npg

SHORT REPORT

Biochemical screening and *PTEN* mutation analysis in individuals with autism spectrum disorders and macrocephaly

Judith A Hobert[1], Rebecca Embacher[2], Jessica L Mester[1,3], Thomas W Frazier II[1,2] and Charis Eng*[,1,3,4,5]

Unlike some other childhood neurodevelopmental disorders, no diagnostic biochemical marker has been identified in all individuals with an autism spectrum disorder (ASD). This deficit likely results from genetic heterogeneity among the population. Therefore, we evaluated a subset of individuals with ASDs, specifically, individuals with or without macrocephaly in the presence or absence of *PTEN* mutations. We sought to determine if amino or organic acid markers could be used to identify individuals with ASDs with or without macrocephaly in the presence or absence of *PTEN* mutations, and to establish the degree of macrocephaly in individuals with ASDs and *PTEN* mutation. Urine, blood and occipital–frontal circumference (OFC) measurements were collected from 69 individuals meeting DSM-IV-TR criteria. Urine and plasma samples were subjected to amino and organic acid analyses. *PTEN* was Sanger-sequenced from germline genomic DNA. Germline *PTEN* mutations were identified in 27% (6/22) of the macrocephalic ASD population. All six *PTEN* mutation-positive individuals were macrocephalic with average OFC + 4.35 standard deviations (SDs) above the mean. No common biochemical abnormalities were identified in macrocephalic ASD individuals with or without *PTEN* mutations. In contrast, among the collective ASD population, elevation of urine aspartic acid (87%; 54/62), plasma taurine (69%; 46/67) and reduction of plasma cystine (72%; 46/64) were observed. *PTEN* sequencing should be carried out for all individuals with ASDs and macrocephaly with OFC ≥ 2SDs above the mean. A proportion of individuals with ASDs may have an underlying disorder in sulfur amino acid metabolism.
European Journal of Human Genetics (2014) 22, 273–276; doi:10.1038/ejhg.2013.114; published online 22 May 2013

Keywords: *PTEN*; autism spectrum disorder; macrocephaly; Cowden syndrome

INTRODUCTION

Autism spectrum disorders (ASDs) are a group of neurodevelopmental disorders with a strong genetic and genomic component.[1,2] They are characterized by impairments in social interaction and communication, and restricted, repetitive behaviors. These disorders are often associated with substantial functional and language impairments that inhibit independent living and frequently necessitate lifelong care.[3] ASDs include autistic disorder, Rett syndrome, childhood disintegrative disorder, pervasive developmental disorder – not otherwise specified (PDD-NOS) and Asperger syndrome. Estimates from the Centers for Disease Control indicate that as many as 1 in 88 children may be affected by some form of ASD.[4]

Several well-described childhood neurodevelopmental disorders, such as phenylketonuria and maple syrup urine disease, are monogenic and are directly linked to specific metabolic/biochemical abnormalities. In contrast, ASDs are both clinically and genetically heterogeneous, and several different metabolic/biochemical abnormalities have been observed. However, no single biochemical abnormality has been consistently identified in all affected individuals. Therefore, it has been difficult to identify biochemical markers that could aid in the diagnosis and biological understanding of these disorders.

It is clear that phenotypic subgroups exist within the ASDs, such as macrocephaly, abnormal ear structure and unusual hair growth

pattern.[5,6] Among these subgroups, macrocephaly is the most frequently identified physical finding and has been reported in 15–35% of individuals.[7] In addition, germline *PTEN* (OMIM #601728) mutations have been identified in a subset of individuals having both ASDs and macrocephaly.[8–11] We recently identified a biochemical alteration (elevated plasma succinate) in *PTEN* and *SDH* mutation-positive individuals meeting full or partial criteria for Cowden Syndrome (CS, OMIM #158350), in which macrocephaly is one major diagnostic criterion.[12] Therefore, we hypothesized that a biochemical marker(s) may exist that can be used to identify individuals within similar phenotypic (and in some cases genetic) subgroups. The aim of our current study was to ascertain whether common biochemical alterations could be identified in subgroups of individuals with ASDs, including those with *PTEN* mutations.

