OPEN ACCESS

ORIGINAL ARTICLE

Whole exome sequencing in family trios reveals *de novo* mutations in *PURA* as a cause of severe neurodevelopmental delay and learning disability

David Hunt,[1] Richard J Leventer,[2] Cas Simons,[3] Ryan Taft,[3,4,5] Kathryn J Swoboda,[6] Mary Gawne-Cain,[7] the DDD study,[8] Alex C Magee,[9] Peter D Turnpenny,[10] Diana Baralle[1,11]

For numbered affiliations see end of article.

Correspondence to
Dr Diana Baralle and David Hunt, Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton, SO16 5YA, UK; david.hunt2@uhs.nhs.uk, d.baralle@soton.ac.uk

Received 30 September 2014
Revised 6 October 2014
Accepted 8 October 2014

ABSTRACT

Background *De novo* mutations are emerging as an important cause of neurocognitive impairment, and whole exome sequencing of case-parent trios is a powerful way of detecting them. Here, we report the findings in four such trios.

Methods The Deciphering Developmental Disorders study is using whole exome sequencing in family trios to investigate children with severe, sporadic, undiagnosed developmental delay. Three of our patients were ascertained from the first 1133 children to have been investigated through this large-scale study. Case 4 was a phenotypically isolated case recruited into an undiagnosed rare disorders sequencing study.

Results Protein-altering *de novo* mutations in *PURA* were identified in four subjects. They include two different frameshifts, one inframe deletion and one missense mutation. *PURA* encodes Pur-α, a highly conserved multifunctional protein that has an important role in normal postnatal brain development in animal models. The associated human phenotype of *de novo* heterozygous mutations in this gene is variable, but moderate to severe neurodevelopmental delay and learning disability are common to all. Neonatal hypotonia, early feeding difficulties and seizures, or 'seizure-like' movements, were also common.

Additionally, it is suspected that anterior pituitary dysregulation may be within the spectrum of this disorder. Psychomotor developmental outcomes appear variable between patients, and we propose a possible genotype–phenotype correlation, with disruption of Pur repeat III resulting in a more severe phenotype.

Conclusions These findings provide definitive evidence for the role of *PURA* in causing a variable syndrome of neurodevelopmental delay, learning disability, neonatal hypotonia, feeding difficulties, abnormal movements and epilepsy in humans, and help clarify the role of *PURA* in the previously described 5q31.3 microdeletion phenotype.

INTRODUCTION

Neurodevelopmental disorders are common and encompass a broad range of intellectual, behavioural and motor disabilities. Learning disability alone affects 1%–3% of the population and, for the most part, has a complex genetic basis.[1] Indeed, it is this complex genetic heterogeneity and variability of expression that has previously posed a significant barrier to the investigation and molecular diagnosis of neurodevelopmental disorders.

However, with the advent of next-generation sequencing technology, extensive interrogation of the exome has become possible. The use of whole exome sequencing (WES) has enabled the identification of pathogenic mutations in patients with well-characterised neurodevelopmental phenotypes, such as Kabuki syndrome[2] and Schinzel–Giedion syndrome.[3]

In many cases, however, there may be no consistent physical characteristics to help group patients with sporadic neurodevelopmental disorders for molecular genetic investigation. For this reason, such cases are inherently more challenging. One paradigm that has proved to be extremely effective is WES in family trios.[4 5] This approach has helped to successfully identify numerous pathogenic *de novo* mutations as the cause of sporadic neurodevelopmental delay,[4 6 7] and forms the basis of the Deciphering Developmental Disorders (DDD) study, through which the mutations in three of our four patients were identified.

The enrichment for *de novo* mutations as a cause of sporadic neurodevelopmental disorders is not surprising given the overall association with reduced fecundity and the baseline rate of DNA replication errors, which has been reported from detailed genomic studies as $\sim 10^{-8}$ *de novo* germline base substitutions per base pair per generation.[8]

We report four unrelated children with significant neurodevelopmental delay who have been investigated by WES in family trios and found to have pathogenic *de novo* mutations in *PURA* (MIM 600473).

METHODS

Of our four patients, three were referred to regional Clinical Genetics services across the UK, where they were recruited to the DDD study (http://www.ddduk.org). DDD has so far investigated 1133 children with severe, undiagnosed developmental delay, and their parents, using a combination of genome-wide assays to detect all major classes of genetic variation in the protein-coding portion of the genome. They have recorded clinical information and phenotypes using the Human Phenotype Ontology[9] via a secure web portal within the DECIPHER database.[10]

To cite: Hunt D, Leventer RJ, Simons C, *et al. J Med Genet* Published Online First: [*please include* Day Month Year] doi:10.1136/jmedgenet-2014-102798

DNA samples from patients and their parents were analysed by the Wellcome Trust Sanger Institute using high-resolution microarray analysis (array-comparative genomic hybridisation (CGH) and SNP-genotyping) to investigate CNVs in the child, and exome sequencing to investigate SNPs and small insertions/deletions (indels). Putative *de novo* sequence variants were validated using targeted Sanger sequencing. The population prevalence (minor allele frequency) of each variant in nearly 15 000 samples from diverse populations was recorded, and the effect of each genomic variant was predicted using the Ensembl Variant Effect Predictor.[11] Likely diagnostic variants in known developmental disorder genes were fed back to the referring clinical geneticists for validation and discussion with the family via the patient's record in DECIPHER, where they can be viewed in an interactive genome browser. Full genomic datasets were also deposited in the European Genome–Phenome Archive (http://www.ebi.ac.uk/ega).

Patient 4 was referred to paediatric neurology. She underwent extensive neurological and metabolic investigations in Australia. The exomes of Patient 4 and both parents were sequenced in an n=1 family trio study used for diagnostic exploration by Ambry Genetics using SureSelect Target Enrichment System (Agilent Technologies) followed by 2×100 nt paired-end sequencing on a Illumina HiSeq 2000. Raw sequence reads for Patient 4 and her parents were aligned to the reference human genome (GRCh37), and pedigree-informed variant calling was performed using the Real Time Genomics integrated analysis tool rtgFamily V.3.2.[12] All variants were annotated using SnpEff V.3.4[13] using data from dbNSFP2.4[14] and dbSNP138.[15] Subsequent analysis and identification of candidate variants was performed with an in-house workflow incorporating the annotated variant data and pedigree information.

