



Restless Legs Syndrome-associated intronic common variant in *Meis1* alters enhancer function in the developing telencephalon

Derek Spieler, Maria Kaffe, Franziska Knauf, et al.

Genome Res. published online March 18, 2014

Access the most recent version at doi:[10.1101/gr.166751.113](https://doi.org/10.1101/gr.166751.113)

Supplemental Material <http://genome.cshlp.org/content/suppl/2014/02/10/gr.166751.113.DC1.html>

P<P Published online March 18, 2014 in advance of the print journal.

Open Access Freely available online through the *Genome Research* Open Access option.

Creative Commons License This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported), as described at <http://creativecommons.org/licenses/by-nc/3.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Genome Research* go to:
<http://genome.cshlp.org/subscriptions>

Research

Restless Legs Syndrome-associated intronic common variant in *Meis1* alters enhancer function in the developing telencephalon

Derek Spieler,^{1,25} Maria Kaffe,^{1,2,25} Franziska Knauf,^{1,25} José Bessa,³ Juan J. Tena,³ Florian Giesert,⁴ Barbara Schormair,^{1,2} Erik Tilch,^{1,2} Heekyoung Lee,^{5,6,7,8} Marion Horsch,⁹ Darina Czamara,¹⁰ Nazanin Karbalai,¹⁰ Christine von Toerne,¹¹ Melanie Waldenberger,¹² Christian Gieger,¹² Peter Lichtner,¹ Melina Claussnitzer,^{5,6,7,8,24} Ronald Naumann,¹³ Bertram Müller-Myhsok,^{10,14,15} Miguel Torres,¹⁶ Lillian Garrett,¹⁷ Jan Rozman,^{9,7,18} Martin Klingenspor,^{5,6,9,18} Valérie Gailus-Durner,⁹ Helmut Fuchs,⁹ Martin Hrabě de Angelis,^{9,7,19} Johannes Beckers,^{7,9,19} Sabine M. Hölter,¹⁷ Thomas Meitinger,^{1,2,14} Stefanie M. Hauck,^{7,11} Helmut Laumen,^{5,6,7,8,9} Wolfgang Wurst,^{4,10,14,17,20,21} Fernando Casares,³ Jose Luis Gómez-Skarmeta,³ and Juliane Winkelmann^{1,14,22,23,26}

^{1–23}[Author affiliations appear at the end of the paper.]

Genome-wide association studies (GWAS) identified the *MEIS1* locus for Restless Legs Syndrome (RLS), but causal single nucleotide polymorphisms (SNPs) and their functional relevance remain unknown. This locus contains a large number of highly conserved noncoding regions (HCNRs) potentially functioning as *cis*-regulatory modules. We analyzed these HCNRs for allele-dependent enhancer activity in zebrafish and mice and found that the risk allele of the lead SNP rs12469063 reduces enhancer activity in the *Meis1* expression domain of the murine embryonic ganglionic eminences (GE). CREB1 binds this enhancer and rs12469063 affects its binding in vitro. In addition, *MEIS1* target genes suggest a role in the specification of neuronal progenitors in the GE, and heterozygous *Meis1*-deficient mice exhibit hyperactivity, resembling the RLS phenotype. Thus, in vivo and in vitro analysis of a common SNP with small effect size showed allele-dependent function in the prospective basal ganglia representing the first neurodevelopmental region implicated in RLS.

[Supplemental material is available for this article.]

With an age-dependent prevalence of up to 10% in western societies, RLS is a sensorimotor neurological disorder characterized by an urge to move the legs and uncomfortable sensations in the lower limbs. These symptoms occur exclusively during rest, in the evening, or at night, and improve only with movement (Szentkiralyi et al. 2011). Therefore, patients are forced to stand up and walk, sometimes calling themselves “nightwalkers.” Consequences are severe sleep disturbances, depression, anxiety, and increased cardiovascular risk (Winkelmann et al. 2008). Severely affected patients require medication, with dopaminergic compounds being the first line treatment, but side effects such as augmentation limit their long-term use (Hornyak et al. 2012). So far, very little is known about the underlying neurobiology of this disorder.

RLS has a heritability of ~50% and is a complex genetic disorder. GWAS identified common genetic variants within six risk

loci (Winkelmann et al. 2007, 2011; Schormair et al. 2008). These convey only small effects on the phenotype; and generally, the molecular mechanism and function behind low-effect-size common genetic variants identified in complex genetic diseases provide a formidable challenge. The strongest RLS association signal identified delineates a 32-kb linkage disequilibrium (LD) block in intron 8 of *MEIS1* (2p14). A common haplotype completely tagged by a pair of highly correlated intronic SNPs, rs12469063 and rs2300478 (LD $r^2 = 0.97$), and rs6710341 in this block was shown to confer a significantly increased risk for RLS (odds ratio [OR] of 2.8) (Winkelmann et al. 2007). *MEIS1* belongs to the TALE family of homeobox transcription factors and is involved in the development and homeostasis of numerous organs and diseases such as leukemia or neuroblastoma (Azcoitia et al. 2005). In the nervous system, multiple expression domains suggest essential roles for *MEIS1* (Toresson et al. 2000); however, its spatial, temporal, and functional involvement in RLS pathogenesis has remained unknown. The *MEIS1* locus comprises a cluster of HCNRs, and it has been shown that these indicate potential *cis*-regulatory elements

²⁴Present address: Hebrew SeniorLife Institute for Aging Research, Harvard Medical School, Boston, MA 02131, USA

²⁵These authors contributed equally to this work.

²⁶Corresponding author

E-mail winkelmann@stanford.edu

Article published online before print. Article, supplemental material, and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.166751.113>. Freely available online through the *Genome Research* Open Access option.

© 2014 Spieler et al. This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported), as described at <http://creativecommons.org/licenses/by-nc/3.0/>.

for genes implicated in transcriptional regulation and development (Woolfe et al. 2005). Since disease-associated variants identified by means of GWAS are more likely located in regulatory DNA (Maurano et al. 2012), we aim to study the causal role of *MEIS1* in RLS by analyzing the allele-dependent *cis*-regulatory function of the common intronic RLS-associated DNA variants of *MEIS1* and differential binding of upstream factors as well as RLS-related behavioral consequences of altered *Meis1* function.

Results

Identification of common candidate causal variants in the *MEIS1* GWAS locus

To capture common variation found in the *MEIS1* GWAS locus, we screened the entire intronic 32-kb LD block associated with RLS in 188 RLS patients for variants using high-resolution melting curve analysis (Herrmann et al. 2006). Our discovery sample consisted of carriers of the known risk haplotype in order to enrich candidate causal variants residing on this haplotype. We focused our analysis on common variants with a minor allele frequency (MAF) ≥ 0.05 due to our small sample size limiting the power for rare and low frequency variants. Of 136 variants identified, 54 were common SNPs in our data set (MAF ≥ 0.05) (Supplemental Table S1). To account for sampling bias in our discovery sample, we also included five variants reported as common in the general population (dbSNP Build 130), but with lower frequency in our data set, yielding a total of 59 variants. We genotyped these in independent 735 cases and 735 age-, sex-, and ethnicity-matched population-based controls (KORA cohort) (Wichmann et al. 2005). Based on strong LD, a total of 16 variants were assayed using tagging SNPs. We then imputed missing genotypes in additional samples, leading to a final set of high-quality genotypes of 38 variants for statistical analysis in 1302 cases and 1259 controls. Logistic regression followed by permutation-based adjustment for multiple testing revealed 27 associated SNPs ($P_{\text{corr}} < 0.05$) (Supplemental Table S2). The two strongest signals were the known highly correlated rs12469063 ($P_{\text{nom}} = 7.7 \times 10^{-18}$; OR = 1.7) and rs2300478 ($P_{\text{nom}} = 4.2 \times 10^{-18}$; OR = 1.7) (Fig. 1B). Conditional analysis with either rs12469063 or rs2300478 as covariate did not reveal additional indepen-

