Microdeletions of *ELP4* are associated with language impairment, autism spectrum disorder and epilepsy.

| | of Cytogenetics
Marshall, Christian; The Hospital for Sick Children, Program in Genetics and Genome Biology, The Centre for Applied Genomics; The Hospital for Sick Children, Department of Molecular Genetics, Paediatric Laboratory Medicine
Scherer, Stephen; The Hospital for Sick Children, Program in Genetics and Genome Biology, The Centre for Applied Genomics; University of Toronto, McLaughlin Centre and Department of Molecular Genetics
Strug, Lisa; University of Toronto, Program in Genetics and Genome Biology, The Hospital for Sick Children; Division of Biostatistics, Dalla Lana School of Public Health
Collier, David; Eli Lilly and Company, Neuroscience Discovery Research; Institute of Psychiatry, Psychology and Neuroscience, King's College London, Social, Genetic and Developmental Psychiatry Centre
Pal, Deb; Institute of Psychiatry, Psychology and Neuroscience, King's College London, Department of Basic and Clinical Neuroscience |

Microdeletions of *ELP4* are associated with language impairment, autism

spectrum disorder and epilepsy

Laura Addis[1,2], Joo Wook Ahn[3], Richard Dobson[4], Abhishek Dixit[4], Caroline M Ogilvie[3],

Dalila Pinto[5], Andrea K Vaags[6], Hilary Coon[7], Pauline Chaste[8], Scott Wilson[9,10,11], Jeremy R

Parr[12], Joris Andrieux[13], Bruno Lenne[14], Zeynep Tumer[15], Vincenzo Leuzzi[16], Kristina

Aubell[17], Hannele Koillinen[18], Sarah Curran[3], Christian R Marshall[6,19], Stephen W Scherer[6,20],

Lisa J Strug[21], David A Collier[2,9], Deb K Pal[1]*

1: Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
2: Neuroscience Discovery Research, Eli Lilly and Company, Erl Wood, Surrey, UK
3: Department of Cytogenetics, Guy's and St Thomas' NHS Foundation Trust, London, UK
4: Department of Biostatistics and NIHR BRC for Mental Health, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
5: Departments of Psychiatry, and Genetics and Genomic Sciences, Seaver Autism Center, The Mindich Child Health & Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA
6: Program in Genetics and Genome Biology, The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario, Canada.
7: Utah Autism Research Program, Department of Psychiatry, University of Utah School of Medicine, Salt Lake City, UT, USA
8: Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA
9: Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
10: Department of Endocrinology and Diabetes, Sir Charles Gairdner Hospital, Perth, WA, Australia.
11: School of Medicine and Pharmacology, University of Western Australia, Nedlands, WA, Australia.
12: Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, UK
13: Institut de Génétique Médicale, Hopital Jeanne de Flandre, CHRU de Lille, France
14: Centre de Génétique Chromosomique, GHICL, Hôpital Saint Vincent de Paul, Lille, France
15: Applied Human Molecular Genetics, Kennedy Center, Department of Clinical Genetics, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark
16: Dept of Pediatrics, Child Neurology and Psychiatry, Sapienza Università di Roma, Rome, Italy
17: Institute of Human Genetics, Medical University of Graz, Graz, Austria
18: Department of Clinical Genetics, Helsinki University Hospital, Helsinki, Finland

Laura Addis

19: Department of Molecular Genetics, Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada
20: McLaughlin Centre and Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada
21: Program in Genetics and Genome Biology, The Hospital for Sick Children; Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Correspondence: Prof. Deb K Pal, Department of Basic and Clinical Neuroscience, Institute

of Psychiatry, Psychology and Neuroscience, King's College London, 1-4 Windsor Walk,

London SE5 8AF, London, UK. Tel: +44 20 7848 0636; Fax: +44 20 7848 0988; E-mail:

deb.pal@kcl.ac.uk

Laura Addis

ABSTRACT

Copy number variations (CNV) are important in the aetiology of neurodevelopmental disorders and show broad phenotypic manifestations. We compared the presence of small CNVs disrupting the *ELP4-PAX6* locus in 4,092 U.K. individuals with a range of neurodevelopmental conditions, clinically referred for array comparative genomic hybridisation (aCGH), with WTCCC controls (n=4,783). The phenotypic analysis was then extended using the DECIPHER database. We followed up association using an autism patient cohort (n=3,143) compared with six additional control groups (n=6,469). In the clinical discovery series we identified eight cases with *ELP4* deletions, and one with a partial duplication of *ELP4* and *PAX6*. These cases were referred for neurological phenotypes including language impairment, developmental delay, autism and epilepsy. Six further cases with a primary diagnosis of ASD and similar secondary phenotypes were identified with *ELP4* deletions, as well as another six (out of 9) with neurodevelopmental phenotypes from DECIPHER. CNVs at *ELP4* were only present in 1/11,252 controls. We found a significant excess of CNVs in discovery cases compared with controls, p=7.5x10[-3]; as well as for autism, p=2.7x10[-3]. Our results suggest *ELP4* deletions are highly likely to be pathogenic, predisposing to a range of neurodevelopmental phenotypes from ASD to language impairment and epilepsy.

Key Words: Copy Number Variation (CNV), Epilepsy and seizures, Developmental, Neurology

Laura Addis

INTRODUCTION

Copy number variation (CNV) plays an important role in the aetiology of neurodevelopmental and psychiatric disorders. Both recurrent *de novo* and rare segregating CNVs have begun to explain the overlap of diverse phenotypes in individual cases and families (Cooper, et al., 2011; Malhotra and Sebat, 2012). Copy number variation is a strong risk factor in both focal and generalized epilepsies, and they are also found in 8% of patients with epileptic encephalopathies (Mefford, et al., 2010; Mefford, et al., 2011). Recent findings in the rare epileptic encephalopathies illustrate the connection between epilepsy, language impairment and autism spectrum disorder (ASD) through overrepresentation of novel CNVs containing cell adhesion genes, (e.g. cadherins and contactins) (Lesca, et al., 2012). However, there are also differences between disorders: for example, specific language impairment cases, whilst having an increased burden of CNVs, do not in general show enrichment for novel or *de novo* events (Simpson, et al., 2015), whereas rare copy number variation is an important source of risk in ASD (Pinto, et al., 2014).

