2. Materials and methods

2.1. Patient

This 6-year-old boy is the only child of his nonconsanguineous Japanese parents. He was born with a body weight of 2478 g without asphyxia at 37 weeks. He was noted to have poor weight gain at ~2 months. He exhibited hypertonus and athetosis at 4 months, and was referred to our neurology division for evaluation at 8 months. At first evaluation, he showed failure to thrive with a body weight of 6.0 kg (−2.7 SD) and height of 60.0 cm (−4.4 SD).

He showed no dysmorphic features and behavioral problems. He could not hold his head or follow an object with his eyes. His muscle tone was hypertonic and all extremities displayed exaggerated tendon reflexes and bilateral extensor plantar responses. Dystonia was also evident in his upper and lower extremities. Routine laboratory examinations revealed mild elevations of aminotransferase/alanine aminotransferase (AST/ALT) of ~100 IU/mL; however, there were no biochemical abnormalities in the levels of serum ammonia, lactate, and pyruvate, very long chain fatty acids, or arylsulfatase A. Nerve conduction velocities and electromyographic studies were all normal. No waves were identified by auditory evoked brain responses, even at a maximum stimulation intensity of 115 dB. Magnetic resonance imaging (MRI) revealed abnormal increased signals in the globus pallidi on T2-weighted images (Fig. 1A, upper left). Myelination delay, a decrease in the volume of cerebral white matter, and a very thin corpus callosum were also observed (Fig. 1A, lower left). At 4 years of age, he was admitted to our hospital due to status epileptics with generalized tonic–clonic seizures lasting 1 h, which were controlled by intravenous diazepam. Electroencephalography showed frequent spikes and slow waves in central and mid-temporal areas. He was initiated with phenobarbital and no seizures have been noticed since then. MRI revealed the completion of myelination, persistent signal abnormalities in the globus pallidi, and the progression of cerebellar vermian atrophy (Fig. 1A, right panels). A deficient Cr peak in ^{1}H-magnetic resonance spectroscopy (MRS) (Fig. 1B) and high ratio of creatine/creatinine concentration in urine 4.73 mmol/mmol (normal range; 0.0075-1.51) led us to suspect SLC6A8 deficiency [17]. Now he is still very small and showing severe dystonia at 6 years of ages. He cannot hold his head and does not pursuit object. During upper respiratory infections, he showed increased levels of AST/ALT >10,000 IU/mL without cholestasis. These values returned to his baseline levels without any therapy after a few weeks.

2.2. Cr uptake in fibroblasts

Cr uptake was measured in total cell lysates in triplicate using gas chromatography–mass spectrometry with stable isotope-labeled Cr as the internal standard, as described by Rosenberg et al. [18]. Briefly, fibroblasts from the patient were cultured for 24 h in medium (HAM/F10 supplemented with 10% fetal bovine serum, penicillin, and streptavidin) that contained physiological levels of Cr (25 mM). Cr was added to the medium and obtains final concentrations of 25 and 500 mM [12].

2.3. Genomic DNA sequencing, RT-PCR, and sequencing

Genomic DNA was prepared from the patient's white blood cells using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA). PCR of all exons and exon–intron boundaries of the *SLC6A8* gene was performed as previously described [10]. Subsequent sequencing analyses of the PCR fragments were performed with specific primers using the Ex Taq PCR Kit (version 1.0; Takara, Shiga, Japan) according to the manufacturer's instructions (Supplementary Table). PCR buffers were selected from the GC I, GC II, and EX buffers (Takara) (Supplementary Table). Total RNA was extracted from leukocytes using the TRIzol reagent and reverse transcribed with Prime Script Reverse Transcriptase (Takara, Shiga, Japan) using oligo(dT) primers. RT-PCR was performed using primers that covered from exon 2 to exon 13 according to the manufacturer's instructions (Supplementary Table). The PCR fragments were sequenced using the Big Dye Terminators Cycle Sequencing Kit (v1.1; Applied Biosystems, Foster City, CA, USA).

2.4. Western blotting

Fibroblasts (~3.0×10^{5} cells) were harvested and lysed with 300 μL of SDS lysis buffer (100 mM Tris–HCl (pH 6.8), 2% SDS, 100 mM DTT, 20% sucrose). After boiling for 5 min, aliquots were

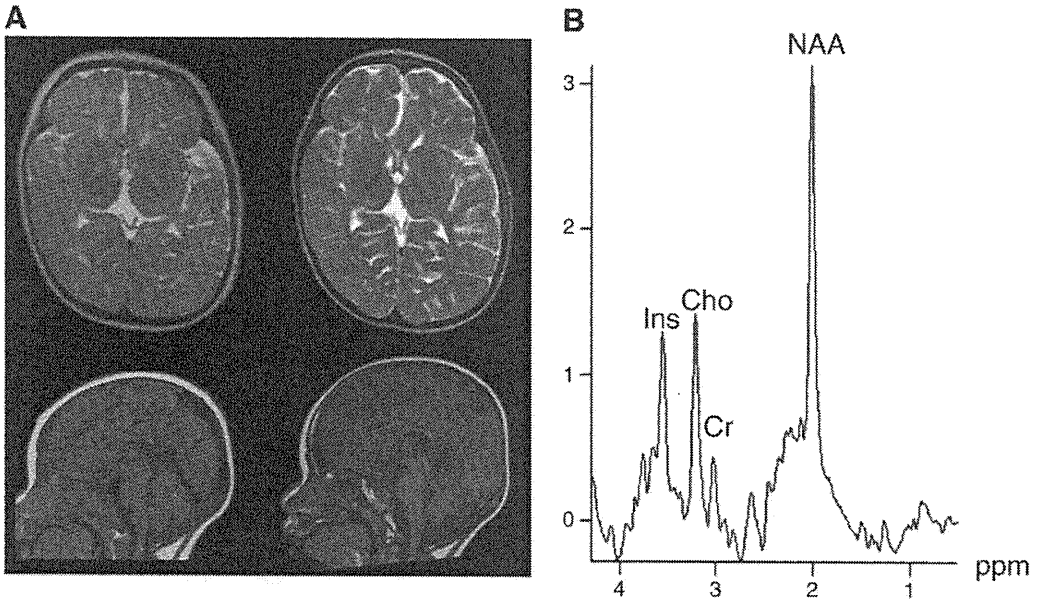

Fig. 1. MRI and MRS. T2-weighted axial image at 8 months shows an abnormal increased signal in the globus pallidi, myelination delay, and a decreased volume of cerebral white matter (A; upper left). T1-weighted sagittal image demonstrating a very thin corpus callosum (A; lower left). T2-weighted axial image (A; upper right) and T1-weighted sagittal image (A; lower right) at 4 years show persistent high signals in the globus pallidi, thin corpus callosum, and progression of cerebellar vermian atrophy. ^{1}H-MRS from the left basal ganglia at 4 years (B). MRS at 4 years shows a very low peak of Cr at 3.0 ppm (arrowhead).

analyzed by western blotting using the XV PANTERA Gel System (DRC, Tokyo, Japan). Detection was performed using the ECL-Advance System (GE Healthcare, Buckinghamshire, UK). The primary antibody was rat monoclonal anti-BAP31 antibody (1:1000, sc-56007; Santa Cruz Biotechnology, Santa Cruz, CA, USA). The secondary antibody was goat anti-rat IgG-HRP conjugate (1:2000, sc-2006; Santa Cruz Biotechnology, Santa Cruz, CA, USA).

3. Results

3.1. Cr uptake in fibroblasts

The uptake of Cr in fibroblasts from the patient was almost undetectable (0.19 pmol Cr/μg protein) when incubated at a physiological Cr concentration (25 μmol/L) (controls: n = 13, 27.8 ± 5.6 pmol Cr/μg protein; Cr transporter deficiency patients: n = 12, 0.58 ± 1.03 pmol Cr/μg protein) [12].

3.2. PCR, RT-PCR analysis, and sequencing

As MRI and the Cr uptake test strongly suggested an SLC6A8 deficiency, we analyzed *SLC6A8*. We could not obtain PCR products from exons 5–13 of *SLC6A8* in the patient. As we could amplify DNA from various exons of *ABCD1*, we performed a long-range PCR using various primers that were spaced between *SLC6A8* and *ABCD1*. We obtained an ~500 bp PCR product using primers at intron 4 of *SLC6A8* and the other one situated ~20 kb-telomeric. Sequencing this fragment revealed breakpoints at intron 4 of *SLC6A8* and intron 4 of *BAP31*, which are situated tail-to-tail (Figs. 2A, B). The deletion encompasses ~19 kb and involves exons 5–13 of *SLC6A8* and exons 5–8 of *BAP31* (Figs. 2B, C). Next, we analyzed the effects of this deletion on the transcript level of both genes. RT-PCR between exon 2 and 3′-UTR of *SLC6A8* revealed an aberrant *SLC6A8* transcript in which intron 4 was incorporated, resulting in a truncated protein of 391 amino acids (a.a.), much shorter than the wild-type protein (635 a.a.) (Fig. 3A). Similarly, we detected *BAP31* mRNA in which intron 4 of *BAP31* remained in the transcript. The deduced sequence predicted a 114-a.a. protein that is much shorter than wild-type *BAP31* (isoform a, 313 a.a.; isoform b, 246 a.a.)(Fig. 3B). Primers flanking the deletion successfully amplified a 507-bp product only from the patient among family members, indicating that the deletion occurred de novo (Fig. 3C).

3.3. Western blotting

We detected BAP31 protein of ~28 kDa in normal control fibroblasts. However, we could not detect a signal for BAP31 in fibroblasts from the patient, even at ~12 kDa, deduced from the possible mRNA predicted size, indicating either nonsense-mediated decay of the aberrant mRNA or the short half-life of truncated BAP31 (Fig. 3D).

4. Discussion

The reported phenotype of Cr transporter deficiency consists of intellectual disability, language delay, seizures, and autistic behavior [13]. Although >40 mutations of *SLC6A8* have been reported, a clear geno-phenotypic correlation has not been established. Two cases, one case with a large genomic deletion of exons 8–13 of *SLC6A8* and another with a deletion of the complete coding region of *SLC6A8*, showed severe developmental delay, seizures, failure to thrive, and dystonia [15]. Cellular Cr uptake was almost completely lost in these two patients. Our case is the third patient to be described with a large *SLC6A8* deletion, and also showed similar severe developmental delay, failure to thrive, and dystonia. A large genomic deletion of exons 8–13 of *SLC6A8* is suggestive for a loss-of-function of SLC6A8 and, indeed, Cr uptake was disrupted in our patient. The symptoms of these three cases are more severe than those of patients with single nucleotide mutations, and it appears that the complete loss of

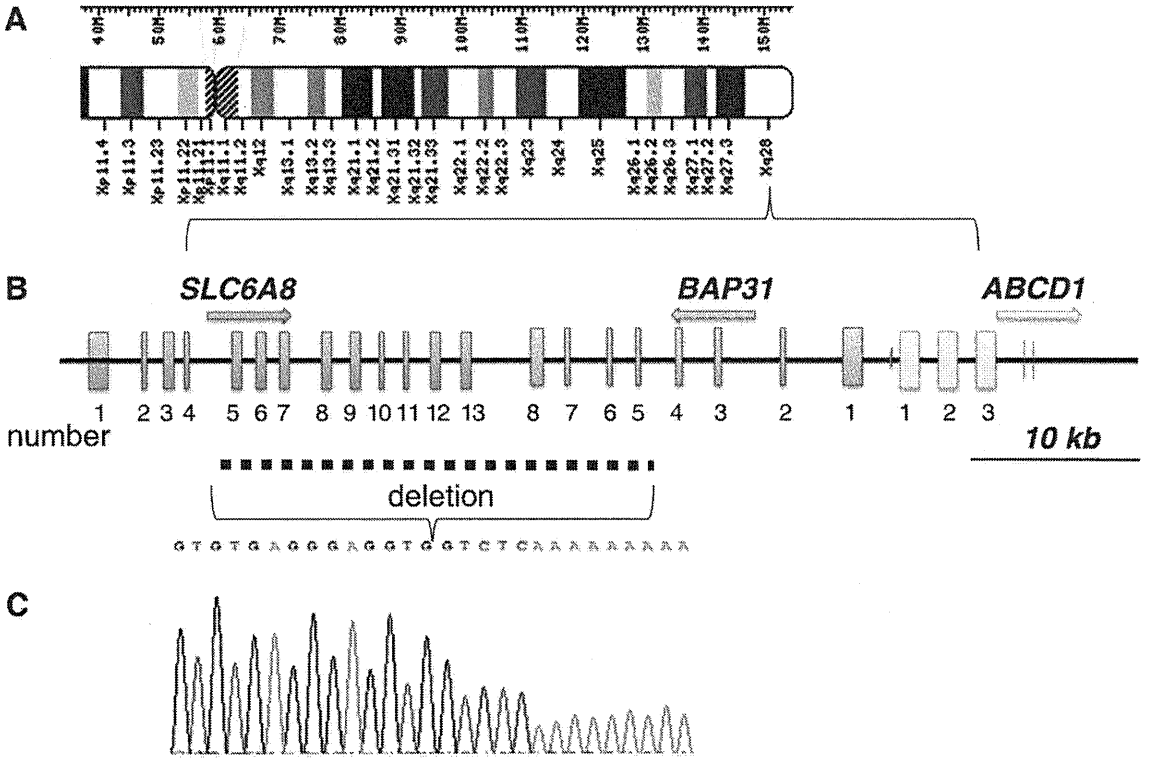

Fig. 2. 19-kb deletion involving *SLC6A8* and *BAP31*.(A) Position of the *SLC6A8* and *BAP31* genes on Chr. Xq28.(B) Position of *SLC6A8*, *BAP31*, and *ABCD1*. *BAP31* and *SLC6A8* are tail-to-tail orientation.(C) Sequence chromatogram around the conjugation point of the two breakpoints. Intron 4 of *SLC6A8* and intron 4 of *BAP31* are joined.

