

*Research Letter***Sequence Variation in Ultraconserved and Highly Conserved Elements Does Not Cause X-Linked Mental Retardation****Armand Bottani,<sup>1</sup> Jamel Chelly,<sup>2</sup> Arjan P.M. de Brouwer,<sup>3</sup> Bruno Pardo,<sup>1</sup> Mandy Barker,<sup>1</sup> Valeria Capra,<sup>4</sup> Lucia Bartoloni,<sup>1</sup> Stylianos E. Antonarakis,<sup>1</sup> and Bernard Conrad<sup>1,5\*</sup>**<sup>1</sup>Department of Genetic Medicine and Development, Geneva University Medical School and University Hospitals, Geneva, Switzerland<sup>2</sup>Institut Cochin, Département de Génétique et Pathologie Moléculaire,

Equipe de Génétique et Physiopathologie des Retards Mentaux, rue du Faubourg St Jacques, Paris, France

<sup>3</sup>Department of Human Genetics 417, University Medical Center Nijmegen, St. Radboud, HB Nijmegen, The Netherlands<sup>4</sup>U.O. Neurochirurgia, Istituto Scientifico Giannina Gaslini, Genova, Italy<sup>5</sup>Division of Human Genetics and Cytogenetics, Bern University Children's Hospital, Bern, Switzerland

Received 3 November 2006; Accepted 24 November 2006

**How to cite this article: Bottani A, Chelly J, de Brouwer APM, Pardo B, Barker M, Capra V, Bartoloni L, Antonarakis SE, Conrad B. 2007. Sequence variation in ultraconserved and highly conserved elements does not cause X-linked mental retardation. *Am J Med Genet Part A* 143A:888–890.****To the Editor:**

About 5% of mammalian genomes show evidence of past purifying selection [Waterston et al., 2002]. Coding regions comprise only ~1.5% of the constrained sequence elements in genomes, while ~3% of sequence space is occupied by conserved non-coding sequences (CNCs) [Dermitzakis et al., 2005]. The conservation of CNCs suggests that these elements are probably functional [Drake et al., 2006]. Two classes of CNCs are of particular interest, namely ultraconserved elements (UCEs) [Bejerano et al., 2004], and highly conserved elements (HCEs) [Siepel et al., 2005]. We decided to study the role of sequence variation within UCEs and HCEs in patients with X-linked mental retardation (XLMR).

Mental retardation affects ~2–3% of children in developed countries, making it a major, still largely unresolved medical problem [Ropers and Hamel, 2005]. X-chromosome linked genes are considered particularly important for the etiology of MR, which is based on an excess of affected males. Monogenic forms of XLMR were estimated to affect ~10% of males with mental retardation [Mandel and Chelly, 2004; Ropers and Hamel, 2005], and 59 genes have to date been causally implicated in familial MR (<http://www.ggc.org/xlmr.htm>). For these reasons, we decided to screen all of the known UCEs and HCEs on the X-chromosome for sequence variants that cosegregate with MR. The X-chromosome contains 27 UCEs, and some of them are located adjacent to

known MR-genes, such as for instance *ARX* [Poirier et al., 2006]. Most other UCEs are located close to developmental regulators, making them candidates for MR. The two HCEs on the X-chromosome that were included in the study are located in the 3' UTRs of the known MR-genes *MECP2* and *FMR1* [Poirier et al., 2006].

Nine patients with idiopathic mental retardation (IMR) were recruited at the Division of Medical Genetics of Geneva University Hospitals, according to a protocol approved by the local ethical committee. Patients were included in the study based on the following features, mild to moderate MR with no or few minor malformations, a normal karyotype at a resolution of >550 bands, and no mutation in *FMR1*. For some of the patients, subtelomeric and interstitial insertions–deletions had been excluded by multiplex ligation-dependent probe amplification (MLPA) and comparative genome hybridization (CGH). The sample collection of the European Mental Retardation Consortium was as outlined (<http://www.eur-omrx.com>). Healthy male and female controls were recruited at the Istituto Scientifico Giannina Gaslini,

Grant sponsor: Swiss National Science Foundation.

