PLOS | GENETICS

RESEARCH ARTICLE

Genome-wide Association Study and Meta-Analysis Identify *ISL1* as Genome-wide Significant Susceptibility Gene for Bladder Exstrophy

Markus Draaken[1,2]*, Michael Knapp[3]*, Tracie Pennimpede[4]*, Johanna M. Schmidt[1], Anne-Karolin Ebert[5], Wolfgang Rösch[6], Raimund Stein[7], Boris Utsch[8], Karin Hirsch[9], Thomas M. Boemers[10], Elisabeth Mangold[1], Stefanie Heilmann[1,2], Kerstin U. Ludwig[1,2], Ekkehart Jenetzky[11,12], Nadine Zwink[11], Susanne Moebus[13], Bernhard G. Herrmann[4], Manuel Mattheisen[14,15,16], Markus M. Nöthen[1,2], Michael Ludwig[17☯], Heiko Reutter[1,18☯]*

1 Institute of Human Genetics, University of Bonn, Bonn, Germany, **2** Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany, **3** Institute of Medical Biometry, Informatics, and Epidemiology, University of Bonn, Bonn, Germany, **4** Department of Developmental Genetics, Max Planck Institute for Molecular Genetics, Berlin, Germany, **5** Department of Urology and Pediatric Urology, University Hospital of Ulm, Germany, **6** Department of Pediatric Urology, St. Hedwig Hospital Barmherzige Brüder, Regensburg, Germany, **7** Department of Urology, Division of Pediatric Urology, University of Mainz, Mainz, Germany, **8** Department of General Pediatrics and Neonatology, Justus Liebig University, Giessen, Germany, **9** Department of Urology, Division of Paediatric Urology, University of Erlangen-Nürnberg, Erlangen, Germany, **10** Department of Pediatric Surgery and Pediatric Urology, Children's Hospital of Cologne, Cologne, Germany, **11** Department of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany, **12** Department of Child and Adolescent Psychiatry and Psychotherapy, Johannes-Gutenberg University, Mainz, Germany, **13** Institute of Medical Informatics, Biometry, and Epidemiology, University Hospital of Essen, University Duisburg-Essen, Essen, Germany, **14** Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts, United States of America, **15** Department of Biomedicine, Aarhus University, Aarhus, Denmark, **16** Department of Genomic Mathematics, University of Bonn, Bonn, Germany, **17** Department of Clinical Chemistry and Clinical Pharmacology, University of Bonn, Bonn, Germany, **18** Department of Neonatology, Children's Hospital, University of Bonn, Bonn, Germany

☯ These authors contributed equally to this work.
* reutter@uni-bonn.de

OPEN ACCESS

Citation: Draaken M, Knapp M, Pennimpede T, Schmidt JM, Ebert A-K, Rösch W, et al. (2015) Genome-wide Association Study and Meta-Analysis Identify *ISL1* as Genome-wide Significant Susceptibility Gene for Bladder Exstrophy. PLoS Genet 11(3): e1005024. doi:10.1371/journal.pgen.1005024

Editor: Wei Cheng, Monash University, AUSTRALIA

Received: August 20, 2014

Accepted: January 26, 2015

Published: March 12, 2015

Copyright: © 2015 Draaken et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: HR is supported by grant RE 1723/1-1 from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). MMN received support from the Alfried Krupp von Bohlen und Halbach-Stiftung, and is a member of the DFG-funded Excellence Cluster Immunosensation. Controls for our GWAS analyses were drawn from the Heinz Nixdorf Recall Study (HNR) cohort, which was established with the support of the Heinz Nixdorf Foundation. The funders had no role in study design,

Abstract

The bladder exstrophy-epispadias complex (BEEC) represents the severe end of the uro-rectal malformation spectrum, and is thought to result from aberrant embryonic morphogenesis of the cloacal membrane and the urorectal septum. The most common form of BEEC is isolated classic bladder exstrophy (CBE). To identify susceptibility loci for CBE, we performed a genome-wide association study (GWAS) of 110 CBE patients and 1,177 controls of European origin. Here, an association was found with a region of approximately 220kb on chromosome 5q11.1. This region harbors the *ISL1* (*ISL LIM homeobox 1*) gene. Multiple markers in this region showed evidence for association with CBE, including 84 markers with genome-wide significance. We then performed a meta-analysis using data from a previous GWAS by our group of 98 CBE patients and 526 controls of European origin. This meta-analysis also implicated the 5q11.1 locus in CBE risk. A total of 138 markers at this locus reached genome-wide significance in the meta-analysis, and the most significant marker

data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

(rs9291768) achieved a P value of 2.13×10^{-12}. No other locus in the meta-analysis achieved genome-wide significance. We then performed murine expression analyses to follow up this finding. Here, Isl1 expression was detected in the genital region within the critical time frame for human CBE development. Genital regions with *Isl1* expression included the peri-cloacal mesenchyme and the urorectal septum. The present study identified the first genome-wide significant locus for CBE at chromosomal region 5q11.1, and provides strong evidence for the hypothesis that *ISL1* is the responsible candidate gene in this region.

Author Summary

The etiology of classic exstrophy of the bladder (CBE) remains unclear. The present genome-wide association study and meta-analysis identified an association between CBE and a region on chromosome 5q11.1. This region contains the gene encoding insulin gene enhancer protein, *ISL-1*. In this region, 138 single nucleotide polymorphisms (SNPs) reached genome-wide significance, with the SNP rs9291768 showing the lowest *P* value ($p = 2.13 \times 10^{-12}$). Our findings, as supported by expression analyses in murine models, suggest that *ISL1* is a susceptibility gene for CBE.

Introduction

The bladder exstrophy-epispadias complex (BEEC; OMIM %600057) is the most severe of all human congenital anomalies of the kidney and urinary tract (CAKUT), and involves the abdominal wall, pelvis, all of the urinary tract, the genitalia, and occasionally the spine and anus. The severity-spectrum of the BEEC comprises the mildest form, epispadias (E); the intermediate form, classic bladder exstrophy (CBE); and the most severe form, exstrophy of the cloaca (CE) [1,2]. Despite advances in surgical techniques and improved understanding of the underlying anatomical defects, in later life many male and female patients experience chronic upper and lower urinary tract infections, sexual dysfunction, and urinary-, or in the case of cloacal exstrophies, urinary and fecal incontinence [3,4]. The estimated overall birth prevalence for the complete BEEC spectrum in children of European descent is 1 in 10 000 [5]. Birth prevalence, as assessed with the inclusion of terminated pregnancies, differs between subtypes. Estimated rates are: 1 in 117,000 in males and 1 in 484,000 in females for E [6]; 1 in 37,000 for CBE [6]; and 1 in 200,000 to 1 in 400,000 for CE [7]. According to the Birth Defects Monitoring Program of the Centers for Disease Control and Prevention, the birth prevalence of CBE among North American ethnic groups varies, with the highest birth prevalence being observed among Native Americans (8 in 100,000), and the lowest among Asians (1 in 100,000) [8]. Although BEEC can occur as part of a complex malformation syndrome, approximately 98.5% of cases are classified as isolated [9]. The reported recurrence risk for CBE among siblings in families with non-consanguineous and non-affected parents ranges between 0.3–2.3%, whereas the reported recurrence risk for the offspring of affected patients is 1.4% [10–12]. Hence, the recurrence risk for the offspring of CBE patients shows an approximate 400-fold increase compared to that observed in the general population [10]. Identification of genetic risk factors for the BEEC has been the subject of extensive recent research, and several lines of evidence support the hypothesis that genetic factors are implicated. These include reports of BEEC-associated chromosomal aberrations [13]; reports of at least 30 families with multiple affected members [13,14]; and observations of high concordance rates in monozygotic twins [5]. Array-based