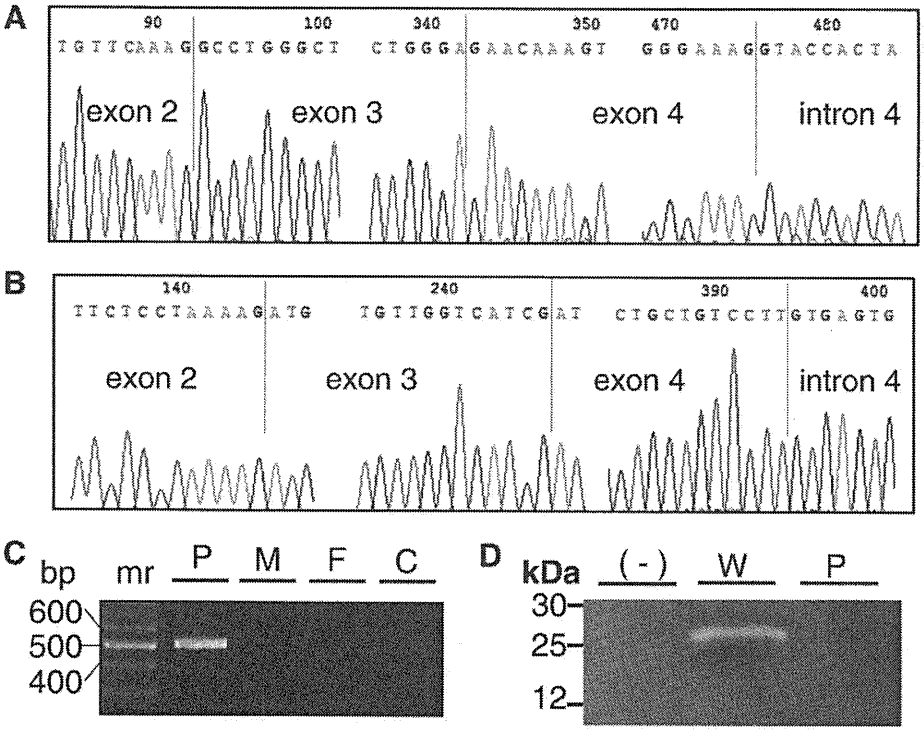

Fig. 3. RT-PCR and sequencing of *SLC6A8* and *BAP31*. RT-PCR with primers at exon 2 and 3′-UTR of *SLC6A8* indicates the transcript involves intron 4 (A). RT-PCR between exon 2 and intron 4 of *BAP31* revealed the transcript contains intron 4 (B). Breakpoint-specific PCR analysis of the patient's family (C). Primers flanking the deletion successfully amplified a 507-bp product from the patient, indicating the deletion occurred *de novo*. P, patient; M, mother; F, father; C, normal control; mr, markers. Western blotting from fibroblasts from wild-type (W) and patient (P) (D). The primary antibody was rat monoclonal anti-BAP31 (1:1000) and the secondary antibody was goat anti-rat IgG-HRP conjugate (1:2000, sc-2006). The band of ~28 kDa shown in the wild-type is unseen in the patient.

transporter function leads to severe neurological dysfunction. As for the two previously reported cases, the breakpoints have not been elucidated and therefore, we cannot exclude the possibility that the severe phenotype of these three cases is influenced by BAP31.

We discovered a large ~19-kb deletion that encompassed exons 5–13 of *SLC6A8* and exons 5–8 of *BAP31*. As far as we know, this report is the first case presentation in which *SLC6A8* and *BAP31* are both mutated. Phenotype of the *BAP31* mutations has been only reported as a part of Xq28 deletion syndrome or (CADDS) [MIM ID #300475]. Since deletion causes complete loss of *ABCD1*, these patients develop cerebral demyelination at more early age compared to classical ALD. Moreover, patients with CADDS also showed liver disease and sensorineural deafness. These two clinical symptoms have been suggested to originate from the loss of BAP31 function. However, other than these 3 cases with CADDS, mutations of *BAP31* have not been reported. Our case, carrying a deletion of *SLC6A8* and *BAP31*, also displayed an overlapping phenotype with CADDS, i.e., liver disease and sensorineural deafness. Our case strengthens the idea that the loss of BAP31 is related to liver disease and sensorineural deafness. Sequencing of *BAP31* in patients with X-linked sensorineural deafness may help to clarify this further.

Recently, BAP31 was shown to be a component of the quality control system of the endoplasmic reticulum (ER). By receiving apoptotic signals from mitochondria, BAP31 is involved in the activation of procaspase-8. Thus, BAP31 is emerging as critical component of ER stress signaling between mitochondria and the ER [19]. Our present case had increased levels of AST/ALT during an infection that induced ER stress; therefore, it is possible that dysfunction of mitochondrial-ER signaling is involved in the recurrent liver dysfunction observed during infection in our case. Wolfram syndrome-1 is a severe autosomal recessive neurodegerative disease characterized by diabetes mellitus, optic atrophy, diabetes insipidus, and deafness (WFS1; OMIM #222300). Recently, WFS1 was shown to negatively regulate a key transcription factor involved in ER stress signaling, activating transcription factor 6α (ATF6α), through the ubiquitin-proteasome pathway[20]. It is tempting to speculate the defect in ER stress signaling accounts for association of BAP31 to deafness and liver disease.

Supplementary materials related to this article can be found online at doi:10.1016/j.ymgme.2012.02.018.

Conflict of interest statement

We have no conflict of interest to disclose.

Acknowledgments

This work was supported by research grants from the Ministry of Health, Labour and Welfare (H.O., T.W., and N. M), the Japan Science and Technology Agency (N. M), the Strategic Research Program for Brain Sciences (N. M), a Grant-in-Aid for Scientific Research on Innovative Areas-(Foundation of Synapse and Neurocircuit Pathology)-from the Ministry of Education, Culture, Sports, Science and Technology of Japan (N. M), and a Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science (H.O., N. M).

References

[1] S. Stockler, P.W. Schutz, G.S. Salomons, Cerebral creatine deficiency syndromes: clinical aspects, treatment and pathophysiology, Subcell. Biochem. 46 (2007) 149–166.
[2] N. Longo, O. Ardon, R. Vanzo, E. Schwartz, M. Pasquali, Disorders of creatine transport and metabolism, Am. J. Med. Genet. C Semin. Med. Genet. 157 (2011) 72–78.
[3] P. Gregor, S.R. Nash, M.G. Caron, M.F. Seldin, S.T. Warren, Assignment of the creatine transporter gene (SLC6A8) to human chromosome Xq28 telomeric to G6PD, Genomics 25 (1995) 332–333.
[4] E.H. Rosenberg, L.S. Almeida, T. Kleefstra, R.S. deGrauw, H.G. Yntema, N. Bahi, C. Moraine, H.H. Ropers, J.P. Fryns, T.J. deGrauw, C. Jakobs, G.S. Salomons, High prevalence of SLC6A8 deficiency in X-linked mental retardation, Am. J. Hum. Genet. 75 (2004) 97–105.

[5] A. Newmeyer, K.M. Cecil, M. Schapiro, J.F. Clark, T.J. Degrauw, Incidence of brain creatine transporter deficiency in males with developmental delay referred for brain magnetic resonance imaging, J. Dev. Behav. Pediatr. 26 (2005) 276–282.
[6] A.J. Clark, E.H. Rosenberg, L.S. Almeida, T.C. Wood, C. Jakobs, R.E. Stevenson, C.E. Schwartz, G.S. Salomons, X-linked creatine transporter (SLC6A8) mutations in about 1% of males with mental retardation of unknown etiology, Hum. Genet. 119 (2006) 604–610.
[7] L. Lion-Francois, D. Cheillan, G. Pitelet, C. Acquaviva-Bourdain, G. Bussy, F. Cotton, L. Guibaud, D. Gerard, C. Rivier, C. Vianey-Saban, C. Jakobs, G.S. Salomons, V. des Portes, High frequency of creatine deficiency syndromes in patients with unexplained mental retardation, Neurology 67 (2006) 1713–1714.
[8] A. Arias, M. Corbella, C. Fons, A. Sempere, J. Garcia-Villoria, A. Ormazabal, P. Poo, M. Pineda, M.A. Vilaseca, J. Campistol, P. Briones, T. Pampols, G.S. Salomons, A. Ribes, R. Artuch, Creatine transporter deficiency: prevalence among patients with mental retardation and pitfalls in metabolite screening, Clin. Biochem. 40 (2007) 1328–1331.
[9] O.T. Betsalel, J.M. van de Kamp, C. Martinez-Munoz, E.H. Rosenberg, A.P. de Brouwer, P.J. Pouwels, M.S. van der Knaap, G.M. Mancini, C. Jakobs, B.C. Hamel, G.S. Salomons, Detection of low-level somatic and germline mosaicism by denaturing high-performance liquid chromatography in a EURO-MRX family with SLC6A8 deficiency, Neurogenetics 9 (2008) 183–190.
[10] H. Puusepp, K. Kall, G.S. Salomons, I. Talvik, M. Mannamaa, R. Rein, C. Jakobs, K. Ounap, The screening of SLC6A8 deficiency among Estonian families with X-linked mental retardation, J. Inherit. Metab. Dis. (2009).
[11] O. Ardon, C. Amat di San Filippo, G.S. Salomons, N. Longo, Creatine transporter deficiency in two half-brothers, Am. J. Med. Genet. A 152A (2010) 1979–1983.
[12] E.H. Rosenberg, C. Martinez Munoz, O.T. Betsalel, S.J. van Dooren, M. Fernandez, C. Jakobs, T.J. deGrauw, T. Kleefstra, C.E. Schwartz, G.S. Salomons, Functional characterization of missense variants in the creatine transporter gene (SLC6A8): improved diagnostic application, Hum. Mutat. 28 (2007) 890–896.
[13] G.S. Salomons, S.J. van Dooren, N.M. Verhoeven, D. Marsden, C. Schwartz, K.M. Cecil, T.J. DeGrauw, C. Jakobs, X-linked creatine transporter defect: an overview, J. Inherit. Metab. Dis. 26 (2003) 309–318.
[14] O.T. Betsalel, E.H. Rosenberg, L.S. Almeida, T. Kleefstra, C.E. Schwartz, V. Valayannopoulos, O. Abdul-Rahman, N. Poplawski, L. Vilarinho, P. Wolf, J.T. den Dunnen, C. Jakobs, G.S. Salomons, Characterization of novel SLC6A8 variants with the use of splice-site analysis tools and implementation of a newly developed LOVD database, Eur. J. Hum. Genet. 19 (2011) 56–63.
[15] I.A. Anselm, F.S. Alkuraya, G.S. Salomons, C. Jakobs, A.B. Fulton, M. Mazumdar, M. Rivkin, R. Frye, T.Y. Poussaint, D. Marsden, X-linked creatine transporter defect: a report on two unrelated boys with a severe clinical phenotype, J. Inherit. Metab. Dis. 29 (2006) 214–219.
[16] D. Corzo, W. Gibson, K. Johnson, G. Mitchell, G. LePage, G.F. Cox, R. Casey, C. Zeiss, H. Tyson, G.R. Cutting, G.V. Raymond, K.D. Smith, P.A. Watkins, A.B. Moser, H.W. Moser, S.J. Steinberg, Contiguous deletion of the X-linked adrenoleukodystrophy gene (ABCD1) and DXS1357E: a novel neonatal phenotype similar to peroxisomal biogenesis disorders, Am. J. Hum. Genet. 70 (2002) 1520–1531.
[17] C. Valongo, M.L. Cardoso, P. Domingues, L. Almeida, N. Verhoeven, G. Salomons, C. Jakobs, L. Vilarinho, Age related reference values for urine creatine and guanidinoacetic acid concentration in children and adolescents by gas chromatography-mass spectrometry, Clin. Chim. Acta 348 (2004) 155–161.
[18] G.S. Salomons, S.J. van Dooren, N.M. Verhoeven, K.M. Cecil, W.S. Ball, T.J. Degrauw, C. Jakobs, X-linked creatine-transporter gene (SLC6A8) defect: a new creatine-deficiency syndrome, Am. J. Hum. Genet. 68 (2001) 1497–1500.
[19] R. Iwasawa, A.L. Mahul-Mellier, C. Datler, E. Pazarentzos, S. Grimm, Fis1 and Bap31 bridge the mitochondria-ER interface to establish a platform for apoptosis induction, EMBO J. 30 (2011) 556–568.
[20] S.G. Fonseca, S. Ishigaki, C.M. Oslowski, S. Lu, K.L. Lipson, R. Ghosh, E. Hayashi, H. Ishihara, Y. Oka, M.A. Permutt, F. Urano, Wolfram syndrome 1 gene negatively regulates ER stress signaling in rodent and human cells, J. Clin. Invest. 120 (2010) 744–755.

Journal of Human Genetics (2012) 57, 73–77

www.nature.com/jhg

SHORT COMMUNICATION

Concomitant microduplications of *MECP2* and *ATRX* in male patients with severe mental retardation

Shozo Honda[1], Shigeko Satomura[2], Shin Hayashi[1,3], Issei Imoto[1,4], Eiji Nakagawa[5,6], Yu-ichi Goto[5,6] and Johji Inazawa[1,3,7], and the Japanese Mental Retardation Consortium[8]

Investigations of chromosomal rearrangements in patients with mental retardation (MR) are particularly informative in the search for genes involved in MR. Here we report a family with concomitant duplications of methyl CpG binding protein 2 (*MECP2*) at Xq28 and *ATRX* (the causative gene for X-linked alpha thalassemia/mental retardation) at Xq21.1 detected by array-comparative genomic hybridization. The alterations were observed in a 25-year-old man who inherited them from his mother, who showed a normal phenotype and completely skewed X-chromosome inactivation, and also in his cousin, a 32-year-old man. The proband and his cousin showed severe MR, muscular hypotonia, recurrent respiratory infections and various other features characteristic of *MECP2* duplication syndrome. However, the proband also had cerebellar atrophy never reported before in *MECP2* duplication syndrome, suggesting that his phenotypes were modified through the *ATRX* duplication in an additive or epistatic manner.
Journal of Human Genetics (2012) **57**, 73–77; doi:10.1038/jhg.2011.131; published online 1 December 2011

Keywords: array CGH; ATRX; duplication; MECP2; X-linked mental retardation

Duplication at Xq28 involving methyl CpG binding protein 2 (*MECP2*) has been detected at high frequency (1–2%) in males with unexplained X-linked mental retardation (XLMR).[1,2] *MECP2* duplication syndrome is now recognized as a clinical entity showing severe MR, muscular hypotonia, absence of speech, a history of recurrent infection and mild dysmorphic features.[3] In the course of a program to screen possible patients with XLMR for copy-number aberrations by array-comparative genomic hybridization (aCGH) using a bacterial artificial chromosome (BAC)-based X-tiling array (MCG X-tiling array),[2,4] we detected an ~0.4-Mb duplication at Xq28 involving *MECP2* together with an ~0.3-Mb duplication at Xq21.1 that included *ATRX*, the causative gene for ATR-X (X-linked alpha thalassemia/mental retardation) syndrome, in a 25-year-old man and his cousin, a 32-year-old man (Figure 1a).

The proband (III-1, Figure 1b) was born at 41 weeks after an uneventful pregnancy as the first child of non-consanguineous healthy parents. At birth, his weight and occipital–frontal circumference (OFC) were 3280 g (±0 s.d.) and 33.5 cm (0.3 s.d.), respectively. He was developmentally retarded: first smiling at 3 months, holding up his head at 5 months, rolling over at 7 months, sitting by himself at 12 months and crawling at 13 months. At 25 years, his height, weight and OFC were 160.8 cm (−1.7 s.d.), 50 kg (−1.2 s.d.) and 56.3 cm (−0.9 s.d.), respectively. The proband exhibited hypertelorism, microcephaly and synophrys (Figure 1c). At 28 years, magnetic resonance imaging (MRI) showed cerebral atrophy, cerebellar atrophy and a thin corpus callosum (Figure 1d). He could walk and communicate until he was 14 years old, but became unable to do either of this after developing epilepsy. At the age of 4 years and 10 months, his total Developmental Quotient was 22, calculated by using the Kyoto Scale of Psychological Development. A blood investigation showed that his IgA level was low. The HbH inclusion body that is detected frequently in patients with *ATRX* mutation was not found by brilliant cresyl staining. His younger brother (III-2) had intrapartum asphyxia and two maternal uncles (II-3, II-4) died immediately after birth.