\*Correspondence to: Dr. Bernard Conrad, Division of Human Genetics and Cytogenetics, Bern University Children's Hospital, CH-3010 Bern, Switzerland. E-mail: [bernard.conrad@insel.ch](mailto:bernard.conrad@insel.ch)

DOI 10.1002/ajmg.a.31651

Genova, Italy, and the Division of Medical Genetics, Geneva.

The 27 UCEs and 2 HCRs were PCR-amplified with primers spanning the core element plus at least 75 flanking nucleotides, followed by sequencing of both strands on an automated sequencer (ABI Genetic Analyzer 3100). The frequency of the A>G UCE460 variant in healthy controls and XLMR samples was determined by PCR-sequencing and by PCR-pyrosequencing assays (oligonucleotide sequences and protocols available upon request).

We sequenced all 27 UCEs and two HCEs in IMR patients 1–9, and found only one position that deviated from the consensus sequence, at nucleotide 275 of UC460 [Bejerano et al., 2004] in IMR5 (Fig. 1). However, we observed two known polymorphisms at the expected frequencies, 74 bases upstream of UC471, and 8 bases upstream of UC473. The healthy mother of the propositus was a heterozygous A/G carrier. IMR5 had mild MR, absent speech, and stereotypic behavior. His karyotype was normal at a resolution of 550 bands, he had no mutation in *FMRI*, MLPA analysis for subtelomeric deletions and CGH were normal. Deletion of the region at 22q13 and

Angelman syndrome were excluded as possible reasons for absent speech. Thus, the UCE460 G variant allele represented a candidate for the phenotype of IMR5.

We next sought to determine the population frequency of the variant G allele by sequencing UC460 in healthy male and female control individuals. The UCE460 G allele was identified in 3/127 (2.3%) healthy males and in 11/207 (5.3%) X-chromosome healthy heterozygous females. We concluded that based on its population allele frequency the variant UCE460 G allele was unlikely to be pathogenic. However, it could have a modifying effect on other MR-genes, analogously to known modifier genes [Badano and Katsanis, 2002], and be weakly associated with MR. We further screened the euromrx sample collection (<http://www.euromrx.com>) for the variant UC460 G allele using a pyrosequencing assay. We found the variant G allele in 7/245 (2.8%), and the wt A allele in 238/245 (97.1%) XLMR-patients. As a control, we typed position 601 of UCE462, using an analogous pyrosequencing assay, and found the wt G allele in 238/238 XLMR-patients. In conclusion, these results