molecular karyotyping and regional association studies have implicated micro-duplications on chromosome 22q11.21 and polymorphisms in the *TP63* (*Tumor protein p63*) gene [15–19]. However, in the vast majority of cases, the genetic contribution to the BEEC remains elusive, and the molecular basis of the disruption of the respective developmental processes is poorly understood.

The aim of the present study was to identify susceptibility loci for CBE. Firstly, we conducted a genome-wide association study (GWAS) of 110 isolated CBE patients and 1,177 controls of European descent. Secondly, we performed a meta-analysis using the data from step 1 and data from our previous GWAS of 98 CBE patients and 526 controls [20]. Thirdly, we followed up our main finding by: (i) re-sequencing *ISL-1* (*ISL LIM homeobox 1*), the main candidate gene within the region of genome wide significance on chromosome 5q11.1, in all patients; and (ii) performing murine expression analyses.

Results and Discussion

In the subsequent text, our previous GWAS [20] is termed GWAS1 and the present GWAS is termed GWAS2.

The post quality control data set of GWAS2 comprised 110 CBE patients and 1,177 controls. The GWAS2 analyses identified a region of approximately 220 kb on chromosome 5q11.1. This region harbors the gene *ISL1*. Multiple markers in this region showed evidence for association with CBE (S1 Table). The most significant marker, rs6874700, showed a P value of 6.27×10^{-11}. The significance of this marker was supported by the presence of 172 surrounding markers with P values of $< 10^{-5}$. A total of 84 markers at this locus, including rs6874700, reached genome-wide significance, i.e. $P < 5 \times 10^{-8}$. No other locus in the GWAS2 analyses achieved this level of significance.

Next, we combined the effect estimates of GWAS1 and GWAS2 in a fixed effect meta-analysis. This meta-analysis also implicated the 220 kb region on chromosome 5q11.1. In the meta-analysis, multiple markers in this region showed evidence for association with CBE (Fig. 1). The most significant marker, rs9291768, had a P value of 2.13×10^{-12}. The possible relevance of rs9291768 in CBE was supported by the presence of 137 surrounding markers with P values of $< 5 \times 10^{-8}$. No other locus in the meta-analysis achieved this level of significance (Fig. 2). All markers with P values of $< 10^{-5}$ are listed in S2 Table.

The genotype-specific relative risks (RRs) for allele T of rs9291768 were: (i) RR_het = 2.00 for heterozygotes (95%-CI = 1.33–3.02); and (ii) RR_hom = 4.77 (95%-CI = 3.06–7.45) for homozygotes. This is compatible with neither a recessive ($P = 3.9 \times 10^{-5}$), nor a dominant mode of inheritance ($P = 1.6 \times 10^{-4}$). According to the genotype data from Ensembl release 74—December 2013©, the frequency of the CBE allele T at rs9291768 is highest in African populations (0.534), intermediate in European populations (0.425), and lowest in Asian populations (0.083).

Our previous study [20] failed to identify the possible relevance of both marker rs9291768 and the region comprising *ISL1*. In that report, rs9291768 obtained a P value of 1.1×10^{-3}, which is not considered worthy of note in the context of a GWAS. In the present meta-analysis, the estimated relative risk for this SNP was 2.18 (see Table 1). The GWAS1 sample comprised 98 cases and 526 controls, and the power to achieve genome-wide significance (i.e. $< 5 \times 10^{-8}$) for a SNP with RR = 2.18 was only 31% under the assumption of a multiplicative model and a minor allele frequency of 0.377. For GWAS2, which comprised 110 cases and 1177 controls, the power was higher at 53%. However, the combination of GWAS1 and GWAS2 provides a power of 98%. This substantial increase in power is the central motivation for conducting