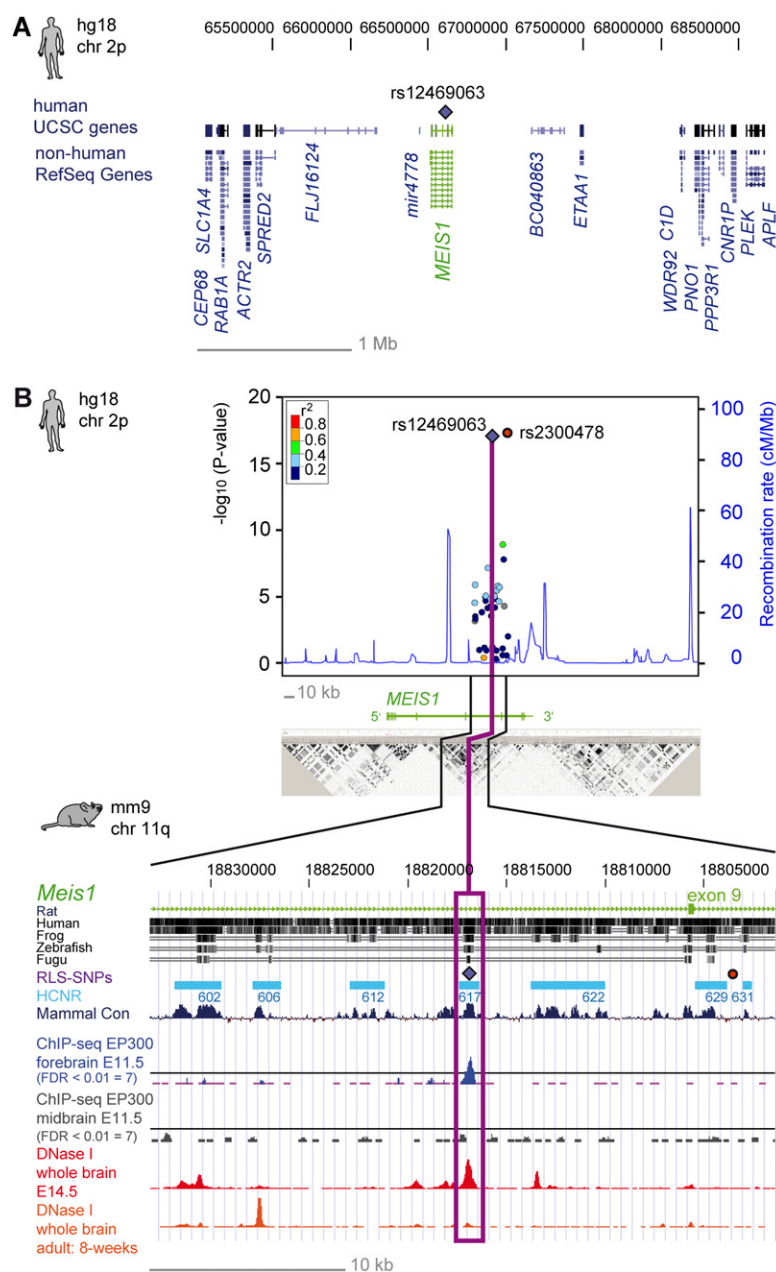


Figure 1. Association signals, sequence conservation, and epigenetic signatures in embryonic and adult stages of the *MEIS1/Meis1* locus (A) RLS-associated locus on 2p14 with the lead SNP rs12469063 (violet diamond) and its genetic environment. (B) The x-axis represents the genomic position, referring to the hg18 genome annotation. The right-hand y-axis represents the recombination rate, and the negative log₁₀ of the nominal P-values of all SNPs genotyped are given on the left-hand y-axis. The color-coded LD between SNPs is based on the lead SNP rs12469063 (violet diamond) as a reference. rs2300478, which lies in a nonconserved region, is depicted with a red circle. Recombination rate and r^2 values were estimated with the CEU population from HapMap II (release 22). The LD block below is based on HapMap data, measured in r^2 and visualized with Haploview. Highlighting refers to the magnitude of pairwise LD, with a white-to-black shading indicating lower to higher LD values. The lower part of Figure 1B uses mouse genome annotation mm9 and lists from top to bottom: (1) conservation between the mouse genome and the genome of the rat, human, frog, zebrafish, and Fugu; (2) corresponding localization of the human RLS-SNPs rs12469063 (violet diamond) and rs2300478 (red circle); (3) seven analyzed HCNRs in turquoise boxes; (4) the conservation in mammals (dark blue); (5) DNA binding of EP300 (ChIP-seq) in E11.5 forebrain (blue) and midbrain (gray); and (6) DNase I hypersensitivity in whole brain of E14.5 embryos and adult mice (red and orange peaks, respectively).

dent association signals within the 32-kb block after permutation-based adjustment but confirmed the known risk haplotype configuration rs12469063/rs6710341. This haplotype showed an increased effect size with an OR of 2.33 compared to all other haplotypes with a frequency $\geq 1\%$ in our data set combined (95% confidence interval: 1.85–2.94 with an overall frequency [f] of 0.138 defined by rs6710341 – allele A [$f = 0.857$] and rs12469063 – allele G [$f = 0.282$]). Taken together, the common DNA variants rs12469063 and rs2300478 remained the lead SNP pair within the 32-kb block with the increased haplotype risk being compatible with additional small effect contributors not detectable by our approach.

rs12469063 impairs enhancer activity in zebrafish

Given that *cis*-regulatory function is often found in HCNRs of developmental genes (Woolfe et al. 2005), we focused on variants found within these elements. Thus, we systematically examined all HCNRs in the associated LD block (Figs. 1, 2) in a zebrafish transgenic EGFP-reporter assay (i.e., 24/48 h post-fertilization [hpf]). Altogether, eight HCNRs showed high conservation in mouse,

chicken, and frog, five of which were conserved also in the *Fugu* genome (HCNR 602, 606, 617, 628, and 629). Since one HCNR did not contain any common variants, we subjected the remaining seven HCNRs to further analysis (HCNR 602, 606, 612, 617, 622, 629, and 631) (Fig. 2). None of the examined HCNRs contained any rare variants. While SNP rs12469063 resided in one of the HCNRs (617), rs2300478 was located in a nonconserved stretch and therefore was not further examined (Fig. 1B).

Of all HCNRs assayed, solely HCNR 617 exhibited both a reproducible neural expression pattern and a genotype-specific expression comparing the protective and risk allele reporter constructs (Fig. 2; Supplemental Fig. S1). With the protective allele, an enhancer signal was detected in the retina, the fore-, mid-, and hindbrain, and the spinal cord. In contrast, zebrafish embryos carrying the risk allele showed highly reduced enhancer activity (Fig. 2B,C). Together, these findings indicate that rs12469063 lies within a region of high interspecies conservation with neural enhancer activity and has an allele-specific functional impact.

rs12469063 alters enhancer activity in the murine embryonic ganglionic eminences

Next, HCNR 617 was analyzed in a transgenic mouse beta-galactosidase (*lacZ*) enhancer assay (Pennacchio et al. 2006). We found a reproducible *lacZ* gene expression pattern for both alleles in the forebrain of all embryos (E12.5) within the ganglionic eminences (Fig. 3A,D; Supplemental Fig. S2), giving rise to the basal ganglia of the adult brain, in addition to parts of the amygdala and several interneuron populations (Medina and Abellán 2012). Notably, *lacZ* reporter activity matched the endogenous *Meis1* telencephalic expression domain (Fig. 3D). In addition, transcripts of all further RLS-associated loci, *Btbd9*, *Ptprd*, *Map2k5*, and *Tox3* (Winkelmann et al. 2011) were expressed in this telencephalic area (Fig. 3E). Given the allele-specific impaired EGFP expression in zebrafish, we looked for allele-specific differences of *lacZ* activity, relying on the assumption that both constructs have the same bias for copy number as suggested by the small variance of both groups. Transgenic embryos of Theiler stages 19 and 20 (Theiler 1989) were included in the evaluation of beta-galactosidase signal intensity and spatial expression measured according to Cavalieri ($n = 4$ for both the protective and risk allele) (Fig. 3A,B; Supplemental Figs. S2, S3). For risk allele carriers, analyses showed a significant reduction of 65% in signal intensity ($P = 0.029$) and of 76% in the stereological analysis ($P = 0.029$) (Fig. 3A,B).

Direct evidence for an allele-specific function of rs12469063 was further provided by electrophoretic mobility shift assays (EMSA). Incubation with nuclear protein extracts from forebrain of E12.5 mice showed differential allele-specific gel shifts (Fig. 3C). The gel shifts were specific, as competition with the unlabeled allele in excess attenuated the differential bands. Combined with the results of the transgenic animal models, where the risk allele reduces enhancer function, this implies an enhanced binding of an activator protein complex to the protective allele or a repressor protein complex to the risk allele. The lead SNP rs12469063 was the only DNA sequence variation between the two oligonucleotides examined with EMSA, and thus held responsible for the differential DNA–protein complex formation (Fig. 3C).

Since binding of the enhancer-associated epigenomic marker EP300 allows the prediction of tissue-specific enhancer activity (Visel et al. 2009), we assessed the 32-kb LD block for EP300 binding in ChIP-seq data of murine brain (E11.5). The only EP300 peak corresponding to HCNR 617 was detected in forebrain tissue

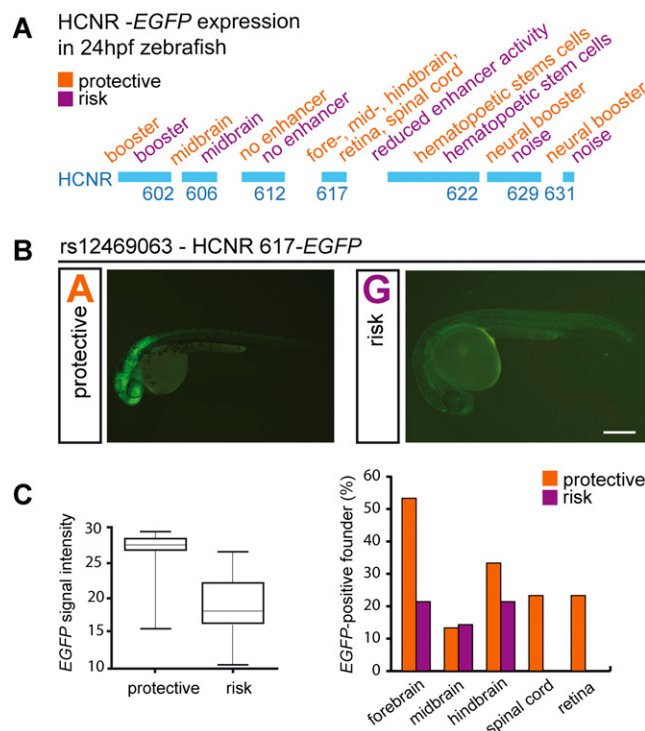


Figure 2. In vivo allele-specific enhancer function of HCNR 617 in zebrafish. (A) Results of the zebrafish enhancer screen with the expression domains for the protective and risk allele in orange and purple, respectively. (B) Representative embryos of the F1 generation with the protective ([A] adenine) and the risk allele ([G] = guanine) construct of HCNR 617 showed an allele-specific difference in reporter EGFP expression in the neural tube. The protective allele of HCNR 617 drove EGFP expression in the retina, fore-, mid-, and hindbrain and spinal cord. The risk allele significantly reduced the expression almost to background levels (bar represents 250 μ m). (C) Boxplot (left) The risk allele reduced the fluorescence intensity to 70% compared to the protective allele (100%) (Wilcoxon rank sum test, $P = 0.08$; EGFP-brain signal measured, $n_{\text{protective}} = 5$, $n_{\text{risk}} = 5$). The bar chart (right) represents the percentage of all EGFP-positive founders in the respective area (y-axis) with respect to the protective (orange) and risk (purple) allele of HCNR 617. The risk allele leads either to a significant reduction or a complete loss of the EGFP signal.

