Role of the LF-SINE–Derived Distal *ISL1* Enhancer in Patients with Classic Bladder Exstrophy

Rong Zhang¹ Michael Knapp² Franziska Kause¹ Heiko Reutter^{1,3} Michael Ludwig⁴

¹Institute of Human Genetics, University of Bonn, Bonn, Germany

²Institute of Medical Biometry, Informatics, and Epidemiology, University of Bonn, Bonn, Germany

³Department of Neonatology and Pediatric Intensive Care, Children's Hospital, University of Bonn, Bonn, Germany

⁴Department of Clinical Chemistry and Clinical Pharmacology, University of Bonn, Bonn, Germany

Address for correspondence Michael Ludwig, PhD, Department of Clinical Chemistry and Clinical Pharmacology, University of Bonn, Sigmund-Freud-Street 25, 53105 Bonn, Germany (e-mail: mludwig@uni-bonn.de).

J Pediatr Genet 2017;6:169–173.

Abstract

Keywords

- ▶ *cis*-regulatory element
- ▶ classic bladder exstrophy
- ▶ enhancer
- ▶ *ISL1*
- ▶ LF-SINE

A genome-wide association study and meta-analysis identified *ISL1* as the first genome-wide significant susceptibility gene for classic bladder exstrophy (CBE). A short interspersed repetitive element (SINE), first detected in lobe-finned fishes (LF-SINE), was shown to drive *Is1* expression in embryonic mouse genital eminence. Hence, we assumed this enhancer a conclusive target for mutations associated with CBE formation and analyzed a cohort of 200 CBE patients. Although we identified two enhancer variants in five CBE patients, their clinical significance seems unlikely, implying that sequence variants in the *ISL1* LF-SINE enhancer are not frequently associated with CBE.

Introduction

Classic bladder exstrophy (CBE) represents the most common form of the bladder exstrophy-epispadias complex (BEEC), presenting as the most severe congenital anomaly of the kidney and urinary tract.^{1,2} CBE is thought to result from aberrant embryonic morphogenesis of the urorectal septum and the cloacal membrane^{1,2} and has an estimated birth prevalence of 1 in 37,000.³ The recurrence risk for the offspring of CBE patients shows an approximately 400-fold increase compared with that in the general population,⁴ and further lines of evidence implicate, that genetic factors contribute to CBE etiology. These include reports of more than 30 families with multiple affected members^{5,6} and the observation of high concordance rates in monozygotic twins.⁷ Moreover, at least 30 BEEC-associated chromosomal aberrations and copy number variations have been observed^{6,8} and array-based analyses, and regional association studies suggested microduplications on chromosome 22q11.21 and polymorphisms in the *TP63* gene as causally

related.^{9–13} Consequently, a genome-wide association study (GWAS) and meta-analysis identified *ISL1* as the first genome-wide significant susceptibility gene for CBE.¹⁴

ISL1 encodes the insulin gene enhancer protein ISL-1, a member of the LIM zinc-binding/homeobox-domain transcription factor family, first identified as a regulator of insulin expression.¹⁵ ISL-1 plays an essential role in cell specification, differentiation, and survival in embryogenesis and its expression has been detected in various tissues and cell types.¹⁶ Of interest, *Is1* expression has also been observed within the genital region, including the pericloacal mesenchyme and the urogenital septum, of embryonic mice within the critical time frame for human CBE formation.^{14,17}

As *ISL1* coding sequence analysis in 207 CBE patients failed to detect any direct disease-causing mutation,¹⁴ we searched for noncoding elements that might be involved in *ISL1* expression regulation and, when mutated, can explain CBE formation.

Three enhancers were shown to drive the motor and sensory neuron-specific *Is1* expression in chick, mouse,

received

February 14, 2017

accepted after revision

March 20, 2017

published online

April 21, 2017

Copyright © 2017 by Georg Thieme Verlag KG, Stuttgart · New York

DOI <https://doi.org/10.1055/s-0037-1602387>

ISSN 2146-4596.