The cousin of the proband (III-3) was born in 41 weeks after an uneventful pregnancy to non-consanguineous healthy parents by normal delivery. At birth, his weight and OFC were 2850 g (−1.2 s.d.) and 37 cm (+2.4 s.d.), respectively. He was characterized by macrocephaly. He had started smiling at 2–3 months, holding up his head at 4 months, sitting by himself at 12 months and walking at 40 months. At 32 years of age, his height, weight and OFC were in the normal range (164.5 cm, −1.1 s.d.; 57 kg, −0.5 s.d.; 59.4 cm, +1.8 s.d.). Information on his Developmental Quotient was unavailable. A blood investigation showed that his IgA level was low. He had been affected

[1]Department of Molecular Cytogenetics, Tokyo Medical and Dental University, Tokyo, Japan; [2]Department of Pediatrics, Japan Red Cross Tokushima Hinomine Rehabilitation Center for Children with Severe Mental and Physical Disabilities, Tokushima, Japan; [3]Hard Tissue Genome Research Center, Tokyo Medical and Dental University, Tokyo, Japan; [4]Department of Human Genetics and Public Health, The University of Tokushima Graduate School, Tokushima, Japan; [5]Division of Child Neurology, National Center Hospital of Neurology and Psychiatry, Tokyo, Japan; [6]Department of Mental Retardation and Birth Defect Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo, Japan and [7]Global Center of Excellence (GCOE) Program for 'International Research Center for Molecular Science in Tooth and Bone Diseases', Tokyo Medical and Dental University, Tokyo, Japan
[8]See the Appendix for details regarding the members of the Consortium.
Correspondence: Professor J Inazawa, Department of Molecular Cytogenetics, Medical Research Institute, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan.
E-mail: johinaz.cgen@mri.tmd.ac.jp
Received 6 May 2011; revised 18 October 2011; accepted 25 October 2011; published online 1 December 2011

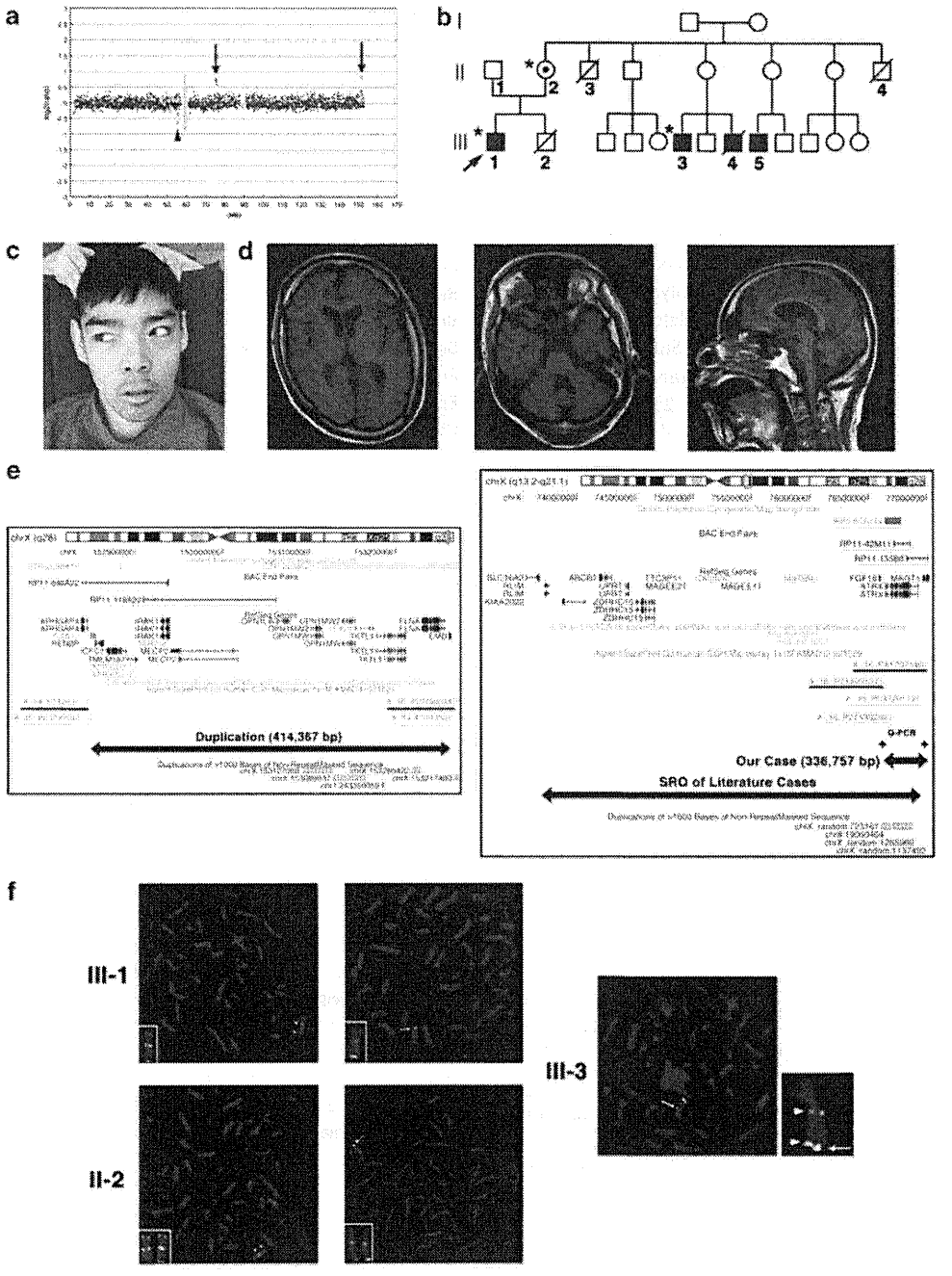

Figure 1 (**a**) Profile of the copy-number ratio on chromosome X in the proband (III-1) detected with aCGH using an MCG X-tiling array. Each dot represents the test/reference value after normalization and $\log_2$ transformation in each BAC clone, and arrows indicate duplications (ratio > 0.4). The gray vertical lines represent the centromeric region for which no clones were available. Arrowheads indicate benign CNVs (Supplementary Table 1). (**b**) Three-generation genealogy of the studied family. Closed squares and circles and dotted circles indicate MR and carriers, respectively. The proband (III-1) indicated by an arrow was used for aCGH with the X-tiling array. Asterisks indicate persons having the duplications at Xq21.1 and Xq28. A slash indicates a death. (**c**) The proband (III-1) at 27 years showed hypertelorism, microcephaly and synophrys. (**d**) Brain MRI findings of the proband (III-1) at age 27. Coronal (left and middle) and (right) sagittal T1w sequences show cerebral atrophy, cerebellar atrophy and a thin corpus callosum. (**e**) Mapping of the duplications at Xq28 (left) and Xq21.1 (right) on the basis of the UCSC Genome Browser according to NCBI Build 36.1, March 2006, hg18 (http://www.genome.ucsc.edu/). A chromosome ideogram is presented. The track setting in the UCSC genome browser was set up to look for 'Base Positions', 'FISH Clones', 'BAC End Pairs', 'RefSeq Genes', 'sno/miRNA', 'Agilent Array' and 'Segmental Duplications'. Underlines below the BAC clone ID and oligonucleotide probe ID show a high (green bars) and normal (black bars) ratio detected with the MCG X-tiling array and Agilent array 244 K. Plus signs in the right panel indicate the duplication at Xq21.1 confirmed by quantitative genomic PCR. Duplicated regions in our case and the smallest region of overlap in reported cases[9] in the right panel are indicated with closed arrows. (**f**) In the proband (III-1) and the carrier mother (II-2), representative results of FISH using the clone RP11-42M11 at Xq21.1 (left) and the clone RP11-119A22 at Xq28 (right) showed separate green signals (arrowheads) and strong green signals (arrows), respectively. The red signals are of clone RP11-16H4 at Xp22.12 (left) RP11-13M9 at Xq13.2 (right) as a reference. Enlarged images of chromosome X are shown in the lower left insets in each panel. Similarly, in the affected proband's maternal cousin (III-3), representative results of FISH demonstrated separate red signals of the clone RP11-42M11 at Xq21.1 and strong green signals of the clone RP11-119A22 at Xq28. An enlarged image of chromosome X is shown in the lower right panel, indicating that the duplicated sequence at Xq21.1 inserted into the duplication at Xq28 together with the original Xq21.1 (arrowheads), whereas the duplicated sequence at Xq28 was inserted in close proximity (arrow). A full color version of this figure is available at the *Journal of Human Genetics* journal online.

by pneumonia frequently since 3 months after birth. No MRI analysis had been performed. His younger brother (III-4) died because of disseminated intravascular coagulation at the age of 29, and his other cousin (III-5) shows a similar clinical manifestation to the proband.

On the basis of the results of precise mapping with an oligonucleotide array (Agilent array 244K, Palo Alto, CA, USA; data not shown), these aberrations are as follows: arr Xq21.1 (76646979–76983735)×2, arr Xq28 (152847991–153262357)×2 (Figure 1e). Although some copy-number variants (CNVs) were detected in other regions simultaneously, all of them have been registered in the Database of Genomic Variants (http://projects.tcag.ca/variation/ assembly, March 2006, Supplementary Table 1) and in part in our CNV database (MCG CGH database, http://www.cghtmd.jp/CNVdatabase). Subsequent real-time quantitative genomic PCR (qPCR) using primer sets recognizing around dup(X)(q21.1) (Supplementary Table 2) narrowed down dup(X)(q21.1) to between positions 76646868 and 76973049, including all of *ATRX* and part of *MAGT1* (Figure 1d). Fluorescence *in situ* hybridization (FISH) detected these duplications in the proband's unaffected mother (II-2) and his affected maternal male cousin (III-3) (Figures 1b and f), indicating maternally inherited duplications in these patients. In addition, the duplicated segment at Xq21.1 inserted into the duplicated region at Xq28, by contrast the segment at Xq28 was duplicated in tandem (Figure 1f). Our finding that the mother, a presumptive obligate carrier, had completely skewed X inactivation (dup(X):X=50:0) in a lymphoblastoid cell line , as shown by the androgen receptor X-inactivation assay described previously[5] and a late replication assay[6] with FISH (Supplementary Figure 1), supported our assumption that skewed X-chromosome inactivation appears to be characteristic of carriers of *MECP2* duplication such as other reported cases.[3]

The two affected men showed severe MR, muscular hypotonia, recurrent respiratory infections and various other features characteristic of *MECP2* duplication syndrome (Table 1). Moreover, they did not show short stature, hypoplastic genitalia and early life feeding issues, which were reported to be characteristic of MR in patients with duplications encompassing *ATRX* (Table 1).[9] The smallest region of overlap (SRO) of the reported *ATRX* duplication cases contains 11 genes, including *ATRX* and two miRNAs,[9] whereas the duplicated region of the present family includes only *ATRX* (Figure 1e), suggesting that genes other than *ATRX* within the SRO contribute to phenotypes observed in previously reported cases (Table 1).[9]

ATRX interacts with MECP2 *in vitro* and colocalizes at pericentromeric heterochromatin in mature neurons of the mouse brain.[10] Recently, it was reported that ATRX, MECP2 and cohesin cooperate to silence a subset of imprinted genes in the postnatal mouse brain.[11] Those experimental findings suggest that abnormally expressed ATRX with MECP2 through their simultaneous duplications may modify the phenotypes usually observed in *MECP2* duplication syndrome. Although our patients showed neither notably different nor more severe phenotypes compared with reported patients with *MECP2* duplication syndrome, the proband was found to have cerebellar atrophy by MRI (Figure 1e), which has never been reported before in *MECP2* duplication syndrome.[1,3] It is possible that these phenotypes in the proband were modified through ATRX duplication in an additive or epistatic manner.

The mutations in *ATRX* give rise to changes in the pattern of methylation of several highly repeated sequences, including the ribosomal DNA (rDNA) arrays[12] and significantly altered mRNA expression in four ATRX targets (*NME4*, *SLC7A5*, *RASA3* and *GAS8*) relative to normal controls.[13] Although a Southern blot hybridization method reported previously[12] showed no change in the pattern of methylation at rDNA arrays compared with normal controls (Figure 2a), quantitative RT-PCR revealed that the expression of *ATRX* was upregulated in the present cases. Although *SLC7A5* expression showed no previous change compared with that in the healthy control (Figure 2b) and the expression of *GAS8* was too low for quantitative RT-PCR (data not shown), the expression of *NME4* and *RASA3* was similar to that in the patients with *ATRX* mutations. The alteration to the expression may be influenced by *MECP2* duplication or additive/epistatic effect between *ATRX* and *MECP2* duplication.

Table 1 Phenotype comparisons between our cases and MECP2 duplication syndrome patients or patients with *ATRX* duplication

Phenotype	*MECP2* duplication syndrome[3,7,8]	*ATRX* duplication cases[9]	Our Cases III-1	Our Cases III-3
Mental retardation	118/119	11/11	+	+
Hypotonia	86/93	7/11	+	+
Absent speech	63/72	NA	+	+
Lack ambulation	20/71	NA	+	+
Recurrent infection	82/111	6/7	+	+
Breathing abnormalities	6/18	NA	–	+
Stereotyped hand movements	15/33	NA	+	+
Autistic features/autism	13/17	NA	+	+
Epilepsy	57/110	NA	+	+
GU abnormalities	29/67	7/10 (Hypoplastic genitalia)	–	+ (Bladder distention)
Death before 25 years	25/66	NA	–	–
Spasticity	42/71	NA	–	+
Ataxia	20/37	NA	+	+
GER	15/25	NA	–	+
Swallowing difficulty	23/45	NA	–	+
IPO or constipation	25/33	NA	–	–
IgA deficiency	4/10	NA	+	+
Short stature	NA	11/11	–	–
Early life feeding issues	NA	7/9	–	–
Failure to thrive	16/31	7/9	–	+
Broad thorax	NA	4/4	–	+
Pectus excavatum	NA	3/7	–	–
Short neck	NA	4/8	–	–
Simian crease	NA	4/5	–	–
Digital findings	22/52	6/7	–	–
Microcephaly	24/71	8/11	+	–
Hypertelorism	8/72	2/6	+	+
Epicanthal folds	4/72	6/8	–	–
Down-slanted palpebral fissures	NA	1/8	–	–
Ptosis	2/72	6/9	–	–
Flat nasal bridge	15/72	9/10	–	+
Down-turned corners of the mouth	NA	8/10	–	–
High-arched palate	3/72	4/4	–	–
Micro/retrognathia	NA	4/7	–	–
Low set ears	NA	4/10	–	–
Simple ears	NA	2/10	–	–
Cryptorchidism	2/4	9/10	+	–
Impaired social interaction	NA	5/6	+	+

Abbreviations: ATR-X, the causative gene for X-linked alpha thalassemia/mental retardation; GER, gastroesophageal reflux; GU, genito-urinary system; IPO, intestinal pseudo-obstruction; *MECP2*, methyl CpG binding protein 2; NA, not available.