RESULTS

Clinical photographs of the patients are shown in figure 1. Clinical features are shown in detail in table 1.

The mutations found by WES for each case are described below. No other causative mutations were indentified in the exomes.

Patient 1

Patient 1 has a *de novo* frameshift mutation (p.Phe243Tyrfs*50) in *PURA*.

She had neonatal hypotonia and was nasogastric (NG) tube-fed for the first week of life. Her swallow has remained poor. A single apnoeic/hyponoeic episode occurred while she was a neonate. Abnormal 'seizure-like' movements were investigated at 7 months of age by EEG, which was normal. Visual evoked potentials revealed broadened wave forms consistent with neurological nystagmus (with preserved optokinetic nystagmus).

At her last clinical assessment, aged 4 years 7 months, she was not walking, and remained non-verbal. Patient 1 had hypotonic facies with a prominent forehead, epicanthic folds and mild telecanthus. Lower limb posture was abnormal with feet held in plantar flexion. There was restricted ankle movement, mild hypotonia and generalised weakness. Coordination was poor, but not grossly ataxic. Intermittent dysconjugate gaze was noted.

Prominent early breast bud development led to endocrine investigations that revealed this to be gonadotropin-dependent. She has been treated with intramuscular decapeptyl from age 3 years. MRI brain scans show delayed myelination.

Patient 2

Patient 2 has a *de novo* frameshift mutation (p.Glu283Argfs*45) in *PURA*.

She did not have neonatal hypotonia, respiratory difficulties or feeding problems. She achieved unsupported sitting at 12 months, independent walking at 24 months and first words at 2 years 6 months. At 14 years 3 months, she was independently mobile, able to dress herself and feed herself. She communicates in sentences, although with limited vocabulary. Patient 2 has an anxious disposition and lacks awareness of danger. Clinical examination revealed microcephaly, tall forehead, hypotonic facies, mild facial asymmetry, upslanting palpebral fissures and large central incisors. She has long thin fingers and toes, with 5th finger clino/camptodactyly bilaterally, over-riding 2nd toes and deep palmar creases.

Patient 3

Patient 3 has a non-synonymous missense mutation (p.Ile206Phe). This amino acid substitution arises in Pur repeat II, a very highly conserved region of sequence within Purα (figure 2). Crystallography studies suggest that Pur repeat I and Pur repeat II interact to form a functional Pur domain.[16] *In silico* analysis with SIFT and PolyPhen produced scores of 0.01 and 0.969, respectively, supporting pathogenicity.

Respiratory distress at birth necessitated supplementary oxygen. Neonatal hypotonia and hypoglycaemia were present, and NG tube feeding was necessary. She sat unsupported at 12–14 months, walked at 22 months and said her first words at ∼3 years 6 months. At 12 years 10 months, she was able to run unsteadily with a wide-based gait. She communicates at a basic level with short phrases, repetitive speech and limited comprehension. Her behaviour can be obsessional and attention-seeking, with limited awareness of others. Patient 3 has a long face, full cheeks, high forehead and telecanthus. Neurological examination revealed hypotonia, mild weakness and poor coordination. EEG at 3 years for possible seizures showed: occasional paroxysmal discharges in the form of spikes and sharp waves over the right frontal and left mid-temporal region in sleep.

Patient 4

Patient 4 has an inframe deletion (p.Phe233del), affecting a very highly conserved phenylalanine residue within Pur repeat III, a presumed functional domain of Purα which is necessary for homodimerisation in crystallography studies[16] (figure 1), and is present even in very distantly related organisms such as *Caenorhabditis elegans*. It is, therefore, highly likely to be of functional significance.

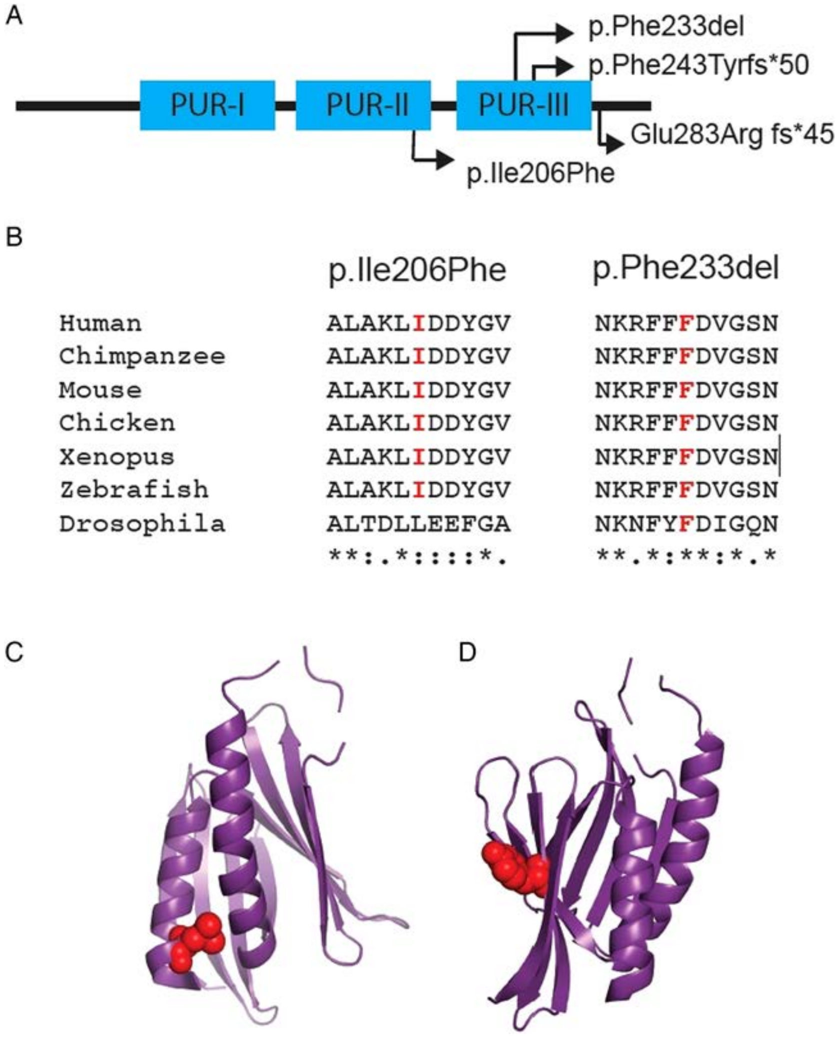
She developed central apnoea, hypothermia and severe hypotonia from day 2 of life. There was absent suck and gag reflex requiring early NG feeding. At 6 years 9 months, unsupported sitting had not been achieved. There is little language development.