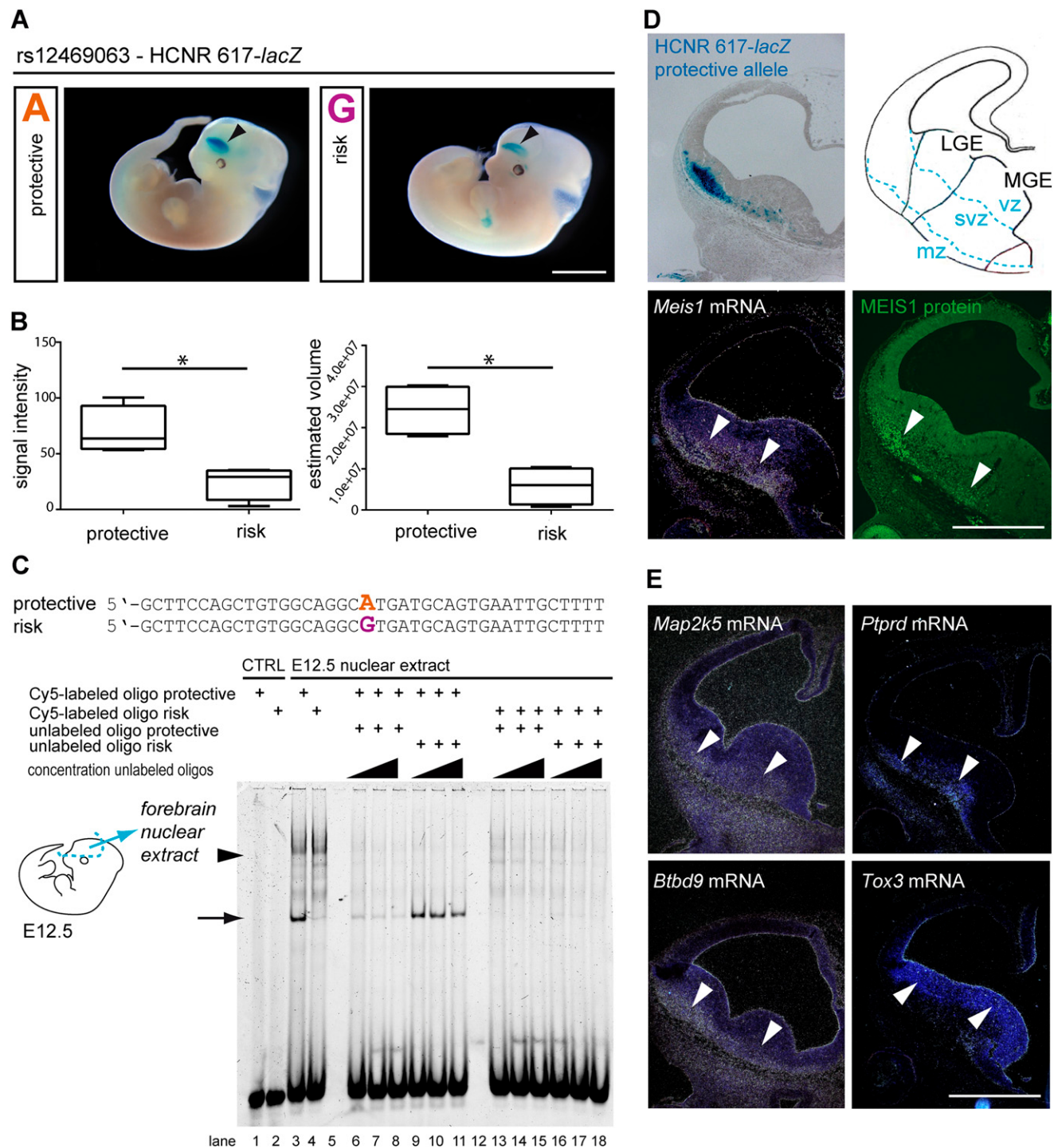


Figure 3. In vivo and in vitro identification of allele-specific HCNR 617 enhancer function in the ganglionic eminences of the mouse. (A) Representative transgenic mouse embryos after pronucleus injection of HCNR 617 protective (left) and risk allele constructs (right) at stage E12.5. Blue color indicates regions expressing the reporter gene beta-galactosidase (*lacZ*). Arrowhead indicates the reproducible telencephalic signal (bar represents 2.5 mm). (B) Analysis of the signal intensity in Theiler-staged embryos (stages 19 and 20; $n_{\text{protective}} = 4$, $n_{\text{risk}} = 4$) showed a reduction down to 35% (Wilcoxon rank sum test, $P = 0.029$). Stereological volume estimation according to Cavalieri revealed a significant volume reduction down to 24% for the risk allele (Wilcoxon rank sum test, $P = 0.029$). (C) Electrophoretic mobility shift assay using oligonucleotides encompassing rs12469063 showed allele-specific differences of DNA-protein complex formation using E12.5 forebrain nuclear extract (lanes 3 and 4, arrow and arrowhead). Specificity is proven by competition with unlabeled risk and protective oligonucleotide in excess (lanes 6–18). CTRL, control (no protein). (D) Frontal sections through the forebrain reveal *lacZ* reporter activity in the mantle zone (mz) of the ganglionic eminences, in the same region as *Meis1* transcript and MEIS1 protein (arrowheads; bar represents 500 μm); no transcripts are visible in the ventricular (vz) and subventricular (svz) zones. (E) Transcripts of all additional GWAS-RLS risk loci map to the embryonic ganglionic eminences. mRNA of *Ptprd*, *Btd9*, and *Map2k5* was detected in the mantle zone (mz) of the LGE and MGE. *Tox3* was expressed in the adjacent ventricular (vz) and subventricular (svz) zone (arrowheads; bar represents 500 μm).

(Fig. 1B). Next, we analyzed whether HCNR 617 is also an active enhancer during adulthood by using DNase I hypersensitivity maps, indicating active regulatory regions (Sabo et al. 2006). This analysis further supported the active status of HCNR 617 during development (E14.5), whereas in adult brain tissue (8 wk), no activity of HCNR 617 could be identified (Fig. 1B). Together, these observations underline the spatiotemporal enhancer specificity of HCNR 617 and strongly suggest *MEIS1* regulation in RLS during development rather than in adult tissue homeostasis.

rs12469063 affects a cAMP response element-binding protein/CREB1 transcription factor binding site

To identify upstream transcription factors binding the sequence motif altered by rs12469063, we carried out affinity chromatography using allele-specific DNA oligonucleotides and nuclear protein extracted from E12.5 mouse embryonic telencephalon. Nuclear proteins bound to the affinity matrix were identified by mass spectrometry, and protein levels were quantified allowing calculation of a ratio reflecting binding preference (Fig. 4A). A total of 51 identified proteins demonstrated allele-specific differential binding (mean fold change ≥ 1.5) (Supplemental Table S3), with 33 showing consistent results across three independent experiments. Among these, CREB1 was the protein with the highest fold change, also known to bind DNA in a sequence-specific manner and of functional relevance in the CNS. We observed a strong similarity of the genomic sequence spanning rs12469063 and the consensus CREB binding site (Fig. 4B) and found *Creb1*/CREB1 to colocalize with *Meis1*/MEIS1 in the mantle zone of the ganglionic eminences at both transcript and protein levels (Fig. 4C,D). Moreover, rs12469063 influenced the predicted CREB1 transcription factor interaction based on in silico analysis (Genomatix, SNPInspector). Furthermore, supershift experiments with a CREB-specific antibody attenuated the gel shift (Fig. 4E). As the supershifted band showed a higher intensity signal for the risk allele, and mass spectrometry results showed that CREB1 binds 3.2-fold higher to the risk allele (Supplemental Table S3), we suggest that CREB1 binds more strongly to the risk allele. Finally, disruption of the entire predicted CREB1 binding site resulted in the loss of the specific gel shift (Fig. 4E). Combined, these data sets identified CREB1 as an upstream factor differentially binding rs12469063.

MEIS1 downstream target genes in the E12.5 ganglionic eminences

To model the reduced expression of *MEIS1* in RLS, we chose mice heterozygous for the *Meis1*^{tm1Mtor} allele. These mice express only one functional allele of *Meis1* and should thus recapitulate reduced *MEIS1* expression in RLS patients. Whereas homozygous *Meis1*^{tm1Mtor} mice are embryonic lethal by E14.5, heterozygous *Meis1*^{tm1Mtor} mice are viable and fertile (Azcoitia et al. 2005). To identify genes that are directly or indirectly regulated due to reduced *Meis1* expression in the LGE/MGE, we performed whole genome transcriptomics of forebrains from wild-type and heterozygous *Meis1*^{tm1Mtor} embryos (E12.5) (Fig. 5). Based on significance analysis of microarrays (SAM), we found 126 significantly regulated genes in male and 161 in female embryos (mean fold change > 1.5 ; false discovery rate [FDR] $< 10\%$). Hierarchical cluster analysis of genes differentially regulated (unpaired two class analysis [SAM] of 285 genes; FDR $< 7.9\%$; fold change > 1.5) revealed that the expression profiles clearly cluster according to the respective sex (Fig. 5; Supplemental Table S4). Using pathway analysis software

(Ingenuity), we identified significantly overrepresented functional annotations in male and female embryos. Despite the differences, we found that in both sexes the functional categories “neurological disease” and “nervous system development” were among the most significant overrepresented biological processes (Table 1). In addition, the occurrence of the “cardiovascular disease” category was in agreement with the reported function of *Meis1* during vasculogenesis (Azcoitia et al. 2005). Of note, no transcripts of the currently known RLS-associated risk loci, *Btbd9*, *Ptprd*, *Map2k5*, and *Tox3*, were regulated, although transcripts belonging to homologous genes, such as *Ptprv*, *Ptprc*, and *Map3k12*, were differentially expressed. Interestingly, the homeobox gene *Gbx2*, which is essential for the proper development of cholinergic interneurons in the striatum (Chen et al. 2010), was found to be down-regulated (Table 1; Supplemental Table S4). Altogether, these data correlate with known *MEIS1* function regulating transcription of developmental genes rather than housekeeping genes (Penkov et al. 2013).

Motor restlessness/hyperactivity in heterozygous *Meis1*^{tm1Mtor} mice

Given the reduced reporter gene activity in the endogenous *Meis1* telencephalic expression domain, we also aimed to model the functional effect of reduced *Meis1* expression on behavior and analyzed adult heterozygous *Meis1*^{tm1Mtor} mice (Azcoitia et al. 2005). In the open field test measured over 20 min, we found a significantly increased total distance traveled and average speed of forward locomotion in heterozygous compared to wild-type mice ($P = 0.003$ and $P = 0.004$) (Fig. 6A,B). No changes were observed with respect to anxiety-related behavior (Supplemental Fig. S4). Moreover, we found a consistent trend of elevated average distance traveled measured over a period of 21 h (up to 16.3%, $P = 0.06$) (Fig. 6D) with corresponding enhanced energy metabolism in the mutants measured by indirect calorimetry (Fig. 6C; Supplemental Table S5). The oxygen consumption adjusted to the body mass differences was significantly increased in the mutants (adjusted mean VO_2 : males plus 8.5%, females plus 5.5%, $P_{\text{genotype}} = 0.007$; adjusted minimum VO_2 : males plus 7.8%, females plus 4.3%, $P_{\text{genotype}} = 0.04$; adjusted maximum VO_2 : males plus 6.5%, females plus 10.7%, $P_{\text{genotype}} = 0.003$) (Fig. 6C). Taken together, we consider these findings as an in vivo demonstration that reduced expression of *Meis1* is leading to an alteration of the behavior toward hyperactive locomotion.

Discussion

In the present study, we demonstrated the allele-specific impairment of an intronic *cis*-regulatory element by the common, RLS-associated SNP rs12469063 in two independent in vivo experimental systems. In mice, the risk allele led to a significant reduction of the enhancer activity in the LGE and MGE, suggesting this brain area as RLS-relevant. We concentrated on genetic variation in highly conserved noncoding regions assuming subsequent alteration in gene expression contributing to complex disease (Maurano et al. 2012). Noncoding sequence conservation has been shown to be capable of identifying human *cis*-regulatory elements (Prabhakar et al. 2006). Therefore, we might have missed functional effects of SNPs mediated through other mechanisms such as microRNA regulation, coding SNPs as well as noncoding regulatory elements lacking conservation.