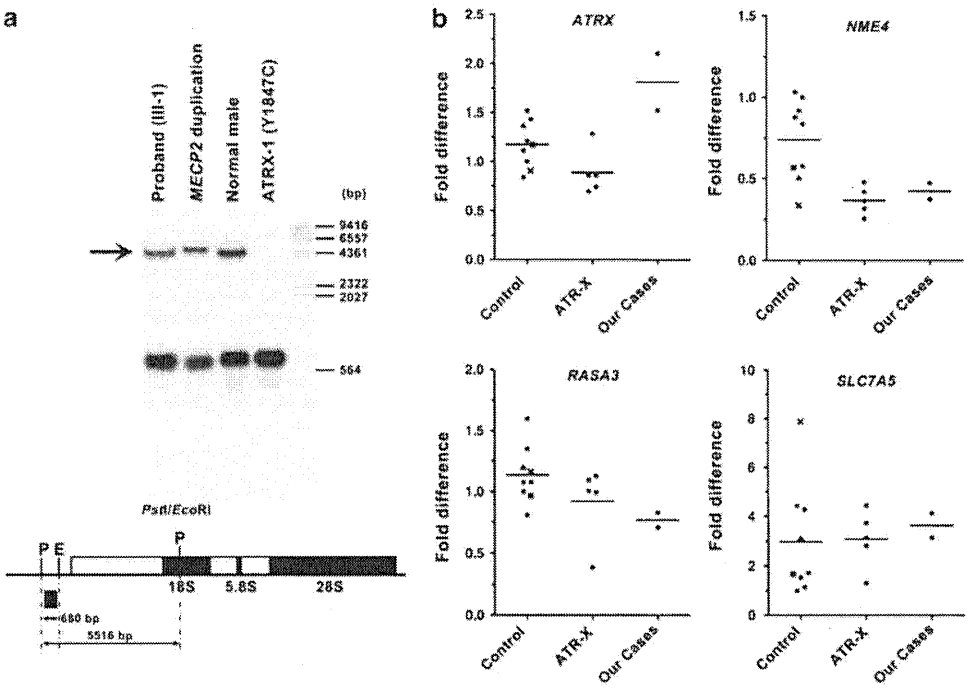

Figure 2 (**a**) Investigation of the methylation pattern of human rDNA repeats using Southern blotting. Genomic DNA from the lymphoblastoid cell line of our proband (III-1) with the *ATRX* duplication and the *MECP2* duplication, an MR male patient with *MECP2* duplication, an unrelated normal male, and an ATR-X patients with a missense mutation resulting in Y1847C in ATRX (Supplementary Table 3). DNA samples were digested with *PstI* followed by the methylation-sensitive enzyme *Eco*RI. Hybridization is shown for probes corresponding to the region between restriction sites of the two enzymes in the 3′ end of the non-transcribed spacer. The methylated, uncut band is indicated (arrow). A restriction map of part of the rDNA repeat unit shown with the 18S, 5.8S and 28S genes in order and transcribed spacer as filled and open boxes, respectively, represents the sites for *PstI* (P) and *EcoRI* (E). A black bar indicates the probe for the Southern hybridization. The size of the DNA segment resulting from the restriction enzymes is represented by closed arrows. (**b**) Real-time quantitative RT-PCR analysis of the mRNA expression of *ATRX* and three *ATRX* target genes (*NME4*, *RASA3* and *SLC7A5*) but not *GAS8*, the expression of which was too low to be estimated, in lymphoblastoid cells of our two patients, ATR-X patients whose *ATRX* mutations were identified through routine screening in a set of known XLMR genes by the Japanese Mental Retardation Consortium (unpublished data, n=5; Supplementary Table 3) and controls, including six healthy samples, the proband's parents, and a patient and a carrier with the *MECP2* duplication.[2] All the subjects provided written informed consent for the use of their phenotypic and genetic data. The proband's carrier mother, the patient and the carrier with the *MECP2* duplication are represented by a cross, triangle and square, respectively, in the control column. Data show the average values for fold differences relative to a normal male. Black bars represent mean values of each group.

The result of FISH suggests that *ATRX* duplication and *MECP2* duplication were occurred simultaneously resulting in complex genomic rearrangement. The proximal breakpoint of dup(X)(q21.1) and distal breakpoint of dup(X)(q28) were located on segmental duplications (Figure 1e) and the duplicated sequence at Xq21.1 existed near dup(X)(q28) (Figure 1f). Fork Stalling and Template Switching (FoSTeS) has been proposed as a replication-based mechanism that produces nonrecurrent rearrangements potentially facilitated by the presence of segmental duplications.[14] Previous reports suggested that complex genomic rearrangements at Xq28 such as an embedded triplicated segment and stretches of non-duplicated sequence within dup(X)(q28) were probably mediated by FoSTeS,[7,15] and a particular genomic architecture, especially low copy repeats at distal breakpoints of dup(X)(q28), may render the *MECP2* region unstable. Thus, the dup(X)(q28) and dup(X)(q21.1) detected in our patients might be generated simultaneously by FoSTeS or other mechanism in a segmental duplication-dependent manner, suggesting the structural analysis of the entire X chromosome in patients with dup(X)(q28) to be important for understanding their correct clinical condition and providing appropriate education.

ACKNOWLEDGEMENTS

This work is part of an ongoing study by the Japanese Mental Retardation Research Consortium. We thank the patients and their families for their generous participation in this study, N Murakami for cell culture and EBV-transformation, and M Kato, A Takahashi and R Mori for their technical assistance. This work is supported by grants-in-aid for Scientific Research on Priority Areas and the Global Center of Excellence Program for Frontier Research on Molecular Destruction and Reconstitution of Tooth and Bone from the Ministry of Education, Culture, Sports, Science, and Technology, Japan; by a grant from the New Energy and Industrial Technology Development Organization (NEDO); and in part by a research grant for Nervous and Mental Disorders from the Ministry of Health, Labour and Welfare, Japan. This work is also supported by Joint Usage/Research Program of Medical Research Institute, Tokyo Medical and Dental University. S.Honda is supported by a Research Fellowship of the Japan Society for the Promotion of Science (JSPS) for Young Scientists.

1 Lugtenberg, D., Kleefstra, T., Oudakker, A. R., Nillesen, W. M., Yntema, H. G., Tzschach, A. *et al.* Structural variation in Xq28: MECP2 duplications in 1% of patients with unexplained XLMR and in 2% of male patients with severe encephalopathy. *Eur. J. Hum. Genet.* **17,** 444–453 (2009).
2 Honda, S., Hayashi, S., Imoto, I., Toyama, J., Okazawa, H., Nakagawa, E. *et al.* Copy-number variations on the X chromosome in Japanese patients with mental retardation detected by array-based comparative genomic hybridization analysis. *J. Hum. Genet.* **553,** 590–599 (2010).
3 Ramocki, M. B., Tavyev, Y. J. & Peters, S. U. The MECP2 duplication syndrome. *Am. J. Med. Genet. A* **152A,** 1079–1088 (2010).
4 Hayashi, S., Honda, S., Minaguchi, M., Makita, Y., Okamoto, N., Kosaki, R. *et al.* Construction of a high-density and high-resolution human chromosome

X array for comparative genomic hybridization analysis. *J. Hum. Genet.* **52,** 397–405 (2007).
5 Kubota, T., Nonoyama, S., Tonoki, H., Masuno, M., Imaizumi, K., Kojima, M. *et al.* A new assay for the analysis of X-chromosome inactivation based on methylation-specific PCR. *Hum. Genet.* **104,** 49–55 (1999).
6 Inazawa, J., Azuma, T., Ariyama, T. & Abe, T. A simple G-banding technique adaptable for fluorescent *in situ* hyblidization (FISH) and physical ordering of human renin (REN) and cathepsin E (CTSE) genes by multi-color FISH. *Acta Histochem. Cytochem.* **26,** 319–324 (1993).
7 Bartsch, O., Gebauer, K., Lechno, S., van Esch, H., Froyen, G., Bonin, M. *et al.* Four unrelated patients with Lubs X-linked mental retardation syndrome and different Xq28 duplications. *Am. J. Med. Genet. A* **152A,** 305–312 (2009).
8 Friez, M. J., Jones, J. R., Clarkson, K., Lubs, H., Abuelo, D., Bier, J. A. *et al.* Recurrent infections, hypotonia, and mental retardation caused by duplication of MECP2 and adjacent region in Xq28. *Pediatrics* **118,** 1687–1695 (2006).
9 Lugtenberg, D., de Brouwer, A. P., Oudakker, A. R., Pfundt, R., Hamel, B. C., van Bokhoven, H. *et al.* Xq13.2q21.1 duplication encompassing the ATRX gene in a man with mental retardation, minor facial and genital anomalies, short stature and broad thorax. *Am. J. Med. Genet. A* **149A,** 760–766 (2009).
10 Nan, X., Hou, J., Maclean, A., Nasir, J., Lafuente, M. J., Shu, X. *et al.* Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. *Proc. Natl Acad. Sci. USA* **104,** 2709–2714 (2007).
11 Kernohan, K. D., Jiang, Y., Tremblay, D. C., Bonvissuto, A. C., Eubanks, J. H., Mann, M. R. *et al.* ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain. *Dev. Cell* **18,** 191–202 (2010).
12 Gibbons, R. J., McDowell, T. L., Raman, S., O'Rourke, D. M., Garrick, D., Ayyub, H. *et al.* Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. *Nat. Genet.* **24,** 368–371 (2000).
13 Law, M. J., Lower, K. M., Voon, H. P., Hughes, J. R., Garrick, D., Viprakasit, V. *et al.* ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell* **143,** 367–378 (2010).
14 Lee, J. A., Carvalho, C. M. & Lupski, J. R. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell* **131,** 1235–1247 (2007).
15 Carvalho, C. M., Zhang, F., Liu, P., Patel, A., Sahoo, T., Bacino, C. A. *et al.* Complex rearrangements in patients with duplications of MECP2 can occur by fork stalling and template switching. *Hum. Mol. Genet.* **18,** 2188–2203 (2009).

APPENDIX

Yu-ichi Goto, Department of Mental Retardation and Birth Defect Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo, Japan; Johji Inazawa, Department of Molecular Cytogenetics, Medical Research Institute and School of Biomedical Science, Tokyo Medical and Dental University, Tokyo, Japan; Mitsuhiro Kato, Department of Pediatrics, Yamagata University School of Medicine, Yamagata, Japan; Takeo Kubota, Department of Epigenetic Medicine, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Yamanashi, Japan; Kenji Kurosawa, Division of Medical Genetics, Kanagawa Children's Medical Center, Yokohama, Japan; Naomichi Matsumoto, Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan; Eiji Nakagawa, Department of Mental Retardation and Birth Defect Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo, Japan; Eiji Nanba, Division of Functional Genomics, Research Center for Bioscience and Technology, Tottori University, Yonago, Japan; Hitoshi Okazawa, Department of Neuropathology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan; Shinji Saitoh, Department of Pediatrics, Hokkaido University Graduate School of Medicine, Sapporo, Japan; and Takahito Wada, Department of Medical Genetics, Shinshu University School of Medicine, Matsumoto, Japan.

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

nature genetics

ANALYSIS

Systematic documentation and analysis of human genetic variation in hemoglobinopathies using the microattribution approach

Belinda Giardine[1,37*], Joseph Borg[2–4,37], Douglas R Higgs[5], Kenneth R Peterson[6], Sjaak Philipsen[7], Donna Maglott[8], Belinda K Singleton[9], David J Anstee[9], A Nazli Basak[10], Barnaby Clark[11], Flavia C Costa[6], Paula Faustino[12], Halyna Fedosyuk[6], Alex E Felice[3,4], Alain Francina[13], Renzo Galanello[14], Monica V E Gallivan[15], Marianthi Georgitsi[16], Richard J Gibbons[5], Piero C Giordano[17], Cornelis L Harteveld[17], James D Hoyer[18], Martin Jarvis[19], Philippe Joly[13], Emmanuel Kanavakis[20], Panagoula Kollia[21], Stephan Menzel[11], Webb Miller[1], Kamran Moradkhani[22], John Old[23], Adamantia Papachatzopoulou[24], Manoussos N Papadakis[25], Petros Papadopoulos[7], Sonja Pavlovic[26], Lucia Perseu[27], Milena Radmilovic[26], Cathy Riemer[1], Stefania Satta[14], Iris Schrijver[28], Maja Stojiljkovic[26], Swee Lay Thein[11], Jan Traeger-Synodinos[20], Ray Tully[8], Takahito Wada[29], John S Waye[30,31], Claudia Wiemann[32], Branka Zukic[26], David H K Chui[33,34], Henri Wajcman[22,35], Ross C Hardison[1,36] & George P Patrinos[16]

We developed a series of interrelated locus-specific databases to store all published and unpublished genetic variation related to hemoglobinopathies and thalassemia and implemented microattribution to encourage submission of unpublished observations of genetic variation to these public repositories. A total of 1,941 unique genetic variants in 37 genes, encoding globins and other erythroid proteins, are currently documented in these databases, with reciprocal attribution of microcitations to data contributors. Our project provides the first example of implementing microattribution to incentivise submission of all known genetic variation in a defined system. It has demonstrably increased the reporting of human variants, leading to a comprehensive online resource for systematically describing human genetic variation in the globin genes and other genes contributing to hemoglobinopathies and thalassemias. The principles established here will serve as a model for other systems and for the analysis of other common and/or complex human genetic diseases.

Since completion of the human genome project, a major aim in the field of genetics has been to determine how individual genomes differ from each other and how these differences explain variation in phenotype. However, it often remains unclear which variants cause changes in phenotype and which are phenotype neutral; furthermore, in many instances, the mechanisms by which variants cause changes in gene expression and phenotypes remain unknown. To address this, DNA sequence data will need to be matched with well-defined phenotypes to make meaningful connections between structure, function and mechanism.

A potential hurdle to this approach is how to encourage 'phenotypers' to report their observations. After the initial excitement during the 1980s and 1990s of identifying disease-causing molecular defects and the mechanisms by which they arise, enthusiasm in this area has declined such that it has become increasingly difficult to report small numbers of human variants in scientific journals. Consequently, many new variants associated with well-defined phenotypes and, equally important, variants which cause no change in phenotype remain unreported. Inevitably, a large amount of potentially valuable information remains inaccessible.

To overcome this problem, we implemented a process for capturing such information with the incentive of microattribution, whereby the contribution of those individuals collecting new detailed genotype and phenotype data is positively encouraged and appropriately acknowledged[1]. We have applied the microattribution approach to inherited disorders affecting either the structure of hemoglobin (such as sickle cell disease (SCD)) or the levels and balance of globin chain production (the thalassemias). We also included variants that cause hereditary persistence of fetal hemoglobin (HPFH), a condition associated with increased production of γ-globin which ameliorates the clinical endpoints of SCD and β-thalassemia. The hemoglobinopathies and thalassemias are among the commonest inherited disorders in humans. Variants of the globin-encoding genes, residing in the α-like and β-like globin gene clusters, have provided key insights into the principles underlying human molecular genetics since the discipline was established in the 1950s (ref. 2).

*A full list of author affiliations appear at the end of the paper.