She has hypotonic facies, frontal bossing and thin upper lip. Neurological examination revealed generalised hypotonia and dystonic/dyskinetic facial and limb movements. There was generalised weakness but no ataxia. Eye movements were dysconjugate.

Seizures commenced at 14 months with infantile spasms, and progressed to tonic seizures and focal dyscognitive seizures. Seizures have proved difficult to control.

EEG recordings have been normal or mildly slow when seizures are under control, but highly abnormal during seizures with near-continuous multifocal and bisynchronous sharp/slow activity maximal posteriorly. Video telemetry has revealed epileptic spasms and/or tonic seizure activity. Inborn errors of neurotransmitter biosynthesis and metabolism have been excluded by both Sanger and WES.

Figure 1 Depictions of the four mutations identified within *PURA*. (A) Schematic of Purα, depicting sites of the different mutations identified in our patients with respect to the Pur repeat regions (I–III). (B) Sequence alignment illustrating the high level of conservation of amino acids affected by the p.Ile206Phe and p.Phe233del mutations. Asterisk denotes complete conservation, colon denotes high conservation and single dot denotes moderate conservation. (C) and (D) illustrate the respective locations of the p.Ile206Phe and p.Phe233del mutations within the tertiary structure of the protein.



Serial MRI brain scans have been performed since birth (figure 3). These showed a right frontal horn cyst, which subsequently resolved. There was also patchy high attenuation throughout the white matter. Myelination was delayed but complete by 5 years, by which time there was evidence of excessive extra-axial fluid spaces and possible parenchymal volume loss. MR spectroscopy has demonstrated decreased N-acetyl aspartate within the frontal lobes and basal ganglia.

DISCUSSION

PURA encodes a ubiquitously expressed protein, Purα, which contains an N-terminal glycine-rich region, three Pur repeats (I–III) and a C-terminal glutamine–glutamate rich domain[16] (figure 1). The full-length protein is 322aa in humans and gives rise to a 28 kDa product.[17 18]

Purα is very highly conserved across the phylogenetic tree (figure 1), with regulatory roles in DNA replication, gene transcription, RNA transport and mRNA translation. Originally, it was identified in mouse due to its ability to bind to a sequence within the myelin basic protein promoter.[19 20] The human form was identified through its binding to a purine-rich element within an origin of DNA replication upstream of the human *c-MYC* gene.[17] A consensus sequence for the purine-rich single strand of the so-called PUR element, to which Purα binds, was subsequently derived. It has since become apparent that Purα's preferential recognition sequence comprises GGN repeats.

In order to initiate DNA replication and gene transcription, Purα first destabilises the DNA helix so that it may then bind its target sequence on a single DNA strand.[21] It is able to bind both linearised and supercoiled DNA. However, mutation studies have revealed that the carboxy terminal segment of Purα, which includes Pur repeat III, is necessary for destabilisation of linearised DNA.[22]

Purα has been shown to be important in controlling gene transcription from an array of different genes. Interestingly, it has gene-specific roles as either an activator or repressor of transcription. Purα activates transcription for a large number of cellular genes including those encoding myelin basic protein,[19] tumour necrosis factor α,[23] BC1[24] and the neuron-specific TATA-less gene FE65.[25] Furthermore, proteomics studies suggest that both Purα and its paralog, Purβ, may have an important regulatory role in control of the gene expression of myelin proteolipid protein (Plp1), which is the most abundant protein in central nervous system myelin and is developmentally regulated. Expression of Plp1 peaks in oligodendrocytes during active myelination.[26] By contrast, Purα represses expression from a wide range of genes including amyloid-β precursor protein,[27] α-actin[28] and gata2.[29] There is also evidence that Purα is involved in controlling its own transcription through a process of autoregulation.[30]

There are two independently generated *Pura*$^{-/-}$ mice which have helped our understanding of Purα's role in normal