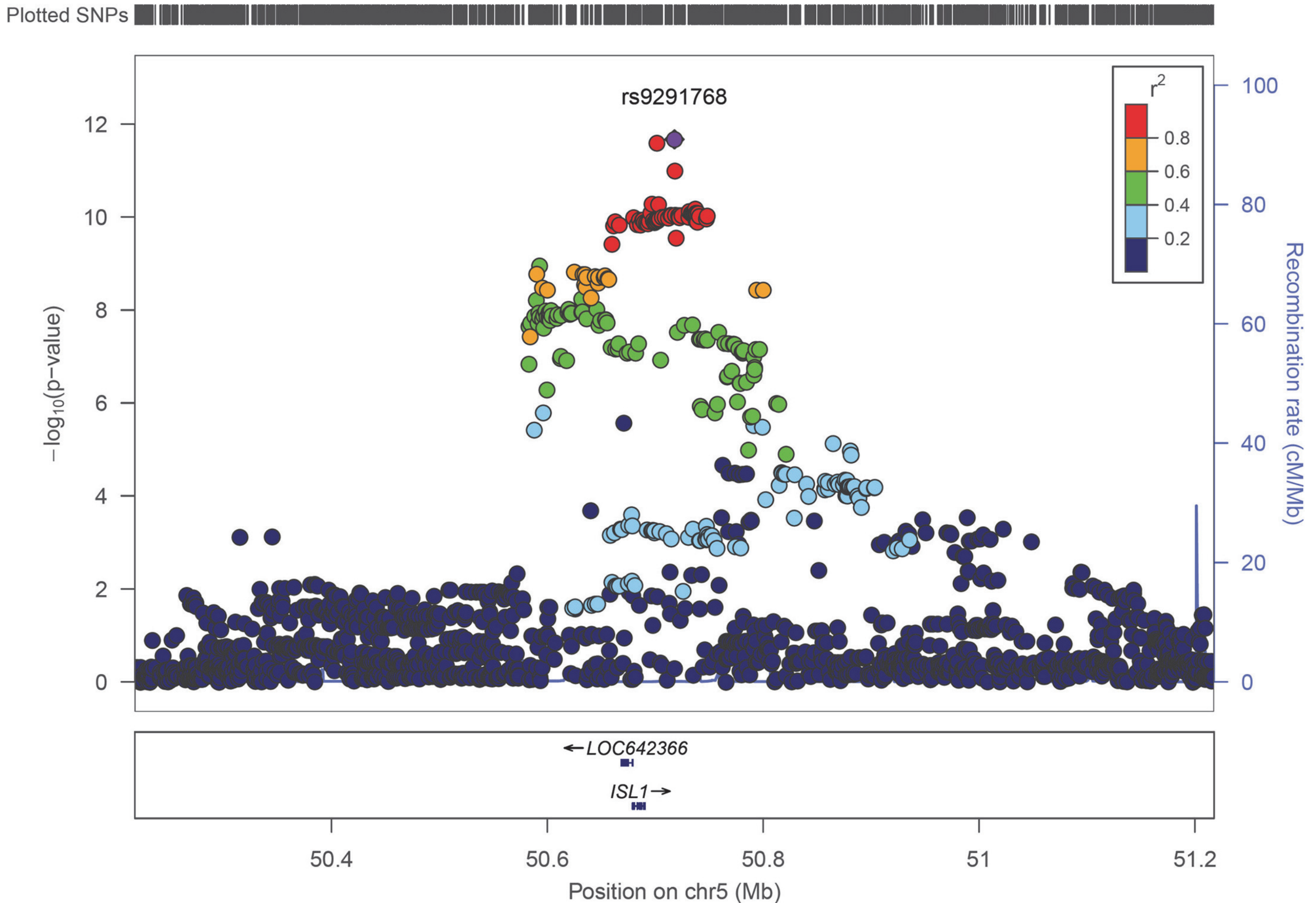

Fig 1. Regional association plot for *ISL1* across a 1.0 Mb window. Association with classic bladder exstrophy of individual SNPs in the meta-analysis GWAS is plotted as $-\log_{10}(p)$ against chromosomal position. The y-axis on the right shows the recombination rate estimated from the 1000 Genomes (Mar 2012) EUR populations. All *P* values (y-axis on the left) are from the meta-analysis. The purple diamond indicates the most significant marker.

doi:10.1371/journal.pgen.1005024.g001

meta-analyses. We therefore assume that the non-identification of rs9291768 in GWAS1 was attributable to the issues of power and random sample variation.

The marker rs9291768 is a non-coding variant, which is located 27.2 kb downstream of the *ISL1* gene. The associated 220 kb haplotype block contains the gene *ISL1*. The only other transcript encoded in the regions flanking rs9291768 (500 kb on either side) is LOC642366, which also resides within the associated haplotype block. LOC642366 encodes an uncharacterized non-coding RNA, which has no ortholog in mouse, zebrafish, drosophila, *C. elegans*, or *S. cerevisiae*. The second most proximal gene to rs9291768 is *Homo sapiens* poly (ADP-ribose) polymerase family, member 8 (*PARP8*), which is located 575 kb proximal to rs9291768. The third and fourth genes are *Homo sapiens* pelota homolog (Drosophila) (*PELO*), and the integrin, alpha 1 subunit of integrin receptors (*ITGA1*), which are both located ∼1.4 Mb distal to rs9291768. According to Mouse Genome Informatics (http://www.informatics.jax.org/), neither *PARP8* nor *ITGA1* is expressed in the genital tubercle or the cloacal membrane during the CBE critical time frame in mouse embryos. Furthermore, mice with complete invalidation of *PARP8* or *ITGA1* display neither CBE features nor CBE-related phenotypes. Studies of the Drosophila pelota gene have implicated *dPelo* in spermatogenesis, mitotic division, and

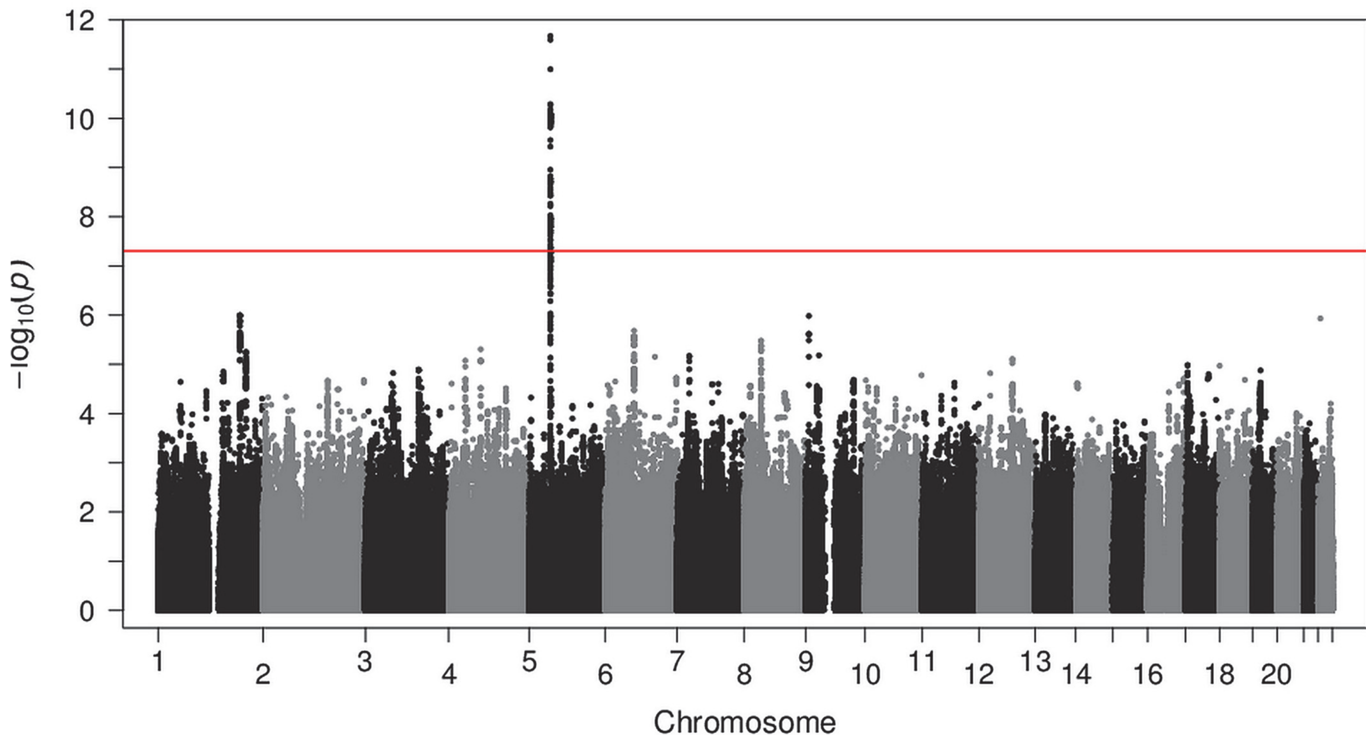

Fig 2. Genome-wide association scan in classic bladder exstrophy patients. Association of SNPs is plotted as $-\log_{10}$(p) against chromosomal position. The y-axis shows the negative $\log_{10}$ *P* values of the logistic regression for SNPs from the meta-analysis that passed quality control. Chromosomes are shown in alternating colors along the x-axis. The genome-wide significance level is indicated by a red line.