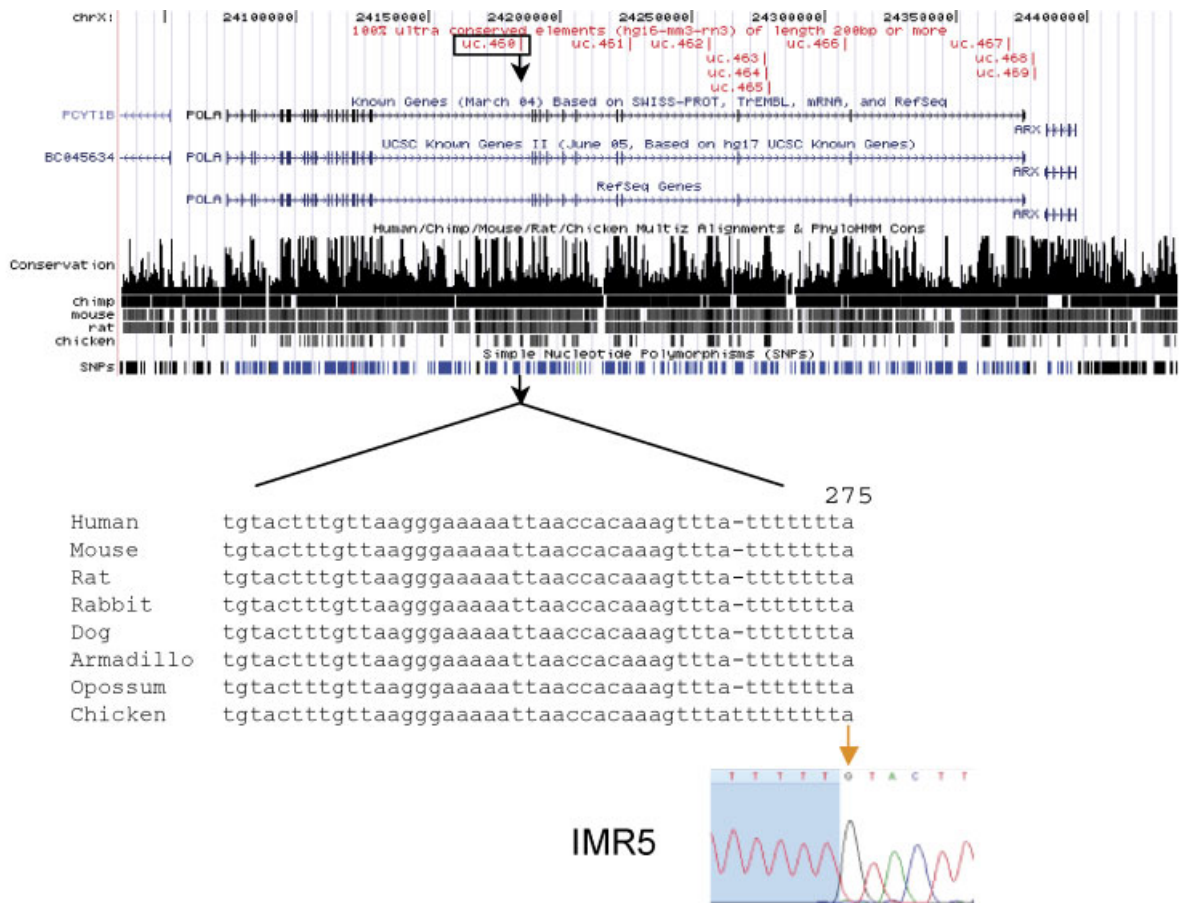


FIG. 1. The UCSC-Genome Browser view of the region on the X-chromosome containing the POLA/ARX genes is shown. The UCE460-sequence is magnified, and the variant identified in IMR5 is shown at the bottom. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

make it unlikely for the UCE460 G allele to play a significant role in XLMR.

We presented here the results of a first study addressing the potential relevance of sequence variation in X-chromosomal ultraconserved CNCs for XLMR. We detected a single sequence variant in UCE460, a 275 basepair long element associated with the 3' ends of the *POLA* gene and of the developmental regulator *ARX*. This group of UCEs was proposed to function as *cis*-regulatory elements controlling *ARX* transcription [Bejerano et al., 2004]. Four heterogeneous entities of isolated MR and MR-malformation disorders are known to be associated with *ARX* mutations [Poirier et al., 2006]. The UCE460 variant G allele therefore represented an attractive candidate for MR-phenotypes. Our study, however, does not support an important role played by this sequence variant for XLMR.

In analogy to nonsynonymous substitutions in protein coding regions, at least some of the sequence variation in CNCs is expected to be pathogenic. However, while our understanding of the consequences of sequence variation within coding regions is well advanced, we are only at the beginning of apprehending the biological effects of sequence variation within CNCs, and its role in human disease. The precise function of most CNCs is not known, and the insights into the biological effects of CNC mutations are only beginning to emerge [Richler et al., 2006]. In conclusion, our study represents an initial step towards the promising characterization of the biological effects of sequence variation within CNCs for human phenotypes and diseases.

#### ACKNOWLEDGMENTS

We thank all the families and patients from the European Mental Retardation Consortium. This work was supported by the Swiss National Science Foundation (BC). LB is supported by a MV-fellowship from the SNSF.