MATERIALS AND METHODS

Research participants

Between April 2010 and April 2012, 69 individuals with ASDs identified through the Cleveland Clinic Center for Autism (diagnostic clinic, outreach programs, Autism School), and the Center for Personalized Genetic Healthcare of the Genomic Medicine Institute at the Cleveland Clinic were recruited for study. Inclusion in the study required documentation of a clinical diagnosis of an ASD from a medical or mental-health professional. Macrocephalic and

[1]Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, OH, USA; [2]Center for Autism, Pediatrics Institute, Cleveland Clinic, Cleveland, OH, USA; [3]Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH, USA; [4]Department of Genetics and Genome Sciences, Case Western Reserve University School of Medicine, Cleveland, OH, USA; [5]CASE Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, OH, USA.
*Correspondence: Professor C Eng, Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, 9500 Euclid Avenue, NE-50, Cleveland, OH 44195, USA.
Tel: + 1 216 444 3440; Fax: + 1 216 636 0655. E-mail: engc@ccf.org

normocephalic participants aged 2–50 years were recruited. Written informed consent was obtained from all adult participants not requiring a legal guardian, or from the parents or guardians of minors or those unable to provide consent. Adults (ages 18–50 years) requiring legal guardians and youth (ages 2–17 years), who were able, were asked to provide assent. The study was approved by the Cleveland Clinic Institutional Review Board for Human Subjects' Protection. Plasma samples for amino and organic acid analyses and occipital–frontal circumference measurements from five *PTEN* mutation-positive individuals who do not exhibit ASD-related symptoms were used as controls.

Procedures

Occipital–frontal circumference (OFC) measurements were obtained from participants by a member of the patients' health-care team using standard procedures. Macrocephaly was defined as ≥ 2 standard deviations (SD) above the mean for normal individuals. SD data for patients were calculated using CDC norms from birth to age 3 and published data from ages 3 to 18 years.[13,14] ASD diagnoses were made using DSM-IV-TR criteria based on expert clinical judgment supplemented by the Autism Diagnostic Interview-Revised and, in less-affected cases, the Autism Diagnostic Observation Schedule.[15] All individuals aged >3 were also evaluated using the Social Responsiveness Scale (preschool, standard, and adult versions).[16,17] Evaluation also included behavioral observations of autism symptoms collected during neurocognitive testing.

Random urine and/or blood samples were obtained during scheduled visits. Following collection, samples were de-identified. Participants were not required to fast prior to sample collection. Genomic DNA was isolated from collected blood specimens and *PTEN* (NM_000314.4; genomic context NC_000010.10 (89623195..89728532)) mutation analysis was performed with a combination of light scanner technology and Sanger sequencing (ABI3730xl). Pathogenicity of variants is determined by bioinformatic and functional analysis by western blot of downstream read outs P-AKT and P-ERK1/2 (MAPK1/2) using standard techniques.

Urine and plasma were aliquoted and frozen within 1 h of collection, and stored at $-80\,^{\circ}$C for future amino acid and organic acid analyses. Amino acid concentrations were determined using ion-exchange chromatography. Organic acid concentrations were determined using gas chromatography-mass spectrometry. Both amino and organic acid analyses were performed in the Biochemical Genetics Laboratory of ARUP, Salt Lake City, UT, USA. Positive and negative controls are run by the clinical laboratory with each batch of samples. Measured amino acids in urine included alanine, arginine, asparagine, aspartic acid, citrulline, cystine, glutamine, glutamic acid, glycine, histidine, homocystine, hydroxyproline, isoleucine, leucine, lysine, methionine, ornithine, phenylalanine, proline, serine, taurine, threonine, tyrosine, and valine. Amino acids measured in plasma included alanine, allo-isoleucine, arginine, aspartic acid, citrulline, cystine, glutamic acid, glutamine, glycine, histidine, homocystine, hydroxyproline, isoleucine, leucine, lysine, methionine, ornithine, phenylalanine, proline, serine, taurine, threonine, tyrosine, and valine. ARUP does not measure tryptophan due to technical constraints.