The examples above indicate that a given genomic alteration can sustain broad susceptibility to several phenotypes depending on the genetic background of the subject. So called 'hotspot' CNVs also manifest this phenotypic variability. The recurrent 15q13.3 microduplication increases the risk for intellectual disability, idiopathic generalised epilepsy, ASD and schizophrenia and (Helbig, et al., 2009; Poot, et al., 2011), and deletions at 16p13.11 contribute to a diverse spectrum of epilepsy disorders (Heinzen, et al., 2010). The 16p11.2 hotspot is also pleiotropic; deletions are common in ASD and developmental delay, (Marshall, et al., 2008) and duplications have been associated with seizures and speech delay (Shinawi, et al., 2010). Other notable examples of pleiotropy are CNVs of the *CNTNAP2*

Laura Addis

gene, which are implicated in ASD, Gilles de la Tourette syndrome, schizophrenia and

epilepsy, and *AUTS2* with ASD and mental retardation. Interestingly *AUTS2* and *CNTNAP2*

may interact with each other on a molecular level, (Poot, et al., 2011), indicating emerging

convergent pathways for neurodevelopment. A recent report of a deletion of the *ELP4* gene at

11p13, and adjacent 3' *PAX6* enhancer elements has been described in a case with aniridia,

autism and mental retardation. This case differs from the 'classical' *PAX6* gene deletions

causing aniridia alone, as only the 3' enhancer elements are deleted in this case and *ELP4* is

included (Davis, et al., 2008). *ELP4* has previously been associated with the

electroencephalographic (EEG) signature of the common childhood epilepsy Rolandic

epilepsy (RE) (Strug, et al., 2009), and such EEG abnormalities as well as epilepsy are well

established in autism spectrum disorder (ASD) and language impairments, (Nasr, et al., 2001;

Parmeggiani, et al., 2010). These examples illustrate again that genomic alterations can show

broad phenotypic manifestations during neurodevelopment, as well as incomplete penetrance.

In the present study we report the presence of a number of deletions of *ELP4* and the

regulatory elements of *PAX6* in the U.K. database of individuals with a childhood onset

developmental condition referred for clinical genetic testing (BB-GRE). We test the

hypothesis that the burden of *ELP4* CNVs is increased in those with neurodevelopmental

conditions compared to controls. This phenotypic analysis is then extended using the

DECIPHER database of chromosomal imbalances in over 10,000 cases of developmental

disorders. Using a CNV-led approach we then further expand the phenotype associated with

ELP4 microdeletions to cases with ASD and varying comorbidities, and carry out a second

case-control analysis of frequency. This data supports our hypothesis that disruption of *ELP4*

and the regulatory regions of *PAX6* contained within its introns, lead to a range of

neurodevelopmental conditions.

Laura Addis

METHODS

Study Design

We used a three-stage design; first testing the hypothesis of CNV enrichment at *ELP4* in a

clinical discovery sample of developmental disorders (Brain and Body Genetics Research

Exchange (BBGRE, https://bbgre.iop.kcl.ac.uk) and control dataset (WTCCC), both from the

U.K.; second, extending the phenotypic analysis to a larger dataset of developmental

disorders (DECIPHER) and finally, replicating the association to neurodevelopmental

disorders in ASD cases (Autism Genome Project and two Canadian ASD cohorts) compared

to a large multi-centre control sample set.

Samples

U.K. Clinical Dataset – BB-GRE

4092 children referred to Guy's and St Thomas NHS Foundation Trust, southeastern UK from

paediatricians and regional hospitals, https://bbgre.iop.kcl.ac.uk. Individuals referred for

array-CGH testing for a range of developmental problems including developmental delay

(DD), ASD, speech or language delay or congenital defects. Individuals had clinical

diagnoses made prior to genetic testing, which was part of standard clinical care. Genomic

data and referral phenotype information were anonymised and recorded in a clinical database,

(63% males; August 2014).

Global Clinical Dataset - DECIPHER

We performed a search in the DECIPHER database (Firth, et al., 2009) in order to identify

additional cases with small CNVs which included and/or disrupted *ELP4*. DECIPHER

(Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources,

http://decipher.sanger.ac.uk) is an interactive web-based database of over 10,000 cases which

enables clinical scientists to maintain records of phenotype and chromosome rearrangement,

to aid patient diagnosis by linking to other bioinformatics resources and interactive tools, and

to share this information with the clinical research community.

Canadian ASD Samples

The cohort contained 349 probands previously published (Lionel, et al., 2011) and 350

additional patients diagnosed with ASD from Canada described below, totaling 699.

Individuals were recruited from four different Canadian sites: The Hospital for Sick Children,

Toronto; McMaster University, Hamilton; Memorial University, St. John's, Newfoundland

and University of Alberta, Edmonton. All had a clinical diagnosis of ASD, using the Autism

Diagnostic Interview-Revised (ADI-R) and/or Autism Diagnostic Observation Schedule

(ADOS).

Autism Genome Project (AGP) Samples

2,147 European ASD cases were genotyped as part of a study by the AGP Consortium for

rare CNVs affecting autism and are formally described in the following reference, (Pinto, et

al., 2014). All cases had a clinical diagnosis of autism rated using the ADI-R and/or the

ADOS.

Control Populations

A total sample of 11,252 controls from six different datasets were included this study. Group

1 was compared with the BBGRE cases, and groups 2-6 with the AGP cases: (1) WTCCC,

Wellcome Trust Case Control Consortium controls - 4,783 population controls from the UK

(Consortium, et al., 2010) (2) Ottawa Heart Institute (OHI) controls - A cohort of 1,234

control individuals collected as part of a large case control GWA study (Stewart, et al., 2009);

(3) German POPGEN controls – a sample of 1,123 individuals of northern German origin

(Schleswig-Holstein) (Krawczak, et al., 2006); (4) Ontario Population Genomics Platform

(OPGP) controls – a Canadian sample of 416 control individuals of European ancestry

(http://www.tcag.ca/facilities/cyto_population_control_DNA.html); (5) HapMap3 controls - a

sample of 1,056 individuals from populations from around the world from the International

HapMap Project (http://www.hapmap.org/); (6) Controls from the AGP project –consisting of

2,640 of European ancestry assembled from three studies in which subjects had no obvious

psychiatric history: 'Study of Addiction Genetics and Environment (SAGE)', 'Ontario

Colorectal Cancer case-control study (OC)', and the 'Health, Aging, and Body Composition

(HABC)'.

Genotyping and CNV Analysis

Array CGH analysis of BB-GRE Samples

Array CGH testing was carried out at the Guys and St Thomas' Services cytogenetics CPA

accredited laboratory. We have previously described the protocols, analysis and interpretation

using an Agilent oligonucleotide array 60K platform (AMAID 028469 and 017457) and a

patient vs. patient hybridization strategy and 3-probe minimum aberration call in (Ahn, et al.,

2013; Ahn, et al., 2010). The average probe density over *ELP4* is 8.5Kb, giving a limit of

around 25Kb for detection. CNVs in this population are available by application to BB-GRE;

https://bbgre.iop.kcl.ac.uk/.