Meis1 expression corresponding to the region of *lacZ* reporter activity strongly points to *Meis1* being regulated by the identified enhancer. Moreover, a bioinformatic approach for delineating

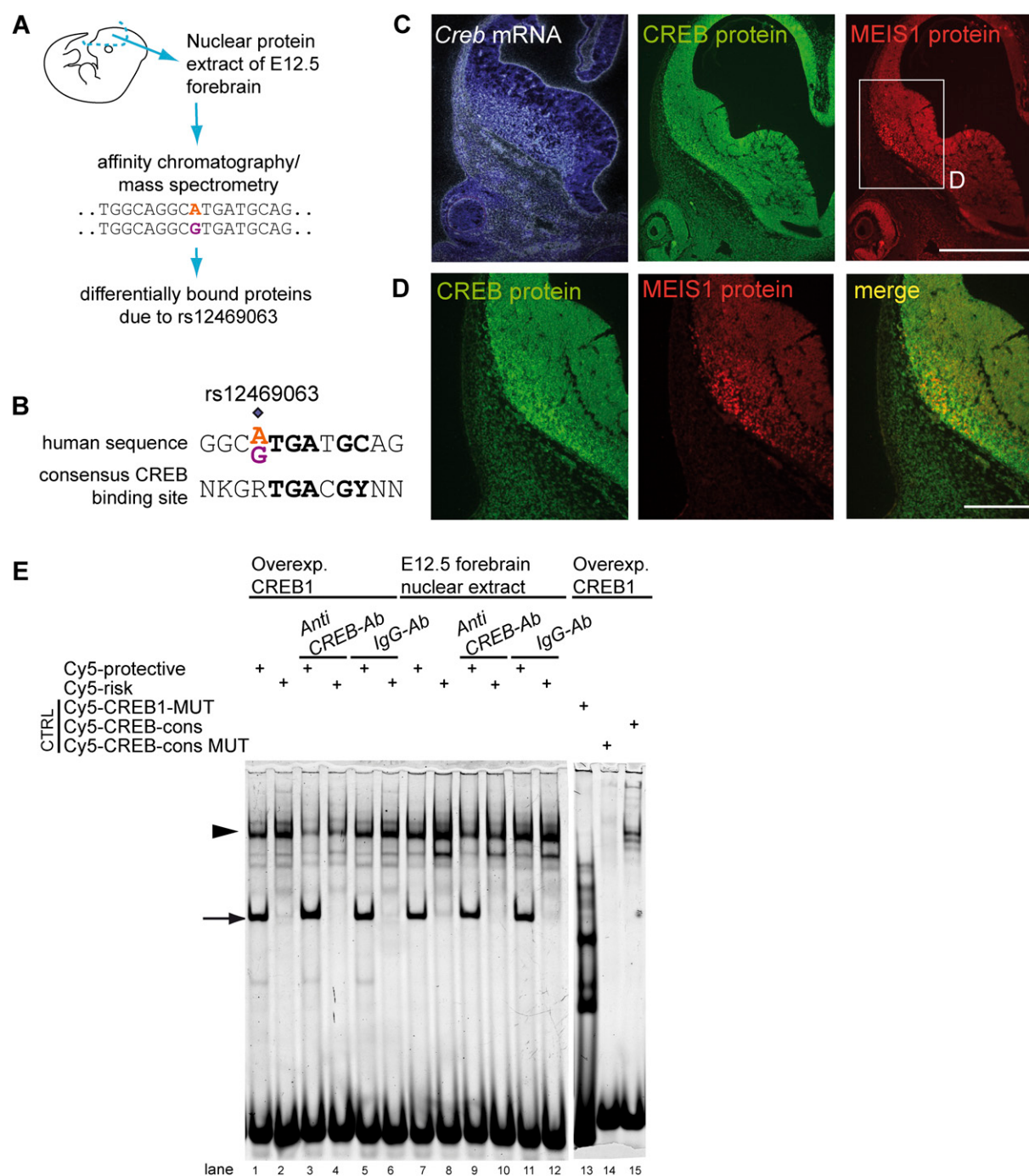


Figure 4. rs12469063 alters a CREB1 transcription factor binding site. (A) Principle work flow for the identification of upstream binding factors by affinity chromatography and mass spectrometry. (B) CREB consensus sequence compared to the human sequence spanning rs12469063. (C,D) *Creb*/CREB and *Meis1*/MEIS1 were both detected in the mantle zone of the ganglionic eminences and show colocalization (bar in C represents 500 μ m; bar in D represents 250 μ m). (D) Magnification of rectangle in C. (E) Supershift EMSA assay with a specific antibody against CREB showed significant reduction of DNA-protein complex formation compared to unspecific IgG antibody (control) for overexpressed CREB1 in 293T nuclear cell lysate (lanes 1–6) and E12.5 forebrain nuclear extract (lanes 7–12, arrowhead). Complete abolishment of specific binding with an oligonucleotide, in which the entire CREB motif is deleted (lane 13). CREB consensus oligonucleotide shows specific binding, being absent with the mutated CREB consensus oligonucleotide (lanes 14, 15). Arrow indicates additional allele-specific band.

enhancer-gene interactions (“PreSTIGE”) (Corradin et al. 2014) and Poll II ChIA-PET and Hi-C experiments in two human cell lines (MCF7 and IMR-90) (Dixon et al. 2012; Li et al. 2012) demon-

strated an exclusive interaction of the RLS-associated intronic region with the regulatory landscape of *MEIS1* (Supplemental Table S6; Supplemental Fig. S5; Supplemental Methods).

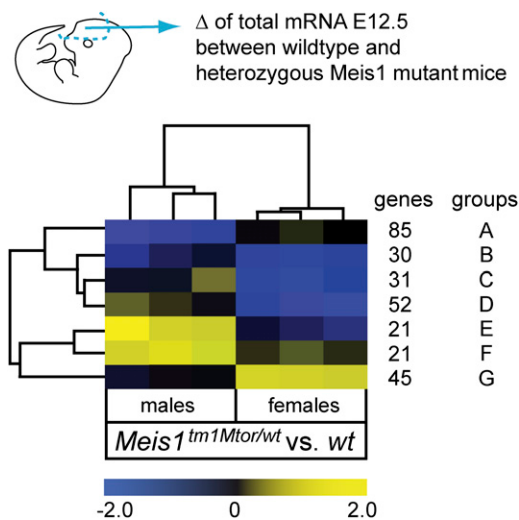


Figure 5. MEIS1 transcriptional target genes in the E12.5 ganglionic eminences. Summarized heat map from unpaired two class analysis (SAM) of genes regulated between heterozygous *Meis1*^{tm1Mtor} and wild-type forebrain tissues at E12.5. Expression profiles of male and female embryos clearly separate. Similar gene profiles are grouped together (rows) and color code gives the mean fold changes of the respective genes for each mutant embryo. Yellow represents up-regulation and blue represents down-regulation in comparison to the respective wild-type control group. The number of genes that are included in each group (A–G) with similar patterns is given. Individual genes are listed in Supplemental Table S4.

The LGE/MGE give rise to the prospective basal ganglia, i.e., pallidum and striatum with several projection neuron and interneuron populations, selected parts of the amygdala, cortical, and olfactory interneurons (Medina and Abellán 2012). Concerning RLS, numerous neuroanatomical regions have been suggested to be involved (Dauvilliers and Winkelmann 2013). The functional relevance of the RLS-SNP rs12469063 in the LGE/MGE suggests that RLS can be regarded as a “basal ganglia disorder,” although further neural structures and pathways might be involved.

The activity of the enhancer during development, indicated by the loss of DNase I hypersensitivity signal in HCNR 617 in adult brain, argues for a role restricted to embryonic development. Our data emphasizes a new focus for RLS research and suggests that the pathophysiology of RLS includes aspects of a neurodevelopmental disorder. As aging is the strongest risk factor for RLS, it is possible that early neurodevelopmental alterations in the LGE/MGE might not be compensated later in life (Somel et al. 2010). The reduced *Meis1* expression might lead to altered neuronal pathway formation, including either striatal projections or interneurons, as suggested by down-regulation of *Gbx2* in heterozygous *Meis1*^{tm1Mtor} mice. *Gbx2* is required for the development of cholinergic interneurons in the striatum (Chen et al. 2010). Transcriptome profiling identified further MEIS1 downstream targets involved in specification of progenitor cells in the ganglionic eminences (Tucker et al. 2008), as well as protein tyrosine phosphatases (*Ptprv*, *Ptpcr*), which might have redundant functions next to the RLS-associated *PTPRD* on 9p23–24 (Schormair et al. 2008).

By affinity chromatography and mass spectrometry, we identified the transcription factor CREB1 as a potential upstream factor, binding stronger to the risk allele of rs12469063 in *in vitro* assays. Since sensitivity of mass spectrometry-based detection of differentially binding proteins is limited due to low protein abundances, it is possible that our analysis did not cover all dif-

ferential binding proteins. Along this line, an additional specific gel shift in the EMSA (Fig. 5E) provides evidence for a further protein complex with increased binding to the protective allele. CREB1 either activates or inactivates the transcription of downstream target genes (Bartsch et al. 1998). Here, we observed an enrichment of CREB1 binding and a reduction of the enhancer's activity for the risk allele, implying a function of CREB1 as a transcriptional repressor affecting rs12469063.

Based on the zebrafish reporter screen, EP300 ChIP-seq, DNase I hypersensitivity data, and the association study, we focused on HCNR 617. Both animal models showed allele-dependent altered enhancer function, but the domain in the telost central nervous system covered additional caudal structures such as the spinal cord. It has been shown that orthologous sequences with a high evolutionary distance do not necessarily retain the same activity in their respective species (Schmidt et al. 2010). Besides HCNR 617, additional *cis*-regulatory elements might be involved in the regulation of *MEIS1* in relation to RLS. Thus, our HCNR zebrafish screen identified two additional transcriptionally active enhancers (HCNR 629 and 631) showing neural allele-dependent, but spatially not reproducible reporter activity, previously referred to as booster activity (Fig. 2; Supplemental Fig. S1; Royo et al. 2012), as well as one HCNR with general, yet allele-independent booster activity (HCNR 602). Moreover, the DNase I hypersensitivity data provided evidence for potential *cis*-regulatory activity of further HCNRs (HCNR 602, 606, and 622), two of which comprise RLS-associated variants (HCNR 602 and 622). It has been shown that genetic variants in strong LD map to clusters of enhancers, which cooperatively affect gene expression and thus confer concerted susceptibility to disease (Corradin et al. 2014). Given our results, such modular organization with multiple enhancers next to the identified enhancer HCNR 617 might play a role in the pathophysiology of RLS. This concept is in line with the identified higher-risk haplotype for RLS, defined by rs12469063 and rs6710341. For all associated common SNPs in all seven HCNRs, this haplotype carries the respective risk alleles, suggesting a possible functional role of these variants next to rs12469063 within the *MEIS1* locus.