Fig. 1 Localization of the human *ISL1* gene (top; adapted from the Ensembl Genome Browser, hg38)²¹ and enhancer sequences in the vicinity of *ISL1* with its transcription start and direction of transcription indicated by an arrow (middle). Location of enhancers CHREST (conserved regulatory element for islet 1) 1–3 and the short interspersed repetitive element (SINE) transposon from a family in lobe-finned fish (LF-SINE) are taken from the literature.^{17–20} Location of the p300 binding motif in LF-SINE and position of rare nucleotide variants detected is indicated below. Enlargements are not drawn to scale.

and zebrafish.^{18–20} These enhancers termed CHREST (conserved regulatory element for *islet 1*) 1–3, are remarkably conserved between these species, and locate either 520 kb distal (CHREST3) or proximal (CHREST1, 266 kb; CHREST2, 302 kb) to the human *ISL1* gene (→Fig. 1). Interestingly, a fourth enhancer resides 499 kb proximal to the *ISL1* transcription start site. Since this sequence represents a member of a recently active short interspersed repetitive element (SINE) transposon family in lobe-finned fishes, it was termed LF-SINE.¹⁷ Remarkably, this LF-SINE has been shown to drive *Isl1* expression in mouse genital eminence at embryonic day (E) 11.5.¹⁷ According to ENCODE,²² this enhancer also comprises a binding site for histone acyltransferase p300, a transcriptional coactivator that, together with its paralog, the CREB-binding protein (CBP), regulates transcription via chromatin remodeling, as well as crosstalk and interference between numerous signaling pathways.²³

All of these findings suggested the LF-SINE enhancer a conclusive target for mutations associated with CBE formation. Hence, we analyzed its sequence in 200 CBE patients.

Material and Methods

Subjects

The sample included 200 isolated CBE patients of European descent. The study was approved by the ethics committee of the Medical Faculty of the University of Bonn and was conducted in accordance with the principles of the Declaration of Helsinki. Before inclusion, a written informed consent was obtained from all participants, or from their proxies in case of legal minors. All patients were recruited in person by

experienced physicians trained in the diagnosis of CBE and details of the recruitment process were outlined elsewhere.^{14,24}

Sequence Analysis

Blood or saliva samples were obtained from patients and isolation of genomic DNA from blood was performed using a Chemagic Magnetic Separation Module I (Chemagen, Baesweiler, Germany). Genomic DNA from saliva samples was extracted by using the Oragene DNA Kit (DNA Genotek Inc., Kanata, Canada).

For DNA analysis, polymerase chain reaction-amplified DNA products were subjected to direct automated sequencing (3130XL Genetic Analyzer, Applied Biosystems, Foster City, California, United States). Primer sequences flanking the LF-SINE were adopted from Bejerano et al¹⁷ and, according to UCSC Genome Browser (hg38), covered 881 bp at genomic region chr5:51,882,318–51,883,198.

Statistical Analysis

All data were obtained as outlined in detail in our GWAS and meta-analysis study.¹⁴

Results

Database research revealed that the LF-SINE, in the vicinity of the respective *Isl1* gene, is conserved throughout birds and mammals (data not shown). Also, data from 100 vertebrates basewise conservation by PhyloP, and Multiz alignments revealed its evolutionary conservation.^{25,26} This finding suggested an important function of this enhancer in the

regulation of gene expression and supported our initial thesis, that mutations in the LF-SINE may be associated with CBE.

Sequence analysis of the LF-SINE in our series of 200 CBE patients revealed only five patients to be heterozygous (C/T) for rs35665267, a single nucleotide polymorphism (SNP), deposited in the SNP database (dbSNP Build 147) with a frequency for the rare T-allele among Europeans ($n = 1,006$ alleles) of 0.0159. Our observed frequency in CBE patients was 0.0125, and hence not significantly different. In a single patient, one additional, novel heterozygous C-to-T-transition was found 137 bp upstream (position chr5:51,882,583) of rs35665267 (►Fig. 1), and interestingly, this patient also carried variant rs35665267 in the heterozygous state. However, the de novo occurrence of the novel variant could not be investigated as the parents were unavailable.

