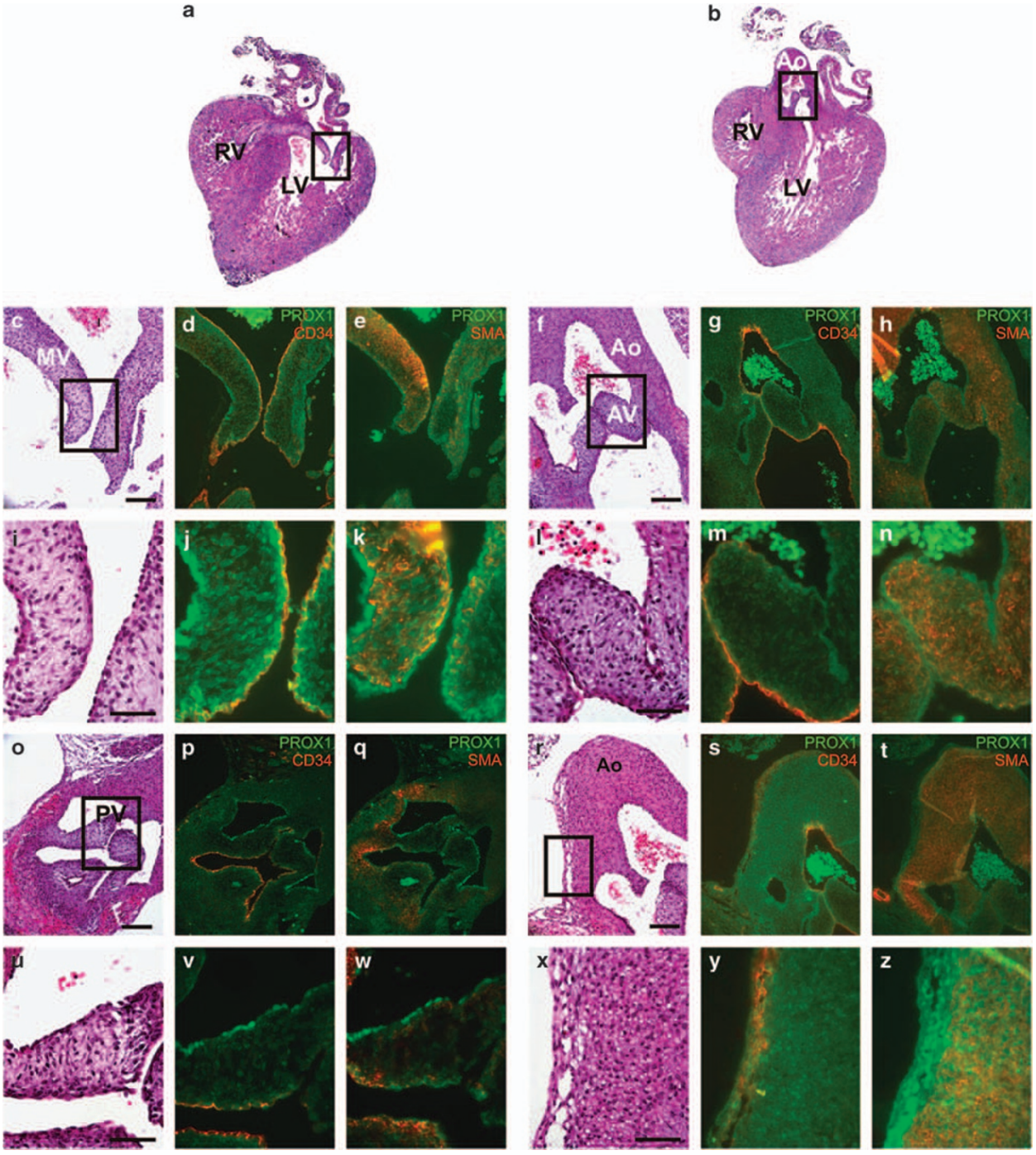

Figure 2 Expression of PROX1 in the developing human heart. Double immunofluorescent microscopy showing the localization of PROX1 within the mitral valve (**d**, **e**, **j** and **k**), aortic valve (**g**, **h**, **m** and **n**), pulmonary valve (**p**, **q**, **v** and **w**) and aortic wall (**s**, **t**, **y** and **z**) of a human heart 54 days after conception. Haematoxylin and eosin-stained sections are shown as reference (**a**, **b**, **c**, **f**, **i**, **l**, **o**, **r**, **u** and **x**); boxes indicate regions of higher magnification. PROX1 is detected in each of the valves and in the wall of the aorta, colocalizing with endothelial marker CD34 and smooth muscle actin (SMA). Ao: aorta; AV: aortic valve; LV: left ventricle; MV: mitral valve; PV: pulmonary valve; RV: right ventricle. Scale bar = 100 μm (**c**, **f**, **o** and **r**) or 50 μm (**i**, **l**, **u** and **x**).