#### REFERENCES

- Badano JL, Katsanis N. 2002. Beyond Mendel: An evolving view of human genetic disease transmission. *Nat Rev Genet* 3:779–789.
- Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, Haussler D. 2004. Ultraconserved elements in the human genome. *Science* 304:1321–1325.
- Dermitzakis ET, Reymond A, Antonarakis SE. 2005. Conserved non-genic sequences—an unexpected feature of mammalian genomes. *Nat Rev Genet* 6:151–157.
- Drake JA, Bird C, Nemesh J, Thomas DJ, Newton-Cheh C, Reymond A, Excoffier L, Attar H, Antonarakis SE, Dermitzakis ET, Hirschhorn JN. 2006. Conserved noncoding sequences are selectively constrained and not mutation cold spots. *Nat Genet* 38:223–227.
- Mandel JL, Chelly J. 2004. Monogenic X-linked mental retardation: Is it as frequent as currently estimated? The paradox of the *ARX* (Aristaless X) mutations. *Eur J Hum Genet* 12:689–693.
- Poirier K, Lacombe D, Gilbert-Dussardier B, Raynaud M, Desportes V, de Brouwer AP, Moraine C, Fryns JP, Ropers HH, Beldjord C, Chelly J, Bienvenu T. 2006. Screening of *ARX* in mental retardation families: Consequences for the strategy of molecular diagnosis. *Neurogenetics* 7:39–46.
- Richler E, Reichert JG, Buxbaum JD, McInnes LA. 2006. Autism and ultraconserved non-coding sequence on chromosome 7q. *Psychiatr Genet* 16:19–23.
- Ropers HH, Hamel BC. 2005. X-linked mental retardation. *Nat Rev Genet* 6:46–57.
- Stapel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, Weinstock GM, Wilson RK, Gibbs RA, Kent WJ, Miller W, Haussler D. 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 15:1034–1050.
- Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Esvara P, Eyas E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, Karolchik D, Kasprzyk A, Kawai J, Keibler E, Kells C, Kent WJ, Kirby A, Kolbe DL, Korf I, Kucherlapati RS, Kulbokas EJ, Kulp D, Landers T, Leger JP, Leonard S, Letunic I, Levine R, Li J, Li M, Lloyd C, Lucas S, Ma B, Maglott DR, Mardis ER, Matthews L, Mauceli E, Mayer JH, McCarthy M, McCombie WR, McLaren S, McLay K, McPherson JD, Meldrim J, Meredith B, Mesirov JP, Miller W, Miner TL, Mongin E, Montgomery KT, Morgan M, Mott R, Mullikin JC, Muzny DM, Nash WE, Nelson JO, Nhan MN, Nicol R, Ning Z, Nusbaum C, O'Connor MJ, Okazaki Y, Oliver K, Overton-Larty E, Pachter L, Parra G, Pepin KH, Peterson J, Pevzner P, Plumb R, Pohl CS, Poliakov A, Ponce TC, Ponting CP, Potter S, Quail M, Reymond A, Roe BA, Roskin KM, Rubin EM, Rust AG, Santos R, Sapojnikov V, Schultz B, Schultz J, Schwartz MS, Schwartz S, Scott C, Seaman S, Searle S, Sharpe T, Sheridan A, Shownkeen R, Sims S, Singer JB, Slater G, Smit A, Smith DR, Spencer B, Stabenau A, Stange-Thomann N, Sugnet C, Suyama M, Tesler G, Thompson J, Torrents D, Trevaskis E, Tromp J, Ucla C, Ureta-Vidal A, Vinson JP, Von Niederhausern AC, Wade CM, Wall M, Weber RJ, Weiss RB, Wendl MC, West AP, Wetterstrand K, Wheeler R, Whelan S, Wierzbowski J, Willey D, Williams S, Wilson RK, Winter E, Worley KC, Wyman D, Yang S, Yang SP, Zdobnov EM, Zody MC, Lander ES. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520–562.