The finding of the LF-SINE-derived enhancer prompted us to reevaluate the data from our GWAS and meta-analysis for this genomic region (►Fig. 2).¹⁴ Here, we had obtained a genotype-specific relative risk ratio (RR) for the risk/other allele (T/C) of 1.25 (95% confidence interval = 0.33–4.74) with a p value of 0.741. The individual sequence data from the LF-SINE analysis reflected the imputed data for rs35665267 in our GWAS,¹⁴ where a frequency for the rare T-allele of 0.0139 was estimated in patients ($n = 208$) versus 0.0108 in controls ($n = 1,703$).

Discussion

With the advent of GWAS, a large number of common SNPs associated with complex diseases have been identified.²⁷ As most of these SNPs locate to intergenic regions, their asso-

ciation with disease risk should be based on the regulation of expression levels of neighboring genes. This may depend on altered activities of *cis*-regulatory elements such as promoters, silencers, enhancers, or insulators, as has been shown in association with many diseases such as Holt–Oram syndrome, Hirschsprung disease, MonoMAC syndrome, or osteoarthritis.²⁸

Our GWAS and meta-analysis identified *ISL1* as a genome-wide significant susceptibility gene for CBE with a haplotype block comprising a total of 138 SNPs.¹⁴ However, analysis of the *ISL1* coding sequence in 207 CBE patients detected no disease-causing mutation¹⁴ indicating, that rather impaired *ISL1* expression regulation might be involved in CBE formation. As enhancers CHREST1–3 have been shown to control selective *Isl1* expression in motor neurons in various species,^{18–20} we focused on the LF-SINE that was found to be capable of driving *Isl1* expression in mouse embryonic genital eminence.¹⁷ This LF-SINE locates around 250 kb proximal of the haplotype block around the *ISL1* gene detected in GWAS.

The sequence analysis of the LF-SINE in 200 CBE patients however revealed only 5 patients heterozygous for rs35665267 and 1 of these 5 patients also carried a novel heterozygous variant. Whereas the novel variant locates outside the p300 motif, rs35665267 does not affect an AT-rich motif, shown to be sufficient for p300 binding in the human α -actin promoter.²⁹ We also observed no evolutionary conservation at both SNP positions. Although sequence conservation points to functionality, Schmidt et al³⁰ reported that transcription factor bound regions show only limited conservation across species. Also, we cannot exclude, that the rare T-allele of rs35665267 and/or the novel variant

Fig. 2 Refined regional association plot for rs35665267 in the LF-SINE across a 10 kb window. Association with CBE of individual SNPs in the initial meta-analysis GWAS¹⁴ is plotted as $-\log_{10}(p)$ against the chromosomal position (hg19). The y-axis on the right shows the recombination rate estimated from the 1,000 Genomes (Mar 2012) EUR populations. All p values (y-axis on the left) are from the meta-analysis. The purple diamond indicates SNP rs35665267. Gray dots denote missing data due to their absence in LocusZoom. CBE, classic bladder exstrophy; GWAS, genome-wide association studies; LF, lobe-finned fish; SINE, short interspersed repetitive element; SNP, single nucleotide polymorphism.

may represent *cis*-regulatory variants that exhibit altered binding affinities in the LF-SINE enhancer. Such mechanism has been detected at the endometriosis risk locus, where the “protective” G-allele of rs17761446 was associated with preferential binding affinities to p300 and transcription factor 7 like-2 (TF7L2) and an increased expression of the *ANRIL* (antisense non-coding RNA in the INK4 locus) gene.³¹

Screening of the LF-SINE identified no direct disease-causing mutation, and it seems therefore unlikely, that such enhancer mutations are a frequent cause of CBE. However, our sample size might have been too small to detect rare causal mutational events. On the other side, we might have missed an enhancer deletion on one allele, a feature not detectable by the method applied. It can also be speculated that the close paralog of *ISL1*, the *ISL2* gene on chromosome 15, can at least partially compensate for impaired *ISL1* function, since homo- and heterodimerization of *ISL1* and *ISL2* has been reported.³² Interestingly, the *ISL2* gene is also flanked by an LF-SINE at similar distance, supporting the importance of this enhancer sequence.³³ Contrary, the *ISL1* enhancer may exert a relevant function in mouse but not in human, an interpretation that relies on the studies of Dermitzakis and Clark.³⁴ From their studies, these authors estimated that 32 to 40% of human functional transcription factor binding sites are not functional in rodents, also indicating the converse situation of functionality in mouse but not in human.