doi:10.1371/journal.pgen.1005024.g002

patterning. Homozygous *Pelo*-null embryos fail to develop beyond embryonic day 7.5, and exhibit no early CBE-related features, such as diastases of the symphysis. Whether rs9291768 per se, or a variant in linkage disequilibrium with it, confers the functional effect underlying the association remains unclear. The rs9291768 marker shows no association with any predicted regulatory sequence (according to ENCODE, TFSEARCH, or FAS-ESS), or splicing motif. Of the other 136 markers at this chromosome 5 locus, only one (rs2303751) is located in a coding region, and none affect a splice site. Marker rs2303751 is in linkage disequilibrium ($r^2 = 0.932$) with the most significant marker rs9291768, and represents a synonymous A/G substitution in exon four of *ISL1*. Furthermore, none of the public eQTL (expression quantitative trait loci) databases contains evidence to suggest that rs9291768, or a SNP in perfect linkage disequilibrium with it, would affect gene expression levels (RegulomeDB, http://regulome.stanford.edu; eQTL browser, eqtl.uchicago.edu).

The present meta-analysis generated no evidence in support of the (non-genome-wide significant) association between CBE and an intergenic region on chromosome 17q21.31-q21.32 identified in GWAS1 [20]. This region is located between the genes *WNT3* (wingless-

Table 1. Most strongly associated SNP in the bladder exstrophy susceptibility locus 5q11.1.

SNP	Chr	Position	Risk/other allel	RAF cases	RAF controls	RAF combined	RR [95% CI]	*P* value
rs9291768	5	50717793	T/C	0.576	0.377	0.399	2.18 [1.75–2.71]	2.13×10^{-12}

Relative risks (RRs) are given with the risk allele set as baseline. Chr, chromosome; RAF, risk allele frequency.

doi:10.1371/journal.pgen.1005024.t001

type MMTV integration site family, member 3) and *WNT9b* (wingless-type MMTV integration site family, member 9b).

Since the CBE-associated region harbors the gene *ISL1*, we performed *ISL1* re-sequencing in 207 CBE patients included in the present meta-analysis. As well as allowing mutation detection, this approach should provide genotype data for polymorphisms in the exons and exon-flanking regions of *ISL1*. Using the results of Sanger sequencing, we compared genotype information from four SNPs with the imputed data. We calculated the allelic accuracy, i.e. the aggregate difference between the actual number of alleles observed and the number of imputed alleles [21]. This yielded accuracy values of 96.9% (rs150104955); 97.3% (rs2288468); 97.3% (rs2303751); and 99.5% (rs3917084). Two of these SNPs (rs2288468, rs2303751) achieved genome-wide significance in the meta-analysis

Although sequencing identified no nonsense or probably pathogenic *ISL1* variant, the following variants were all detected in a heterozygote state in single patients: intron 3, rs2303750; synonymous in exon 5, rs41268419 (p.Ser275 =); non-synonymous in exon 4, rs200209474 (p. Thr181Ser); unreported variants in intron 4, +21delG, -19delT, and -64A>G. Pathogenicity prediction using several publicly available algorithms (SNPs&Go, MutPred, SIFT) predicted that the p.Thr181Ser variant is neutral. Only PolyPhen-2 estimated it as possibly damaging. Furthermore, all of the observed intronic variants can be assumed to be benign. Hence, our patient sample size may have been too small to detect rare causal mutational events. We cannot exclude the possibility that some mutations were overlooked, i.e. mutations located in the promoter region, in as-yet-unknown regulatory sequences, or in non-coding regions that were not present within the covered sequence.

ISL1 encodes the insulin gene enhancer protein ISL1, a LIM zinc-binding/homeobox-domain transcription factor which was initially identified as a regulator of insulin expression [22]. Research in rodents suggests that Isl1 plays a fundamental role in the embryogenesis of multiple tissue types: Isl1 affects cell differentiation and survival, cell fate determination, the generation of cell diversity, and segmental patterning during mouse development [23]. Isl1 binds and regulates the promoters of the glucagon and somatostatin genes, and activates insulin gene transcription in pancreatic beta cells in synergy with *NEUROD1* (neuronal differentiation 1) [24]. A previous study found an association between a heterozygous *ISL1* premature termination mutation (p.Gln310*) and diabetes type II in a large Japanese kindred [25]. Furthermore, in a classic linkage analysis of 186 Swedish multiplex families with diabetes type I, linkage was observed with chromosomal region 5q11-q13, which harbors *ISL1* [26]. This finding supports the hypothesis that ISL1 is implicated in pancreatic function and development, as reported in *Isl1* knockout mice [27]. In the mouse, research at E (embryonic day) 8.5 to E9.5 has shown that ISL1 acts upstream of the sonic hedgehog (Shh) signaling pathway [28], which may be involved in other processes besides the coordination of heart and lung co-development [29]. Interestingly, a recent report by Matsumaru et al. [30] showed that SHH is also important for ventral body wall formation, and that ectopic SHH signaling induces omphalocele, a feature which is associated with CE, the severest form of the BEEC.

A previous study in mice also showed that a homozygous *Isl1* null mutation (*Isl1*$^{-/-}$) induced growth retardation at E9.5 and severe cardiac malformations at E10.5 [31]. Embryos exhibiting these severe cardiac malformations at E10.5 died at E11.5 due to the developmental arrest of spinal motor neurons [32]. Research has demonstrated a further role for Isl1 in mice at E11.5, i.e. in hindlimb-specific patterning and growth in combination with both SHH and the helix-loop-helix transcription factor HAND2 (heart and neural crest derivatives expressed 2) [33]. This interplay is also necessary for normal cardiac development in mice [23]. Recently, Jurberg et al. reported that specific activity of mouse *Isl1* in the progenitors of the ventral lateral mesoderm promotes formation of the cloaca-associated mesoderm as the most posterior derivatives

of lateral mesoderm progenitors [34]. This observation provides independent evidence that *ISL1* is a promising candidate gene for human CBE.