Table 1 Clinical phenotype descriptions of the four index patients

PURA phenotype		Patient 1	Patient 2	Patient 3	Patient 4
Background	Mutation	c.726_727delGT (p.Phe243Tyrfs*50)	c.847delG (p. Glu283Arg fs*45)	c.616A>T (p.Ile206Phe)	c.697_699delTTC (p.Phe233del)
	Inheritance	AD—*de novo*	AD—*de novo*	AD—*de novo*	AD—*de novo*
	Ethnicity	Caucasian	Caucasian	Caucasian	Caucasian
	Sex	F	F	F	F
	Family history of note	None	None	None	None
	Pregnancy	Natural, uneventful	Natural, uneventful	Natural, uneventful	ICSI, uneventful
	Delivery	NVD	NVD	NVD	Elective caesarian
	Gestational age (weeks)	41	41	42	38
	Birth weight	3.74 kg (75th)	3.50 kg (50th)	3.73 kg (75th)	3.012 kg (50th)
	OFC at birth	–	34 cm (25th)	-	35 cm (91st)
	Neonatal respiratory difficulty?	Single apnoeic/hyponoeic episode as a neonate	No	Supplementary oxygen required at birth	Central apnoea and hypothermia from day 2 of life
	Neonatal hypotonia?	Yes	No	Yes	Yes (severe from day 2 of life)
	Neonatal feeding difficulties?	Yes (required NG tube feeding)	No	Yes (required NG tube feeding)	Yes (required NG tube feeding)
	Age at last assessment	4 years 7 months	14 years 3 months	12 years 10months	6 years 9 months
Developmental milestones	Sitting unsupported (age)	2 years 6 months	12 months	13 months	Not reached
	Walking independently (age)	Not reached	24 months	22 months	Not reached
	First words (age)	Not reached	2 years 6 months	3 years 6 months	11 months
Current developmental level	Gross motor	Unable to stand	Fully mobile, broad-based gait	Able to run; ataxic and broad-based gait	Non-ambulatory with central hypotonia, dystonia and dyskinesia
	Fine motor	Pincer grip	Manages buttons and can use fork/spoon	Manages buttons	Pincer grip
	Language	No expressive language, appears to understand some words	Sentences; limited vocabulary	Short phrases, repetitive, difficult to understand, limited comprehension	Essentially non-verbal
	Behaviour	Startles easily	Anxiety, no sense of danger	Relatively easy child, some obsessional and attention-seeking behaviours, poor awareness of others	Startles easily, anxiety and behaviour consistent with global delay
Growth parameters	Height	111 cm (91st)	145 cm (0.4th–2nd)	154 cm (50th)	123 cm (75%)
	Weight	19 kg (75th)	33 kg (<0.4th)	46.5 kg (50th)	25 kg (75%)
	Head circumference	50.6 cm (9th–25th)	51 cm (<0.4th)	56.0 cm (75th–91st)	53 cm (75th)
Facial features		Hypotonic facies, prominent forehead, epicanthic folds, mild telecanthus	Hypotonic facies, microcephaly, tall forehead, facial asymmetry (R<L), upslanting palpebral fissures, large central incisors	Slightly long face and full cheeks, high forehead, telecanthus	Hypotonic facies, mild frontal bossing, thin upper lip, some deciduous teeth possibly malformed
Neurological	Limb posture	Feet held in plantar flexion, restricted movement at ankles	Normal	Normal	Hypotonic/dystonic
	Tone	Hypotonic	Normal	Hypotonic	Hypotonic/dystonic
	Power	Mild generalised weakness	Normal	Mild generalised weakness	Mild generalised weakness
	DTRs/plantar responses	Diminished DTRs/normal plantar response	Normal	Difficult to elicit	Normal DTRs/normal plantar response
	Coordination	Poor	Poor	Poor	Poor
	Gait	Non-ambulatory	Broad-based	Broad-based	Non-ambulatory
	Cranial nerve anomalies?	Dysconjugate gaze	No	No	Dysconjugate gaze and intermittent ocular deviations
	Movement disorder?	No	No	No	Dystonic and choreoathetoid limb movements
	Seizures or 'seizure-like' episodes	'Seizure-like' episodes (at 7 months)	No	'Seizure-like' episodes (at 3years)	Epilepsy (onset at 14 months) with epileptic spasms
Other neurological features/findings?		Normal CSF neurotransmitters	Nil	Nil	Early cortical visual impairment Elevated CSF di-hydro

Continued

Table 1 Continued

PURA phenotype		Patient 1	Patient 2	Patient 3	Patient 4
					biopterin levels Sleep study—cyclical central apnoea
Investigation results	MRI brain	Delayed myelination (at 3 years 5 months)	Normal (at 7 years 8 months)	Normal (at 10 years 3 months)	Multiple abnormalities (see figure 3)
	EEG	Normal	Not performed	Abnormal	Abnormal
Endocrine		Gonadotropin-dependent precocious puberty (early thelarche), on treatment with IM decapeptyl	Nil	Nil	Elevated prolactin level soon after birth Blunted cortisol response to stress Chronically low vitamin D levels despite treatment

AD, autosomal dominant; CSF, cerebrospinal fluid; DTR, deep tendon reflex; ICSI, intracytoplasmic sperm injection; IM, intramuscular; NVD, normal vaginal delivery; OFC, occipital-frontal circumference; NG, nasogastric.

development.[31 32] Both knockout mice are reported to appear normal at birth and develop neurological features at approximately 2 weeks of age, which include continuous and increasingly severe tremor. Khalili *et al*[31] reported that their mice appeared to feed well but did not gain weight normally and died at 1 month of age. They also noted that their heterozygous mice were prone to seizures on routine handling. Hokkanen *et al*[32] reported that their null mutant mice lived up to 6 months. They reported that these animals did not gain weight normally after onset of tremor. They also observed an ataxic gait in these animals with an apparent hind limb weakness.

Both groups found that there was a marked reduction in the expression of the dentritic protein MAP2. This is interesting because Purα binds to mouse BC1 RNA in complex with other