Setting the focus of our study on common variants within HCNRs increased its power but also entailed some limitations. We focused strictly on the 32-kb block of strong LD identified in the initial RLS GWAS (Winkelmann et al. 2007), which is embedded in a larger 64-kb block of weaker LD, and consequently have not considered variants in long-distance LD with the lead SNPs. In addition, rare and low-frequency variants as well as other types of genetic variation have not been analyzed. However, our functional zebrafish screen clearly identified HCNR 617 containing the RLS-associated SNP rs12469063 as the only allele-dependent enhancer, highlighting its functional importance in RLS.

Behavioral consequences of reduced *Meis1* expression showed hyperactivity and increased energy expenditure in mutant mice. Assessing however, a murine phenotype for alterations of a behavioral pattern is challenging. Here, the observed behavior of mutant mice correlates well with the human phenotype hallmarked by an enhanced activity and motor restlessness. We cannot exclude that additional factors modify the observed phenotype as MEIS1 is a pleiotropic protein and also expressed in other movement-related CNS regions. Yet, our findings support a loss-of-function model as a disease mechanism. This is in line with loss-of-function models in mouse and *Drosophila* for the RLS-associated locus *BTBD9*, also demonstrating hyperactive behavior (DeAndrade et al. 2012; Freeman et al. 2012).

Table 1. MEIS1 target genes in the ganglionic eminences of male and female E12.5 heterozygous *Meis1*^{tm1Mtor} mice

GO biological process			
Category	Subcategory	P-value	Genes
Male			
Neurological disease	Progressive motor neuropathy	4.72×10^{-03}	<i>Mapt, Nrnx1, Ptpcr, Rbfox1, Slc1a1, Slc24a3, Sv2a, Tcea1</i>
	Schizophrenia	6.68×10^{-03}	<i>Aldh3b1, Chrn3, Nrnx1, Rbfox1, Rit2, Slc1a1, Slc25a27, Stmn2, Vsnl1</i>
	Amyotrophic lateral sclerosis	9.93×10^{-03}	<i>Nrnx1, Rbfox1, Slc1a1, Tcea1</i>
	Neurological signs	1.19×10^{-02}	<i>Chrn3, Fbxw7, Map3k12, Mapt, Pde4dip, Ptpcr, Rbfox1, Slc1a1, Sv2a, Vsnl1</i>
	Multiple sclerosis	1.74×10^{-02}	<i>Mapt, Ptpcr, Slc24a3, Sv2a</i>
	Neurodegeneration	1.84×10^{-02}	<i>Idua, Mapt, Slc1a1, Ube3a</i>
	Seizure disorder	2.52×10^{-02}	<i>Mapt, Nrnx1, Pde4dip, Ptpcr, Rbfox1, Sv2a, Ube3a</i>
	Dyskinesia	2.64×10^{-02}	<i>Chrn3, Fbxw7, Map3k12, Pde4dip, Ptpcr, Rbfox1, Slc1a1, Sv2a, Vsnl1</i>
	Rett Syndrome	2.88×10^{-02}	<i>Pde4dip, Ube3a, Zfp361l1</i>
	Movement disorders	2.92×10^{-02}	<i>Chrn3, Fbxw7, Hctr1, Idua, Map3k12, Mapt, Pde4dip, Ptpcr, Rbfox1, Slc1a1, Sv2a, Ube3a, Vsnl1</i>
Nervous system development	Neuromuscular disease	3.10×10^{-02}	<i>Fbxw7, Map3k12, Mapt, Pde4dip, Ptpcr, Ptpcr, Rbfox1, Slc1a1, Slc24a3, Sv2a, Vsnl1</i>
	Synaptic transmission	1.50×10^{-03}	<i>Chrn3, Hctr1, Mapt, Nrnx1, Rit2, Slc1a1, Sv2a</i>
	Long-term potentiation	3.46×10^{-02}	<i>Iapp, Mapt, Slc1a1, Ube3a</i>
	Morphology of cerebellum	4.38×10^{-02}	<i>Gbx2, Idua, Ube3a</i>
	Morphology of central nervous system	4.62×10^{-02}	<i>Gbx2, Idua, Irx6, Map3k12, Mapt, Ptpcr, Ube3a</i>
Cardiovascular disease	Vascular disease	3.39×10^{-04}	<i>Chd13, Chrn3, Gbx2, Gp5, Mapt, Mmp13, Mpl, Pcsk2, Rbfox1, Stmn2, Sv2a, Tfpi, Tnfs4, Vsnl1</i>
	Stroke	6.12×10^{-03}	<i>Chrn3, Mapt, Sv2a, Vsnl1</i>
	Thrombosis	8.06×10^{-03}	<i>Gp5, Mpl, Stmn2, Tfpi</i>
	Occlusion of blood vessel	2.59×10^{-02}	<i>Cdh13, Gp5, Mapt, Mmp13, Pcsk2, Rbfox1, Tnfs4</i>
	Infarction	2.64×10^{-02}	<i>Hpgds, Mapt, Slc16a7, Tfpi, Tnfs4</i>
Female			
Neurological disease	Encephalitis	3.73×10^{-03}	<i>Abcg2, Cblb, Esr1, Hspa5, Ncf1, Ppp3ca, Rorc</i>
Nervous system development	Abnormal morphology of vertebrae	1.27×10^{-02}	<i>Esr1, Hells, Hoxa3, Plagl1</i>
Behavior	Maternal nurturing	2.68×10^{-03}	<i>Esr1, Foxb1, Npas3</i>
	Memory	1.35×10^{-02}	<i>Esr1, Foxb1, Ncf1, Npas3, Ppp3ca</i>
Cardiovascular system and disease	Mass of heart	1.17×10^{-02}	<i>Esr1, Klk3, Plagl1, Ppp3ca</i>
	Myocardial infarction	3.70×10^{-02}	<i>Esr1, Itgb2, Masp2, Ppp3ca</i>
	Vasculogenesis	3.94×10^{-02}	<i>Ada, Bmx, Esr1, Folh1, Itgb2, Klk3, Ltbp1, Slc9a3r2, Tiparp, Wars2</i>

Significantly overrepresented functional annotations among regulated genes of heterozygous mutant embryos are listed for male and female mice. Analysis was done using the Ingenuity pathway analysis tool.

Recently, thousands of common variants have been identified for common traits causing a big debate about the underlying biology of common genetic variants conveying only small effect sizes. The greatest challenge in the “post-GWAS” era is to understand the functional consequences of these loci. Here, we showed a SNP-dependent enhancer impairment, suggesting RLS-relevant differences in the regulation of downstream target genes in the embryonic ganglionic eminences, the prospective basal ganglia. This supports the concept of RLS having elements of a neurodevelopmental disorder. Now, for the first time, we are getting closer to a better understanding of the neuronal pathways involved in RLS, enabling us to explore new drugs and treatment strategies in the future.

Methods

Mutational screening, SNP discovery panel

One hundred eighty-eight unrelated RLS-affected individuals (72.8% female, mean age 60.0 ± 11.2 yr) were screened for variants; 94 patients were homozygote for the risk allele of rs12469063 (G/G), and 94 patients were heterozygote (G/A). We used high-

resolution melting curve analysis with LightScanner (IDAHO Technology, Inc.). The targeted region was the 32-kb LD block (hg18; chr2: 66,601,989–66,635,212) as defined by the fine-mapping of the original GWAS signal using tagging-SNPs of the *MEIS1* gene and ± 10 kb 5' and 3' of sequence in 903 cases and 891 controls of European ancestry (Winkelmann et al. 2007). Oligonucleotide sequences for 145 amplicons are available in Supplemental Table S7. DNA was analyzed in doublets to prevent false-positive signals. Samples with aberrant melting patterns were Sanger sequenced.

Association study: case/control population, imputation, and statistical analysis

Genotyping was performed in 735 unrelated RLS-affected individuals (70.8% female, mean age 61.5 ± 14.2 yr) from Germany, and 735 age-, sex-, and ethnicity-matched population-based controls of the KORA cohort (74.5% female, mean age 59.8 ± 11.3 yr) using the MassARRAY system and Sequenom iPLEX Gold technology (Sequenom, Inc.). Oligonucleotide sequences are available in Supplemental Table S8. A total of 59 variants were selected for genotyping (Supplemental Table S1; Supplemental Methods).

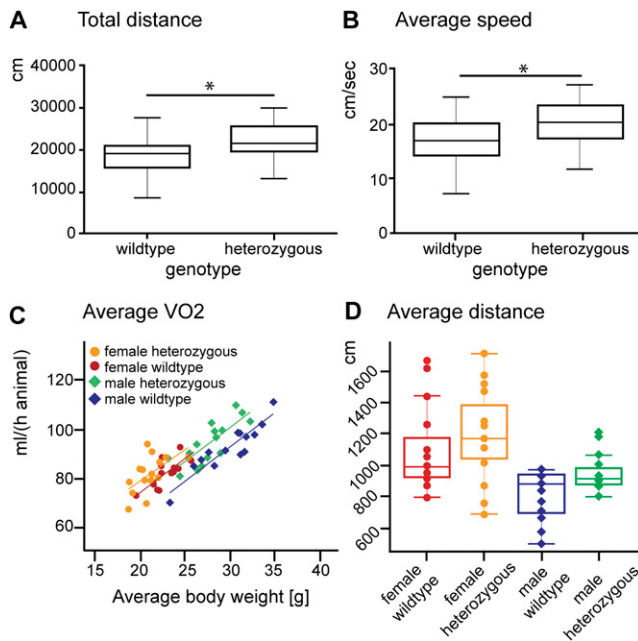


Figure 6. Motor restlessness/hyperactivity in adult heterozygous *Meis1^{tm1Mtor}* mice. Heterozygous *Meis1^{tm1Mtor}* (het) traveled a higher total distance ($P = 0.003$, two-tailed t -tests, $n = 29$) (A) and moved with a higher speed on average ($P = 0.004$, two-tailed t -tests, $n = 30$) (B) in a 20-min open field test for spontaneous locomotor activity in a novel environment. (C) Oxygen consumption (mL/h) plotted versus body mass (g). Scatter-plot with regression lines split by sex and genotype to adjust for body mass variation. The shift in regression lines indicates higher energy turnover in heterozygous *Meis1^{tm1Mtor}* mice (linear regression model, $P = 0.0005$). (D) Mean distance traveled (cm/20 min) monitored by infrared light beams during the indirect calorimetry trial. Heterozygous *Meis1^{tm1Mtor}* mice tended to show increased locomotor activity under home cage conditions (two-way ANOVA, $P = 0.06$).