Canadian ASD and Control Groups 1-5

Canadian ASD cases, and control populations 1-5, were genotyped using the Affymetrix

Genome-Wide Human SNP Array 6.0 with standard protocols. Arrays meeting Affymetrix

quality control guidelines of Contrast QC > 0.4 were further analysed. Raw data analysis was carried out using a multiple-algorithm approach to maximize sensitivity and specificity of CNV calling, as described previously (Lionel, et al., 2011; Silversides, et al., 2012). Briefly, arrays were analyzed for CNVs with Birdsuite (Korn, et al., 2008) iPattern (Pinto, et al., 2011) and Affymetrix Genotyping Console and merged into a single dataset. A CNV call was considered high confidence if it was detected by at least two of the calling algorithms and spanned at least 10kb and >5 consecutive array probes. Average probe density over *ELP4* was 2.0Kb, giving a limit of around 10Kb for detection.

Autism Genome Project Samples and Control Group 6

2,147 ASD cases and 2,640 controls were genotyped with the Illumina Infinium 1M SNP microarray. CNV calling was performed using a multi-algorithm approach incorporating PennCNV, iPattern and QuantiSNP (Pinto, et al., 2010). Subsequent analyses focused on those CNVs spanning five or more array probes and detected by at least two algorithms. The analysis is formally described in (Pinto, et al., 2014). The average probe density over *ELP4* was 2.1Kb, giving a limit of around 10Kb for detection.

Association analysis

A two-tailed Fishers exact test was used to compare frequencies of *ELP4* CNVs in the 4,092 cases in BB-GRE with the 4,783 controls from the WTCCC. Subsequently another two-tailed Fishers exact test was used to compare the frequency in 2,845 unrelated ASD cases compared to 6,469 controls from control sets 2-6 combined.

<u>Limitations</u>

Laura Addis

A limitation of this study is that the CNVs were not identified on the same platform or by the

same analysis method between the sample sets. Therefore there is a chance of false CNV

enrichment related to probe density, data quality and analysis methods. However, all

platforms are high-density, with probe coverage shown in Figure 1. This ensures *ELP4* and

the surrounding region is well covered, and indeed the control data is generated on higher

density platforms than the BBGRE cases, resulting in a higher CNV detection power for

controls. We have also ensured that all reported CNVs can be called using all three

methodologies. The Canadian ASD cases and control groups 1-5, also use the same platform

and analysis methods as each other. Cases and controls from the AGP study also used the

same array and analysis methods as each other. To reduce the chance of error, all of the CNV

calling methods from each data set employ published, rigorous quality control measures as

detailed above, ensuring that CNVs called are highly unlikely to be false positives. All of the

ASD CNVs have also been validated by orthogonal methods such as qPCR. By also using

data sets with different platforms, we have shown that our results are consistent even between

the different methods used.

Laura Addis

RESULTS

Copy Number Variation of *ELP4* in the BB-GRE Database

Out of 4092 individuals referred for neurodevelopmental disorders, we identified nine patients

with small (<1Mb) CNVs disrupting *ELP4* that could also have be detected by the other array

methods. Eight CNVs were deletions, Figure 1 and Table 1, varying in size from 26Kb to

101Kb. The ninth CNV was a 232Kb duplication of the first 7 exons of *ELP4* and the *PAX6*

gene. This patient also carried a 'hotspot' deletion of 1.2Mb at 16p13.11, which is also

implicated in several neuropsychiatric disorders (Heinzen, et al., 2010). One deletion

(117374) is intronic, but does however disrupt regulatory enhancers of *PAX6* and so is

included in our analysis. Four of the deletions were maternally inherited and two were

paternally inherited, one arose *de-novo* and one had unknown inheritance. The inheritance

pattern of the duplication was also unknown. Clinical information was not available for the

parents as it is not collected for BBGRE and referring clinicians cannot be contacted. Three of

the deletion patients carried a second CNV, Table 1, none of which are predicted to affect the

phenotype; 119460 had a deletion of unknown inheritance of 77.6Kb at 5q21 with no genes

present in the region, and 112601 had a maternally inherited 226Kb duplication at 5q15

disrupting *FAM172A*, a potential tumor suppressor. Patient 130693 carried a maternally

inherited duplication of 23Kb at 6p22.2, disrupting the MHC-associated genes *BTN3A3* and

BTN2A1.

All cases were diagnosed with a neurodevelopmental phenotype; five had speech and

language delay or disorder, with one also diagnosed with epilepsy, two had social

communication difficulties and two had a diagnosis of autism, with one further case showing

emerging autistic traits. Six of the patients also had a range of cognitive delays, Table 1.

Laura Addis

Unfortunately we do not know the age at last neurological assessment for BBGRE cases; only

age at aCGH testing is recorded. Therefore some cases may be too young for some

phenotypes to manifest and be reported, e.g. 117003.

Only one CNV involving *ELP4* was found in the WTCCC control set; a 221Kb

microdeletion, Supplementary Information Table 1. On comparison of the U.K. BB-GRE

samples with the WTCCC controls, the difference in CNV frequency disrupting *ELP4* was

significant; p-value=$7.5x10^{-3}$.

Microdeletions of *ELP4* in the DECIPHER Database

We identified nine individuals with a small (< 1Mb) CNV encompassing *ELP4* in the

DECIPHER database (https://decipher.sanger.ac.uk) (Firth, et al., 2009). All were deletions,

Table 2, with at least one breakpoint within the gene. Detailed phenotypic information was

available for eight of the nine patients; six individuals were diagnosed with developmental

delay or intellectual disability. Several cases had speech delay; two had behavioral disorders,

one was diagnosed with a pervasive developmental disorder (PDD), most likely ASD, and

one further case had ASD. Another case was also diagnosed with ADHD and epilepsy

(257614). Three cases were too young at the age of last clinical visit (263619, 265704 and

287341) for a full assessment of neurodevelopmental phenotypes such as ASD.

Two cases, 289275 and 270752, had deletions that disrupted *PAX6* exons. Most likely due to

PAX6 enhancer or exon disruption, these two cases, as well as 265704 and 263741, also have

aniridia, an abnormality of the iris. Two others cases, 257614 and 249728, have congenital

eye malformations but deletion breakpoints much further from *PAX6*. The remaining three

Laura Addis

cases, whilst having breakpoints very similar to those with aniridia, do not share that

phenotype, indicating a complex genotype-phenotype relationship.