translocation have arisen *de novo*, implying that the events occurred either sequentially in one generation or as a single concerted event. If the events were sequential, then it is far more parsimonious that the inversion happened first; the subsequent occurrence of a translocation within this inverted region may be truly independent or may be

causally linked to difficulties of meiotic synapsis between the inverted and non-inverted homologous chromosomes. We have previously reported a complex *de novo* deletion–inversion–deletion event on the X chromosome, which involved four breakpoints spread over a 2.2-Mb region, for which we posited a complex recombination-mediated resolution of a topologically looped intermediate.[36] It may be that such a topological complication triggered the events that took place in the current patient. For large inversions, the formation of 'inversion loops' is postulated, crossing over within which gives rise to recombinant chromosomes. However, for small regions of inversion, it is thought more likely that the unmatched regions 'bud out' of the synapsed chromosomes, perhaps leaving the bud vulnerable to crossing over with a different chromosome. However, although heterozygosity for inversion polymorphisms has been implicated as a predisposing factor for deletions,[37,38] it has been excluded in at least one instance for translocations[39].

The functions of evolutionarily conserved regions in untranslated or non-coding regions (sometimes known as CNEs) have been the subject of speculation since their first recognition. A recent broad study of the properties of CNEs across the vertebrate clade[40] shows that transcription factor and developmental genes are highly over-represented as nearest genomic neighbours to CNEs, and that ~90% of CNEs behave as tissue-specific enhancers when introduced into zebrafish. Clustered CNEs are often found in gene deserts,[41] as is the case with *PROX1*, and gene deserts have been shown to be involved in three-dimensional higher order interphase chromatin structure.[42] In addition, there are many well-characterized examples of pathogenic rearrangements, which disrupt the expression of genes lying tens or hundreds of kilobases away;[43] this effect, known as the position effect, almost always involves gene deserts 5′ or 3′ to genes encoding transcription factors (such as *SOX9*[44] in campomelic dysplasia and *POU3F4*[45] in X-linked deafness) or other developmental genes (such as *SHH*[46] in polydactyly or holoprosencephaly). In most cases where the mechanism is known, the rearrangement removes transcriptional regulatory sequences and results in some degree of loss of expression of the disease gene and a dominantly inherited (haploinsufficient) disorder.[43]

Our *PROX1* upstream CNEs show degrees of conservation similar to or exceeding those examined in the enhancer and ontological study mentioned above.[40] We assume, as do others, that such sequences either bind constellations of transcription factors or participate in some other way in the epigenetic structuring of this region of the genome. Analysis of our aligned *PROX1* upstream CNEs for conserved transcription factor binding sites using MULAN[47] shows absolute conservation of many potential binding sites between all species examined in most of the CNEs (an example output of MULAN is shown for CNE8 in Supplementary Information, Section 3). We propose that some of these CNEs are involved in the transcriptional regulation of *PROX1* and that their separation from the *PROX1* gene in our patient results in loss of a subset of the *PROX1* expression pattern from one allele. This would then result in a ~50% reduction in *PROX1* expression in some or all tissues (including the heart), with consequent haploinsufficiency. To our mind the haploinsufficient *Prox1*$^{+/-}$ phenotype seen in the mouse (ventricular hypoplasia) is strikingly similar to the neonatal consequences of HLH in humans, and *PROX1* should be considered as a candidate gene for other cases of isolated HLH.

Acknowledgements

We are indebted to the proband and his family. HKG was supported by a research fellowship from the Guy's and St Thomas' Charitable Foundation. BLN and NPC are supported by the Wellcome Trust. SRP was supported by a Health Foundation Fellowship. CS, VJK and DIW are supported by the British Heart Foundation and Little Heart Matters. RGR received funding from the British Heart Foundation and the Muscular Dystrophy Campaign. We are not aware of any conflicts of interest.