Received 2 August 2010; accepted 11 February 2011; published online 20 March 2011; doi:10.1038/ng.785

Although most hemoglobinopathies are classic monogenic disorders affecting structural genes, globin gene expression is the end product of a complex regulatory network (transcriptional and epigenetic) that emerges during terminal erythroid differentiation. Consequently, globin gene expression may also be affected by *trans*-acting mutations. Examples of such mutations were initially found in families with rare syndromal disorders, of which α-thalassemia was one component (for example, ATR-X (MIM301040) and ATMDS (MIM300448) syndromes)[3,4]. Similarly, trichothiodystrophy (MIM300448) was shown to be associated with β-thalassemia due to mutations in the XPD component of the general transcription factor complex TFIIH[5]. The association of X-linked thrombocytopenia with β-thalassemia identified a mutation of the erythroid-specific transcription factor GATA-1 (ref. 6), and recently, systematic analysis of subjects with unexplained HPFH has identified mutations in the KLF1 erythroid transcription factor[7]. Finally, the implementation of genome-wide association studies searching for quantitative trait loci that influence the level of fetal hemoglobin (HbF) has revealed several important regulators of *HBG1* and *HBG2* gene expression, including the *HBS1L-MYB*[8] and *BCL11A* loci[9,10] on chromosomes 6 and 2, respectively. As genetic variations in the genes within the erythroid network are investigated in further detail, we anticipate many more discoveries of *trans*-acting mutations that may provide target pathways for manipulating globin gene expression to ameliorate the symptoms of thalassemia and SCD. Therefore it is important that an effective database be created to accommodate all of the mutations affecting the globin genes and the network regulating their expression.

Here we report the first example of implementing microattribution to systematically document genetic variation leading to human genetic disorders, using hemoglobinopathies and thalassemias as an example. Furthermore, we demonstrate that microattribution can incentivise data contribution and, importantly, show how an integrated human variant database (including the recently acquired microattribution data) can provide key insights into human genetic diseases. Microattribution provides an important mechanism and incentive for researchers to report all variants within a specific gene or disease network. Following the principles established for the globin disorders, these databases should provide a key resource for understanding the molecular pathology of human genetic diseases.

Developing the microattribution process

To ensure that all natural mutations and their associated phenotypes are accurately and efficiently recorded, we comprehensively documented genotype and phenotype information in individuals with globin disorders in a series of interrelated locus-specific databases (LSDBs). Traditionally, credit has been given to discoverers of genetic variants through citations of their publications describing the variants. However, the increased rate of discovery through re-sequencing efforts far exceeds the capacity of citations of individual publications to give adequate credit. In order to be used effectively by the community, published variants are deposited into databases such as those described here; nevertheless, many variants may still not be published. Alternatively, variants may be discovered in large-scale collaborative projects. Credit can be given to the discoverers of the variants deposited in databases through the new process of microattribution[1]. Each variant used in a paper is listed in four microattribution tables with its accession number and with unique IDs for the discoverers, or 'authors', of the variant. In this paper, we have applied 'microcitations' to hemoglobinopathy-associated variants in order to provide incentives to data producers to deposit all of their data in these public resources[1]. Depositing the microattribution tables in a central repository (for example, NCBI) provides a venue for quantitative microcitations for every unique author. Using this approach (first implemented in 2010), there has been a marked increase in the number of reported variants in the globin gene network (**Supplementary Fig. 1**).

Implementing microattribution

All genetic variation data have been collected and documented in the HbVar database of hemoglobin variants and thalassemia mutations[11] and the Leiden Open-Access Variation Database (LOVD)–based LSDBs for the other erythroid proteins[12] (**Supplementary Note**) with appropriate attribution of the data contributors. These variants are reported in publicly available microattribution tables (also provided in **Supplementary Table 1**) that have been centrally deposited in NCBI (**Supplementary Fig. 2**). Each microattribution table has different information related to submission to the central depository, microattribution, phenotype and allele frequency (**Supplementary Note**).

In this protocol, data submitters directly contribute variants leading to hemoglobinopathies to HbVar and in return obtain direct microattribution credit. These variants have been recorded with researcher IDs and in the case of previously published variants, the corresponding PubMed ID was also used (**Supplementary Fig. 2**). To date, 232 variants have been directly submitted to HbVar without being published in a peer-reviewed journal, some of which have been deposited with more than one researcher ID. Seventy-six variants were 'orphan', that is, variants for which there was neither a PubMed ID nor a researcher ID, all of which were variants initially deposited to HbVar in the year 2000 and for which either valid contact details for the variant contributors was lacking or the contributor(s) failed to respond to our invitation. These variants have been deposited with an HbVar researcher ID.

For all unpublished variants directly contributed to HbVar by the microattribution process, a very stringent evaluation of the information submitted takes place. Contributed variant data are evaluated by curators, all of whom are senior scientists with extensive editorial experience, especially in the field of hemoglobinopathies. The curators directly contact the data contributors, if needed, for clarifications related to issues pertaining to phenotypic description, method of variant identification, ethnicity of the individual with the variant, allele frequency and so on. Upon acceptance, contributed data become part of the main HbVar data collection recorded with the contributor(s) researcher ID.

Although microattribution can operate locally (within journals and databases each reporting quantitative citation of accessions), depositing the microattribution tables in a central repository of cited accessions (for example, NCBI or European Bioinformatics Institute (EBI)) allows the central registry to be mined for citations associated with unique author identities and with each author's publications and database entries. For the purpose of our project, we have chosen to deposit the microattribution tables in NCBI, and a copy of these tables is also deposited in Nature Publishing Group's central database.

Mining the databases

In the case of globin gene disorders, many variants were conventionally reported in genetics journals, and these variants identified and/or elucidated many mechanisms underlying key aspects of gene regulation in *cis* (for example, promoters, enhancers, silencers, mRNA processing signals and translational signals) and in *trans* (for example, transcription factors, chromatin remodeling factors and protein chaperones)[2]. Furthermore, these variants helped to establish the molecular mechanisms underlying human genetic

disease. Implementation of the microattribution approach has substantially added to the repository of variants, and use of this expanded database will continue to provide an important resource for generating and testing new hypotheses in the globin field. Below, we provide some recent examples illustrating the value of the microattribution approach. The value of the comprehensive globin variant database (pre- and post-microattribution) clearly emphasizes the importance of developing similar databases for other genes and disease systems for which microattribution will become the main route to publication.

The first example of the value of the microattribution approach is the finding that the distribution of promoter mutations differs among globin genes. Although a great deal has been learned about mammalian promoters from previous analyses of the globin genes, the discovery of additional variants continues to develop our knowledge of how these genes are normally activated and how they are altered in human genetic disease. Globin gene promoter mutations contributing to β-like thalassemias and HPFH comprise approximately 10% of the total variants and result in various phenotypes, from the asymptomatic non-deletional HPFH conditions to the mild forms of β- and δ-thalassemia. The *HBB* promoter region harbors several genetic variants associated with β^+ (expressing lower than normal levels of β-globin) and β^0 (expressing no β-globin) thalassemia; these variants cluster in *cis*-regulatory elements known to bind transcription factors (**Fig. 1**). Many of these variants have been published, but an increasing number of unpublished variants have been contributed to HbVar by investigators around the world. The unpublished variants provide a more complete view of the contribution of genetic variants to phenotypes. In this particular case, they reveal phenotypic consequences of variants in more positions of well-known transcription factor binding sites (the CACC box and the TATA box) and show that additional substitutions in other binding sites contribute to phenotype (for example, positions c.-80, c.-81 and c.-138). The *HBB* c.-121C>T transition is adjacent to the CCAAT box. This motif was recognized 30 years ago as a component of some promoters, but the newly reported mutation here is the first indication that genetic variation close to this motif affects *HBB* gene expression in humans.

In contrast to the promoters for *HBB* and *HBD*, variants are not found in the first 100 bp of the *HBG1* and *HBG2* promoters, but instead, variants occur in the upstream region from approximately −100 to −200 bp (**Fig. 2a**). The *HBG1* and *HBG2* gene promoters have several *cis*-regulatory elements in common with *HBB* and *HBD* promoters, such as a TATA box and a proximal CCAAT box, but no variants have been found in these elements. However, the CCAAT box is duplicated in the promoters of *HBG1* and *HBG2*, and the upstream CCAAT box (and the nucleotides very close to it) carries variants associated with HPFH. A newly discovered, unpublished variant, c.-250C>T, calls attention to a tight cluster of mutations all associated with HPFH. An HPFH-associated variant has now been reported at each nucleotide from position c.-251 to c.-248 (198 to 195 bp from the gene transcription start site), and a variant at c.-255 (202 bp from the transcription site) is associated with a similar phenotype (**Fig. 2a**). Given these phenotypes, this cluster of variants within the motif CCCTTCCC delineates a response element important for the silencing of the *HBG1* and presumably *HBG2* genes in adult erythroid cells (the same c.-250C>T mutation has been found in the promoter of *HBG2*; data not shown).

To test the hypothesis, derived from the documented variants, that this motif delineates a response element important for silencing of the *HBG1* and *HBG2* genes, we generated human β-globin locus (β-yeast artificial chromosome (β-YAC)) transgenic mice containing the *HBG1* c.-248C>G variation (the Brazilian non-deletional HPFH mutation), which directly alters the CCCTTCCC sequence at the 3′ C. Adult mutant β-YAC mice showed an HPFH phenotype with an increased number of HbF-containing cells (**Fig. 2b**), and real-time quantitative RT-PCR analyses showed an 8- to 34-fold increase of *HBG1* gene expression relative to wild-type β-YAC mice (**Fig. 2c**). By comparison, β-YAC transgenic mice bearing the Greek type of non-deletional HPFH (*HBG1* c.-170G>A)[13] showed a 56-fold increase of *HBG1* gene expression relative to wild-type β-YAC mice. Future experiments will examine the mechanism of repression at this region. Recent studies have shown that the transcription factor BCL11A acts to repress *HBG1* and *HBG2* expression in adult erythroid cells, acting with the protein SOX6 (ref. 14). Although BCL11A showed no binding to the *HBG1* and *HBG2* proximal promoters, SOX6 showed strong binding that overlapped with GATA1 binding in these regions. In this way, the database has posed a new testable hypothesis. The CCCTTCCC element, which is adjacent to a GATA binding site, may bind a currently unknown protein that acts in concert with BCL11A to repress the production of γ-globins.

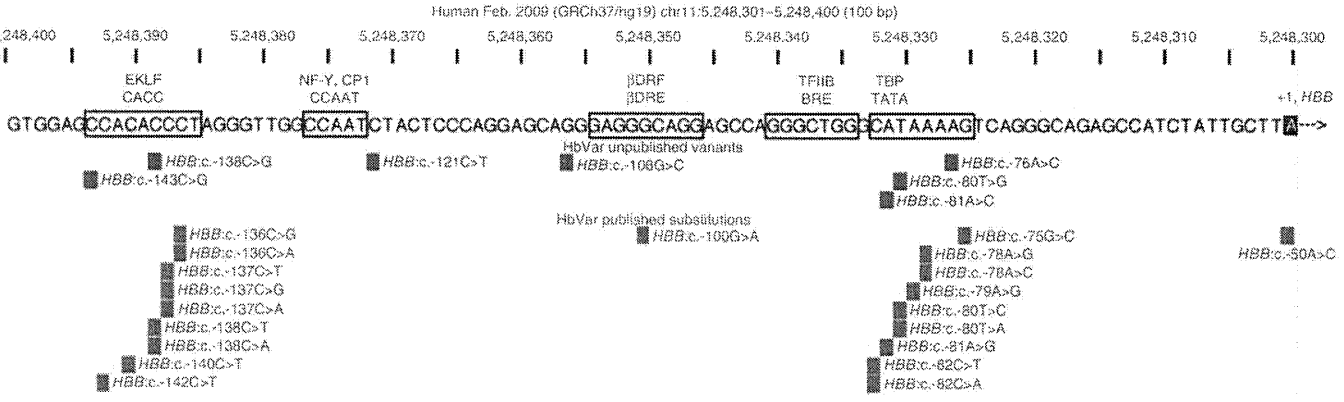

Figure 1 Graphical display of the *HBB* promoter variants recorded in HbVar, partitioned into unpublished variants contributed by investigators (blue) and published variants (purple). The genomic position, sequence change and associated phenotype (β^+ or β^0 thalassemia) are given for each variant. Known protein-binding sites in the DNA sequence are boxed, with the name of the site and the binding protein above it. The transcription start site (+1) is in reverse type. The reverse complement of the genomic sequence is shown so that the gene is in the conventional left-to-right transcriptional orientation. The image was generated by displaying the results of a query on HbVar in the Pennsylvania State University genome browser followed by editing for clarity. Variants are given using the conventional nomenclature.

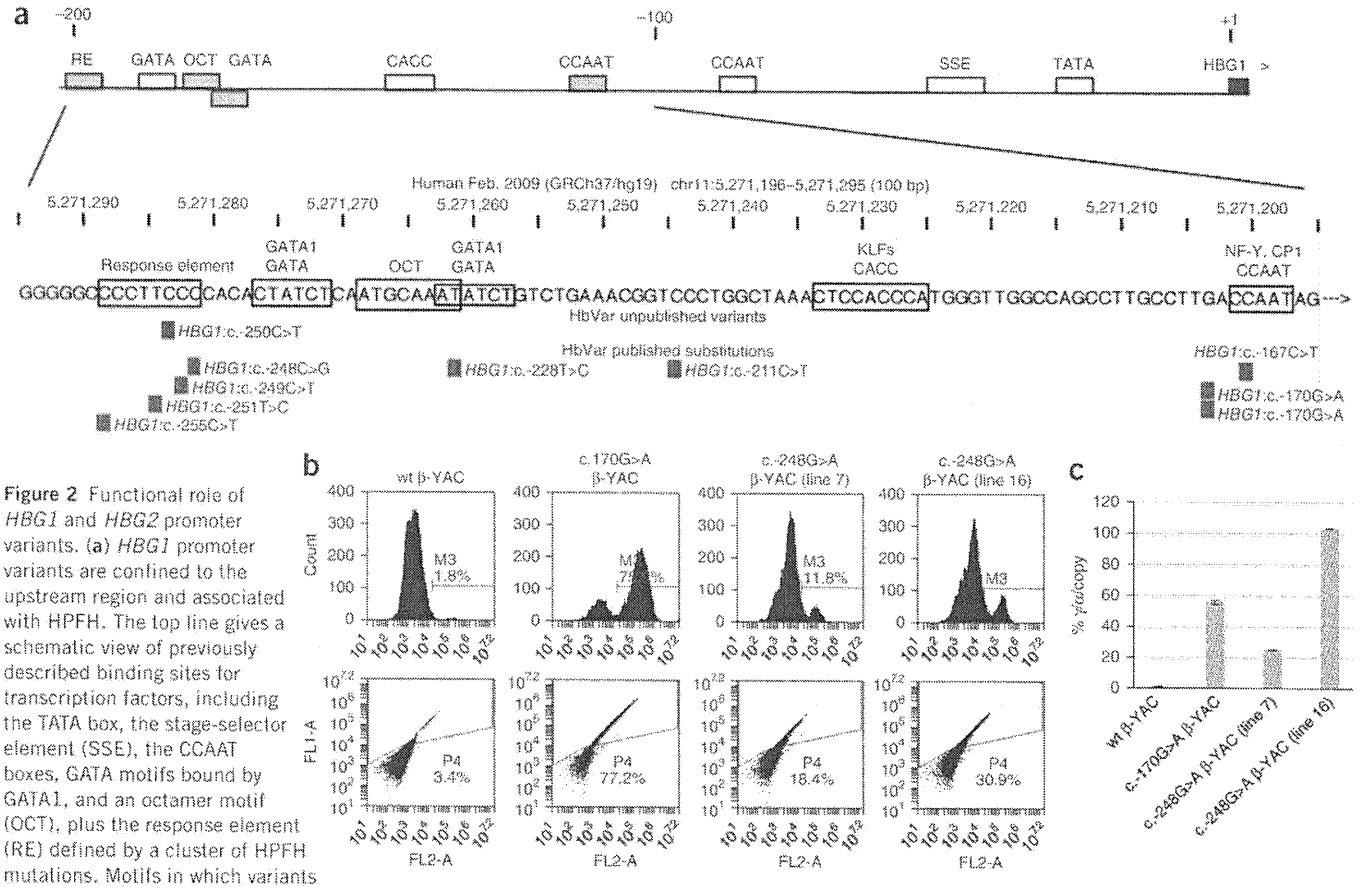

Figure 2 Functional role of *HBG1* and *HBG2* promoter variants. (**a**) *HBG1* promoter variants are confined to the upstream region and associated with HPFH. The top line gives a schematic view of previously described binding sites for transcription factors, including the TATA box, the stage-selector element (SSE), the CCAAT boxes, GATA motifs bound by GATA1, and an octamer motif (OCT), plus the response element (RE) defined by a cluster of HPFH mutations. Motifs in which variants have been found are colored gray. The transcription start site (+1) is in reverse type. The image was generated by displaying the results of a query on HbVar in the Pennsylvania State University genome browser followed by editing for clarity. Variants are given using the conventional nomenclature. (**b**) Flow cytometry analysis of γ-globin$^+$ erythrocytes from adult *HBG1* c.-248C>G HPFH β-YAC transgenic lines. A mouse monoclonal γ-globin antibody was used to determine the percentage of F cells. Line and individual numbers are indicated at the top of the panels. Percent γ-globin–positive cells are indicated within each plot (see also Online Methods). Wild-type (wt) β-YAC mice served as negative controls, and *HBG1* c.-170G>A HPFH β-YAC mice[13] were used as positive controls. In parallel experiments, human β-globin was expressed in 92–97% of the cells analyzed for all lines (data not shown). (**c**) Human γ-globin gene expression in *HBG1* c.-248C>G HPFH β-YAC transgenic lines. Percent γ-globin gene expression, copy number-corrected and normalized to per-copy mouse α-globin gene expression, is shown on the *y* axis. β-YAC construct and line numbers, where appropriate, are indicated at the bottom of the plot. Error bars represent standard deviation of triplicate experiments.