Two cases carried a second CNV: 289275 had an intronic duplication of *ZNF674* at Xp11.3,

and 261471 a 245Kb deletion at 10p11.21, disrupting *CUL2*, *CREM* and *CCNY,* genes not

involved in neuronal development.

Microdeletions of *ELP4* in Autism Cases

Given that several individuals from BB-GRE and DECIPHER have ASD, PDD or social

communication difficulties, we decided to investigate the prevalence of *ELP4* CNVs in two

autism cohorts. Out of 2,446 cases from the Autism Genome Project, AGP, (Pinto, et al.,

2014) three had microdeletions of *ELP4*, Table 2. All three fulfilled the criteria for a strict

definition of autism, and were verbal (verbal IQ >70), but experienced language delay of first

words and phrases. None of the cases had a history of seizures or epilepsy. Case 8596_201

also carried a 500Kb maternally inherited duplication disrupting the collagen gene *COL27A1*

that is highly unlikely to contribute to the neurological phenotype.

Three out of 699 individuals from the Canadian autism study also carried *ELP4* deletions,

Table 3. An affected sister pair both had a 112Kb deletion of half of the gene, and a male

case carried a 130Kb deletion of the 3' (but proximal due to reverse gene orientation) part of

ELP4 and neighboring *IMMP1L*. Again, all three had speech and language delay and the

sister-pair had mild developmental delay. Interestingly, the sisters also both carried a deletion

of one copy of exon 2 of *TMLHE*, an enzyme involved in carnitine biosynthesis, on Xq28.

Laura Addis

A case-control analysis of the frequency of *ELP4* CNVs from unrelated individuals in these

2,845 ASD cases compared with 6,469 control individuals from groups 2-6, where no *ELP4*

CNVs were found, yielded a highly significant p value of 2.7×10^{-3}.

DISCUSSION

In this study we have found a strong and consistent pleiotropic association between CNVs disrupting *ELP4* and neurodevelopmental conditions over several experimental platforms. We have described CNVs that can be captured and called from all three high density platforms/methods. We have also addressed the potential problem of enrichment bias of CNVs in cases, as all control data is generated on higher density platforms than the BBGRE cases, or the same as ASD/DECIPHER cases, resulting in a higher CNV detection power for controls. CNVs disrupting *ELP4* appear to be rare in the general population given that we found only one CNV in the six control groups studied (total n = 11,252), and that there are no regions of segmental duplication around the gene (UCSC Segmental Duplication track, (Bailey, et al., 2002)). *ELP4* now joins the growing list of genes such as *CNTNAP2*, *SHANK3* and *NRXN1*, where heterozygous copy number are repeatedly associated with a wide range of neuropsychiatric disorders (Gregor, et al., 2011; Lesca, et al., 2012; Poot, et al., 2011).

We have extended the phenotype associated with disruptions of *ELP4* from the EEG signature of Rolandic epilepsy and speech sound disorder (Pal, et al., 2010; Strug, et al., 2009) to ASD, social communication difficulties, developmental delay, and epilepsy. This corroborates the findings of Davis et al, who found a deletion of *ELP4* and *PAX6* enhancer elements in a patient with autism, aniridia and mental retardation that was inherited from an affected mother (Davis, et al., 2008). The *ELP4* locus may influence the development of language function, as a frequent trait across almost half of the 24 patients described here are speech and language difficulties. There appears to be a genetic crossroads between childhood epilepsy, autism, and speech and language disorders. Several genes and pathways provide a common link such as the cell adhesion genes cadherins and catenins, glutamate receptors *GRIN2A* and *2B*, brain-

expressed nuclear proteins such as *AUTS2,* and the transcription factor *FOXP2* (Graham and Fisher, 2012; Lesca, et al., 2013; Lesca, et al., 2012; Poot, et al., 2011).

ELP4 is one of six subunits (*ELP1-6*) of the Elongator complex, which plays a role in transcriptional elongation (Wittschieben, et al., 1999), tRNA modification and polarized exocytosis (Huang, et al., 2005). This complex also regulates the migration of multiple cell types; e.g. ELP1 co-localises with filamin A in membrane ruffles, and when depleted creates a disorganised actin cytoskeleton, contributing to motility defects (Johansen, et al., 2008). Impairment of Elongator may be involved in several different neurological disorders (Nguyen, et al., 2009) e.g. variants within *ELP3* are associated with cases of sporadic ALS, a progressive motor-neuron disease (Simpson, et al., 2009). Furthermore, mutations of *ELP1* cause familial dysautonomia (Slaugenhaupt, et al., 2001), a neurodevelopmental and neurodegenerative disorder with EEG abnormalities and seizures, characterized by defects in neironal development and survival. Elongator also underlies the migration and branching of cortical projection neurons during development and memory consolidation (Creppe, et al., 2009). Thus there are several mechanisms through which disruption of *ELP4* could result in altered neuronal development and migration, as well as the balance of neuronal excitatory and inhibitory circuits. These changes may disrupt Elongator function in a temporal and regional manner depending on cellular context and the different array of Elongator targets available.

It is of note that the large intronic regions between exons 9 and 12 of *ELP4* are ultraconserved. They contain long-range *cis*-regulatory enhancers for downstream *PAX6,* which are tissue- or developmental stage specific in their expression (McBride, et al., 2011). PAX6 is a transcription factor crucial for the correct development of the eyes, spinal cord, several brain regions and other organs. Deletions of *PAX6* with *WT1* cause Wilms tumor,

aniridia, genital anomalies, and intellectual disability (WAGR syndrome). Loss-of-function

mutations in *PAX6* also cause aniridia. A rare case of duplication of *PAX6* and the last two

exons/introns of *ELP4* has been reported with frontotemporal neonatal seizures,

developmental delay, microcephaly and minor ocular findings, (Aradhya, et al., 2011).

Recently *PAX6* has also been proposed as the foremost transcription factor governing

glutamatergic neuronal differentiation (Kim, et al., 2014), linking it with the major idiopathic

focal epilepsy gene glutamate receptor *GRIN2A*. Therefore disruption of *PAX6* and/or its

regulatory elements within *ELP4* and its link via the glutamatergic neurotransmission system

described above also make it a prime candidate for involvement in the neurodevelopmental

disorders described in some cases here. However, whilst it is difficult to untangle which gene

is causing which penotype in a few cases, for the majority, *PAX6* is not disrupted, indicating

the phenotypic consequences of *ELP4* disruption alone.