References

1 Grossfeld PD: The genetics of hypoplastic left heart syndrome. *Cardiol Young* 1999; **9**: 627–632.
2 Hoffman JI, Kaplan S: The incidence of congenital heart disease. *J Am Coll Cardiol* 2002; **39**: 1890–1900.
3 Simpson JM: Hypoplastic left heart syndrome. *Ultrasound Obstet Gynecol* 2000; **15**: 271–278.
4 Phillips HM, Renforth GL, Spalluto C *et al*: Narrowing the critical region within 11q24-qter for hypoplastic left heart and identification of a candidate gene, JAM3, expressed during cardiogenesis. *Genomics* 2002; **79**: 475–478.
5 Jacobsen P, Hauge M, Henningsen K, Hobolth N, Mikkelsen M, Philip J: An (11;21) translocation in four generations with chromosome 11 abnormalities in the offspring. A clinical, cytogenetical, and gene marker study. *Hum Hered* 1973; **23**: 568–585.
6 Brackley KJ, Kilby MD, Wright JG *et al*: Outcome after prenatal diagnosis of hypoplastic left-heart syndrome: a case series. *Lancet* 2000; **356**: 1143–1147.
7 Allan LD, Sharland GK, Milburn A *et al*: Prospective diagnosis of 1,006 consecutive cases of congenital heart disease in the fetus. *J Am Coll Cardiol* 1994; **23**: 1452–1458.
8 Natowicz M, Chatten J, Clancy R *et al*: Genetic disorders and major extracardiac anomalies associated with the hypoplastic left heart syndrome. *Pediatrics* 1988; **82**: 698–706.
9 Sletten LJ, Pierpont ME: Variation in severity of cardiac disease in Holt-Oram syndrome. *Am J Med Genet* 1996; **65**: 128–132.
10 Baujat G, Le Merrer M: Ellis-van Creveld syndrome. *Orphanet J Rare Dis* 2007; **2**: 27.
11 Hanauer D, Argilla M, Wallerstein R: Rubinstein-Taybi syndrome and hypoplastic left heart. *Am J Med Genet* 2002; **112**: 109–111.
12 Maestri NE, Beaty TH, Boughman JA: Etiologic heterogeneity in the familial aggregation of congenital cardiovascular malformations. *Am J Hum Genet* 1989; **45**: 556–564.
13 Grobman W, Pergament E: Isolated hypoplastic left heart syndrome in three siblings. *Obstet Gynecol* 1996; **88**: 673–675.
14 Kojima H, Ogimi Y, Mizutani K, Nishimura Y: Hypoplastic-left-heart syndrome in siblings. *Lancet* 1969; **2**: 701.