SNPs were imputed in a larger RLS case/control (1302/1259) set using MACH 1.0 (<http://www.sph.umich.edu/csg/abecasis/MACH/index.html>) and standard settings. Post-imputation quality control exclusion criterion was an Rsq score < 0.3 . Statistical analysis was performed as logistic regression with disease status as the dependent and allele dosage as the independent variable using gender as a covariate. All calculations were done using R 3.0.1. For the conditional analysis, either rs2300478 or rs12469063 together with rs6710341 was entered as an additional covariate. Westfall-Young minimal-P method was used to correct for multiple testing with $n = 10,000$ permutations, and the significance level was set at 0.05 (Westfall and Young 1993). The plot for Figure 1 was generated using the UCSC Genome Browser (<http://genome.ucsc.edu/>), LocusZoom v1.1 (Pruim et al. 2010), and Haploview 4.2 based on HapMap data (CEU) (The International HapMap Consortium 2003). Haplotype analysis was performed in the genotyped data set only, using both Haploview 4.2 (Barrett et al. 2005) (<http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview>) and UNPHASED 3.1.7 (Dudbridge 2008) (<https://sites.google.com/site/fdudbridge/software/unphased-3-1>). Haploview 4.2 using the method by Gabriel et al. (2002) showed that there was one risk haplotype significantly enriched (or depleted) (assessed at a 5% level after 20,000 permutations) in these cases. This haplotype was perfectly tagged by rs2300478/rs12469063 and rs6710341. For statistical testing and estimation of ORs, we used UNPHASED 3.1.7 including gender as a covariate.

Genomatix software

MatInspector and SNPInspector software (Cartharius et al. 2005) was employed to predict putative transcription factor binding sites and to analyze in silico SNP dependent differences, respectively. GeneRanker software (Berriz et al. 2003) was used for characterization of large sets of genes using annotation data from various sources, such as Gene Ontology or Genomatix proprietary annotation.

EP300 ChIP-seq and DNase I hypersensitivity profile

EP300 ChIP-seq data for forebrain of E11.5 murine embryos was mapped to all HCNRs of the LD block examined (Visel et al. 2009). We used the ENCODE/University of Washington "UW DNase I HS track," available online in the UCSC Genome Browser (<http://genome.ucsc.edu/>) (Sabo et al. 2006).

HCNR selection and cloning of HCNR reporter vectors

For determination of HCNRs, we used the VISTA Genome Browser v2.0 (<http://pipeline.lbl.gov/cgi-bin/gateway2>) (Dubchak et al. 2000). HCNR selection settings are given in the Supplemental Methods. HCNRs were amplified from human DNA. For each HCNR, two constructs with the protective and risk alleles of the associated SNPs were cloned. For common SNPs located in HCNRs, but not genotyped successfully, the respective construct included either the allele enriched in the discovery sample (rs113851554, allele T) or the allele correlated with the risk allele of a tagging SNP (rs11897119, allele T) as risk allele. All vectors were Sanger sequenced and did not include any additional rare variants. Primer sequences and vectors are detailed in the Supplemental Methods.

Generation and analysis of transgenic animals

In zebrafish, microinjections and screening was performed as previously reported (Royo et al. 2011). Transgenic mouse lines (C57BL/6, Harlan) were generated by pronucleus injection of linearized vectors. Primer sequences used for genotyping are detailed in the Supplemental Methods. Intensity of beta-galactosidase staining was measured using ImageJ software (<http://rsbweb.nih.gov/ij/index.html>) on pictures of entire embryos ($n_{\text{protective}} = 4$; $n_{\text{risk}} = 4$). High-precision design-based stereology was used to quantify the volume of the stained area of the ganglionic eminences ($n_{\text{protective}} = 4$; $n_{\text{risk}} = 4$). Volume was determined on 8- μ m sections on a stereology workstation (Stereo Investigator, MicroBrightField). On every fourth section, the area of interest was counted using the Cavalieri estimator probe (scan grid size = 20 μ m; height of the unbiased virtual counting spaces = 12 μ m).

Histological methods

Embryos were immersion-fixed in 4% paraformaldehyde/PBS. Radioactive in situ hybridization was performed on 8- μ m paraffin sections as previously described (Giesert et al. 2013). Immunohistochemistry on paraffin sections was performed according to standard procedures. Templates for in vitro transcription of riboprobes and antibodies used are detailed in the Supplemental Methods.

Electrophoretic mobility shift assay (EMSA) and overexpression of CREB1 in 293T cells

For EMSA experiments, forebrain nuclear extracts of E12.5 mouse embryos and CREB1 protein overexpressed in 293T cells were used.

Experimental details and 5'-Cy5-labeled oligonucleotides are given in the Supplemental Methods.

Affinity chromatography and mass spectrometry

Nuclear protein extracts from forebrain of E12.5 mouse embryos were incubated with double-stranded biotinylated oligonucleotide probes. The sequence of the double-stranded biotin-labeled oligomers used was the same as in EMSA, replacing Cy5 with biotin labeling. Proteins were eluted with an ascending salt concentration (200–1250 mM NaCl). The affinity chromatography was performed three times, and the differential elution fractions were analyzed with mass spectrometry. Proteins were proteolysed with trypsin, and LC-MS/MS analysis was performed with an HPLC system (Ultimate3000, Dionex), directly coupled to an LTQ Orbitrap XL (ThermoFisher Scientific) as described previously (Merl et al. 2012). For quantitative analyses, spectra were loaded into the Progenesis LC-MS software (version 4.0, Nonlinear) and analyzed as previously described (Hauck et al. 2010). Peptide identifications were performed using Mascot (Matrix Science, version 2.3) and the Ensembl mouse database (Release 62; 54,576 sequences) with one missed cleavage allowed, a parent ion tolerance of 10 parts per million (ppm), and a fragment ion mass tolerance of 0.6 Da. Carbamidomethylation was set as fixed; methionine oxidation, and asparagine or glutamine deamidations were allowed as variable modifications. A Mascot-integrated decoy database search calculated an average peptide FDR of < 2%. Peptide assignments were reimported into Progenesis LC-MS. Normalized abundances of all unique peptides were summed up and allocated to the respective protein, comparing the protective (NR) with the risk (R) condition.

RNA isolation and transcriptome analysis

Total mRNA was extracted from forebrains of E12.5 mouse embryos (wild-type and *Meis1^{tm1Mtor/-}* mice), and transcriptomes were analyzed using MouseRef-8v2 Expression BeadChips (Illumina). Details of statistical, cluster, and pathway analyses are given in the Supplemental Methods.

Behavioral and metabolic phenotyping of *Meis1^{tm1Mtor/-}* mice

Previously described *Meis1^{tm1Mtor}* mice were phenotyped at the German Mouse Clinic (Azcoitia et al. 2005; Gailus-Durner et al. 2009). Experimental settings of reported findings are detailed in the Supplemental Methods.

Ethics

The KORA study was reviewed and approved by the local ethical committee (Bayerische Landesärztekammer). All animal work was performed in accordance with the German Animal Welfare Act and the Spanish Ethical Committee for Animal Research from Consejo Superior de Investigaciones (CSIC). Approval for the generation of the transgenic mice was obtained from the Landesdirektion Dresden (Az. 24-9168.11-9/2005-1).

Data access

The complete cRNA array data set from this study has been submitted to the NCBI Gene Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE44592. All novel variants have been submitted to NCBI dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>). Their ss-numbers are given in Supplemental Table S1.

List of affiliations

¹Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany; ²Institute of Human Genetics, Technische Universität München, 81675 Munich, Germany; ³Centro Andaluz de Biología del Desarrollo, Consejo Superior de Investigaciones Científicas/Universidad Pablo de Olavide, 41013 Seville, Spain; ⁴Institute of Developmental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany; ⁵Else Kroener-Fresenius-Centre for Nutritional Medicine, Technische Universität München, 85350 Freising-Weihenstephan, Germany; ⁶ZIEL-Research Centre for Nutrition and Food Sciences, Technische Universität München, 85350 Freising-Weihenstephan, Germany; ⁷German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany; ⁸Clinical Cooperation Group Nutrigenomics and Type 2 Diabetes, Helmholtz Zentrum München and Technische Universität München, 85350 Freising-Weihenstephan, Germany; ⁹German Mouse Clinic/Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany; ¹⁰Max Planck Institute of Psychiatry, 80804 Munich, Germany; ¹¹Research Unit Protein Science, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany; ¹²Research Unit Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany; ¹³Transgenic Core Facility, Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany; ¹⁴Munich Cluster for Systems Neurology (SyNergy), 80336 Munich, Germany; ¹⁵University of Liverpool, Institute of Translational Medicine, Liverpool, L69 3GA, United Kingdom; ¹⁶Centro Nacional de Investigaciones Cardiovasculares (CNIC), 28029 Madrid, Spain; ¹⁷German Mouse Clinic/Institute of Developmental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany; ¹⁸Molecular Nutritional Medicine, Technische Universität München Center of Life and Food Sciences, 85354 Weihenstephan, Germany; ¹⁹Technische Universität München, Experimental Genetics, 85354 Freising, Germany; ²⁰Technische Universität München, Developmental Genetics, 85764 Neuherberg, Germany; ²¹Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), 80336 Munich, Germany; ²²Department of Neurology, Technische Universität München, 81675 Munich, Germany; ²³Department of Neurology and Neurosciences, Stanford Center for Sleep Medicine and Sciences, Stanford University, Palo Alto, California 94304, USA.

Acknowledgments

We thank Len A. Pennacchio for the Hsp68-*lacZ*-plasmid and Irmgard Zaus, Jelena Golic, Silvia Naranjo, Rocio Morales, Xavier Ruiz, Bettina Sperling, Jan Einicke, Ann-Elisabeth Schwarz, Anke Bettenbrock, and the TCF Dresden team for technical assistance. The project was supported by Fritz-Thyssen-Stiftung, Cologne, Germany (10.09.2.146; 10.12.2.183), KKF-TUM (8766156), DAAD (0811963), and COST ("HOX and TALE homeoproteins in Development and Disease"). B.S. was partially supported by DFG grants (WI 1820/4-1; WI 1820/5-1) and a TUM-Excellence stipend. The KORA study was financed by the Helmholtz Zentrum München, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. J.L.G.-S. and F.C. acknowledge funding of the Spanish and the Andalusian Governments and the Feder program for grants

(BFU2010-14839, BFM2009-07044, CSD2007-00008, and Proyectos de Excelencia CVI-3488 and CVI 2658). This work was funded in part by a grant from the German Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD), to the German Mouse Clinic (Infrafrontier: 01KX1012), to the German Center for Neurodegenerative Diseases (DZNE), Germany; by the Initiative and Networking Fund of the Helmholtz Association in the framework of the Helmholtz Alliance for Mental Research in an Ageing Society (HA-215); and the Munich Cluster for Systems Neurology (EXC 1010 SyNergy) and its Collaborative Research Center (CRC) 870/2 "Assembly and Function of Neuronal Circuits."