The genetic model of disease described here is clearly not monogenic: in 14/24 patients the

ELP4 CNVs were inherited (13 unrelated events due to the ASD sister pair); four occurred *de*

novo and six were of unknown inheritance. The phenotypic status of most parents is unknown

and therefore a precise estimation of penetrance will require further segregation studies.

However, presuming that the majority of the parents are unaffected, these inherited CNVs are

unlikely to cause a phenotype by reduced expression from haploinsufficiency alone. It is most

likely that an interacting model of disease is in action and screening of the second allele of

ELP4 and its regulatory regions can rule out the unmasking of recessive mutations. We note

that sequencing of *ELP4* exons has failed to find mutations within RE patients (Reinthaler, et

al., 2014) and postulate that disruption of the regulatory elements of *ELP4* and/or of *PAX6*

within its introns could be causal in the developmental disorders described here.

A two-hit hypothesis can explain CNVs that are non-syndromic, i.e. those that are associated

with variable phenotypes and not always inherited, such as the deletions described here,

(Girirajan and Eichler, 2010). One hit may reach a threshold to induce some sub-clinical

features and create a sensitized background, onto which the second hit (mutation or second

CNV) occurs producing a more severe phenotype. If we assume that these disorders share

common neurodevelopmental pathways, the final disease outcome will then differ depending

on the combination of genes affected. Interestingly, a sister pair with ASD in our study who

shared the same *ELP4* microdeletion, also shared a microdeletion of exon 2 of the carnitine

biosynthesis enzyme gene *TMLHE*, on Xq28. Deletions of *TMLHE* are important in non-

dysmorphic autism in male-male multiplex families, although with low penetrance (Celestino-

Soper, et al., 2012). However the significance in females is unclear. It is possible that

deletion of the only copy of *TMLHE* is enough of a risk factor for some males to develop

ASD, but for females (who normally have two copies of *TMLHE*), further 'hits' are

necessary, such as the loss of *ELP4* in these sisters. Several other patients also carry a second

CNV, as described earlier, but it is unlikely that these specific CNVs contribute to the

neurological phenotype. Exome sequencing of the patients without a second causal CNV may

uncover coding mutations that would contribute to the developmental burden of *ELP4* loss.

The predominance in our datasets of deletions verses duplications is unlikely to be a platform

bias as both aCGH and SNP arrays were used. Deletion enrichment could be a consequence of

undiagnosed duplications, but as this study was not driven by a particular diagnosis this is

unlikely. When CNVs are generated by non-allelic homologous recombination (NAHR)

between low-copy repeats, a deletion and reciprocal duplication are generated (Malhotra and

Sebat, 2012). A possibility is that the duplications could be selected against due to negative

genetic selection, i.e. a lower viability or fecundity of carriers. However, since the breakpoints

for *ELP4* CNVs differ between cases and there are no low-copy repeats that could explain the generation of CNVs, this mechanism is also unlikely. Instead, there is more in common with deletions seen at *NRXN1*, which may occur by a mechanism involving inverted repeats of variable sizes, or a significantly higher AT nucleotide content at the breakpoints, generating a rearrangement hotspot of genome instability. These non-recurrent breakpoints could be generated by a non-homologous end joining mechanism of double strand breaks or by replication errors and may be influenced by the genomic architecture of a region in particular people, (Enggaard Hoeffding, et al., 2014).

Examination of the data from the copy number variation morbidity map of developmental delay (Cooper, et al., 2011) shows four microdeletions (<1Mb) with breakpoints within *ELP4* (n=15,767), and five microduplications, Supplementary Information Table 1. All duplication cases had neurological deficits, and two deletion cases and one duplication had ASD. However, seven microdeletions of *ELP4* were also found in the 8,329 control individuals, one of which is the WTCCC sample reported here. This increase in frequency of smaller CNVs among controls compared to all of the other control datasets used in our study, indicates that they may be due to an artifact from the less dense Illumina arrays used by Cooper *et al*, compared to the more rigorous platform and methods used to analyse their cases. Indeed, the authors commented that their detection power is substantially higher for cases, the reverse of our study, and will manifest itself as false positive enrichment for CNVs in controls. However, more information (not publicly available) is needed about the specific arrays used for each control with an *ELP4* deletion, their LRR and BAF images and probe coverage over *ELP4* to draw further conclusions about potential false positives and array bias in their investigation.

Future work will focus on the functional consequences of the *ELP4* deletions by investigation

of expression levels of the gene in these cases. Work with cellular and animal models with

ELP4 deletions will help to cement the role of *ELP4* in neurodevelopment through

identification of altered interaction networks and developmental pathways such as neuronal

migration, branching and survival.

ACKNOWLEDGEMENTS

Compliance with Ethical Standards

The BBGRE project was approved by the Cambridgeshire Central Research Ethics Committee. In the UK, DECIPHER has been approved by the Eastern MREC 04/MRE05/50 and the project has also been notified to the Information Commissioner in accordance with the

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Data Protection Act. The DECIPHER Ethical Framework is detailed here:

https://decipher.sanger.ac.uk/assets/pdfs/decipher_ethical_framework.pdf

Ethical statements for the AGP and control populations and information on informed consent

can be found on their published references. Informed consent was obtained from all individual

participants included in this study.

REFERENCES

Ahn JW, Dixit A, Johnston C, Ogilvie CM, Collier DA, Curran S, Dobson RJ. 2013. BBGRE: brain and body genetic resource exchange. Database (Oxford) 2013:bat067.

Ahn JW, Mann K, Walsh S, Shehab M, Hoang S, Docherty Z, Mohammed S, Mackie Ogilvie C. 2010. Validation and implementation of array comparative genomic hybridisation as a first line test in place of postnatal karyotyping for genome imbalance. Mol Cytogenet 3:9.

Aradhya S, Smaoui N, Marble M, Lacassie Y. 2011. De novo duplication 11p13 involving the PAX6 gene in a patient with neonatal seizures, hypotonia, microcephaly, developmental disability and minor ocular manifestations. Am J Med Genet A 155A(2):442-4.

Bailey JA, Gu Z, Clark RA, Reinert K, Samonte RV, Schwartz S, Adams MD, Myers EW, Li PW, Eichler EE. 2002. Recent segmental duplications in the human genome. Science 297(5583):1003-7.

Celestino-Soper PB, Violante S, Crawford EL, Luo R, Lionel AC, Delaby E, Cai G, Sadikovic B, Lee K, Lo C, Gao K, Person REet al. 2012. A common X-linked inborn error of carnitine biosynthesis may be a risk factor for nondysmorphic autism. Proc Natl Acad Sci U S A 109(21):7974-81.