Measured organic acids in urine included, but were not limited to, lactic acid, pyruvic acid, succinic acid, fumaric acid, 2-ketoglutaric acid, methylmalonic acid, 3-hydroxybutyric acid, acetoacetic acid, 2-keto-3-methylvaleric acid, 2-ketoisocaproic acid, 2-ketoisovaleric acid, ethylmalonic acid, adipic acid, suberic acid, sebacic acid, 4-hydroxyphenylacetic acid, 4-hydroxyphenyllactic acid, 4-hydroxy-phenylpyruvic acid, and succinylacetone. Measured organic acids in plasma included, but were not limited to, lactic acid, pyruvic acid, succinic acid, 3-hydroxybutyric acid, acetoacetic acid, 2-keto-3-methyl-valeric acid, 2-ketoisocaproic acid, 2-ketoisovaleric acid, and citric acid. Reference ranges for urine and plasma amino and organic acids were established in an age-matched population by the Biochemical Genetics Laboratory at ARUP. Urinary amino acids were reported as µmol/creatinine and plasma amino acids as µmol/l. Urinary organic acids were reported as mmol of acid/mole of creatinine and plasma values as µmol/l.

RESULTS
Study summary

A total of 69 individuals (49 autistic, 13 PDD-NOS, 7 Aspergers) were enrolled in the study. Enrolled participants were predominantly male

(90%) and ranged in age from 2 to 45 years. Forty-five individuals were found to be normocephalic (45/68; 66%), 22 were macrocephalic (22/68; 32%), 1 was microcephalic and in one study participant, OFC was not determined (Table 1 and additional data file 1). All macrocephalic individuals had 'absolute macrocephaly', meaning they were of normal height for their age (data not shown).

PTEN genotyping and OFC

Overall, germline pathogenic *PTEN* mutations were detected in 6 (6/62; 10%) males and no females in the study population. The six detected pathogenic *PTEN* mutations included one frameshift, three amino-acid substitutions, one premature termination and one splice-site mutation (Table 2). All *PTEN* mutation-positive individuals were macrocephalic, with OFCs ranging from $+3.6$ to $+5.4$ SDs above the mean. All mutation-positive individuals exhibited phenotypes consistent with a diagnosis of Bannayan–Riley–Ruvalcaba syndrome (BRRS, OMIM #153480) (Table 2).

Macrocephaly was observed in individuals with ASDs in the absence of germline *PTEN* mutations (16/62; 26%) and in individuals with ASD and germline *PTEN* mutations (6/6, 100%) ($P<0.05$). Overall, 27% (6/22) of the macrocephalic population harbored a germline *PTEN* mutation. The prevalence and degree of macrocephaly was most striking in *PTEN* mutation-positive individuals. Macrocephaly was identified in all *PTEN* mutation-positive individuals with ASDs (6/6; 100%; OFC $+4.35$), as well as the majority of *PTEN* mutation-positive individuals without ASDs (4/5; 80%; OFC $+3.96$). (Table 3 and Supplementary data). No significant difference in OFC was observed when comparing these two populations ($P=0.7$).

Biochemical studies

A total of 68 plasma samples and 62 urine samples from 69 individuals were collected and subjected to biochemical analyses. Urine amino acid analysis showed elevations in aspartic acid (54/62; 87%) and glycine (23/62; 37%). Plasma amino acid analysis revealed an elevation in taurine (46/67; 69%) and reductions in cystine (46/64; 72%) and methionine (15/67; 22%) (Table 4). None of the observed amino acid changes were thought to be related to reported medications (data not shown), and were not associated with specific groups within this study population (see Supplementary Table). Plasma and urine organic acid analyses did not identify any common abnormalities among the study population (data not shown).

Table 1 Patient characteristics

	Autistic, n = 49	PDD-NOS, n = 13	Asperger's, n = 7	Total, n = 69
Gender				
Female	5	2	0	7
Male	44	11	7	62
Mean age, years	8.3	13	11	9.6
Head circumference				
Normal	30	11	4	45
Macrocephaly	18	2	2	22
Microcephaly	0	0	1	1
Unknown	1	0	0	1

Abbreviations: PDD-NOS, pervasive developmental disorder-not otherwise specified.
Note: the three individuals designated as 'unknown' were diagnosed using the Autism Diagnostic Interview-Revised (ADI-R) or expert clinician observation over multiple sessions. They all meet DSM-IV-TR criteria for an ASD using the CDC catchment definition. Macrocephaly ≥ 2SDs above the mean; microcephaly ≤ 2SDs below the mean.