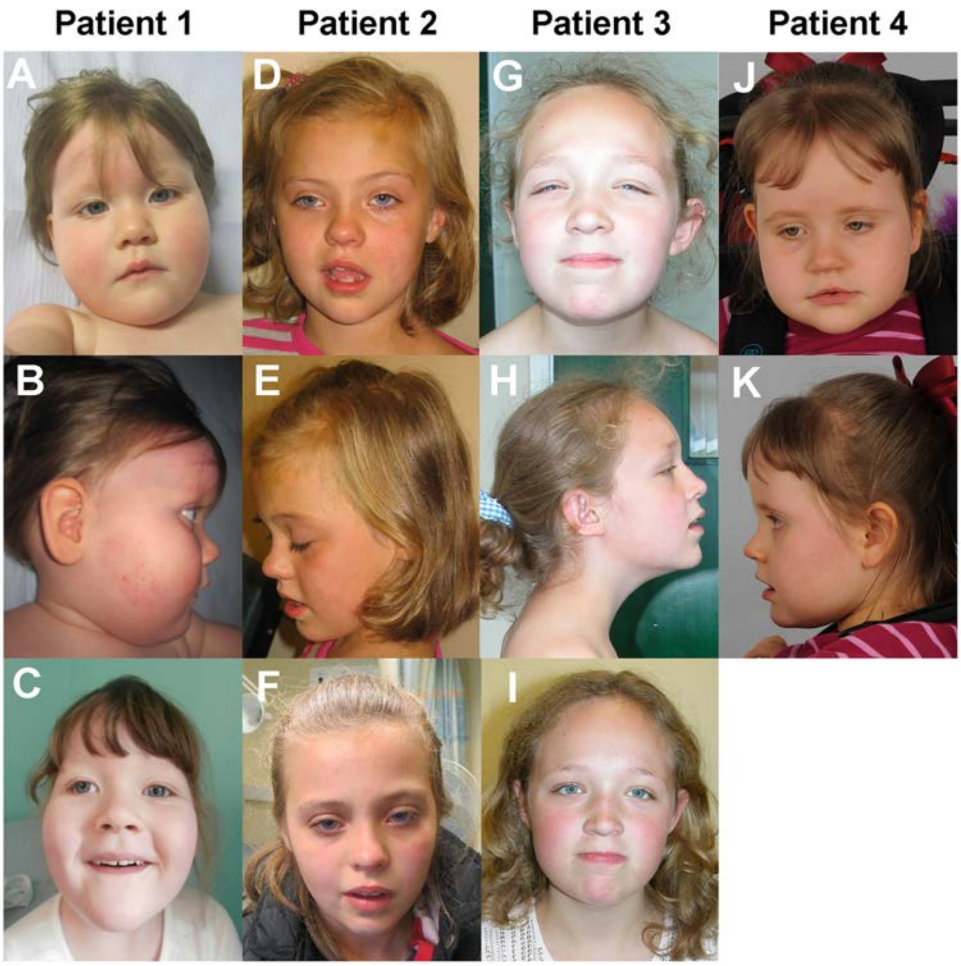


Figure 2 Clinical photographs. Patient 1 is shown at age 1 year 7 months (A and B) and 4 years 7 months (C). Patient 2 is shown at 7 years 3 months (D and E) and 14 years 3 months (F). Patient 3 is shown at 9 years 7 months (G and H) and 12 years 10 months (I). Patient 4 is shown at 6 years 9 months (J and K). While there is no obvious gestalt, all four patients were noted to have quite prominent foreheads with high anterior hairlines. Patients 1, 2 and 4 have mildly hypotonic facies.

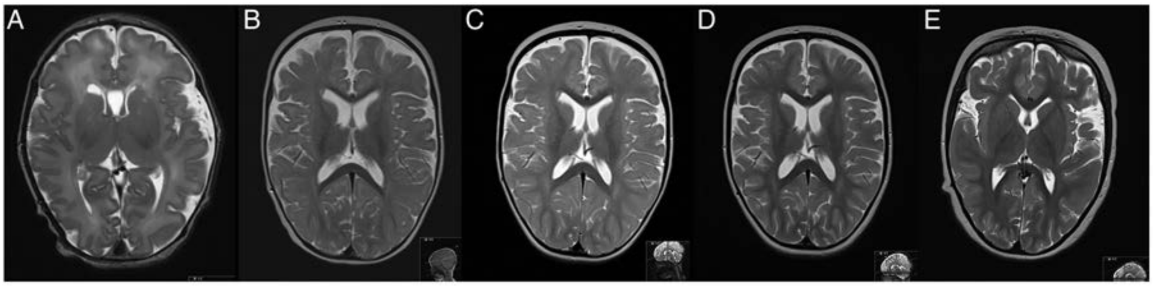


Figure 3 Serial MRI brain scans from Patient 4. (A) At 1 week, there is patchy high attenuation within white matter and a right frontal horn cyst, which is not evident on subsequent scans. (B) At 14 months, the white matter appears normal but thickening of tissue at the ependymal margin of the right frontal horn is apparent. (C) At 2 years 2 months, subtle hypomyelination is apparent in that there is poor definition of the grey-white matter boundary in the frontal lobes. (D) At 3 years 10 months, subtle hypomyelination persists. (E) At 5 years, myelination is complete. However, there are excessive extra-axial fluid spaces and there is possible cerebral atrophy.

proteins such as Fragile X Mental Retardation Protein and Staufen, as well as various mRNA species.[33] These form so-called messenger ribonucleoprotein granules, which are critical to normal dendritic function.

Purα binds to the $(CGG)_n$ sequence in *FMR1* that is pathologically expanded in Fragile X syndrome. Intriguingly, it has been suggested that Fragile X-associated tremor/ataxia syndrome (FXTAS), which may arise in premutation carriers, is due to the sequestration of Purα and other rCGG repeat binding proteins, thereby preventing them from fulfilling their normal cellular function.[34] The movement disorder observed in knockout mice might therefore be functionally related to FXTAS.

Until now, there have been no specific reports of mutations within *PURA* as a cause of human disease. It is, however, noteworthy that a 5q31.3 microdeletion phenotype has recently emerged[35] and *PURA*, which lies within the shared deletion interval of the seven patients described to date, and has been proposed as a candidate gene for the associated phenotype.[36 37] The shared phenotype of all patients reported thus far includes hypotonia, feeding difficulty and developmental delay. Additionally, respiratory problems, such as apnoea, and seizures or 'seizure-like' movements are reported in the majority of these patients.

While *NRG2*, a member of the neuregulin family, is highly likely to be contributory to the 5q31.3 microdeletion phenotype,[36] one of the two most recent patients to have been described in the literature with a similar phenotype has a microdeletion that has narrowed down the shortest region of overlap (SRO) to a 101 kb region encompassing only three genes: *PURA*, *C5orf53* and *C5orf32*. Given that the function of the latter two genes is yet to be characterised, Brown *et al*[37] have proposed that this 'lends further support for *PURA* as the likely primary candidate gene for the core neurodevelopmental features of this (5q31.3 microdeletion) syndrome'.

In this report, we provide the first evidence, that mutations limited to *PURA* are indeed sufficient to cause significant neurodevelopmental delay and learning disability in humans.

The four unrelated index patients have different *de novo* mutations in *PURA*. Patients 1 and 2 have frameshift mutations (p.Phe243Tyrfs*50 and p.Glu283Argfs*45, respectively). Given that *PURA* is a single exon gene, these altered gene products would not be subject to nonsense-mediated decay. As such, there is potential for these translated proteins to have dominant negative or gain-of-function effects or, alternatively, result in functional haploinsufficiency. Patient 3 has a missense mutation (p.Ile206Phe) and Patient 4 has an inframe deletion (p.Phe233del). Both of these mutations occur within highly conserved regions of sequence, giving rise to the Pur repeat II and Pur repeat III regions, respectively. These repeat regions are unique to Purine-rich element-binding proteins, and are of functional significance.

In all four affected individuals (figure 1), there was a shared core phenotype (table 1) of moderate to severe neurodevelopmental delay. Central hypotonia and early feeding difficulties were also common, as were respiratory difficulties ranging from distress at birth to single or recurrent central hyponoeic/apnoeic episodes in the newborn period. Three of our patients have a history of seizures or 'seizure-like' movements.