Consortium WTCC, Craddock N, Hurles ME, Cardin N, Pearson RD, Plagnol V, Robson S, Vukcevic D, Barnes C, Conrad DF, Giannoulatou E, Holmes Cet al. 2010. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature 464(7289):713-20.

Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, Williams C, Stalker H, Hamid R, Hannig V, Abdel-Hamid H, Bader Pet al. 2011. A copy number variation morbidity map of developmental delay. Nat Genet 43(9):838-46.

Creppe C, Malinouskaya L, Volvert ML, Gillard M, Close P, Malaise O, Laguesse S, Cornez I, Rahmouni S, Ormenese S, Belachew S, Malgrange Bet al. 2009. Elongator controls the migration and differentiation of cortical neurons through acetylation of alpha-tubulin. Cell 136(3):551-64.

Davis LK, Meyer KJ, Rudd DS, Librant AL, Epping EA, Sheffield VC, Wassink TH. 2008. Pax6 3' deletion results in aniridia, autism and mental retardation. Hum Genet 123(4):371-8.

Enggaard Hoeffding LK, Hansen T, Ingason A, Doung L, Thygesen JH, Moller RS, Tommerup N, Kirov G, Rujescu D, Larsen LA, Werge T. 2014. Sequence analysis of 17 NRXN1 deletions. Am J Med Genet B Neuropsychiatr Genet 165B(1):52-61.

Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, Van Vooren S, Moreau Y, Pettett RM, Carter NP. 2009. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet 84(4):524-33.

Girirajan S, Eichler EE. 2010. Phenotypic variability and genetic susceptibility to genomic disorders. Hum Mol Genet 19(R2):R176-87.

Graham SA, Fisher SE. 2012. Decoding the genetics of speech and language. Curr Opin Neurobiol 23(1):43-51.

Gregor A, Albrecht B, Bader I, Bijlsma EK, Ekici AB, Engels H, Hackmann K, Horn D, Hoyer J, Klapecki J, Kohlhase J, Maystadt Iet al. 2011. Expanding the clinical spectrum associated with defects in CNTNAP2 and NRXN1. BMC Med Genet 12:106.

Heinzen EL, Radtke RA, Urban TJ, Cavalleri GL, Depondt C, Need AC, Walley NM, Nicoletti P, Ge D, Catarino CB, Duncan JS, Kasperaviciute Det al. 2010. Rare deletions at 16p13.11 predispose to a diverse spectrum of sporadic epilepsy syndromes. Am J Hum Genet 86(5):707-18.

Helbig I, Mefford HC, Sharp AJ, Guipponi M, Fichera M, Franke A, Muhle H, de Kovel C, Baker C, von Spiczak S, Kron KL, Steinich Iet al. 2009. 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. Nat Genet 41(2):160-2.

Huang B, Johansson MJ, Bystrom AS. 2005. An early step in wobble uridine tRNA modification requires the Elongator complex. RNA 11(4):424-36.

Johansen LD, Naumanen T, Knudsen A, Westerlund N, Gromova I, Junttila M, Nielsen C, Bottzauw T, Tolkovsky A, Westermarck J, Coffey ET, Jaattela Met al. 2008. IKAP localizes to membrane ruffles with filamin A and regulates actin cytoskeleton organization and cell migration. J Cell Sci 121(Pt 6):854-64.

Kim KC, Lee DK, Go HS, Kim P, Choi CS, Kim JW, Jeon SJ, Song MR, Shin CY. 2014. Pax6-dependent cortical glutamatergic neuronal differentiation regulates autism-like behavior in prenatally valproic Acid-exposed rat offspring. Mol Neurobiol 49(1):512-28.

Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, Hubbell E, Veitch J, Collins PJ, Darvishi K, Lee C, Nizzari MMet al. 2008. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat Genet 40(10):1253-60.

Krawczak M, Nikolaus S, von Eberstein H, Croucher PJ, El Mokhtari NE, Schreiber S. 2006. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. Community Genet 9(1):55-61.

Lesca G, Rudolf G, Bruneau N, Lozovaya N, Labalme A, Boutry-Kryza N, Salmi M, Tsintsadze T, Addis L, Motte J, Wright S, Tsintsadze Vet al. 2013. GRIN2A mutations in acquired epileptic aphasia and related childhood focal epilepsies and encephalopathies with speech and language dysfunction. Nat Genet 45(9):1061-6.

Lesca G, Rudolf G, Labalme A, Hirsch E, Arzimanoglou A, Genton P, Motte J, de Saint Martin A, Valenti MP, Boulay C, De Bellescize J, Keo-Kosal Pet al. 2012. Epileptic encephalopathies of the Landau-Kleffner and continuous spike and waves during slow-wave sleep types: Genomic dissection makes the link with autism. Epilepsia 53(9):1526-38.

Lionel AC, Crosbie J, Barbosa N, Goodale T, Thiruvahindrapuram B, Rickaby J, Gazzellone M, Carson AR, Howe JL, Wang Z, Wei J, Stewart AFet al. 2011. Rare copy number variation discovery and cross-disorder comparisons identify risk genes for ADHD. Sci Transl Med 3(95):95ra75.

Malhotra D, Sebat J. 2012. CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell 148(6):1223-41.

Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, Thiruvahindrapduram B, Fiebig Aet al. 2008. Structural variation of chromosomes in autism spectrum disorder. Am J Hum Genet 82(2):477-88.

McBride DJ, Buckle A, van Heyningen V, Kleinjan DA. 2011. DNaseI hypersensitivity and ultraconservation reveal novel, interdependent long-range enhancers at the complex Pax6 cis-regulatory region. PLoS One 6(12):e28616.

Mefford HC, Muhle H, Ostertag P, von Spiczak S, Buysse K, Baker C, Franke A, Malafosse A, Genton P, Thomas P, Gurnett CA, Schreiber Set al. 2010. Genome-wide copy number variation in epilepsy: novel susceptibility loci in idiopathic generalized and focal epilepsies. PLoS Genet 6(5):e1000962.

Mefford HC, Yendle SC, Hsu C, Cook J, Geraghty E, McMahon JM, Eeg-Olofsson O, Sadleir LG, Gill D, Ben-Zeev B, Lerman-Sagie T, Mackay Met al. 2011. Rare copy number variants are an important cause of epileptic encephalopathies. Ann Neurol 70(6):974-85.

Nasr JT, Gabis L, Savatic M, Andriola MR. 2001. The Electroencephalogram in Children with Developmental Dysphasia. Epilepsy Behav 2(2):115-118.