15 Hajdu J, Marton T, Toth-Pal E *et al*: Severe left heart developmental disorder and severe fetal arrhythmia in the same family – a coincidental association? *Orv Hetil* 1998; **139**: 767–769.
16 Brownell LG, Shokeir MH: Inheritance of hypoplastic left heart syndrome (HLHS): further observations. *Clin Genet* 1976; **9**: 245–249.
17 Loffredo CA, Chokkalingam A, Sill AM *et al*: Prevalence of congenital cardiovascular malformations among relatives of infants with hypoplastic left heart, coarctation of the aorta, and d-transposition of the great arteries. *Am J Med Genet A* 2004; **124A**: 225–230.
18 Brenner JI, Berg KA, Schneider DS, Clark EB, Boughman JA: Cardiac malformations in relatives of infants with hypoplastic left-heart syndrome. *Am J Dis Child* 1989; **143**: 1492–1494.
19 Gill HK, Splitt M, Sharland GK, Simpson JM: Patterns of recurrence of congenital heart disease: an analysis of 6,640 consecutive pregnancies evaluated by detailed fetal echocardiography. *J Am Coll Cardiol* 2003; **42**: 923–929.
20 Male A, Davies A, Bergbaum A *et al*: Delineation of an estimated 6.7 MB candidate interval for an anophthalmia gene at 3q26.33-q28 and description of the syndrome associated with visible chromosome deletions of this region. *Eur J Hum Genet* 2002; **10**: 807–812.
21 Iafrate AJ, Feuk L, Rivera MN *et al*: Detection of large-scale variation in the human genome. *Nat Genet* 2004; **36**: 949–951.
22 Carter NP: Bivariate chromosome analysis using a commercial flow cytometer. *Methods Mol Biol* 1994; **29**: 187–204.
23 Piper K, Brickwood S, Turnpenny LW *et al*: Beta cell differentiation during early human pancreas development. *J Endocrinol* 2004; **181**: 11–23.
24 Hearn T, Spalluto C, Phillips VJ *et al*: Subcellular localization of ALMS1 supports involvement of centrosome and basal body dysfunction in the pathogenesis of obesity, insulin resistance, and type 2 diabetes. *Diabetes* 2005; **54**: 1581–1587.
25 Oliver G, Sosa-Pineda B, Geisendorf S, Spana EP, Doe CQ, Gruss P: Prox 1, a prospero-related homeobox gene expressed during mouse development. *Mech Dev* 1993; **44**: 3–16.
26 Tomarev SI, Sundin O, Banerjee-Basu S, Duncan MK, Yang JM, Piatigorsky J: Chicken homeobox gene Prox 1 related to Drosophila prospero is expressed in the developing lens and retina. *Dev Dyn* 1996; **206**: 354–367.
27 Wigle JT, Oliver G: Prox1 function is required for the development of the murine lymphatic system. *Cell* 1999; **98**: 769–778.
28 Wigle JT, Chowdhury K, Gruss P, Oliver G: Prox1 function is crucial for mouse lens-fibre elongation. *Nat Genet* 1999; **21**: 318–322.
29 Tada T, Kishimoto H: Ultrastructural and histological studies on closure of the mouse ductus arteriosus. *Acta Anat (Basel)* 1990; **139**: 326–334.
30 Harvey NL, Srinivasan RS, Dillard ME *et al*: Lymphatic vascular defects promoted by Prox1 haploinsufficiency cause adult-onset obesity. *Nat Genet* 2005; **37**: 1072–1081.
31 Risebro CA, Searles RG, Melville AA *et al*: Prox1 maintains muscle structure and growth in the developing heart. *Development* 2009; **136**: 495–505.
32 Duncan MK, Cui W, Oh DJ, Tomarev SI: Prox1 is differentially localized during lens development. *Mech Dev* 2002; **112**: 195–198.
33 Lavado A, Oliver G: Prox1 expression patterns in the developing and adult murine brain. *Dev Dyn* 2007; **236**: 518–524.
34 Wilting J, Buttler K, Schulte I, Papoutsi M, Schweigerer L, Manner J: The proepicardium delivers hemangioblasts but not lymphangioblasts to the developing heart. *Dev Biol* 2007; **305**: 451–459.
35 Rodriguez-Niedenfuhr M, Papoutsi M, Christ B *et al*: Prox1 is a marker of ectodermal placodes, endodermal compartments, lymphatic endothelium and lymphangioblasts. *Anat Embryol (Berl)* 2001; **204**: 399–406.
36 Wheway JM, Yau SC, Nihalani V *et al*: A complex deletion-inversion-deletion event results in a chimeric IL1RAPL1-dystrophin transcript and a contiguous gene deletion syndrome. *J Med Genet* 2003; **40**: 127–131.
37 Koolen DA, Vissers LE, Pfundt R *et al*: A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat Genet* 2006; **38**: 999–1001.
38 Shaw-Smith C, Pittman AM, Willatt L *et al*: Microdeletion encompassing MAPT at chromosome 17q21.3 is associated with developmental delay and learning disability. *Nat Genet* 2006; **38**: 1032–1037.
39 Zollino M, Lecce R, Murdolo M *et al*: Wolf-Hirschhorn syndrome-associated chromosome changes are not mediated by olfactory receptor gene clusters nor by inversion polymorphism on 4p16. *Hum Genet* 2007; **122**: 423–430.
40 Woolfe A, Goodson M, Goode DK *et al*: Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol* 2005; **3**: e7.
41 Ovcharenko I, Loots GG, Nobrega MA, Hardison RC, Miller W, Stubbs L: Evolution and functional classification of vertebrate gene deserts. *Genome Res* 2005; **15**: 137–145.
42 Shopland LS, Lynch CR, Peterson KA *et al*: Folding and organization of a contiguous chromosome region according to the gene distribution pattern in primary genomic sequence. *J Cell Biol* 2006; **174**: 27–38.
43 Kleinjan DA, van Heyningen V: Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet* 2005; **76**: 8–32.
44 Pfeifer D, Kist R, Dewar K *et al*: Campomelic dysplasia translocation breakpoints are scattered over 1 Mb proximal to SOX9: evidence for an extended control region. *Am J Hum Genet* 1999; **65**: 111–124.
45 de Kok YJ, Vossenaar ER, Cremers CW *et al*: Identification of a hot spot for microdeletions in patients with X-linked deafness type 3 (DFN3) 900 kb proximal to the DFN3 gene POU3F4. *Hum Mol Genet* 1996; **5**: 1229–1235.
46 Lettice LA, Horikoshi T, Heaney SJ *et al*: Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *Proc Natl Acad Sci USA* 2002; **99**: 7548–7553.
47 Ovcharenko I, Loots GG, Giardine BM *et al*: Mulan: multiple-sequence local alignment and visualization for studying function and evolution. *Genome Res* 2005; **15**: 184–194.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)