Table 2 Germline *PTEN* mutations identified in patients with ASDs and macrocephaly

ID	M/F	Age	Mutation	Proteinª	OFC SD	PHTS features
5130-01-001	M	7	c. 420_421insA	p.H141TFS*39	+5.37	CAL, CNSH, DD, VN, PF
5556-01-001	M	13	c. 208C>G	p.L70V	+4.21	Seiz, DD, Hash, L, OMP
5724-01-002	M	11	c. 3G>T	p.M1I	+3.94	CAL, Seiz, OMP, PF
5909-01-001	M	7	c. 1003C>T	p.R335*	+3.58	BN, DD, PF
5708-01-001	M	11	c. 209 + 5G>A	Splice-site mutation, intron 3	+4.59	CAL, DD, L, PF, TN
5724-01-001	M	12	c. 3G>T	p.M1I	+4.41	Seiz, OMP, PF

Abbreviations: BN, Benign skin neoplasm, CAL, Café-au-lait macules; CNSH, CNS hemangioma; DD, Developmental delay; FS, Frameshift; Hash, Hashimoto disease; L, Lipoma; VN, Vascular Neoplasm; OMP, Oral mucosal papillomatosis; PF, Penile freckling; Seiz, Seizure history; TN, Thyroid nodule.
ªNote that protein lysates show upregulation of the AKT and MAPK pathway reflected by increased P-AKT and P-MAPK42/44 on western blot[27] (Lei and Eng, unpublished data).

Table 3 OFC status by *PTEN* genotype

Genotype	Average OFC	OFC std dev	Macrocephalic (%)	n
PTEN M− ASD+	+1.46	1.91	10/31 (32)	32
PTEN Promoter −903G>A; c. 1026+32T>G	+0.03	0.27	0/2 (0)	2
PTEN c. 1026+32T>G	+1.22	1.22	5/28 (18)	28
PTEN c. 1026+32T>G homozygous	+6.07	—	1/1 (100)	1
PTEN M+ ASD+	+4.35	0.61	6/6 (100)	6
PTEN M+ ASD−	+3.96	2.24	4/5 (80)	5

Abbreviations: ASD, Autism spectrum disorder; M-, mutation negative; M+, mutation positive; OFC, occipital-frontal circumference.

Table 4 Summary of biochemical findings

	n	%
Elevated		
Asp, urine	54/62	87
Gly, urine	23/62	37
Gly, plasma	0/67	0
Tau, plasma	46/67	69
Reduced		
Cys, plasma	46/64	72
Met, plasma	15/67	22

Abbreviations: Asp, aspartic acid; Cys, cystine; Gly, glycine; Met, methionine; Tau, taurine.

DISCUSSION

Rapid brain growth during early childhood has been strongly associated with ASDs, particularly for male individuals with regressive autism.[7,18] Although macrocephaly has been reported in 15–35% of children with an ASD, and is the most frequent physical finding,[7] it is also observed in ∼2% of healthy individuals and others who do not meet DSM-IV-TR criteria.

Several genetic and non-genetic conditions are associated with macrocephaly and include benign familial macrocephaly, cancer predisposition syndromes, and metabolic disorders.[11] Germline *PTEN* mutations are positively associated with the occurrence and degree of macrocephaly both in humans and in *Pten* knock-in mice.[19] Germline *PTEN* mutations molecularly define *PTEN* hamartoma tumor syndrome (PHTS), a cancer predisposition syndrome, and are associated with an elevated risk of female breast, epithelial thyroid, renal, and endometrial cancers.[20] PHTS is a molecular-based umbrella term that encompasses *PTEN* mutation-positive individuals that may be affected with clinically distinct syndromes, chief of which are CS and BRRS. Previous studies have reported germline *PTEN* mutations in 1–20% of macrocephalic individuals with an ASD, intellectual disability or developmental delay.[8–10,21] These studies reported OFCs for *PTEN* mutation-positive individuals ranging from +2.5–8.0,[8] +2.9–5.8,[10] and +3.3–6.1[9] SDs above the mean. Similarly, we have identified germline *PTEN* mutations in 27% of macrocephalic individuals with mean OFC +4.4 SDs above the mean. This finding corroborates a recent study that examined the prevalence and degree of macrocephaly in *PTEN* mutation-positive individuals and reported the prevalence and mean OFC (≥18 years) to be 94% and +4.89 SDs above the mean, respectively.[19]