Author contributions: D.S. and J.W. conceived and designed the study. Wet-lab experiments were performed by D.S., M. Kaffé, F.K., J. Bessa, J.J.T., F.G., H. Lee, and M.C., and supervised by T.M., H. Laumen, W.W., F.C., J.L.G.-S., and J.W. Mutation screen and sequencing was performed by F.K. and supervised by P.L. and J.W. SNP data was imputed by B.M.-M., and B.M.-M., B.S., D.C., and N.K. performed the statistical analysis. D.S. performed the ChIP-seq and DNase I hypersensitivity data analysis. F.K. performed the stereological volume estimation. M. Kaffé, C.v.T., and S.M. Hauck conceived, designed, and performed the mass spectrometry experiments; C.v.T. and S.M. Hauck analyzed and quantified the mass spectrometry data. R.N. performed the pronucleus injections. E.T. performed the PreSTIGE analysis. M.T. generated the *Meis1*-deficient mice. L.G., J.R., M. Klingenspor, and S.M. Höltér conceived, designed, and performed the phenotyping experiments; and J.R., L.G., and S.M. Höltér analyzed and interpreted the phenotyping data. The study design of the GMC (German Mouse Clinic) was conceived by V.G.-D., H.F., and M.H.A. KORA cohort design was conceived by C.G.; M.W. provided the population-based controls. Transcriptome data was obtained by M.H. and analyzed by M.H. and J. Beckers. The manuscript was written by D.S., M. Kaffé, F.K., B.S., and J.W. All authors took part in the revision of the manuscript and approved the final version.

References

- Azcoitia V, Aracil M, Martínez AC, Torres M. 2005. The homeodomain protein *Meis1* is essential for definitive hematopoiesis and vascular patterning in the mouse embryo. *Dev Biol* **280**: 307–320.
- Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**: 263–265.
- Bartsch D, Casadio A, Karl KA, Serodio P, Kandel ER. 1998. *CREB1* encodes a nuclear activator, a repressor, and a cytoplasmic modulator that form a regulatory unit critical for long-term facilitation. *Cell* **95**: 211–223.
- Berriz GF, King OD, Bryant B, Sander C, Roth FP. 2003. Characterizing gene sets with FuncAssociate. *Bioinformatics* **19**: 2502–2504.
- Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T. 2005. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* **21**: 2933–2942.
- Chen L, Chatterjee M, Li JY. 2010. The mouse homeobox gene *Gbx2* is required for the development of cholinergic interneurons in the striatum. *J Neurosci* **30**: 14824–14834.
- Corradin O, Saiakhova A, Akhtar-Zaidi B, Myeroff L, Willis J, Cowper-Sal Lari R, Lupien M, Markowitz S, Scacheri PC. 2014. Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. *Genome Res* **24**: 1–13.
- Dauvilliers Y, Winkelmann J. 2013. Restless legs syndrome: update on pathogenesis. *Curr Opin Pulm Med* **19**: 594–600.
- DeAndrade MP, Johnson RL Jr, Unger EL, Zhang L, van Groen T, Gamble KL, Li Y. 2012. Motor restlessness, sleep disturbances, thermal sensory alterations and elevated serum iron levels in *Btd9* mutant mice. *Hum Mol Genet* **21**: 3984–3992.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**: 376–380.
- Dubchak I, Brudno M, Loots GG, Pachter L, Mayor C, Rubin EM, Frazer KA. 2000. Active conservation of noncoding sequences revealed by three-way species comparisons. *Genome Res* **10**: 1304–1306.
- Dudbridge F. 2008. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum Hered* **66**: 87–98.
- Freeman A, Pranski E, Miller RD, Radmard S, Bernhard D, Jinnah HA, Betarbet R, Rye DB, Sanyal S. 2012. Sleep fragmentation and motor restlessness in a *Drosophila* model of Restless Legs Syndrome. *Curr Biol* **22**: 1142–1148.
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, et al. 2002. The structure of haplotype blocks in the human genome. *Science* **296**: 2225–2229.
- Gailus-Durner V, Fuchs H, Adler T, Aguilar Pimentel A, Becker L, Bolle I, Calzada-Wack J, Dalke C, Ehrhardt N, Ferwagner B, et al. 2009. Systemic first-line phenotyping. *Methods Mol Biol* **530**: 463–509.
- Giesert F, Hofmann A, Bürger A, Zerle J, Kloos K, Hafen U, Ernst L, Zhang J, Vogt-Weisenhorn DM, Wurst W. 2013. Expression analysis of *Lrrk1*, *Lrrk2* and *Lrrk2* splice variants in mice. *PLoS ONE* **8**: e63778.
- Hauck SM, Dietter J, Kramer RL, Hofmaier F, Zipplies JK, Amann B, Feuchtinger A, Deeg CA, Ueffing M. 2010. Deciphering membrane-associated molecular processes in target tissue of autoimmune uveitis by label-free quantitative mass spectrometry. *Mol Cell Proteomics* **9**: 2292–2305.
- Herrmann MG, Durtschi JD, Bromley LK, Wittwer CT, Voelkerding KV. 2006. Amplicon DNA melting analysis for mutation scanning and genotyping: cross-platform comparison of instruments and dyes. *Clin Chem* **52**: 494–503.
- Hornyak M, Trenkwalder C, Kohlen R, Scholz H. 2012. Efficacy and safety of dopamine agonists in restless legs syndrome. *Sleep Med* **13**: 228–236.
- The International HapMap Consortium. 2003. The International HapMap Project. *Nature* **426**: 789–796.
- Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, Poh HM, Goh Y, Lim J, Zhang J, et al. 2012. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**: 84–98.
- Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, et al. 2012. Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**: 1190–1195.
- Medina L, Abellán A. 2012. Subpallial structures. In *The mouse nervous system* (ed. Watson C, Paxinos G, Puelles L), pp. 173–219. Elsevier, London; Waltham, San Diego.
- Merl J, Ueffing M, Hauck SM, von Toerne C. 2012. Direct comparison of MS-based label-free and SILAC quantitative proteome profiling strategies in primary retinal Müller cells. *Proteomics* **12**: 1902–1911.
- Penkov D, Mateos San Martín D, Fernández-Díaz LC, Rosselló CA, Torroja C, Sánchez-Cabo F, Warnatz HJ, Sultan M, Yaspo ML, Gabrieli A, et al. 2013. Analysis of the DNA-binding profile and function of TALE homeoproteins reveals their specialization and specific interactions with Hox genes/proteins. *Cell Rep* **3**: 1321–1333.
- Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD, et al. 2006. *In vivo* enhancer analysis of human conserved non-coding sequences. *Nature* **444**: 499–502.
- Prabhakar S, Poulin F, Shoukry M, Afzal V, Rubin EM, Couronne O, Pennacchio LA. 2006. Close sequence comparisons are sufficient to identify human *cis*-regulatory elements. *Genome Res* **16**: 855–863.
- Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Glied TP, Boehnke M, Abecasis GR, Willer CJ. 2010. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**: 2336–2337.
- Royo JL, Maeso I, Irimia M, Gao F, Peter IS, Lopes CS, D'Aniello S, Casares F, Davidson EH, García-Fernández J, et al. 2011. Transphyletic conservation of developmental regulatory state in animal evolution. *Proc Natl Acad Sci* **108**: 14186–14191.
- Royo JL, Bessa J, Hidalgo C, Fernández-Miñán A, Tena JJ, Roncero Y, Gómez-Skarmeta JL, Casares F. 2012. Identification and analysis of conserved *cis*-regulatory regions of the *MEIS1* gene. *PLoS ONE* **7**: e33617.
- Sabo PJ, Kuehn MS, Thurman R, Johnson BE, Johnson EM, Cao H, Yu M, Rosenzweig E, Goldy J, Haydock A, et al. 2006. Genome-scale mapping of DNase I sensitivity *in vivo* using tiling DNA microarrays. *Nat Methods* **3**: 511–518.
- Schmidt D, Wilson MD, Ballester B, Schwalie PC, Brown GD, Marshall A, Kutter C, Watt S, Martínez-Jiménez CP, Mackay S, et al. 2010. Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* **328**: 1036–1040.
- Schormair B, Kemlink D, Roeske D, Eckstein G, Xiong L, Lichtner P, Ripke S, Trenkwalder C, Zimprich A, Stiasny-Kolster K, et al. 2008. *PTPRD* (protein tyrosine phosphatase receptor type δ) is associated with restless legs syndrome. *Nat Genet* **40**: 946–948.

- Somel M, Guo S, Fu N, Yan Z, Hu HY, Xu Y, Yuan Y, Ning Z, Hu Y, Menzel C, et al. 2010. MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res* **20**: 1207–1218.
- Szentkiralyi A, Fendrich K, Hoffmann W, Happe S, Berger K. 2011. Incidence of restless legs syndrome in two population-based cohort studies in Germany. *Sleep Med* **12**: 815–820.
- Theiler K. 1989. *The house mouse, atlas of embryonic development*. Springer, New York.
- Toresson H, Parmar M, Campbell K. 2000. Expression of *Meis* and *Pbx* genes and their protein products in the developing telencephalon: implications for regional differentiation. *Mech Dev* **94**: 183–187.
- Tucker ES, Segall S, Gopalakrishna D, Wu Y, Vernon M, Polleux F, Lamantia AS. 2008. Molecular specification and patterning of progenitor cells in the lateral and medial ganglionic eminences. *J Neurosci* **28**: 9504–9518.
- Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, et al. 2009. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**: 854–858.
- Westfall PH, Young SS. 1993. *Resampling-based multiple testing: Examples and methods for p-value adjustment*. Wiley, New York.
- Wichmann HE, Gieger C, Illig T, Group MKS. 2005. KORA-gen—resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* (Suppl 1) **67**: S26–S30.
- Winkelmann J, Schormair B, Lichtner P, Ripke S, Xiong L, Jalilzadeh S, Fulda S, Pütz B, Eckstein G, Hauck S, et al. 2007. Genome-wide association study of restless legs syndrome identifies common variants in three genomic regions. *Nat Genet* **39**: 1000–1006.
- Winkelman JW, Shahar E, Sharief I, Gottlieb DJ. 2008. Association of restless legs syndrome and cardiovascular disease in the Sleep Heart Health Study. *Neurology* **70**: 35–42.
- Winkelmann J, Czamara D, Schormair B, Knauf F, Schulte EC, Trenkwalder C, Dauvilliers Y, Polo O, Högl B, Berger K, et al. 2011. Genome-wide association study identifies novel restless legs syndrome susceptibility loci on 2p14 and 16q12.1. *PLoS Genet* **7**: e1002171.
- Woolfe A, Goodson M, Goode DK, Snell P, McEwen GK, Vavouri T, Smith SE, North P, Callaway H, Kelly K, et al. 2005. Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol* **3**: e7.

Received September 15, 2013; accepted in revised form February 5, 2014.