Finally, from our sequencing results and although rs35665267 shows a *p* value of 0.741, this does not necessarily exclude this SNP and the LF-SINE from involvement in CBE. Many complex traits, as CBE, are likely to be the result of the interplay of genes and environmental factors.^{7,35} Given this observation, variations in the enhancer may have only a marginal effect, unlikely to be detected from a GWAS. Hence, sequence analysis in more CBE patients is required to address its role in the etiology of CBE.

Acknowledgments

The authors thank the German self-help organization for their assistance in the recruitment of individuals with CBE (Selbsthilfegruppe Blasenexstrophie/Epispadie e.V.). M.L. and H.R. are members of the “Network for the Systematic Investigation of the Molecular Causes, Clinical Implications, and Psychosocial Outcome of Congenital Uro-Rectal Malformations (CURE-NET).” R.Z. and H.R. are supported by a grant from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG; grant number RE 1723/1-1).

References

- Husman DA, Vandersteen DR. Anatomy of the cloacal exstrophy. In: Gearhart JP, Mathews R, eds. *The Exstrophy-Epispadias Complex*. New York, NY: Kluwer Academic/Plenum Publishers; 1999:199–206
- Hurst JA. Anterior abdominal wall defects. In: Firth HV, Hall JG, eds. *Oxford Desk Reference: Clinical Genetics*. New York, NY: Oxford University Press; 2012:566
- Gearhart JP, Jeffs RD. Exstrophy-epispadias complex and bladder anomalies. In: Walsh PC, Retik AB, Vaughan ED, Wein AJ, eds. *Campbell's Urology*. 7th ed. Philadelphia, PA: WB Saunders; 1998: 1939–1990
- Reutter H, Boyadjiev SA, Gambhir L, et al. Phenotype severity in the bladder exstrophy-epispadias complex: analysis of genetic and nongenetic contributing factors in 441 families from North America and Europe. *J Pediatr* 2011;159(05):825–831.e1
- Reutter H, Shapiro E, Gruen JR. Seven new cases of familial isolated bladder exstrophy and epispadias complex (BEEC) and review of the literature. *Am J Med Genet A* 2003;120A(02): 215–221
- Ludwig M, Ching B, Reutter H, Boyadjiev SA. Bladder exstrophy-epispadias complex. *Birth Defects Res A Clin Mol Teratol* 2009; 85(06):509–522
- Reutter H, Qi L, Gearhart JP, et al. Concordance analyses of twins with bladder exstrophy-epispadias complex suggest genetic etiology. *Am J Med Genet A* 2007;143A(22):2751–2756
- von Lowtzow C, Hofmann A, Zhang R, et al. CNV analysis in 169 patients with bladder exstrophy-epispadias complex. *BMC Med Genet* 2016;17(01):35. Doi: 10.1186/s12881-016-0299-x
- Draaken M, Reutter H, Schramm C, et al. Microduplications at 22q11.21 are associated with non-syndromic classic bladder exstrophy. *Eur J Med Genet* 2010;53(02):55–60
- Lundin J, Söderhäll C, Lundén L, et al. 22q11.2 microduplication in two patients with bladder exstrophy and hearing impairment. *Eur J Med Genet* 2010;53(02):61–65
- Draaken M, Baudisch F, Timmermann B, et al. Classic bladder exstrophy: Frequent 22q11.21 duplications and definition of a 414 kb phenocritical region. *Birth Defects Res A Clin Mol Teratol* 2014;100(06):512–517
- Wilkins S, Zhang KW, Mahfuz I, et al. Insertion/deletion polymorphisms in the Δ Np63 promoter are a risk factor for bladder exstrophy-epispadias complex. *PLoS Genet* 2012;8(12):e1003070. Doi: 10.1371/journal.pgen.1003070
- Qi L, Wang M, Yagnik G, et al. Candidate gene association study implicates p63 in the etiology of nonsyndromic bladder-exstrophy-epispadias complex. *Birth Defects Res A Clin Mol Teratol* 2013;97(12):759–763
- Draaken M, Knapp M, Pennimpede T, et al. Genome-wide association study and meta-analysis identify *ISL1* as genome-wide significant susceptibility gene for bladder exstrophy. *PLoS Genet* 2015;11(03):e1005024. Doi: 10.1371/journal.pgen.1005024
- Karlsson O, Thor S, Norberg T, Ohlsson H, Edlund T. Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain. *Nature* 1990;344(6269):879–882
- Zhuang S, Zhang Q, Zhuang T, Evans SM, Liang X, Sun Y. Expression of *Isl1* during mouse development. *Gene Expr Patterns* 2013; 13(08):407–412
- Bejerano G, Lowe CB, Ahituv N, et al. A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* 2006;441(7089):87–90
- Tsuchida T, Ensini M, Morton SB, et al. Topographic organization of embryonic motor neurons defined by expression of LIM homeobox genes. *Cell* 1994;79(06):957–970
- Uemura O, Okada Y, Ando H, et al. Comparative functional genomics revealed conservation and diversification of three enhancers of the *isl1* gene for motor and sensory neuron-specific expression. *Dev Biol* 2005;278(02):587–606
- Kim N, Park C, Jeong Y, Song MR. Functional diversification of motor neuron-specific *Isl1* enhancers during evolution. *PLoS Genet* 2015; 11(10):e1005560. Doi: 10.1371/journal.pgen.1005560
- Hubbard T, Andrews D, Caccamo M, et al. Ensembl 2005. *Nucleic Acids Res* 2005;33(Database issue):D447–D453
- ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;489(7414):57–74
- Dyson HJ, Wright PE. Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators

- CREB-binding protein (CBP) and p300. *J Biol Chem* 2016;291(13): 6714–6722
- 24 Reutter H, Draaken M, Pennimpede T, et al. Genome-wide association study and mouse expression data identify a highly conserved 32 kb intergenic region between WNT3 and WNT9b as possible susceptibility locus for isolated classic exstrophy of the bladder. *Hum Mol Genet* 2014;23(20):5536–5544
 - 25 Siepel A, Bejerano G, Pedersen JS, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 2005;15(08):1034–1050
 - 26 Blanchette M, Kent WJ, Riemer C, et al. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res* 2004;14(04):708–715
 - 27 Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science* 2008;322(5903):881–888
 - 28 Mathelier A, Shi W, Wasserman WW. Identification of altered cis-regulatory elements in human disease. *Trends Genet* 2015;31(02): 67–76
 - 29 Slepak TI, Webster KA, Zang J, et al. Control of cardiac-specific transcription by p300 through myocyte enhancer factor-2D. *J Biol Chem* 2001;276(10):7575–7585
 - 30 Schmidt D, Wilson MD, Ballester B, et al. Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* 2010;328(5981):1036–1040
 - 31 Nakaoka H, Gurumurthy A, Hayano T, et al. Allelic imbalance in regulation of *ANRIL* through chromatin interaction at 9p21 endometriosis risk locus. *PLoS Genet* 2016;12(04):e1005893. Doi: 10.1371/journal.pgen.1005893
 - 32 Jurata LW, Pfaff SL, Gill GN. The nuclear LIM domain interactor NLI mediates homo- and heterodimerization of LIM domain transcription factors. *J Biol Chem* 1998;273(06): 3152–3157
 - 33 McEwen GK, Woolfe A, Goode D, Vavouri T, Callaway H, Elgar G. Ancient duplicated conserved noncoding elements in vertebrates: a genomic and functional analysis. *Genome Res* 2006;16(04): 451–465
 - 34 Dermitzakis ET, Clark AG. Evolution of transcription factor binding sites in Mammalian gene regulatory regions: conservation and turnover. *Mol Biol Evol* 2002;19(07):1114–1121
 - 35 Ludwig M, Reutter H. Genome-wide array data and next generation sequencing unravel the etiology of urogenital malformations. *J Pediatr Genet* 2012;1(04):209–216