In a recent mouse study, Kaku et al. induced conditional *Isl1* deletion in the lateral mesoderm using a *Hoxb6-Cre* driver, and demonstrated that this caused kidney agenesis or hydroureter [35]. The authors observed transient *Isl1* expression between E10.5—E14.5. At early stages, this was observed in the mesenchyme surrounding the ureteric stalk and cloaca. At later stages, expression occurred along the nephric duct, at the base of the ureteric stalk, and in the genital tubercle. This suggests that *Isl1* may be implicated in kidney, ureter, and bladder development. These mice show a variable phenotype, which can include agenesis of the genital tubercle (R. Nishinakamura, personal communication). The variability of this defect is probably due to mosaicism, which arises as a result of the *Hoxb6-Cre* driver [33]. Kaku et al. also reported that conditional loss of *Isl1* resulted in a concomitant reduction in the expression of bone morphogenetic protein 4 (Bmp4) [35]. Using mouse *Isl1Cre;Bmp4*$^{flox/flox}$ mutants, Suzuki et al. showed that BMP4 signaling in the caudal *Isl1* expression domain was required for formation of the anterior peri-cloacal mesenchyme (aPCM) at E10.5 [36]. Rather than decreasing Isl1 function, loss of this signal caused defective pelvic and urogenital organ formation, including kidney and bladder agenesis, with abnormal development of the lower limbs and pelvis. Moreover, tissue lineage analyses suggested that *Isl1*-expressing cells are an essential cell population in terms of caudal body formation, including the pelvic/urogenital organs and hindlimb [36].

In the present mouse analyses, *Isl1* was expressed during the critical timeframe for development of tissues involved in CBE, and strong *Isl1* expression was detected in the developing genital region (Fig. 3). From E9.5, a broad *Isl1* domain was detected in the cloacal region. This was maintained in the outgrowing genital tubercle (including the urorectal septum) until at least E14.5.

Three groups have recently elucidated the molecular basis of the Danforth's short tail (*Sd*) mouse. They reported the insertion of a retrotransposon in the 5' regulatory domain of the murine *Ptf1a* gene, which encodes pancreas specific transcription factor 1A [37–39]. As a consequence, and in contrast to their wildtype littermates, *Sd* mice showed ectopic *Ptf1a* expression in the notochord and hindgut at E8.5 to E9.5, which extended to the cloaca and mesonephros at E10.5 and to the pancreatic bud at E10.5 and E11.5 [38]. The resultant phenotype of this *Sd* mutation mirrors the phenotype observed in human caudal malformation syndromes, a phenotype that is also observed in *Isl1* transgenic mice [40]. Moreover, the BEEC related human Currarino syndrome (MIM: #176450), which comprises hemisacrum, anorectal malformations, and a presacral mass, is caused by mutations of the transcription factor *MNX1*/*HLXB9* (Motor neuron and pancreas homeobox protein 1). The genes *Ptf1a*, *Isl1*, and *Mnx1* have been implicated in pancreas development, and MNX1 has been identified as a direct target of PTF1a [41].

Coordinated development of caudal body structures is necessary for the formation of the bladder, rectum, and the external genitalia [42,43]. These organs are derived from the transient embryonic cloaca and the PCM, an infra-umbilical mesenchyme [44], as well as the anterior PCM [36,42]. SHH-, ISL1-, and BMP4-expressing cells contribute to both the PCM and the anterior PCM [36,42,45]. Perturbation of this morphoregulatory network may thus lead to malformation of caudal structures, including the bladder, rectum, and external genitalia.

In summary, the present report describes a novel association between *ISL1* and human CBE. While previous conventional linkage- and candidate gene studies in humans have suggested the involvement of *ISL1* in diabetes type I and II and in congenital heart defects, to our knowledge, the present study is the first to implicate *ISL1* in the formation of human urogenital malformations. The observed variation in CBE birth prevalence across populations is