Overall, comparative analysis of the globin gene promoter mutations revealed a distinct distribution pattern for each gene. In *HBD*, promoter mutations are widely spread within the proximal promoter region and do not form mutational clusters around *cis*-regulatory elements (**Supplementary Fig. 3**). Notably, the mutations c.-81A>G and c.-80T>C have been found in the TATA boxes of *HBB* and *HBD*, suggesting that they could be the result of genetic recombination events[15].

A second example of the value of the microattribution approach was the discovery of α-thalassemia resulting from inherited or acquired mutations in *ATRX*. The comprehensive database originally identified and defined some of the key *trans*-acting factors in the globin gene system. The expanded database continues to refine our understanding of such *trans*-acting factors. Unlike the common forms of α-thalassemia resulting from *cis*-acting genetic defects, two rare forms of α-thalassemia are caused by *trans*-acting mutations in the X-linked *ATRX*. These mutations cause ATR-X syndrome, which is characterized by a severe form of syndromal mental retardation with characteristic dysmorphic faces, genital abnormalities and a mild but variable form of hemoglobin H disease[3]. In addition, acquired mutations in *ATRX* are seen in individuals who develop ATMDS syndrome, a condition in which α-thalassemia (AT) is associated with myelodysplastic syndrome (MDS)[4]. In both conditions, the levels of α-globin mRNA are reduced, suggesting that *ATRX* is involved in the normal regulation of α-globin gene expression. To date, 107 unique inherited and/or acquired disease-causing missense mutations have been found, which are located predominantly in two highly conserved domains of ATRX (**Supplementary Fig. 4**). These variants cluster within a globular domain that contains a plant homeodomain, which binds the N-terminal tails of histone H3, and the 7 helicase sub-domains, which identify ATRX as a member of the SNF2 family of chromatin-associated proteins. Structure and function studies based on natural mutations in the comprehensive database have elucidated precisely how ATRX is recruited to some of its targets through an interaction with the N-terminal tails of histone H3.

Notably, the degree of α-thalassemia seen in individuals with ATMDS (having acquired *ATRX* gene mutations) is much greater than in individuals with the ATR-X syndrome (having inherited *ATRX* gene mutations), even when, by comparing mutations on the comprehensive database, we can see that the same *ATRX* mutation occurs in both conditions[16]. Again, analysis of the comprehensive variant database poses a new testable hypothesis. These findings suggest that

another component of the ATRX pathway may frequently be mutated in individuals with the common forms of MDS.

A third example of the value of microattribution is the discovery of variants in KLF leading to elevated HbF levels. *KLF1* encodes a key erythroid transcriptional regulator that has many target genes with essential functions in erythroid cells including the globins, membrane proteins and heme synthesis enzymes[17]. The first report on *KLF1* mutations in humans linked them to the rare blood group In(Lu) phenotype[18], in which the expression of the Lutheran blood group antigens is diminished. The reported individuals carried eight different loss-of-function mutations and one mutation abolishing a GATA1 binding site in the *KLF1* promoter. In all cases, the mutant *KLF1* allele occurred in the presence of a normal *KLF1* allele. A subsequent study on a large Maltese pedigree demonstrated that haploinsufficiency for KLF1 causes HPFH[7]. A mutation in *KLF1*, resulting in p.Lys288X, was present exclusively in all individuals in this family with HPFH. This mutation ablates the complete zinc finger domain and therefore abrogates DNA binding of the mutant KLF1 protein (**Fig. 3** and **Supplementary Table 2**). The occurrence of HPFH in the individuals with In(Lu) has not been investigated. An analysis of archived blood samples from a number of these individuals with In(Lu) showed that their HbF levels were raised compared to those observed in control samples. Also, 30 out of 31 Sardinian individuals bearing four different *KLF1* mutations showed raised HbF levels compared to control samples. In addition, two individuals suffering from dyserythropoietic anemia carried a KLF1 p.Glu325Lys alteration and had an HbF level of 40% (**Fig. 3** and **Supplementary Table 2**)[19,20]. Mutations at this position alter the DNA binding specificity of KLF1. We note that the mouse neonatal anemia mutant (Nan) has an alteration in the orthologous amino acid of Klf1, p.Glu339Asp[21,22]. Adult heterozygous Nan animals show increased expression of embryonic globins, a condition akin to HPFH. Collectively, these data support the link between KLF1 and HPFH and highlight the importance of the second DNA-binding zinc finger for normal KLF1 function. This raises the possibility that some of the *KLF1* mutations which result in altered DNA binding specificity may have increased impact on HbF levels. This hypothesis can now be experimentally tested *in vitro* by DNA binding assays and *in vivo* in animal models.

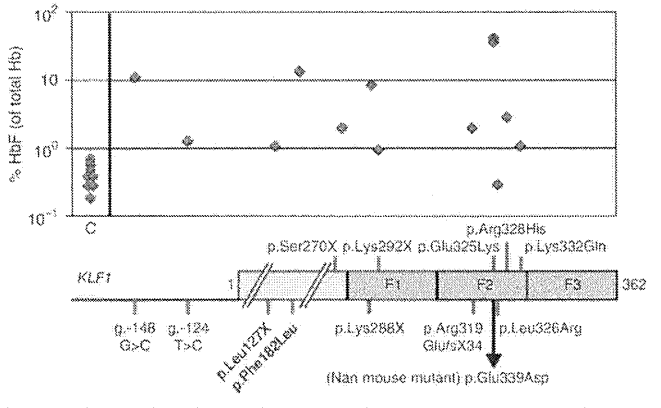

Figure 3 Correlation of the different *KLF1* gene variants deposited into HbVar (shown as blue and red squares, depicting unpublished and published information, respectively) and their corresponding HbF levels (median value in cases of three or more individuals) compared to wild-type individuals (shown as green squares). *KLF1* is not shown to scale. A simplified diagram depicting the *KLF1* promoter and protein is shown underneath. The positions of the zinc fingers are indicated (F1, F2 and F3). For the exact HbF levels corresponding to each *KLF1* gene variant, see **Supplementary Table 2**.

A final example of the value of microattribution is the discovery of hemoglobin variants. A large proportion of genetic variation in the human globin genes leads to hemoglobin variants. Most hemoglobin variants are rare, result from single amino acid substitutions of a globin chain and have a negligible or even no effect on hemoglobin function[2].

The documented hemoglobin variants reside solely within exons and include: (i) structural variants with a pleiotropic effect (for example, HbS (*HBB* c.20A>T), HbE (*HBB* c.79G>A) and HbC (*HBB* c.19G>A); (ii) variants (138 different variants) leading to unstable hemoglobin, where mutations affect the heme pocket of the globin chain; (iii) variants leading to methemoglobinemia, where the ferrous ion (Fe^{2+}) of the heme group is oxidized to the ferric state (Fe^{3+}) (most of these variants involve replacement by tyrosine of the histidine residues that anchor heme); and (iv) variants (92 different variants) with altered oxygen affinity, most of which result in increased oxygen affinity.

Although all of these correlations between structure and function have depended on data from the comprehensive database, new insights and questions continue to arise as new mutants are added to the repository, an initiative that sparked the implementation of the microattribution process for hemoglobinopathies. Notably, 14 hemoglobin variants result from the same mutation, but this mutation occurs on a different α-globin gene paralogue[23], that is, variations involving related genes that have evolved from recent gene duplication and as such are subject to frequent gene conversion events (**Supplementary Fig. 5**). HbF-Sardinia and HbF-Lesvos provide another such example, involving the same mutation (c.227T>C) but on the paralogous *HBG1* and *HBG2* genes, respectively[24].

DISCUSSION

The development of an integrated set of comprehensive LSDBs for a particular spectrum of human genetic diseases with microattribution, as described here for the hemoglobinopathies, provides an example of how such systems might be set up for a wide range of human genetic disorders in the future. Using the microattribution process set out here, datasets which took decades to accumulate for the globin genes could be assembled rapidly for other genes and disease systems. In the past, the description of natural variants has been accommodated by the conventional literature and has made an enormous contribution to the field of human genetics. In addition, it has shown how some of these mutations have reached polymorphic frequencies through natural selection, and detailed analysis of natural mutants has also been invaluable in establishing many of the general principles underlying mammalian gene regulation and human molecular genetics.

The strength of such observations will continue to increase as new mutations enter the databases, even though these might not merit a full publication on their own. Furthermore, new patterns of mutation may emerge; the accumulation of coding mutations in particular regions of a protein often identify a functionally important domain, as illustrated by *ATRX* and *KLF1* gene variants (**Supplementary Fig. 4** and **Fig. 3**, respectively), and conversely, the identification of common neutral variants may rule out a major functional role for other regions. Similarly, DNA variants of key regulatory regions (promoters, enhancers, silencers, boundary elements and locus control regions) are often critical in identifying important *cis* elements and yet other neutral variants may help map regions of little functional importance (**Fig. 1** and **Supplementary Fig. 3**). At the nucleotide level, such variants can even help map transcription factor binding sites[25]. The emergence of patterns of mutation may also point to the mechanisms of mutation, exemplified by gene conversion events identified at the

HBA1 and *HBA2*, and *HBG1* and *HBG2* genes. Additionally, subtle phenotypic differences, for example, between δβ-thalassemia and deletional HPFH[2], can be attributed to the different junction points and the sequences that are removed or juxtaposed as a result of these deletions. Systematic documentation of these deletions in HbVar is currently under way and may allow for the identification of new regulatory elements that lie within the deleted or juxtaposed regions.

Perhaps the most important aspect of such comprehensive interacting databases is that they will pose and answer questions that would otherwise not be addressed, potentially leading to useful new insights. These databases will not only be of value in establishing the phenotypes of natural variants but may also be used in the development of personalized medicine. In the globin field, a great deal of effort is directed toward the development of drugs to increase the level of HbF and thereby ameliorate the clinical severity of β-thalassemia and SCD. Potential therapeutic agents identified to date include hydroxyurea and butyrate. The response to HbF-augmenting therapies is variable in patients with β-thalassemia and SCD, with approximately 25% of these patients being poor responders or non-responders[26]. Therefore, the ability to predict a patient's response to hydroxyurea and/or other HbF-augmenting drugs would help in optimizing therapy. Polymorphisms in genes regulating HbF expression, hydroxyurea metabolism and erythroid progenitor proliferation might modulate a patient's response to HbF-inducing pharmacological agents[27]. Data to support the use of pharmacogenetic testing of hydroxyurea treatment for hemoglobinopathies are currently very limited. Several SNPs in *HAO2*, *ARG2*, *FLT1* and *NOS1* have been associated with variable HbF response to hydroxyurea treatment[27], and genome-wide transcription profiling efforts are expected to shed light on new pathways involved in this process[28].

Since its establishment in 2000, we have witnessed a substantial annual growth in HbVar content, and a fraction of data submitters were subsequently encouraged to submit a full or short report to the scientific journal *Hemoglobin*[29]. The large repository of previously reported data, together with more recent data acquired by microattribution, shows how the comprehensive documentation of human variation will provide key insights into normal biological processes and how these are perturbed in human genetic disease. We anticipate that microattribution will further encourage new data submitters to contribute their observations to HbVar to receive not only credit in the form of microcitations but also coauthorship in a future microattribution update. The microattribution process established here provides a template for similar ventures for other human genes, their associated systems and the variants that cause their associated genetic diseases. The value of the databases may be considerably further enhanced by linking to collections of blood and DNA samples and also cataloged online, as in the case of many other rare diseases in EuroBioBank.

In essence, this project is a well-coordinated multicenter effort to systematically document genetic variation in globin and associated genes relevant to hemoglobinopathies and thalassemias and is the first example of implementing microattribution to provide incentives for submitting data describing genetic variation. As such, it should serve as a model for the comprehensive documentation and analysis of genetic variations in other common or genetically complex disorders, the conduct of a thorough synopsis of other fields, or both.

URLs. HbVar Database of Hemoglobin Variants and Thalassemia Mutations, http://globin.bx.psu.edu/hbvar/; Golden Helix Server, http://www.goldenhelix.org/; Leiden Open-Access Variation Database, http://www.lovd.nl/; Frequencies of Inherited Disorders database, http://www.findbase.org/; dbSNP database, http://www.ncbi.nlm.nih.gov/projects/SNP/; Human Genome Variation Society, http://www.hgvs.org/; ResearcherID System of Thomson ISI, http://www.researcherid.com/; Open ID system, http://openid.net/; Genotype-to-Phenotype database project's Researcher Identification Primer (RIP), http://www.gen2phen.org/.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

This work was supported by the Netherlands Genomics Initiative (NGI), Erasmus MC (MRace; 296088), the Landsteiner Foundation for Blood Transfusion Research (LSBR; 1040), US National Institutes of Health (NIH) (R01-HL073455) and the Netherlands Scientific Organization (NWO DN 82-301 and 912-07-019) to S.P., the NIH grants R01 DK065806, RC HG005573 and U01 HG004695 to R.C.H., the National Institutes for Health Research Biomedical Research Centre (Oxford) to D.R.H. and European Commission grants (FP6-026539 (ITHANET), FP7-200754 (GEN2PHEN)) to G.P.P.