Additionally, some unusual features have been noted that may be part of the phenotypical spectrum for *PURA* mutations. In particular, there are some notable endocrine problems among these patients. Patient 1 has a history of gonadotropin-dependent precocious puberty with persistently elevated luteinizing hormone and follicle-stimulating hormone. She has early breast bud development and is currently on treatment with decapeptyl. None of the other patients are reported to have signs of early puberty. However, Patient 4 does have a history of other endocrine abnormalities including a persistently raised prolactin in the neonatal period and a blunted cortisol response to stress, despite normal baseline levels. She also has persistently low vitamin D levels despite treatment. This suggests that there may be a wider endocrine component, particularly with respect to anterior pituitary function.

Intriguingly, there is some evidence that Purα may be involved in the regulation of gondatropins. One study seeking to identify novel DNA-binding proteins for gonadotropin-releasing hormone 1 (*GNRH1*) promoter, identified both Purα and Purβ as potential regulators of *GNRH1* gene expression.[38] Subsequent *in vivo* studies have confirmed binding of both Purα and Purβ to the upstream region of the *GNRH1* gene. While overexpression of Purβ was shown to significantly downregulate *GNRH1* expression in transiently transfected mouse GT1-7 cells, this could not be demonstrated for Purα. However, it is worth noting that there is evidence that Purα is able to form a functional heterodimer with Purβ.[39]

Curiously, only Patient 1, the youngest, had a head circumference that appeared to be growing at the expected rate. None of the other 3 patients have maintained their projected rate of head growth from early occipital-frontal circumference measurements. This presumably reflects an inadequate growth in underlying brain volume, although no discrete brain structures were noted to be hypoplastic on MRI. This apparent inability to maintain growth velocity is consistent with the observations made in $Pura^{-/-}$ mice by Khalili *et al*.[31]

MRI brain scans were performed on all four patients at various ages. Patients 2 and 3 had normal MRI brain scans at ages 7 and 10 years, respectively. Historical scans were not available to check for early evidence of delayed myelination in these patients. In Patient 1, delayed myelination was detected at 3 years 5 months. In Patient 4, serial MRI brain scans were performed from birth showing a number of abnormalities including a transient right frontal horn cyst and patchy high attenuation of the white matter at birth and delayed myelination, with myelination complete by 5 years of age. There were, however, enlarged extra-axial fluid spaces by this time raising the possibility of mild parenchymal volume loss. As such, it is reasonable to say that evidence of delayed myelination was found only in those patients whose scans were performed early enough to detect it. On the whole, our patients' brain imaging is not entirely typical of the findings reported in the 5q31.3 phenotype, which includes frontotemporal volume loss, simplified frontal gyral pattern with shallow sulcation and delayed or incomplete myelination of the frontotemporal subcortical white matter tracts and anterior limbs of the internal capsules and cyst formation. However, Patient 4's brain imaging bears the greatest overlap with this phenotype.

We are confident that all four patients' phenotypes are secondary to their *de novo* mutations in *PURA*. WES has excluded other significant gene mutations. Furthermore, array CGH has excluded chromosomal microdeletions or duplications that may not necessarily have been detected by WES alone. Additionally, all patients have been thoroughly investigated by multiple physicians of various specialities *en route* to their definitive molecular genetic diagnosis. While there is a core phenotype, there is variability among our patients. Further cases will be useful to assess any distinctive genotype–phenotype correlations, or whether features, such as endocrine disturbance, metabolic abnormalities, epilepsy and a movement disorder might represent rare manifestations within a broad phenotypical spectrum. Regardless, our assumption has been that a functional haploinsufficiency has resulted from all four mutations. Functional studies may be necessary to confirm this hypothesis and exclude other possibilities, such as dominant negative or gain-of-function effects. However, the modular architecture of Purα and functional studies that have been completed to date tell us that truncating frameshift mutations similar to those found in Patient 1 will almost certainly have abolished the functional Pur repeat III sequence that is necessary for dimerisation and binding to linearised DNA. Patient 4's inframe deletion affects a very highly conserved residue within the same Pur repeat and would be expected to cause similar functional problems. Both these children are severely affected, being non-ambulatory and non-verbal. They are, however, the youngest two patients—but they have already exceeded the ages at which Patients 2 and 3 achieved independent ambulation (22–24 months).

Patient 2's frameshift mutation is downstream of the Pur repeats. The functional effect is not clear at a molecular level, but it seems to be associated with a less severe neurodevelopmental phenotype. Patient 3's missense mutation falls with Pur repeat II and affects a highly conserved residue. Again, the functional effect at a molecular level is not yet clear, but it presumably has potential to interfere with the formation of the ssDNA/ssRNA binding domain. Regardless, it too appears to be associated with a less severe neurodevelopmental phenotype.

We believe that our four patients help to resolve the 5q31.3 microdeletion phenotype. In the Brown *et al*[37] study, Patient 2, whose deletion significantly narrowed the SRO, was more mildly affected than the other six patients whose 5q31.3 microdeletions also included *NRG2* (MIM 603818). Additionally, this patient is reported as non-dysmorphic, whereas the other six patients have quite strikingly dysmorphic features. It has, therefore, been suggested that the combined deletion of *PURA* and *NRG2* (and/or other genes within the SRO for these six patients), may account for a more severe phenotype.[37] It has also been suggested that the more dysmorphic appearance of these patients is, in part, due to their more profound state of hypotonia. Our findings support the hypothesis that the deletion of *PURA* contributes to, but is not the sole cause of, the 5q31.3 microdeletion phenotype.

With the exception of mildly hypotonic facies, which are apparent in three of our patients, there are no obvious consistent dysmorphic facial features in this first cohort. We note, however, that all four patients have fairly prominent foreheads with relatively high anterior hairlines. As such, this is not a genetic syndrome that currently lends itself readily to clinical diagnosis based on history and clinical examination findings alone. However, if in time there should prove to be clear associations with discrete clinical problems, such as gondatropin-dependent precocious puberty or consistent brain imaging findings, it may be that the diagnosis can be strongly suspected on clinical grounds. However, based on the patients described herein, we suspect that this will ultimately prove to be a diagnosis that is usually made following investigation by WES or gene panel testing for neurodevelopmental delay. Indeed, as such technology becomes more readily accessible to clinicians, this diagnosis will undoubtedly become recognisable as a rare but important cause of sporadic neurodevelopmental delay.