Nguyen L, Humbert S, Saudou F, Chariot A. 2009. Elongator - an emerging role in neurological disorders. Trends Mol Med 16(1):1-6.

Pal DK, Li W, Clarke T, Lieberman P, Strug LJ. 2010. Pleiotropic effects of the 11p13 locus on developmental verbal dyspraxia and EEG centrotemporal sharp waves. Genes Brain Behav 9(8):1004-12.

Parmeggiani A, Barcia G, Posar A, Raimondi E, Santucci M, Scaduto MC. 2010. Epilepsy and EEG paroxysmal abnormalities in autism spectrum disorders. Brain Dev 32(9):783-9.

Pinto D, Darvishi K, Shi X, Rajan D, Rigler D, Fitzgerald T, Lionel AC, Thiruvahindrapuram B, Macdonald JR, Mills R, Prasad A, Noonan Ket al. 2011. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. Nat Biotechnol 29(6):512-20.

Pinto D, Delaby E, Merico D, Barbosa M, Merikangas A, Klei L, Thiruvahindrapuram B, Xu X, Ziman R, Wang Z, Vorstman JA, Thompson Aet al. 2014. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. Am J Hum Genet 94(5):677-94.

Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, Almeida J, Bacchelli Eet al. 2010. Functional impact of global rare copy number variation in autism spectrum disorders. Nature 466(7304):368-72.

Poot M, van der Smagt JJ, Brilstra EH, Bourgeron T. 2011. Disentangling the myriad genomics of complex disorders, specifically focusing on autism, epilepsy, and schizophrenia. Cytogenet Genome Res 135(3-4):228-40.

Reinthaler EM, Lal D, Jurkowski W, Feucht M, Steinbock H, Gruber-Sedlmayr U, Ronen GM, Geldner J, Haberlandt E, Neophytou B, Hahn A, Altmuller Jet al. 2014. Analysis of ELP4, SRPX2, and interacting genes in typical and atypical rolandic epilepsy. Epilepsia 55(8):e89-93.

Shinawi M, Liu P, Kang SH, Shen J, Belmont JW, Scott DA, Probst FJ, Craigen WJ, Graham BH, Pursley A, Clark G, Lee Jet al. 2010. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. J Med Genet 47(5):332-41.

Silversides C, Lionel A, Costain G, Merico D, Migita, Liu B, Yuen T, Rickaby J, Thiruvahindrapuram B, Marshall C, Scherer S, Bassett A. 2012. Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. PloS Genetics 8(8):e1002843.

Simpson CL, Lemmens R, Miskiewicz K, Broom WJ, Hansen VK, van Vught PW, Landers JE, Sapp P, Van Den Bosch L, Knight J, Neale BM, Turner MRet al. 2009. Variants of the elongator protein 3 (ELP3) gene are associated with motor neuron degeneration. Hum Mol Genet 18(3):472-81.

Simpson NH, Ceroni F, Reader RH, Covill LE, Knight JC, Hennessy ER, Bolton PF, Conti-Ramsden G, O'Hare A, Baird G, Fisher SE, Newbury DF. 2015. Genome-wide analysis identifies a role for common copy number variants in specific language impairment. Eur J Hum Genet.

Slaugenhaupt SA, Blumenfeld A, Gill SP, Leyne M, Mull J, Cuajungco MP, Liebert CB, Chadwick B, Idelson M, Reznik L, Robbins C, Makalowska Iet al. 2001. Tissue-specific expression of a splicing mutation in the IKBKAP gene causes familial dysautonomia. Am J Hum Genet 68(3):598-605.

Stewart AF, Dandona S, Chen L, Assogba O, Belanger M, Ewart G, LaRose R, Doelle H, Williams K, Wells GA, McPherson R, Roberts R. 2009. Kinesin family member 6 variant Trp719Arg does not associate with angiographically defined coronary artery disease in the Ottawa Heart Genomics Study. J Am Coll Cardiol 53(16):1471-2.

Strug LJ, Clarke T, Chiang T, Chien M, Baskurt Z, Li W, Dorfman R, Bali B, Wirrell E, Kugler SL, Mandelbaum DE, Wolf SMet al. 2009. Centrotemporal sharp wave EEG trait in rolandic epilepsy maps to Elongator Protein Complex 4 (ELP4). Eur J Hum Genet 17:1171-1181.

Wittschieben BO, Otero G, de Bizemont T, Fellows J, Erdjument-Bromage H, Ohba R, Li Y, Allis CD, Tempst P, Svejstrup JQ. 1999. A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II holoenzyme. Mol Cell 4(1):123-8.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

FIGURE LEGEND

Figure 1 Deletions (red) and a duplication (blue) identified over the *ELP4-PAX6* locus on 11p13 in 9 patients from the BB-GRE clinical genetic database with neurodevelopmental phenotypes, 6 patients with autism from the AGP and Canadian ASD resource, and 9 patients with neurodevelopmental phenotypes from the DECIPHER database. Hg19 (http://genome.ucsc.edu/). Tracks showing positions of probes genotyped from the Illumina 1M Single Array, the Affymetrix GenomeWide Human SNP6 Array and the Custom Agilent oligonucleotide array used for BBGRE patients are above the UCSC gene tracks. Alternatively spliced gene transcripts are shown.

Figure 1 Deletions (red) and a duplication (blue) identified over the ELP4-PAX6 locus on 11p13 in 9 patients from the BB-GRE clinical genetic database with neurodevelopmental phenotypes, 6 patients with autism from the AGP and Canadian ASD resource, and 9 patients with neurodevelopmental phenotypes from the DECIPHER database. Hg19 (http://genome.ucsc.edu/). Tracks showing positions of probes genotyped from the Illumina 1M Single Array, the Affymetrix GenomeWide Human SNP6 Array and the Custom Agilent oligonucleotide array used for BBGRE patients are above the UCSC gene tracks. Alternatively spliced gene transcripts are shown.
62x20mm (300 x 300 DPI)

Table 1. Microdeletions and a microduplication of *ELP4* on Chr11 identified in 9 patients from the BB-GRE clinical genetic database,

(http://bbgre-dev.iop.kcl.ac.uk). AgeAtTest indicates age at arrayCGH testing.