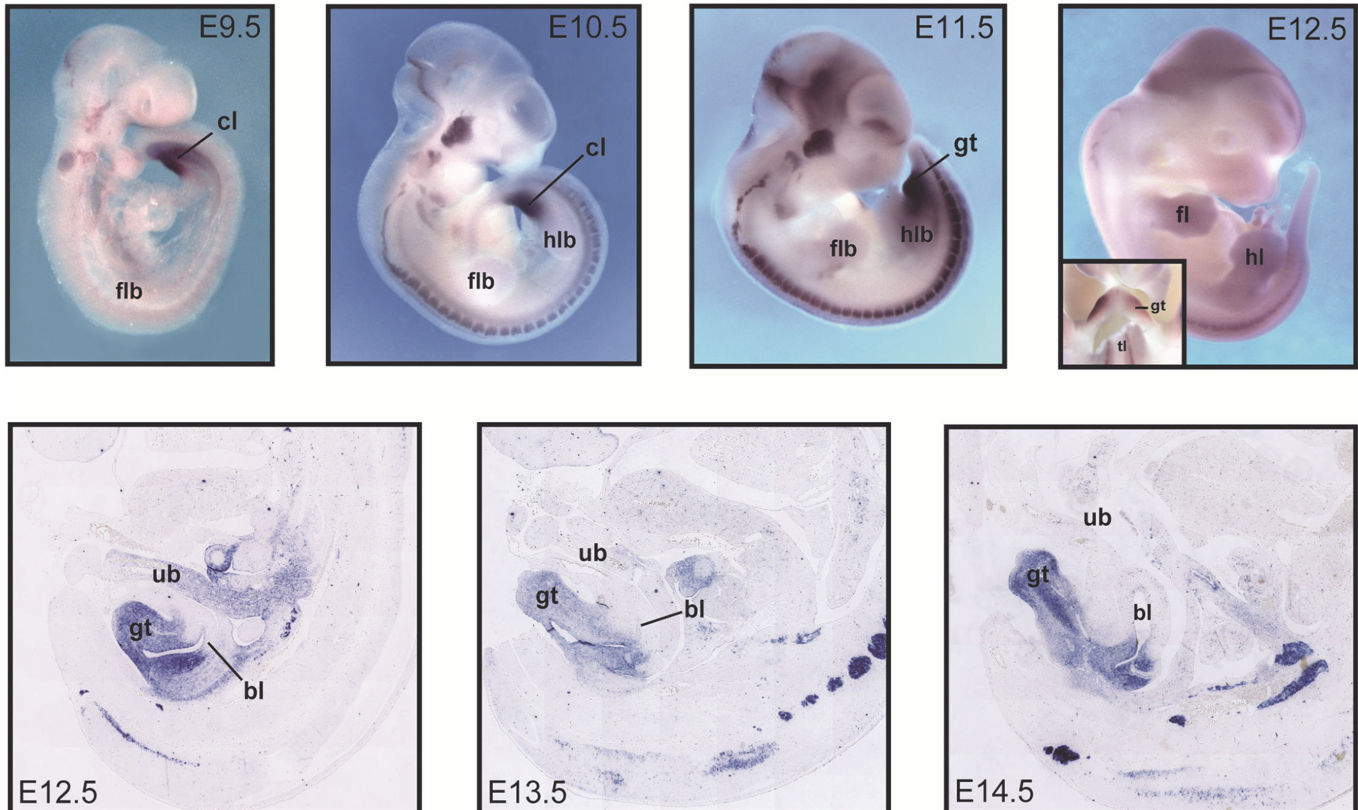

Fig 3. Expression of *Isl1* during mouse development. Whole-mount *in situ* hybridization (ISH) for *Isl1* in wildtype mouse embryos between E9.5-E12.5 revealed strong expression in the developing genital region, including the cloaca, cloacal membrane, and emerging genital tubercle. ISH on mid-sagittal paraffin sections at later embryonic stages (E12.5-E14.5) revealed expression throughout the genital tubercle, within the peri-cloacal mesenchyme and urorectal septum. *Isl1* was also detected in the craniofacial- and spinal ganglia.

doi:10.1371/journal.pgen.1005024.g003

consistent with the cross-population frequencies of the rs9291768 T-allele, thus supporting our finding.

The importance of *Pft1a* and *Isl1* in the formation of murine genital development and caudal regression phenotypes, the involvement of *MNX1* in the BEEC related human Currarino syndrome, and the role of all three genes in pancreatic development suggest that these genes are involved in a common pathway. However, the present data do not exclude the possibility that the association between CBE and the region surrounding *ISL1* is attributable to long-range functional interactions with other regions in the human genome. Future studies are warranted to identify the mechanisms through which genetic variation at *ISL1* contributes to CBE development.

Methods

Subjects GWAS2

The initial GWAS2 sample comprised 123 isolated CBE patients and 1,320 controls of European descent. Prior to inclusion, written informed consent was obtained from all subjects, or from their proxies in the case of legal minors. For patients and controls, demographic

information was collected using a structured questionnaire. This study was approved by the institutional ethics committee of each participating center, and was conducted in accordance with the principles of the Declaration of Helsinki. All CBE patients were recruited in person by experienced physicians trained in the diagnosis of the BEEC. Details of the recruitment process for patients and controls are provided in Reutter et al. [20].

Genotyping

For the 123 isolated CBE patients in GWAS2, genotyping was performed using the Illumina BeadChip HumanOmniExpress (San Diego, California, USA), and DNA was extracted from blood or saliva using standard procedures. Case-control comparisons were made using the genotypes of 1,320 population-based controls, which had been processed using the same array [46]. Genome-wide genotyping of 730,525 markers was conducted using the Infinium HD Ultra Assay from Illumina (Illumina, San Diego, California, USA).

Pre-imputation quality control of GWAS2

Markers were excluded from the analysis if: (i) the minor allele frequency was <1% or the call rate was <95% in either cases or controls; or (ii) the test for Hardy-Weinberg equilibrium resulted in $P<10^{-4}$ in the control sample or $P<10^{-6}$ in the case sample. A total of 616,799 autosomal markers fulfilled these quality criteria. Individuals were excluded if their call rate was <99%, or if they were outliers in a multidimensional scaling (MDS) analysis. Relatedness of individuals within GWAS2, and between GWAS1 and GWAS2, was evaluated using both the KING program [47], and an identity-by-state-based in-house program. The post quality control data set of GWAS2 comprised 110 CBE patients and 1,177 controls.

Imputation

GWAS1 and GWAS2 were imputed separately to the 1000 Genomes Project and HapMap 3 reference panels using IMPUTE2 [48].

Post-imputation quality control

For each of the three data sets, variants were excluded if: (i) the imputation info score was <0.4; (ii) the dosage of the minor allele was <1% in either cases or controls; (iii) the test for Hardy-Weinberg equilibrium (calculated on the basis of the 80% best-guess genotypes) resulted in $P<10^{-4}$ in the control sample; or (iv) the 80% best-guess genotypes were only available for <80% of cases or controls. In total, 7,261,187 SNPs were analyzed in at least one data set.

Statistical analysis

Single-marker analysis was performed using logistic regression. The allele dosage and the first five components obtained from MDS were used as independent variables for the variants in the three data sets. The effect estimates for the data sets were then combined in an inverse variance-weighted fixed-effects meta-analysis. The genomic inflation factor in this meta-analysis was 1.0196.

Power

Power was calculated to enable detection of genome-wide significance ($P<5 \times 10^{-8}$) in the combined analysis of the GWAS1 and GWAS2 samples. Under the assumption of a multiplicative model, this was 80% for an allele frequency of 0.35 (0.20) and a RR of 1.94 (2.05). This is

within the range of RRs observed for other multifactorial, nonsyndromic human malformations. For example, the power of the present study to detect a locus with an effect-strength similar to that of the most strongly associated locus in nonsyndromic cleft lip with or without cleft palate was 99.2% [49].

ISL-1 resequencing

Sequence analysis of the complete *ISL-1* coding regions and their splice consensus motifs was performed in 207 of our 208 CBE patients using standard techniques. Primers are listed in S3 Table. For the remaining patient, no additional DNA sample was available. During this analysis, we also obtained information for several SNPs deposited in dbSNP Build142 (rs3917084, rs150104955, rs2288468, rs2303750, rs2303751, rs200209474, and rs41268419).

In situ hybridization of mouse embryo sections

The expression of Isl1 was analyzed using in situ hybridization, standard procedures, and a $\sim$450bp antisense probe spanning exons 2 and 3 from XM_006517533.1. Details of the in situ hybridization methods are provided elsewhere [50].

Supporting Information

S1 Table. All markers identified in GWAS 2 within chromosomal region 5q11.1 with *P* values $< 10^{-5}$. The most significant marker, rs6874700, showed a *P* value of 6.27×10^{-11}. Relative risks (RRs) are given with the risk allele set as baseline. Chr, chromosome; RAF, risk allele frequency.
(PDF)

S2 Table. All markers identified in the meta-analysis of GWAS 1 and GWAS 2 over the entire genome with *P* values of $< 10^{-5}$. Relative risks (RRs) are given with the risk allele set as baseline. Chr, chromosome; RAF, risk allele frequency.
(PDF)

S3 Table. Primers (5'->3' direction) used for *ISL1* sequence analysis. F, forward; R, reverse.
(PDF)

Acknowledgments

We thank the German self-help organization for their assistance in the recruitment of individuals with BEEC (Selbsthilfegruppe Blasenekstrophie / Epispadie e.V.). MD, TP, AKE, WR, RS, EJ, NZ, BGH, MMN, ML, and HR 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)". We thank Peter Teßmann for his excellent technical assistance.