Author affiliations
[1]Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton, UK
[2]The Royal Children's Hospital Department of Neurology, University of Melbourne Department of Paediatrics and the Murdoch Childrens Research Institute, Melbourne, Victoria, Australia
[3]Institute for Molecular Bioscience, The University of Queensland, St Lucia, Queensland, Australia
[4]Departments of Integrated Systems Biology and of Pediatrics, School of Medicine and Health Sciences, George Washington University, USA
[5]Illumina, Inc., San Diego, California, USA
[6]Pediatric Motor Disorders Research Program, Department of Neurology, University of Utah School of Medicine, Salt Lake City, Utah, USA
[7]Department of Radiology, Southampton General Hospital, Southampton, UK
[8]Wellcome Trust Sanger Institute, Cambridge, UK
[9]Genetic Medicine, Belfast City Hospital, Belfast, Northern Ireland
[10]Peninsula Clinical Genetics Service, Royal Devon and Exeter Hospital (Heavitree), Exeter, UK
[11]Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, Hampshire, UK

Acknowledgements The DDD study presents independent research commissioned by the Health Innovation Challenge Fund (grant number HICF-1009-003), a parallel funding partnership between the Wellcome Trust and the Department of Health, and the Wellcome Trust Sanger Institute (grant number WT098051). The views expressed in this publication are those of the author(s) and not necessarily those of the Wellcome Trust or the Department of Health. The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network. We would like to thank Associate Professor Avihu Boneh and Dr Diana Johnston who contributed clinical data for Patient 4. We thank the patients and their families for their participation. DB is a Hefce senior clinical fellow.

Contributors All the authors contributed significantly to this research and preparation of the manuscript. DH is the first author and along with DB (corresponding and senior author) coordinates the group. DB, ACM, PDT and DH phenotyped the UK clinical cases, put forward the sample trios for exome through the DDD study and interpreted the results obtained. RJL and KS phenotyped the Australian case, with CS and RT undertaking the exoming and interpretation of results. MG-C assessed the clinical findings of all the cases radiologically. All authors have been involved in the drafting, critical revision and final approval of the manuscript for publication. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Competing interests None.

Patient consent Obtained.

Ethics approval The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and GEN/284/12 granted by the Republic of Ireland REC), Royal Children's Hospital Melbourne Human Research Ethics Committee approval (REF: 28097 G) and The University of Queensland Human Research Ethics Committee approval (2013001536).

Provenance and peer review Not commissioned; externally peer reviewed.

Web Resources Online Mendelian Inheritance in Man, http://www.omim.org/; Simple Modular Architecture Research Tool (SMART), http://smart.embl-heidelberg.de; UCSC Genome Browser, http://genome.ucsc.edu

Sequence Data Coding sequence mutations in PURA are referenced against NM_005859.4.

REFERENCES

1 Van Bokhoven H. Genetic and epigenetic networks in intellectual disabilities. *Annu Rev Genet* 2011;45:81–104.

2 Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K-I, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet* 2010;42:790–3.

3 Hoischen A, van Bon BWM, Gilissen C, Arts P, van Lier B, Steehouwer M, de Vries P, de Reuver R, Wieskamp N, Mortier G, Devriendt K, Amorim MZ, Revencu N, Kidd A, Barbosa M, Turner A, Smith J, Oley C, Henderson A, Hayes IM, Thompson EM, Brunner HG, de Vries BBA, Veltman JA. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet* 2010;42:483–5.

4 Vissers LELM, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wieskamp N, del Rosario M, van Bon BWM, Hoischen A, de Vries BBA, Brunner HG, Veltman JA. A de novo paradigm for mental retardation. *Nat Genet* 2010;42:1109–12.

5 Ku CS, Polychronakos C, Tan EK, Naidoo N, Pawitan Y, Roukos DH, Mort M, Cooper DN. A new paradigm emerges from the study of de novo mutations in the context of neurodevelopmental disease. *Mol Psychiatry* 2013;18:141–53.

6 Rivière J-B, van Bon BWM, Hoischen A, Kholmanskikh SS, O'Roak BJ, Gilissen C, Gijsen S, Sullivan CT, Christian SL, Abdul-Rahman OA, Atkin JF, Chassaing N, Drouin-Garraud V, Fry AE, Fryns J-P, Gripp KW, Kempers M, Kleefstra T, Mancini GMS, Nowaczyk MJM, van Ravenswaaij-Arts CMA, Roscioli T, Marble M, Rosenfeld JA, Siu VM, de Vries BBA, Shendure J, Verloes A, Veltman JA, Brunner HG, Ross ME, Pilz DT, Dobyns WB. De novo mutations in the actin genes ACTB and ACTG1 cause Baraitser-Winter syndrome. *Nat Genet* 2012;44:440–4, S1–2.

7 Schuurs-Hoeijmakers JHM, Oh EC, Vissers LELM, Swinkels MEM, Gilissen C, Willemsen MA, Holvoet M, Steehouwer M, Veltman JA, de Vries BBA, van Bokhoven H, de Brouwer APM, Katsanis N, Devriendt K, Brunner HG. Recurrent de novo mutations in PACS1 cause defective cranial-neural-crest migration and define a recognizable intellectual-disability syndrome. *Am J Hum Genet* 2012;91:1122–7.

8 Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA; 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* 2010;467:1061–73.

9 Robinson PN, Mundlos S. The Human Phenotype Ontology. *Clinical Genetics* 2010;77:525–34.

10 Bragin E, Chatzimichali EA, Wright CF, Hurles ME, Firth HV, Bevan AP, *et al.* DECIPHER: database for the interpretation of phenotype-linked plausibly pathogenic sequence and copy-number variation. *Nucleic Acids Research*. 2014.

11 McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* 2010;26:2069–70.

12 Cleary JG, Braithwaite R, Gaastra K, Hilbush BS, Inglis S, Irvine SA, Jackson A, Littin R, Nohzadeh-Malakshah S, Rathod M, Ware D, Trigg L, De La Vega FM. Joint variant and de novo mutation identification on pedigrees from high-throughput sequencing data. *J Comput Biol J Comput Mol Cell Biol* 2014;21:405–19.

13 Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* 2012;6:80–92.

14 Liu X, Jian X, Boerwinkle E. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum Mutat* 2011;32:894–9.

15 Sherry ST, Ward M, Sirotkin K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res* 1999;9:677–9.

16 Graebsch A, Roche S, Niessing D. X-ray structure of Pur-alpha reveals a Whirly-like fold and an unusual nucleic-acid binding surface. *Proc Natl Acad Sci U S A* 2009;106:18521–6.