ID	Sex	AgeAtTest	Phenotype	hg19Start	hg19Stop	Size bp	Inheritance	CNV	Other CNV
129016	F	3 years	Developmental delay (progressing), microcephaly, poor balance	31,573,422	31,674,789	101,368	Unknown	X1	
119460	M	2 years	Social communication difficulties, speech and language delay	31,561,220	31,625,448	64,229	Maternal	x1	Deletion Chr5:97,302,377-97,380,022. No genes.
116589	M	3 years	PDD: Social interaction difficulties, language disorder, behaviour problems.	31,584,329	31,642,325	57,997	Maternal	x1	
108970	M	5 years	Severe cognitive delay (IQ 20-34) speech & language disorder, reading & spelling development disorder, autism spectrum disorder, epilepsy >24 months at age of onset	31,495,260	31,546,276	51,017	Paternal	x1	
117374	M	20 years	Autism, learning difficulties	31,705,076	31,747,631	42,556	Maternal	x1	
112601	F	1 year	Developmental delay, speech & language disorder, microcephaly (<5th centile), mild cognitive delay, motor skills development disorder	31,691,270	31,722,740	31,471	Paternal	x1	Duplication Chr5: 93,197,999-93,424,468. Disrupts *FAM172A*.

117003	M	8 months	Developmental delay, hypotonia, ventriculomegaly	31,601,768	31,632,347	30,580	Maternal	x1	
130693	M	5 years	Moderate developmental delay mainly affecting language, emerging autistic traits	31,760,904	31,786,914	26,010	*De novo*	x1	Deletion Chr6: 26,440,746-26,463,502. Disrupts *BTN3A3* and *BTN2A1*.
112031	F	12 years	Developmental delay, hypotonia	31,616,889	31,849,574	232,686	Unknown	x3	Deletion Chr16: 15,048,750-16,305,736 16p13.11 hotspot.

Table 2. Microdeletions of *ELP4* on Chr11 identified in 9 individuals from the DECIPHER database. Age indicates age at last clinical assessment.

ID	Sex	Age	Phenotype	hg19Start	hg19Stop	Size bp	Inheritance	CNV	Other CNV
265704	M	<1yr	Aniridia. No further information.	31,172,410	31,775,457	603,047	*De novo*	x1	
249728	F	24yrs	Rieger anomaly.	31,118,027	31,710,576	592,549	*De novo*	x1	
257614	M	7yrs	Epilepsy – partial complex seizures with secondary generalization due to cortical dysplasia, mild developmental delay, ADHD, neurinomas, congenital malformation in left eye, fine motor dyspraxia.	30,991,456	31,564,708	573,252	Parent	x1	
292869	F	15yrs	Severe intellectual disability, muscle hypotrophy with severe hypotonia and absent gross motor and fine adaptive motor development; no language; severe dysphagia requiring tube feeding; craniofacial abnormalities.	31,597,322	31,802,120	204,798	*De novo*	x1	
287341	M	2yrs	Partial aniridia. Currently no signs of neurological impairment.	31,605,859	31,783,590	177,731	Maternal	x1	
258970	M	4 yrs	Developmental delay, behavioral disturbances, regression of language at 18mo to absent at age 4, pervasive developmental disorder	31,605,859	31,775,457	169,598	Unknown	x1	

261471	M	4yrs	Behavioral and speech disorders, mild mental retardation.	31,625,389	31,775,457	150,068	Parent	x1	Deletion Chr10:35360169-35605506 disrupting *CUL2*, *CREM*, *CCNY*.
289275	M	24yrs	Aniridia, global developmental delay, autistic behaviour.	31,742,075	31,870,603	128,528	Unknown	x1	Duplication ChrX:46389227-46396390 disrupting intron of *ZNF674*.
270752	F	9yrs	Aniridia, congenital cataract. Mild developmental delay due to processing speed deficiencies largely due to visual impairment.	31,735,689	31,825,698	90,009	Paternal	x1	

Table 3. Microdeletions of *ELP4* on Chr11 identified in 3 patients from the Autism Genome Project and 3 from a Canadian study of autism. * sibling sister pair.

Autism Study ID	Sex	Phenotype	hg19Start	hg19Stop	Size bp	Inheritance	CNV	Other CNV
3617_3	M	Strict autism; no seizures, verbal, language delay; delayed first words (at 24 mo), delayed first phrases (at 36 mo); verbal IQ >70.	31460506	31655108	194,602	Paternal	x1	
NA0285	M	Autism, language delay; delayed first words (at 32 mo), no seizures	31518924	31649475	130,551	Maternal	x1	
8596_201	M	Strict autism, high functioning; no seizures, verbal, language delay; delayed first words (at 25 mo), typical first phrases (at 25 mo); verbal IQ >70.	31488890	31607986	119,096	Maternal	x1	Duplication, chr9:115994263-116495631 disrupting *COL27A1*
MM1259-003*	F	Autism, language delay; delayed first words (at 21 mo), mild developmental delay, motor delay, no seizures	31652219	31764393	112,174	Unknown	x1	Deletion chrX:154772341-154775951 disrupting *TMHLE*
MM1259-004*	F	Autism, language delay; delayed first words (at 18 mo), delayed first phrases (at 36 mo), expressive language problems, mild developmental delay, motor delay, no seizures	31652219	31764393	112,174	Unknown	x1	Deletion chrX:154772341-154775951 disrupting *TMHLE*
20130_6005001	M	Strict autism; no seizures, verbal, language delay; delayed first words (at 36 mo), delayed first phrases (at 48 mo); verbal IQ >70, coordination problems	31576768	31653568	76,800	Maternal	x1	

Supplementary Table 1 Microdeletions and duplications disrupting *ELP4* reported in cases and controls from the copy number variation

morbidity map of developmental delay (Cooper et al., 2011).

Sample ID	Variant ID	hg19Start	hg19Stop	Size bp	CNV
Cases:					
9890931	nsv540979	31,118,026	31,790,388	672,363	x3
9882508	nsv540980	31,152,003	31,751,699	599,697	x1
9896715	nsv540981	31,401,095	31,656,511	255,417	x1
9908285	nsv540986	31,747,371	32,063,394	315,979	x3
9883063	nsv540987	31,775,599	31,804,354	28,756	x1
9882119	nsv540982	31,602,061	31,875,238	273,178	x3
9899815	nsv540983	31,656,450	31,857,797	201,348	x3
9883029	nsv540984	31,703,100	31,751,699	48,600	x1
9889844	nsv540985	31,703,100	31,929,503	226,404	x3
Controls:					
WTCCC	nsv553949	31,317,835	31,539,587	221,753	x1
	nsv553989	31,727,232	31,796,560	69,329	x1
HGDP00106	nsv5553987	31,488,890	31,696,336	207,447	x1
	nsv553986	31,488,890	31,679,048	190,159	x1
	nsv553988	31,535,355	31,651,189	115,835	x1
	nsv553984	31,454,975	31,654,406	199,432	x1
	nsv553985	31,488,890	31,654,406	165,517	x1