Author Contributions

Conceived and designed the experiments: HR MD MK TP MMN. Performed the experiments: MD TP JMS. Analyzed the data: MD TP JMS MK MM ML. Contributed reagents/materials/analysis tools: MK AKE WR RS BU KH TMB SH EM KUL SM EJ NZ BGH HR. Wrote the paper: MD MK TP MMN ML HR.

References

1. Husmann DA, Vandersteen DR (1999) Anatomy of the cloacal exstrophy. In: Gearhart JP, Mathews R (eds.): The exstrophy-epispadias complex, New York, Kluwer Academic/Plenum Publishers, pp. 199–206.
2. Hurst JA (2012) Anterior abdominal wall defects," In: Firth HV, Hall JG (eds.): Oxford desk reference. Clinical genetics, New York, Oxford University Press, p.566.
3. Stein R, Fisch M, Black P, Hohenfellner R (1999) Strategies for reconstruction after unsuccessful or unsatisfactory primary treatment of patients with bladder exstrophy or incontinent epispadias. J Urol 161:1934–1941. PMID: 10332476
4. Ebert AK, Reutter H, Ludwig M, Rösch WH (2009) The exstrophy-epispadias complex. Orphanet J Rare Dis 4: 1–17. doi: 10.1186/1750-1172-4-1 PMID: 19133130
5. Reutter H, Qi L, Gearhart JP, Boemers T, Ebert AK, et al. (2007) Concordance analyses of twins with bladder exstrophy-epispadias complex suggest genetic etiology. Am J Med Genet A 143A: 2751–2756. PMID: 17937426
6. Gearhart JP, Jeffs RD (1998) Exstrophy-epispadias complex and bladder anomalies. In Campbell's Urology. 7th edition, Walsh P.C., Retik A.B., Vaughan E.D. and Wein A,J, ed. (Philadelphia, US: WB Saunders), pp. 1939–1990.
7. Wiesel A, Queisser-Luft A, Clementi M, Bianca S, Stoll C (2005) Prenatal detection of congenital renal malformations by fetal ultrasonographic examination: an analysis of 709,030 births in 12 European countries. Eur J Med Genet 48: 131–144. PMID: 16053904
8. International Clearinghouse for Birth Defects Monitoring Systems (1987) Epidemiology of bladder exstrophy and epispadias: a communication from the International Clearinghouse for Birth Defects Monitoring Systems. Teratology 36: 221–227. PMID: 3424208
9. Boyadjiev SA, Dodson JL, Radford CL, Ashrafi GH, Beaty TH, et al. (2004) Clinical and molecular characterization of the bladder exstrophy-epispadias complex: analysis of 232 families. BJU Int 94: 1337–1343. PMID: 15610117
10. Reutter H, Boyadjiev SA, Gambhir L, Ebert AK, Rösch W, et al. (2011) 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 159: 825–831. doi: 10.1016/j.jpeds.2011.04.042 PMID: 21679965
11. Shapiro E, Lepor H, Jeffs RD (1984) The inheritance of the exstrophy-epispadias complex. J Urol 132: 308–310. PMID: 6737583
12. Messelink EJ, Aronson DC, Knuist M, Heij HA, Vos A (1994) Four cases of bladder exstrophy in two families. J Med Genet 31: 490–492. PMID: 8071977
13. Ludwig M, Ching B, Reutter H, Boyadjiev SA (2009) The bladder exstrophy-epispadias complex. Birth Defects Res Part A Clin Mol Teratol 85:509–522. doi: 10.1002/bdra.20557 PMID: 19161161
14. Reutter H, Shapiro E, Gruen JR (2003) Seven new cases of familial isolated bladder exstrophy and epispadias complex (BEEC) and review of the literature. Am J Med Genet A 120A:215–221. PMID: 12833402
15. Draaken M, Reutter H, Schramm C, Bartels E, Boemers TM, et al. (2010) Microduplications at 22q11.21 are associated with non-syndromic classic bladder exstrophy. Eur J Med Genet 53: 55–60. doi: 10.1016/j.ejmg.2009.12.005 PMID: 20060941
16. Lundin J, Söderhäll C, Lundén L, Hammarsjö A, White I, et al. (2010) 22q11.2 microduplication in two patients with bladder exstrophy and hearing impairment. Eur J Med Genet 53: 61–65. doi: 10.1016/j.ejmg.2009.11.004 PMID: 20045748
17. Draaken M, Baudisch F, Timmermann B, Kuhl H, Kerick M, et al. (2014) Classic bladder exstrophy: Frequent 22q11.21 duplications and definition of a 414 kb phenocritical region. Birth Defects Res Part A Clin Mol Teratol 100: 512–517. doi: 10.1002/bdra.23249 PMID: 24764164
18. Wilkins S, Zhang KW, Mahfuz I, Quantin R, D'Cruz N, et al. (2012) Insertion/deletion polymorphisms in the ΔNp63 promoter are a risk factor for bladder exstrophy epispadias complex. PLoS Genet 8: e1003070. doi: 10.1371/journal.pgen.1003070 PMID: 23284286
19. Qi L, Wang M, Yagnick G, Gearhart JP, Ebert AK, et al. (2013) Candidate gene association study implicates *p63* in the etiology of nonsyndromic bladder-exstrophy-epispadias complex. Birth Defects Res Part A Clin Mol Teratol 97: 759–763. doi: 10.1002/bdra.23161 PMID: 23913486
20. Reutter H, Draaken M, Pennimpede T, Wittler L, Brockschmidt FF, et al. (2014) 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 (in press), doi: 10.1093/hmg/ddu259

21. Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, et al. (2012) Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat Genet 44: 291–296. doi: 10.1038/ng.1076 PMID: 22286218

22. Karlsson O, Thor S, Norberg T, Ohlsson H, Edlund T (1990) 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 344: 879–882. PMID: 1691825

23. Zhuang S, Zhang Q, Zhuang T, Evans SM, Liang X, Sun Y (2013) Expression of Isl1 during mouse development. Gene Expr Patterns 13: 407–412. doi: 10.1016/j.gep.2013.07.001 PMID: 23906961

24. Zhang H, Wang WP, Guo T, Yang JC, Chen P, et al. (2009) The LIM-homeodomain protein ISL1 activates insulin gene promoter directly through synergy with BETA2. J Mol Biol 392: 566–577. doi: 10.1016/j.jmb.2009.07.036 PMID: 19619559

25. Shimomura H, Sanke T, Hanabusa T, Tsunoda K, Furuta H, Nanjo K (2000) Nonsense mutation of islet-1 gene (Q310X) found in a type 2 diabetic patient with a strong family history. Diabetes 49: 1597–1600. PMID: 10969846

26. Holm P, Rydlander B, Luthman H, Kockum I, European Consortium for IDDM Genome Studies (2004) Interaction and association analysis of a type 1 diabetes susceptibility locus on chromosome 5q11-q13 and the 7q32 chromosomal region in Scandinavian families. Diabetes 53: 1584–1591. PMID: 15161765

27. Ahlgren U, Pfaff SL, Jessell TM, Edlund T, Edlund H (1997) Independent requirement for ISL1 in formation of pancreatic mesenchyme and islet cells. Nature 385: 257–260. PMID: 9000074

28. Lin L, Bu L, Cai CL, Zhang X, Evans S (2013) Isl1 is upstream of sonic hedgehog in a pathway required for cardiac morphogenesis. Dev Biol 295: 756–763.