17 Bergemann AD, Johnson EM. The HeLa Pur factor binds single-stranded DNA at a specific element conserved in gene flanking regions and origins of DNA replication. *Mol Cell Biol* 1992;12:1257–65.

18 Bergemann AD, Ma ZW, Johnson EM. Sequence of cDNA comprising the human pur gene and sequence-specific single-stranded-DNA-binding properties of the encoded protein. *Mol Cell Biol* 1992;12:5673–82.

19 Haas S, Gordon J, Khalili K. A developmentally regulated DNA-binding protein from mouse brain stimulates myelin basic protein gene expression. *Mol Cell Biol* 1993;13:3103–12.

20 Haas S, Thatikunta P, Steplewski A, Johnson EM, Khalili K, Amini S. A 39-kD DNA-binding protein from mouse brain stimulates transcription of myelin basic protein gene in oligodendrocytic cells. *J Cell Biol* 1995;130:1171–9.

21 Darbinian N, Gallia GL, Khalili K. Helix-destabilizing properties of the human single-stranded DNA- and RNA-binding protein Puralpha. *J Cell Biochem* 2001;80:589–95.

22 Wortman MJ, Johnson EM, Bergemann AD. Mechanism of DNA binding and localized strand separation by Pur alpha and comparison with Pur family member, Pur beta. *Biochim Biophys Acta* 2005;1743:64–78.

23 Darbinian N, Sawaya BE, Khalili K, Jaffe N, Wortman B, Giordano A, Amini S. Functional interaction between cyclin T1/cdk9 and Puralpha determines the level of TNFalpha promoter activation by Tat in glial cells. *J Neuroimmunol* 2001;121:3–11.

24 Kobayashi S, Agui K, Kamo S, Li Y, Anzai K. Neural BC1 RNA associates with pur alpha, a single-stranded DNA and RNA binding protein, which is involved in the transcription of the BC1 RNA gene. *Biochem Biophys Res Commun* 2000;277:341–7.

25 Zambrano N, De Renzis S, Minopoli G, Faraonio R, Donini V, Scaloni A, Cimino F, Russo T. DNA-binding protein Pur alpha and transcription factor YY1 function as transcription activators of the neuron-specific FE65 gene promoter. *Biochem J* 1997;328(Pt 1):293–300.

26 Dobretsova A, Johnson JW, Jones RC, Edmondson RD, Wight PA. Proteomic analysis of nuclear factors binding to an intronic enhancer in the myelin proteolipid protein gene. *J Neurochem* 2008;105:1979–95.

27 Darbinian N, Cui J, Basile A, Del Valle L, Otte J, Miklossy J, Sawaya BE, Amini S, Khalili K, Gordon J. Negative regulation of AbetaPP gene expression by pur-alpha. *J Alzheimers Dis JAD* 2008;15:71–82.

28 Knapp AM, Ramsey JE, Wang S-X, Godburn KE, Strauch AR, Kelm RJ. Nucleoprotein interactions governing cell type-dependent repression of the mouse smooth muscle alpha-actin promoter by single-stranded DNA-binding proteins Pur alpha and Pur beta. *J Biol Chem* 2006;281:7907–18.

29 Penberthy WT, Zhao C, Zhang Y, Jessen JR, Yang Z, Bricaud O, Collazo A, Meng A, Lin S. Pur alpha and Sp8 as opposing regulators of neural gata2 expression. *Dev Biol* 2004;275:225–34.

30 Muralidharan V, Sweet T, Nadraga Y, Amini S, Khalili K. Regulation of Puralpha gene transcription: evidence for autoregulation of Puralpha promoter. *J Cell Physiol* 2001;186:406–13.

31 Khalili K, Del Valle L, Muralidharan V, Gault WJ, Darbinian N, Otte J, Meier E, Johnson EM, Daniel DC, Kinoshita Y, Amini S, Gordon J. Puralpha is essential for postnatal brain development and developmentally coupled cellular proliferation as revealed by genetic inactivation in the mouse. *Mol Cell Biol* 2003;23:6857–75.

32 Hokkanen S, Feldmann HM, Ding H, Jung CKE, Bojarski L, Renner-Müller I, Schüller U, Kretzschmar H, Wolf E, Herms J. Lack of Pur-alpha alters postnatal brain development and causes megalencephaly. *Hum Mol Genet* 2012;21:473–84.

33 Johnson EM, Kinoshita Y, Weinreb DB, Wortman MJ, Simon R, Khalili K, Winckler B, Gordon J. Role of Pur alpha in targeting mRNA to sites of translation in hippocampal neuronal dendrites. *J Neurosci Res* 2006;83:929–43.

34 Jin P, Duan R, Qurashi A, Qin Y, Tian D, Rosser TC, Liu H, Feng Y, Warren ST. Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome. *Neuron* 2007;55:556–64.

35 Shimojima K, Isidor B, Le Caignec C, Kondo A, Sakata S, Ohno K, Yamamoto T. A new microdeletion syndrome of 5q31.3 characterized by severe developmental delays, distinctive facial features, and delayed myelination. *Am J Med Genet A* 2011;155A:732–6.

36 Hosoki K, Ohta T, Natsume J, Imai S, Okumura A, Matsui T, Harada N, Bacino CA, Scaglia F, Jones JY, Niikawa N, Saitoh S. Clinical phenotype and candidate genes for the 5q31.3 microdeletion syndrome. *Am J Med Genet A* 2012;158A:1891–6.

37 Brown N, Burgess T, Forbes R, McGillivray G, Kornberg A, Mandelstam S, Stark Z. 5q31.3 Microdeletion syndrome: clinical and molecular characterization of two further cases. *Am J Med Genet A* 2013;161:2604–8.

38 Zhao S, Kelm RJ, Fernald RD. Regulation of gonadotropin-releasing hormone-1 gene transcription by members of the purine-rich element-binding protein family. *Am J Physiol Endocrinol Metab* 2010;298:E524–33.

39 Kelm RJ, Cogan JG, Elder PK, Strauch AR, Getz MJ. Molecular interactions between single-stranded DNA-binding proteins associated with an essential MCAT element in the mouse smooth muscle alpha-actin promoter. *J Biol Chem* 1999;274:14238–45.