Biochemical analyses of urine and blood from individuals in our study did not reveal amino acid or organic acid alterations that could be associated with a specific subgroup, such as ASD and macrocephaly or an ASD and germline *PTEN* mutations. The elevations in plasma succinate that we had previously observed among *PTEN* and *SDH* mutation-positive individuals meeting criteria for Cowden Syndrome were not observed among our cohort of *PTEN* mutation-positive individuals with ASD and macrocephaly.[12] This difference might arise from different disease processes or could be related to patient age, as most CS-associated phenotypes do not arise until the second or third decade of life. However, elevated urine aspartic acid (87%), elevated plasma taurine (69%) and a reduced plasma cystine (72%) were observed among the entire group affected with ASDs. Reductions in plasma cystine can result from a delay between sample collection and storage or from prolonged storage at −20 °C. To maintain sample integrity in our study, blood and urine were processed and frozen within 1 hour of collection and were stored at −80 °C until analyses were performed (as described in the MATERIALS AND METHODS section). Limitations of our study include small cohort and mixture of fed and fasted samples (sample collection procedure did not specify that subjects be in a fasted or fed state).

Our current findings are consistent with elevated urinary aspartic acid (11/14; 79%),[22] elevated plasma taurine (7/14; 50%)[22,23] and reduced plasma[24] and urinary[23] cysteine that have been previously reported in children with autism. Furthermore, our observations of elevated plasma taurine and reduced plasma cystine, and in some cases plasma methionine, are consistent with a previous study showing that individuals with ASDs may have a perturbation in sulfur amino acid metabolism.[23] In contrast to our findings, other studies did not identify elevated urinary aspartic acid or reduced urinary cystine[25] or reduction in plasma taurine among their study populations.[24,26] The reason for disparities between these reports is unclear, but may stem from differences in the study populations and the heterogeneity of the ASDs.

To the best of our knowledge, our study is the first to evaluate urine and plasma amino acid and organic acid abnormalities in individuals with ASD, macrocephaly, and *PTEN* mutations. Although we were not able to demonstrate a clear association between specific

biochemical abnormalities and subsets of ASDs, we have demonstrated a strong association connecting macrocephaly and *PTEN* mutations in individuals with ASDs and have confirmed previous reports of elevated urinary aspartic acid, plasma taurine and reduced plasma cystine and methionine in individuals with ASDs. Therefore, because germline *PTEN* mutations are linked to an increased risk of cancer, we recommend *PTEN* mutation analysis for individuals with autism and macrocephaly. Furthermore, cancer surveillance strategies should be implemented for individuals and family members who are found to harbor *PTEN* mutations. Future studies are needed to further examine the potential perturbation in sulfur amino acid metabolism that we have reported here.