29. Peng T, Tian Y, Boogerd CJ, Lu MM, Kadzik RS, et al. (2013) Coordination of heart and lung co-development by a multipotent cardiopulmonary progenitor. Nature 500: 589–593. doi: 10.1038/nature12358 PMID: 23873040

30. Matsumaru D, Haraguchi R, Miyagawa S, Motoyama J, Nakagata N, Meijlink F, Yamada G (2011) Genetic analysis of hedgehog signaling in ventral body wall development and the onset of omphalocele formation. PLoS ONE 6: e16260. doi: 10.1371/journal.pone.0016260 PMID: 21283718

31. Cai CL, Liang X, Shi Y, Chu PH, Pfaff SL, Chen J, Evans S (2003) Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. Dev Cell 5: 877–889. PMID: 14667410

32. Pfaff SL, Mendelsohn M, Stewart CL, Edlund T, Jessell TM (1996) Requirement for LIM homeobox gene *Isl1* in motor neuron generation reveals a motor neuron-dependent step in interneuron differentation. Cell 84: 309–320. PMID: 8565076

33. Itou J, Kawakami H, Quach T, Osterwalder M, Evans SM, Zeller R, Kawakami Y (2012) Islet1 regulates establishment of the posterior hindlimb field upstream of the Hand2-Shh morphoregulatory gene network in mouse embryos. Development 139: 1620–1629. doi: 10.1242/dev.073056 PMID: 22438573

34. Jurberg AD, Aires R, Varela-Lasheras I, Nóvoa A, Mallo M (2013) Switching axial progenitors from producing trunk to tail tissues in vertebrate embryos. Dev Cell 25: 451–462. doi: 10.1016/j.devcel.2013.05.009 PMID: 23763947

35. Kaku Y, Ohmori T, Kudo K, Fujimura S, Suzuki S, et al. (2013) Islet1 deletion causes kidney agenesis and hydroureter resembling CAKUT. J Am Soc Nephrol 24: 1242–1249. doi: 10.1681/ASN.2012050528 PMID: 23641053

36. Suzuki K, Adachi Y, Numata T, Nakada S, Yanagita M, et al. (2012) Reduced BMP signaling results in hindlimb fusion with lethal pelvic/urogenital organ aplasia: a new mouse model of sirenomelia. PLoS ONE 7: e43453. doi: 10.1371/journal.pone.0043453 PMID: 23028455

37. Lugani F, Arora R, Papeta N, Patel A, Zheng Z, et al. (2013) A retrotransposon insertion in the 5' regulatory domain of Ptf1a results in ectopic gene expression and multiple congenital defects in Danforth's short tail mouse. PLoS Genet 9: e1003206. doi: 10.1371/journal.pgen.1003206 PMID: 23437001

38. Semba K, Araki K, Matsumoto K, Suda H, Ando T, et al. (2013) Ectopic expression of Ptf1a induces spinal defects, urogenital defects, and anorectal malformations in Danforth's short tail mice. PLoS Genet 9: e1003204. doi: 10.1371/journal.pgen.1003204 PMID: 23436999

39. Vlangos CN, Siuniak AN, Robinson D, Chinnaiyan AM, Lyons RH Jr, Cavalcoli JD, Keegan CE (2013) Next-generation sequencing identifies the Danforth's short tail mouse mutation as a retrotransposon insertion affecting Ptf1a expression. PLoS Genet 9: e1003205. doi: 10.1371/journal.pgen.1003205 PMID: 23437000

40. Muller YL, Yueh YG, Yaworsky PJ, Salbaum JM, Kappen C (2003) Caudal dysgenesis in Islet-1 transgenic mice. FASEB J 17: 1349–1351. PMID: 12738808

41. Thompson N, Gésina E, Scheinert P, Bucher P, Grapin-Botton A (2012) RNA profiling and chromatin immunoprecipitation-sequencing reveal that PTF1a stabilizes pancreas progenitor identity via the control of MNX1/HLXB9 and a network of other transcription factors. Mol Cell Biol 32: 1189–1199. doi: 10.1128/MCB.06318-11 PMID: 22232429

42. Haraguchi R, Motoyama J, Sasaki H, Satoh Y, Miyagawa S, et al. (2007) Molecular analysis of coordinated bladder and urogenital organ formation by Hedgehog signaling. Development 134: 525–533. PMID: 17202190

43. Haraguchi R, Matsumaru D, Nakagata N, Miyagawa S, Suzuki K, Kitazawa S, Yamada G (2012) The hedgehog signal induced modulation of bone morphogenetic protein signaling: an essential signaling relay for urinary tract morphogenesis. PLoS ONE 7: e42245. doi: 10.1371/journal.pone.0042245 PMID: 22860096

44. Mildenberger H, Kluth D, Dziuba M (1988) Embryology of bladder exstrophy. J Pediatr Surg 23: 166–170. PMID: 3343652

45. Haraguchi R, Mo R, Hui C, Motoyama J, Makino S, et al. (2001) Unique functions of Sonic hedgehog signaling during external genitalia development. Development 128: 4241–4250. PMID: 11684660

46. Schmermund A, Möhlenkamp S, Stang A, Grönemeyer D, Seibel R, et al. (2002) Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. Risk Factors, Evaluation of Coronary Calcium and Lifestyle. Am Heart J 144: 212–218. PMID: 12177636

47. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM (2010) Robust relationship inference in genome-wide association studies. Bioinformatics 26: 2867–2873. doi: 10.1093/bioinformatics/btq559 PMID: 20926424

48. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 5: e1000529. doi: 10.1371/journal.pgen.1000529 PMID: 19543373

49. Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, et al. (2009) Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. Nat Genet 41: 473–477. doi: 10.1038/ng.333 PMID: 19270707

50. Pennimpede T, Proske J, König A, Vidigal JA, Morkel M, et al. (2012) *In vivo* knockdown of Brachyury results in skeletal defects and urorectal malformations resembling caudal regression syndrome. Dev Biol 372: 55–67. doi: 10.1016/j.ydbio.2012.09.003 PMID: 22995555