AUTHOR CONTRIBUTIONS

B.G., J.B., R.C.H. and G.P.P. conceived and designed the study. B.G., J.B. and D.M. implemented the process, built and populated the databases. K.R.P., E.C.C. and H.F. performed experiments. B.K.S., D.J.A., A.N.B., B.C., P.F., A.E.F., A.F., R.G., M.V.E.G., M.G., R.J.G., P.C.G., C.L.H., J.D.H., M.J., P.J., E.K., P.K., S.M., K.M., J.O., A.P., M.N.P., P.P., S. Pavlovic, L.P., M.R., S.S., I.S., M.S., S.L.T., J.T.-S., R.T., T.W., J.S.W., C.W., B.Z. and G.P.P. contributed data. B.G., J.B., D.R.H. and S. Philipsen analyzed results. R.C.H. and G.P.P. supervised data analysis. D.M., W.M., C.R., D.H.K.C. and H.W. provided expertise and infrastructure. B.G., J.B., D.R.H., K.R.P., S. Philipsen, R.C.H. and G.P.P. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Anonymous. Human variome microattribution reviews. *Nat. Genet.* **40**, 1 (2008).
2. Patrinos, G.P. & Antonarakis, S.E. Human hemoglobin. in *Human Genetics: Problems and Approaches* (eds. Speicher, M., Antonarakis, S.E. & Motulsky, A.), 366–401 (Springer-Verlag, Heidelberg, Germany, 2010).
3. Gibbons, R.J., Picketts, D.J., Villard, L. & Higgs, D.R. Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell* **80**, 837–845 (1995).
4. Gibbons, R.J. *et al.* Identification of acquired somatic mutations in the gene encoding chromatin-remodeling factor ATRX in the α-thalassemia myelodysplasia syndrome (ATMDS). *Nat. Genet.* **34**, 446–449 (2003).
5. Viprakasit, V. *et al.* Mutations in the general transcription factor TFIIH result in β-thalassaemia in individuals with trichothiodystrophy. *Hum. Mol. Genet.* **10**, 2797–2802 (2001).
6. Yu, C. *et al.* X-linked thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting DNA binding rather than FOG-1 interaction. *Blood* **100**, 2040–2045 (2002).
7. Borg, J. *et al.* Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. *Nat. Genet.* **42**, 801–805 (2010).
8. Thein, S.L. *et al.* Intergenic variants of *HBS1L-MYB* are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. *Proc. Natl. Acad. Sci. USA* **104**, 11346–11351 (2007).
9. Menzel, S. *et al.* A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat. Genet.* **39**, 1197–1199 (2007).
10. Sankaran, V.G. *et al.* Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* **322**, 1839–1842 (2007).
11. Hardison, R.C. *et al.* HbVar: a relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. *Hum. Mutat.* **19**, 225–233 (2002).
12. Fokkema, I.F., den Dunnen, J.T. & Taschner, P.E. LOVD: easy creation of a locus-specific sequence variation database using an "LSDB-in-a-box" approach. *Hum. Mutat.* **26**, 63–68 (2005).

13. Peterson, K.R. *et al.* Use of yeast artificial chromosomes (YACs) in studies of mammalian development: production of β-globin locus YAC mice carrying human globin developmental mutants. *Proc. Natl. Acad. Sci. USA* **92**, 5655–5659 (1995).
14. Xu, J. *et al.* Transcriptional silencing of γ-globin by BCL11A involves long-range interactions and cooperation with SOX6. *Genes Dev.* **24**, 783–798 (2010).
15. Borg, J., Georgitsi, M., Aleporou-Marinou, V., Kollia, P. & Patrinos, G.P. Genetic recombination as a major cause of mutagenesis in the human globin gene clusters. *Clin. Biochem.* **42**, 1839–1850 (2009).
16. Steensma, D.P., Gibbons, R.J. & Higgs, D.R. Acquired α-thalassemia in association with myelodysplastic syndrome and other hematologic malignancies. *Blood* **105**, 443–452 (2005).
17. Drissen, R. *et al.* The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. *Mol. Cell. Biol.* **25**, 5205–5214 (2005).
18. Singleton, B.K., Burton, N.M., Green, C., Brady, R.L. & Anstee, D.J. Mutations in *EKLF/KLF1* form the molecular basis of the rare blood group In(Lu) phenotype. *Blood* **112**, 2081–2088 (2008).
19. Arnaud, L. *et al.* A dominant mutation in the gene encoding the erythroid transcription factor KLF1 causes a congenital dyserythropoietic anemia. *Am. J. Hum. Genet.* **87**, 721–727 (2010).
20. Singleton, B.K. *et al.* A novel *EKLF* mutation in a patient with dyserythropoietic anemia: the first association of EKLF with disease in man. *Blood* **114**, 72 (2009).
21. Siatecka, M. *et al.* Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Kruppel-like factor. *Proc. Natl. Acad. Sci. USA* **107**, 15151–15156 (2010).
22. Heruth, D.P. *et al.* Mutation in erythroid specific transcription factor *KLF1* causes hereditary spherocytosis in the Nan hemolytic anemia mouse model. *Genomics* **96**, 303–307 (2010).
23. Moradkhani, K. *et al.* Mutations in the paralogous human α-globin genes yielding identical hemoglobin variants. *Ann. Hematol.* **88**, 535–543 (2009).
24. Papadakis, M.N., Patrinos, G.P., Drakoulakou, O. & Loutradi-Anagnostou, A. HbF-Lesvos: an HbF variant due to a novel G gamma mutation (:G gamma 75 ATA→ACA) detected in a Greek family. *Hum. Genet.* **97**, 260–262 (1996).
25. Patrinos, G.P. *et al.* Improvements in the HbVar database of human hemoglobin variants and thalassemia mutations for population and sequence variation studies. *Nucleic Acids Res.* **32**, D537–D541 (2004).
26. Steinberg, M.H. *et al.* Fetal hemoglobin in sickle cell anemia: determinants of response to hydroxyurea. Multicenter study of hydroxyurea. *Blood* **89**, 1078–1088 (1997).
27. Ma, Q. *et al.* Fetal hemoglobin in sickle cell anemia: genetic determinants of response to hydroxyurea. *Pharmacogenomics J.* **7**, 386–394 (2007).
28. Patrinos, G.P. & Grosveld, F.G. Pharmacogenomics and therapeutics of hemoglobinopathies. *Hemoglobin* **32**, 229–236 (2008).
29. Patrinos, G.P. & Wajcman, H. Recording human globin gene variation. *Hemoglobin* **28**, v–vii (2004).

[1]Pennsylvania State University, Center for Comparative Genomics and Bioinformatics, University Park, Philadelphia, Pennsylvania, USA. [2]Department of Applied Biomedical Sciences, University of Malta, Msida, Malta. [3]Laboratory of Molecular Genetics, Department of Physiology and Biochemistry, University of Malta, Msida, Malta. [4]Thalassemia Clinic, Section of Pathology, Mater Dei Hospital, Msida, Malta. [5]Medical Research Council (MRC) Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, Oxford, UK. [6]University of Kansas Medical Center, Department of Biochemistry and Molecular Biology, Kansas City, Kansas, USA. [7]Erasmus University Medical Center, Faculty of Medicine and Health Sciences, Department of Cell Biology, Rotterdam, The Netherlands. [8]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA. [9]Bristol Institute for Transfusion Sciences (BITS), National Health Service (NHS) Blood and Transplant, Bristol, UK. [10]Bogazici University, Department of Molecular Biology and Genetics, Istanbul, Turkey. [11]King's College London, London, UK. [12]Unidade de Investigação e Desenvolvimento, Departamento de Genética, Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisboa, Portugal. [13]Department of Biochemistry, Edouard Herriot University Hospital, Lyon Cedex, France. [14]Dipartimento di Scienze Biomediche e Biotecnologie, University of Cagliari, Cagliari, Sardinia, Italy. [15]Quest Diagnostics Nichols Institute, Chantilly, Virginia, USA. [16]Department of Pharmacy, School of Health Sciences, University of Patras, Patras, Greece. [17]Hemoglobinopathies Laboratory, Human and Clinical Genetics Department, Leiden University Medical Center, Leiden, The Netherlands. [18]Mayo Clinic, Division of Hematopathology, Rochester, Minnesota, USA. [19]North Middlesex University Hospital, London, UK. [20]National and Kapodistrian University of Athens, School of Medicine, Medical Genetics, St. Sophia's Children's Hospital, Athens, Greece. [21]Department of Biology, National and Kapodistrian University of Athens, School of Physical Sciences, Athens, Greece. [22]Hospital Henri-Mondor and Albert-Chenevier Group, Department of Biochemistry and Genetics, Créteil, France. [23]National Haemoglobinopathy Reference Laboratory, Oxford Haemophilia Centre, Churchill Hospital, Oxford, UK. [24]University of Patras, Faculty of Medicine, Laboratory of General Biology, Patras, Greece. [25]Unit of Prenatal Diagnosis, Center for Thalassemia, Laikon General Hospital, Athens, Greece. [26]Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Belgrade, Serbia. [27]Istituto di Neurogenetica e Neurofarmacologia, National Research Council, Cagliari, Cagliari, Sardinia, Italy. [28]Stanford University School of Medicine, Department of Pathology and Pediatrics, Stanford, California, USA. [29]Division of Neurology, Kanagawa Children's Medical Center, Yokohama, Kanagawa, Japan. [30]Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario, Canada. [31]Molecular Diagnostic Genetics, Hamilton Regional Laboratory Program, Hamilton, Ontario, Canada. [32]Medizinisches Versorgungszentrum (MVZ), Laboratory Prof. Seelig, Karlsruhe, Germany. [33]Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA. [34]Department of Pathology, Boston University School of Medicine, Boston, Massachusetts, USA. [35]INSERM, U955, Créteil, France. [36]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Philadelphia, Pennsylvania, USA. [37]These authors contributed equally to this work. Correspondence should be addressed to G.P.P. (gpatrinos@upatras.gr).

ONLINE METHODS

Quantitation of hemoglobin fractions. Twenty microliters of total blood was analyzed using cation-exchange high performance liquid chromatography (VARIANT, Bio-Rad Laboratories).

Construction of the *HBG1* c.-248C>G HPFH β-YAC. A 213-kb yeast artificial chromosome (YAC) carrying the human β-globin locus with the *HBG1* c.-248C>G point mutation ($^{A}\gamma$-195C>G), leading to the Brazilian type of non-deletional HPFH, which directly alters the CCCTTCCC sequence at the 3′ C, was synthesized as follows, using previously described methods[30]. Briefly, a marked *HBG1* gene ($^{A}\gamma^{m}$) contained as a 5.4-kb *Ssp*I fragment (GenBank file U01317, coordinates 38,683-44,077) in the yeast-integrating plasmid (YIP) pRS406 was mutagenized using the QuikChange Site-Specific Mutagenesis Kit (Stratagene). The presence of the *HBG1* c.-248C>G point mutation was confirmed by DNA sequencing, and the mutation was introduced into the β-YAC by 'pop-in,' 'pop-out' homologous recombination in yeast. The mark in the $^{A}\gamma^{m}$-globin gene is a 6-bp deletion at +21 to +26 relative to the $^{A}\gamma$-globin translation start site, allowing preliminary discrimination of the modified β-YAC from the wild-type β-YAC by restriction enzyme digestion following homologous recombination. The presence of the mutation in clones passing this test was confirmed by DNA sequence analysis of a PCR-amplified fragment encompassing the mutated region. Transformation of yeast, screening of positive clones, purification of the β-YAC and mouse transgenesis were performed as described previously[31].

Copy number determination. The relative β-YAC transgene copy number was calculated using the *HBG1* and *HBG2* genes and a standard curve generated from genomic DNA samples from our wild-type β-YAC transgenic mice. Samples of transgenic mouse genomic DNA were serially diluted from 100–0.01 ng and subjected to SYBR PCR with *HBG1* or *HBG2* primers. The copy number for each reaction was estimated by comparing the threshold cycle of each sample to the threshold cycle of the standards and normalizing to the wild-type β-YAC transgenic mouse samples.

Real-time quantitative RT-PCR. Total RNA, isolated from adult peripheral blood, was reverse-transcribed and the resultant complementary DNA was subjected to real-time quantitative RT-PCR analysis with SYBR green using a CFX96 system (Bio-Rad). Human γ-globin expression was normalized to mouse α-globin expression and corrected for transgene and endogenous gene copy number. PCR primer sequences were as previously described[32]. Results are averages of triplicates, with the standard error indicated.

F-cell detection by flow cytometry. We used a protocol adapted from references 32 and 33. Essentially, mouse blood was collected from the tail vein in heparinized capillary tubes. Ten microliters of whole blood was washed in 1 ml PBS, centrifuged at 200*g* at 4 °C for five minutes, and the pellet was resuspended and fixed in 1 ml of 4% fresh paraformaldehyde and PBS at pH 7.5 (Sigma-Aldrich) for 40 min at 37 °C. The cells were centrifuged, and the pellets were resuspended in 1 ml of ice cold acetone and methanol (4:1) and incubated on ice for one minute. Following centrifugation, cells were washed twice in 1 ml ice-cold PBS and 0.1% BSA and resuspended in 800 µl of PBS, 0.1% BSA and 0.1% Triton X-100 (PBT). One microgram of γ-globin antibody (catalog number sc-21756 unconjugated, Santa Cruz Biotechnology) was added to 100 µl of the cell suspension and incubated for 20 min in the dark at room temperature (37 °C). One milliliter of ice-cold PBS and 0.1% BSA was added, the sample was centrifuged and the pellet was resuspended in 100 µl ice-cold PBT. One hundred microliters of Alexa 488 (catalog number 11001, Invitrogen Molecular Probes) secondary antibody, diluted 1:200 in ice-cold PBT, was added to the cell suspension and the sample was incubated at room temperature for 20 min in the dark. Cells were washed with 1 ml of ice-cold PBS and 0.1% BSA and the pellets were resuspended in 200 µl of PBS. Samples were analyzed using an Accuri C6 Flow Cytometer (Accuri Cytometers, Inc.) with a 530/30 nm (FITC/GFP) emission filter. Data from 30,000 cells were acquired for analysis using CFlow Software (Accuri Cytometers, Inc.); cells were gated to exclude dead cells. For, FL1-A, a 530/30 nm (FITC, GFP) filter was used to identify the Alexa 488–positive F cell population; For FL2-A a 585/40 nm (PE, PI) filter was used as a compensation to identify the Alexa 488–negative cell population. For M3, the mean fluorescent intensity, an increase in F cells is reflected by a peak shift and increase in the peak of fluorescence intensity. P4, distinct positive F cells.

30. Harju, S., Navas, P.A., Stamatoyannopoulos, G. & Peterson, K.R. Genome architecture of the human β-globin locus affects developmental regulation of gene expression. *Mol. Cell. Biol.* **25**, 8765–8778 (2005).
31. Harju-Baker, S., Costa, F.C., Fedosyuk, H., Neades, R. & Peterson, K.R. Silencing of Agamma-globin gene expression during adult definitive erythropoiesis mediated by GATA-1-FOG-1-Mi2 Complex binding at the -566 GATA site. *Mol. Cell. Biol.* **28**, 3101–3113 (2008).
32. Böhmer, R.M. Flow cytometry of erythroid cells in culture: bivariate profiles of fetal and adult hemoglobins. *Methods Cell Biol.* **64**, 139–152 (2001).
33. Amoyal, I. & Fibach, E. Flow cytometric analysis of fetal hemoglobin in erythroid precursors of β-thalassemia. *Clin. Lab. Haematol.* **26**, 187–193 (2004).