CONFLICT OF INTEREST

ACKNOWLEDGEMENTS

1 Ronald A, Hoekstra RA: Autism spectrum disorders and autistic traits: a decade of new twin studies. *Am J Med Genet B Neuropsychiatr Genet* 2011; **156B**: 255–274.
2 Abrahams BS, Geschwind DH: Advances in autism genetics: on the threshold of a new neurobiology. *Nat Rev Genet* 2008; **9**: 341–355.
3 Rapin I: Autism. *N Engl J Med.* 1997; **337**: 97–104.
4 Baio J: Prevalence of autism spectrum disorders—Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008. *Morbidity and Mortality Weekly Report (MMWR)*, (30 March 2012/61(SS03)): 1–19.
5 Miles JH, Takahashi TN, Hong J *et al*: Development and validation of a measure of dysmorphology: useful for autism subgroup classification. *Am J Med Genet A* 2008; **146A**: 1101–1116.
6 Ingram DG, Takahashi TN, Miles JH: Defining autism subgroups: a taxometric solution. *J Autism Dev Disord* 2008; **38**: 950–960.
7 Courchesne E, Carper R, Akshoomoff N: Evidence of brain overgrowth in the first year of life in autism. *JAMA* 2003; **290**: 337–344.
8 Butler MG, Dasouki MJ, Zhou XP *et al*: Subset of individuals with autism spectrum disorders and extreme macrocephaly associated with germline PTEN tumour suppressor gene mutations. *J Med Genet* 2005; **42**: 318–321.
9 McBride KL, Varga EA, Pastore MT *et al*: Confirmation study of PTEN mutations among individuals with autism or developmental delays/mental retardation and macrocephaly. *Autism Res* 2010; **3**: 137–141.
10 Varga EA, Pastore M, Prior T, Herman GE, McBride KL: The prevalence of PTEN mutations in a clinical pediatric cohort with autism spectrum disorders, developmental delay, and macrocephaly. *Genet Med* 2009; **11**: 111–117.
11 Williams CA, Dagli A, Battaglia A: Genetic disorders associated with macrocephaly. *Am J Med Genet A* 2008; **146A**: 2023–2037.
12 Hobert JA, Mester JL, Moline J, Eng C: Elevated plasma succinate in PTEN, SDHB, and SDHD mutation-positive individuals. *Genet Med* 2012; **14**: 616–619.
13 Roche AF, Mukherjee D, Guo SM, Moore WM: Head circumference reference data: birth to 18 years. *Pediatrics* 1987; **79**: 706–712.
14 C. f. D. C. a. P. NCfHS: Centers for Disease Control. Clinical growth chars, 2005. Available at http://www.cdc.gov/growthcharts/clinical_charts.htm#Set2.
15 Lord CRM, DiLavore PC, Risi S: *Autism Diagnostic Observation Schedule: ADOS manual.* Los Angeles, CA, USA, 2002.
16 Constantino JN, Todd RD: Intergenerational transmission of subthreshold autistic traits in the general population. *Biol Psychiatry* 2005; **57**: 655–660.
17 Constantino JN, Gruber CP: Social Responsiveness Scale: Manual; In: *Book Social Responsiveness Scale: Manual.* Los Angeles, CA: Western Psychological Services, 2005.
18 Nordahl CW, Lange N, Li DD *et al*: Brain enlargement is associated with regression in preschool-age boys with autism spectrum disorders. *Proc Natl Acad Sci USA* 2011; **108**: 20195–20200.
19 Mester JL, Tilot AK, Rybicki LA, Frazier TW 2nd, Eng C: Analysis of prevalence and degree of macrocephaly in patients with germline PTEN mutations and of brain weight in Pten knock-in murine model. *Eur J Hum Genet* 2011; **19**: 763–768.
20 Tan MH, Mester JL, Ngeow J, Rybicki LA, Orloff MS, Eng C: Lifetime cancer risks in individuals with germline PTEN mutations. *Clin Cancer Res* 2012; **18**: 400–407.
21 Buxbaum JD, Cai G, Chaste P *et al*: Mutation screening of the PTEN gene in patients with autism spectrum disorders and macrocephaly. *Am J Med Genet B Neuropsychiatr Genet* 2007; **144B**: 484–491.
22 Moreno-Fuenmayor H, Borjas L, Arrieta A, Valera V, Socorro-Candanoza L: Plasma excitatory amino acids in autism. *Invest Clin* 1996; **37**: 113–128.
23 Yap IK, Angley M, Veselkov KA, Holmes E, Lindon JC, Nicholson JK: Urinary metabolic phenotyping differentiates children with autism from their unaffected siblings and age-matched controls. *J Proteome Res* 2010; **9**: 2996–3004.
24 Geier DA, Kern JK, Garver CR, Adams JB, Audhya T, Geier MR: A prospective study of transsulfuration biomarkers in autistic disorders. *Neurochem Res* 2009; **34**: 386–393.
25 Evans C, Dunstan RH, Rothkirch T *et al*: Altered amino acid excretion in children with autism. *Nutr Neurosci* 2008; **11**: 9–17.
26 Adams JB, Audhya T, McDonough-Means S *et al*: Nutritional and metabolic status of children with autism *vs* neurotypical children, and the association with autism severity. *Nutr Metab (Lond)* 2011; **8**: 34.
27 Tan MH, Mester J, Peterson C *et al*: A clinical scoring system for selection of patients for *PTEN* mutation testing is proposed on the basis of a prospective study of 3,042 probands. *Am J Hum Genet* 2011; **88**: 42–56.

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