Amino Acids
DOI 10.1007/s00726-011-1146-1

SHORT COMMUNICATION

A simple screening method using ion chromatography for the diagnosis of cerebral creatine deficiency syndromes

Takahito Wada · Hiroko Shimbo · Hitoshi Osaka

Received: 23 August 2011 / Accepted: 1 November 2011

Abstract Cerebral creatine deficiency syndromes (CCDS) are caused by genetic defects in L-arginine:glycine amidinotransferase, guanidinoacetate methyltransferase or creatine transporter 1. CCDS are characterized by abnormal concentrations of urinary creatine (CR), guanidinoacetic acid (GA), or creatinine (CN). In this study, we describe a simple HPLC method to determine the concentrations of CR, GA, and CN using a weak-acid ion chromatography column with a UV detector without any derivatization. CR, GA, and CN were separated clearly with the retention times (mean ± SD, $n = 3$) of 5.54 ± 0.0035 min for CR, 6.41 ± 0.0079 min for GA, and 13.53 ± 0.046 min for CN. This new method should provide a simple screening test for the diagnosis of CCDS.

Keywords Cerebral creatine deficiency syndromes · HPLC · Creatine · Guanidinoacetic acid · Creatinine

Abbreviations

CCDS	Cerebral creatine deficiency syndromes
AGAT	Arginine:glycine amidinotransferase
GAMT	Guanidinoacetate methyltransferase
CR	Creatine
GA	Guanidinoacetic acid
CN	Creatinine

T. Wada · H. Shimbo · H. Osaka (✉)
Division of Neurology, Kanagawa Children's Medical Center, 2-138-4 Mutsukawa, Minami-ku, Yokohama, Kanagawa 232-8555, Japan
e-mail: hosaka@kcmc.jp

Introduction

Creatine (CR) plays an important role in the storage and transmission of ATP-derived energy (Walker 1979; Wyss and Kaddurah-Daouk 2000). Normal levels of cellular CR are maintained through both diet and biosynthesis in the kidney and liver. Recently, it has been shown that brain must synthesize an important part of its CR, due to the very restricted permeability of blood–brain barrier for CR (Braissant et al. 2011). This synthesis requires the action of two enzymes, arginine:glycine amidinotransferase (AGAT; EC 2.1.4.1) and guanidinoacetate methyltransferase (GAMT; EC 2.1.1.2) (Fig. 1). AGAT transfers an amidino group from arginine to glycine, yielding ornithine and guanidinoacetate (GA), while GAMT transfers a methyl group from S-adenosylmethionine to GA, forming CR. However, tissues are incapable of CR biosynthesis such as the brain, the levels of CR are dependent on a creatine transporter 1 encoded by the *SLC6A8* to transport CR against the concentration gradient (Fig. 1).

The cerebral CR deficiency syndromes (CCDS) are a group of disorders that include two recessive conditions that impair the synthesis of CR, namely, AGAT deficiency (OMIM 612718) (Item et al. 2001) and GAMT deficiency (OMIM 612736) (Stockler et al. 1996), as well as one X-linked condition, namely, SLC6A8 deficiency (OMIM 300036) (Salomons et al. 2001).The common clinical features of all CCDS are mental retardation, speech delay, autistic behavior, and seizures.

In GAMT deficiency, the urinary GA/CN concentration ratio increases 2- to 30-fold (Mercimek-Mahmutoglu et al. 2006). AGAT deficiency shows decreased ratios of both CR/CN and GA/CN concentrations in urine (Carducci et al. 2002). The X-linked-SLC6A8 deficiency is characterized by high urinary CR/CN concentration ratios. Therefore,

Published online: 13 November 2011

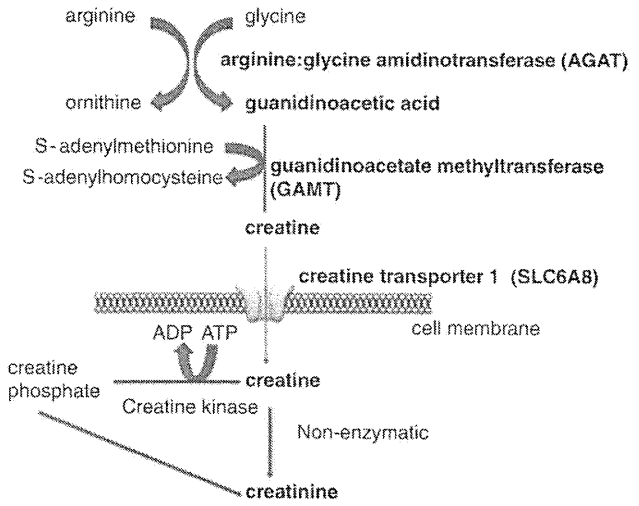

Fig. 1 Synthesis and transport of creatine. L-arginine:glycine amidinotransferase, (AGAT) synthesizes guanidinoacetate from arginine and glycine. Guanidinoacetate methyltransferase (GAMT) transfers a methyl group from S-adenosylmethionine to guanidinoacetate, thus, generating creatine. Creatine enters cells in the brain through the CT1 creatine transporter encoded by the SLC6A8 gene. Creatine is phosphorylated by creatine kinase to phosphocreatine, which is a reversible reaction and contribute to the storage and swift supply of ATP. Both creatine and phosphocreatine, lead non-enzymatically to the formation of creatinine that is excreted in urine

Table 1 Urine creatine and guanidinoacetic acid levels in patients with defects of creatine synthesis and transport

Disease	GA	CR/CN
AGAT deficiency	Low	Low
GAMT deficiency	High	Normal
CR transporter deficiency	Normal	High (4017.5 ± 286.4)

(), values from one patient with CR transporter deficiency (μmol/mmol): HPLC method; 4017.5 ± 286.4; mean ± SD, $n = 2$

Normal values with this study from 15 samples are 798.6 ± 574.8 (CR/CN (μmol/mmol)) (HPLC method; mean ± SD)

AGAT arginine:glycine amidinotransferase, *GAMT* guanidinoacetate methyltransferase, *CR* creatine, *GA* guanidinoacetic acid; *CN* creatinine

evaluation of CR/CN and GA/CN concentrations in urine allows the diagnosis of three types of CCDS (Table 1). So far, 4 cases from two families of AGAT deficiency and ~40 cases with GAMT-related CCDS have been reported (Longo et al. 2011). The SLC6A8 deficiency is the most frequent cause of CCDS. The SLC6A8 deficiency is reported to constitute 0.8–5.4% of the cases of X-linked intellectual disability, and represents the second frequent cause of non-syndromic X-linked intellectual disability following a fragile-X syndrome (Rosenberg et al. 2004; Newmeyer et al. 2005; Clark et al. 2006; Lion-Francois et al. 2006; Arias et al. 2007; Betsalel et al. 2008; Puusepp et al. 2009; Ardon et al. 2010). However, the majority of these patients may remain unrecognized. In this study, we sought to establish an easy and inexpensive high performance liquid chromatography (HPLC) method with UV detection technique to quantify CR, GA, and CN in urine. We chose a weak acidic cation-exchange column of hydrophilic polymer beads with carboxyl groups on their surface and phosphoric acid (H_3PO_4) as the mobile phase. Our results show a complete separation and detection of the three compounds by UV at 210 nm without derivatization.

Materials and methods

All reagents were of analytical grade or better. Acetonitrile, phosphoric acid (85%), creatine monohydrate (98%), guanidinoacetic acid (98%), and creatinine (99%) were obtained commercially (Wako Pure Chemical Industries, Ltd, Osaka, Japan). Ultrapure water was prepared by a Milli-Q system (Millipore, Tokyo, Japan). Standard solutions of CR, GA, and CN were prepared, respectively, by dissolving weighed amounts of the reagents in 2 mM phosphoric acid.

Urine samples collected from healthy individuals or patients were stored at −20°C and thawed just before analyses. Urine (500 μl) of urine was treated with acetonitrile (500 μl) to precipitate proteins, and centrifuged (13,000 rpm, 10 min) after 10 min. The supernatant (50–100 μl) was diluted with 2 mM phosphoric acid (400–950 μl), and a diluted sample (25 μl) was analyzed under the standard condition.

The HPLC set-up comprised a pump, LC-6A, (Shimadzu Ltd., Kyoto, Japan), a Rheodyne injector fitted with a 100 μl loop, an UV–vis spectrophotometric detector, SPD-6A (Shimadzu Ltd.) and a Chromatopack integrator, CR-6A (Shimadzu Ltd.). The separation was performed on a weak acidic cation-exchange column, IC YS-50 (Shodex Ltd., Kawasaki, Japan. 4.6 mm × 125 mm i.d.) using aqueous phosphoric acid as the mobile phase with flow rate of 1.0 ml/min. The analytics were monitored with UV detection at 210 nm.

Standard solutions (25 μl each) of creatine monohydrate (2 mg/l: 134 μmol/l), guanidinoacetic acid (6.7 mg/l: 572 μmol/l), and creatinine (2 mg/l: 177 μmol/l) were loaded into the 100 μl loop, and injected. Analyses were carried out under the following concentrations of mobile phase: 1, 2, 5, and 10 mM H_3PO_4. All analyses were performed at room temperature (32 ± 2°C).

Our new HPLC method was compared to a conventional enzymatic method. Briefly, creatinine amidohydrolase catalyzes CN to CR. Creatine amidinohydrolase and sarcosine oxidase generate sequentially CR to hydrogen peroxide, which is measured at 510 nm in a reaction catalyzed by horseradish peroxidase (Fossati et al. 1983). Fifteen

urine samples were used to determine urinary CR, CN, and the CR/CN ratio. Moreover, one urine sample from a patient with X-linked-SLC6A8 deficiency was analyzed. The results were analyzed by Pearson's correlation coefficient using the PRISM software (La Jolla, CA).

All parents of individuals participating in this study gave a written informed consent after full explanation of the study. The design of this study was approved by the ethical committee of Kanagawa Children's Medical Center.

Results and discussion

CR and GA, not separated with 5 and 10 mM H_3PO_4, were clearly separated with 1 and 2 mM H_3PO_4. However, 2 mM H_3PO_4 produced a complete separation of CR, GA and CN, with retention times (mean ± SD, $n = 3$) of 5.54 min (±0.0035) for CR, 6.41 min (±0.0079) for GA, and 13.53 min (±0.046) for CN using 2 mM H_3PO_4 at a flow rate of 1.0 ml/min (Fig. 2). Since 1 mM H_3PO_4 required a longer retention time (>20 min.), we used 2 mM of H_3PO_4 for subsequent analyses.

Standard solutions (25 µl) of CR (10–1,000 µmol/l), GA (50–4,000 µmol/l), and CN (10–1,000 µmol/l) were analyzed under the above conditions. We obtained a linear correlation between peak areas and concentrations. The linear regression equations for peak area (y; in arbitrary units) and concentration (µmol/l) of the injected calibrator (x) were: $y = 59.821x$ ($R^2 = 0.9971$) for CR; $y = 12.746x$ ($R^2 = 0.9982$) for GA; and $y = 61.604x$ ($R^2 = 0.9968$) for CN. The calibration curves (0–1,000 µmol/l) covered the range of CR, GA, and CN concentrations typically found in urine (when diluted 10- to 40-fold).

Next, we analyzed 15 normal urine samples and compared the values obtained from our new method with those obtained using a conventional enzymatic method. Good correlations were obtained between our new method and the enzymatic method [CR; $r = 0.9276$ ($p < 0.001$, $R^2 = 0.8605$), CN; $r = 0.9370$ ($p < 0.001$, $R^2 = 0.8780$)]. In addition, the CR/CN ratio showed a good correlation between the two methods ($r = 0.984$, $p < 0.001$, $R^2 = 0.9682$). We also obtained a good correlation in a patient with SLC6A8 deficiency of CR/CN (µmol/mmol): Enzymatic method; 4439.5 ± 375.5, HPLC method; 4017.5 ± 286.4; mean ± SD, $n = 2$) The Jaffe reaction and enzymatic methods have been used to determine CN and CR levels (Husdan and Rapoport 1968; Fossati et al. 1983). In addition, Shirokane et al. (1991) described an accurate and simple enzymatic determination for urinary GA (Shirokane et al. 1991). However, this method is not easily accessible because it utilizes guanidinoacetate kinase

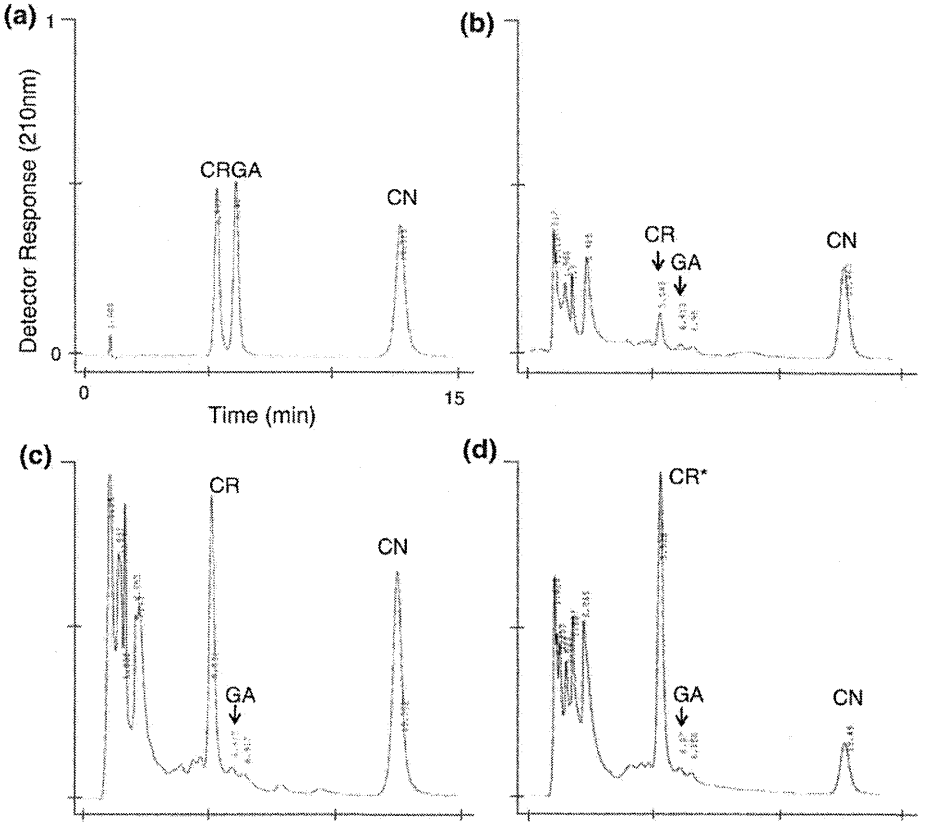

Fig. 2 Examples of chromatograms. Chromatograms of **a** standard samples, and urine sample from **b** a healthy individual-1, **c** a healthy individual-2, and **d** a male patient with SLC6A8 deficiency. Compared to healthy individuals, the patient with SLC6A8 deficiency shows a marked increase in creatine (CR) (*asterisk*)