Restless Legs Syndrome-associated intronic common-variant in *Meis1* alters enhancer function in the developing telencephalon

Supplemental Material

Derek Spieler^{1, 25}, Maria Kaffe^{1,2, 25}, Franziska Knauf^{1, 25}, José Bessa³, Juan J. Tena³, Florian Giesert⁴, Barbara Schormair^{1,2}, Erik Tilch^{1,2}, Heekyoung Lee^{5,6,7,8}, Marion Horsch⁹, Darina Czamara¹⁰, Nazanin Karbalai¹⁰, Christine von Toerne¹¹, Melanie Waldenberger¹², Christian Gieger¹², Peter Lichtner¹, Melina Claussnitzer^{5,6,7,8,13}, Ronald Naumann¹⁴, Bertram Müller-Myhsok^{10,15,16}, Miguel Torres¹⁷, Lillian Garrett¹⁸, Jan Rozman^{9,7,19}, Martin Klingenspor^{5,6,9,19}, Valérie Gailus-Durner⁹, Helmut Fuchs⁹, Martin Hrabě de Angelis^{9,7,20}, Johannes Beckers^{7,9,20}, Sabine M. Höfler¹⁸, Thomas Meitinger^{1,2, 15}, Stefanie M. Hauck^{7,11}, Helmut Laumen^{5,6,7,8,9}, Wolfgang Wurst^{4,10,15,18,21,22}, Fernando Casares³, Jose Luis Gómez-Skarmeta³, Juliane Winkelmann^{1,15,23,24,26}

Supplemental Figures

Figure S1: Results of the enhancer screen in zebrafish.

Figure S2: Theiler-staged transgenic mouse embryos for intensity analysis and volume estimation of the reporter signal.

Figure S3: Volume estimations of the beta-galactosidase signal according to Cavalieri.

Figure S4: Missing difference in hyperactivity due to fear related behavior.

Figure S5: Poll II ChIA-PET and Hi-C experiments in two human cell lines focusing on the RLS-associated intronic region.

Supplemental Tables

Table S1: Results of the high resolution melting curve analysis (“SNP discovery panel”) of the RLS-associated 32 kb LD block (NCBI36/hg18).

Table S2: Results of the case/control association study with the identified variants, covering the RLS-associated 32 kb LD block.

Table S3: List of the 51 proteins identified by affinity chromatography/mass spectrometry as differentially bound by rs12469063.

Table S4: Heat map from unpaired two class analysis (SAM) of genes regulated between *Meis1*^{tm1Mtor/+} and wildtype forebrain tissue at E12.5 expanded to the level of individual genes.

Table S5: Results of calorimetry trial: oxygen consumption and mean distance travelled.

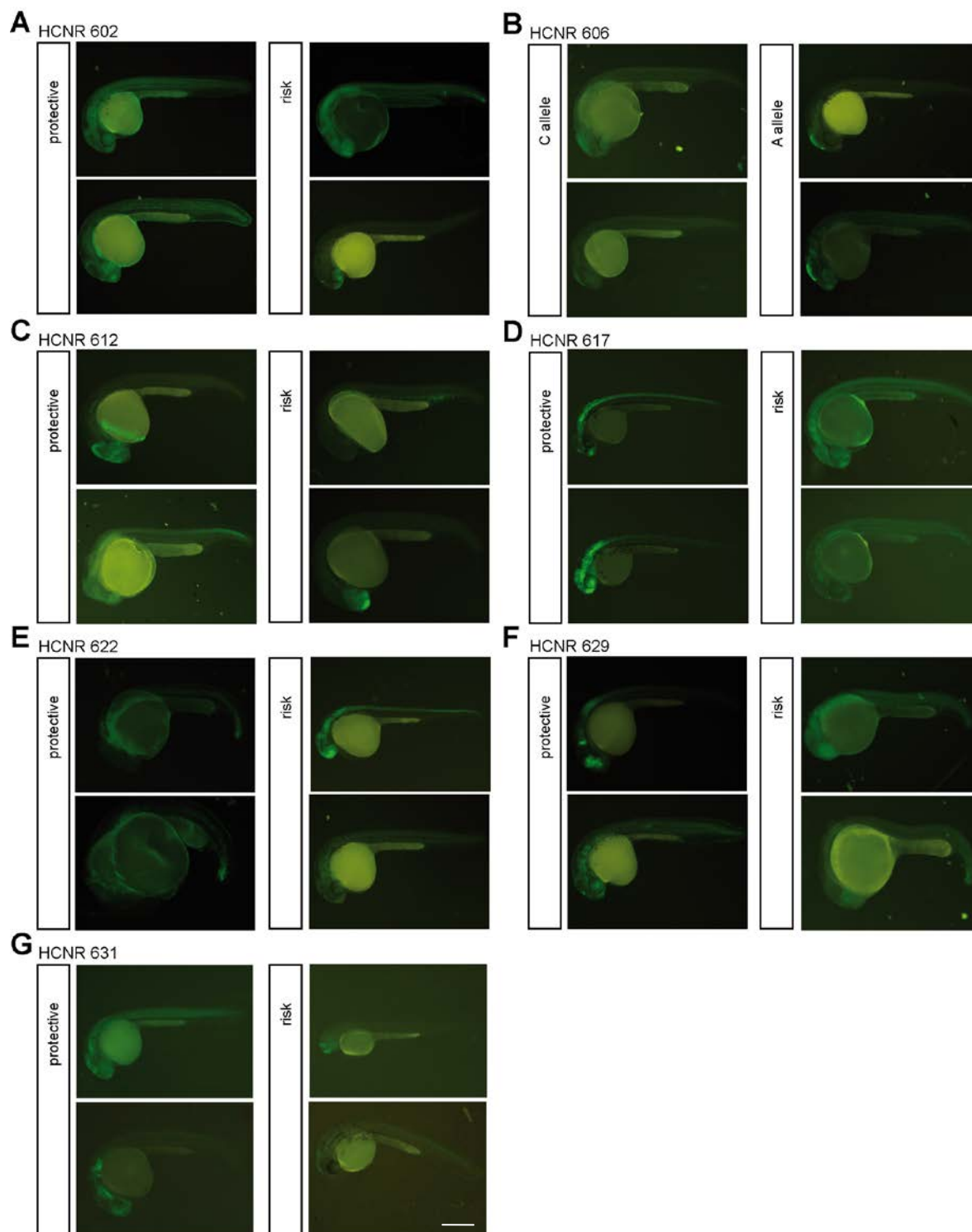
Table S6: PreSTIGE-analysis focusing on the RLS-associated LD block suggests solely *MEIS1* as regulated gene.

Table S7: Oligonucleotide sequences used for "high resolution melting curve analysis" amplicons

Table S8: Oligonucleotide sequences used in the “Sequenom iPLEX Gold” genotyping

Supplemental Experimental Methods

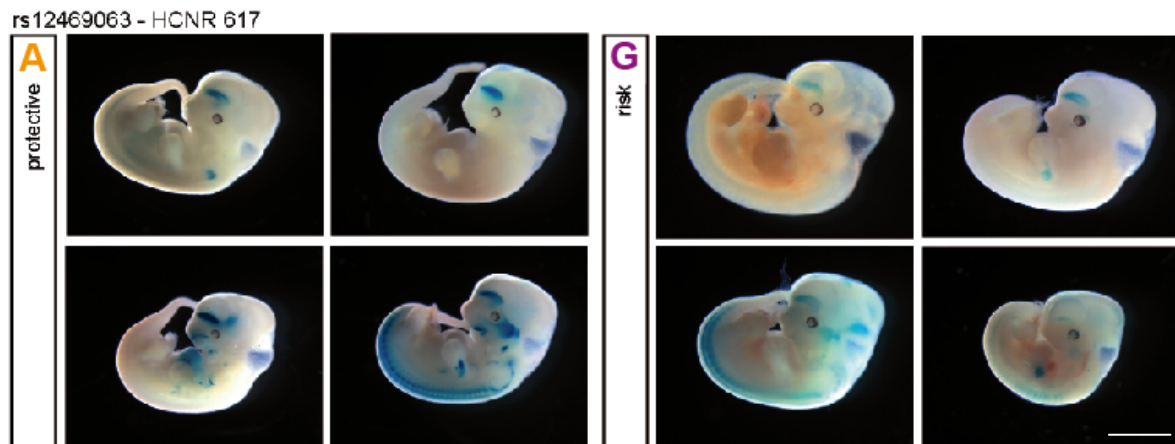
Figure S1: Results of the enhancer screen in zebrafish.



Representative transgenic fish ~24 hpf. (A) Booster activity of HCNR 602 in zebrafish with the protective alleles (left) and the risk alleles (right), respectively. Green EGFP signal refers to variable regions in the fish. (B) Robust enhancer function of HCNR 606 in the midbrain. Representative transgenic fish with the A allele (left) and the C allele (right) in HCNR 606,

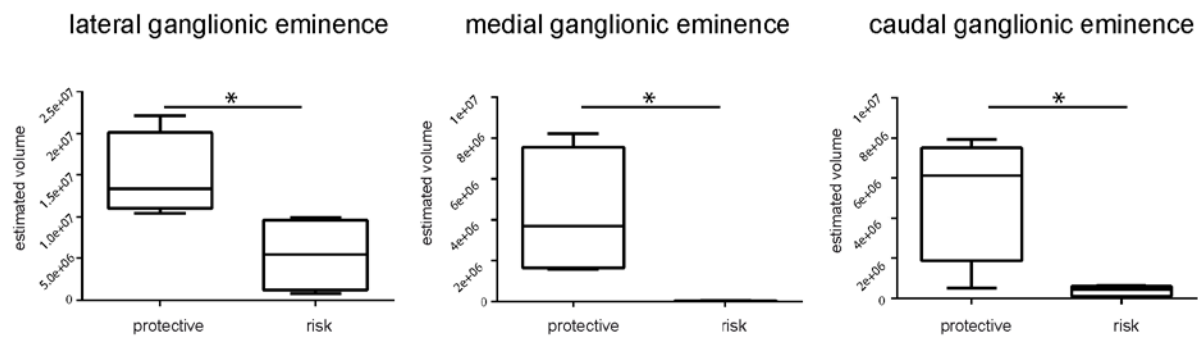
respectively. Green EGFP signal refers to the tectum. (C) Likely lack of enhancer function of HCNR 612. No reproducible pattern could be identified, although the results remained inconclusive due to the low number of identified founders. (D) Allele-specific enhancer HCNR 617 in retina, forebrain, hindbrain, and spinal cord. EGFP signal represents the regions of enhancer function of the protective allele construct – retina, forebrain, hindbrain and spinal cord. Fish in the right column displayed EGFP signal at background level. (E) Inconclusive enhancer function of HCNR 622 in hematopoietic stem cells. EGFP signal in the protective allele carriers was located in the hematopoietic stem cells. (F) Neural booster activity of HCNR 629. Green EGFP signal refers to variable regions in the central nervous system (CNS), risk allele carriers displayed EGFP only at background level. (G) Neural booster activity of HCNR 631. EGFP signal refers to variable regions in the CNS, risk allele carriers displayed EGFP only at background level (bar in G represents 250µm).

Figure S2: Theiler-staged transgenic mouse embryos for intensity analysis and volume estimation of the reporter signal.



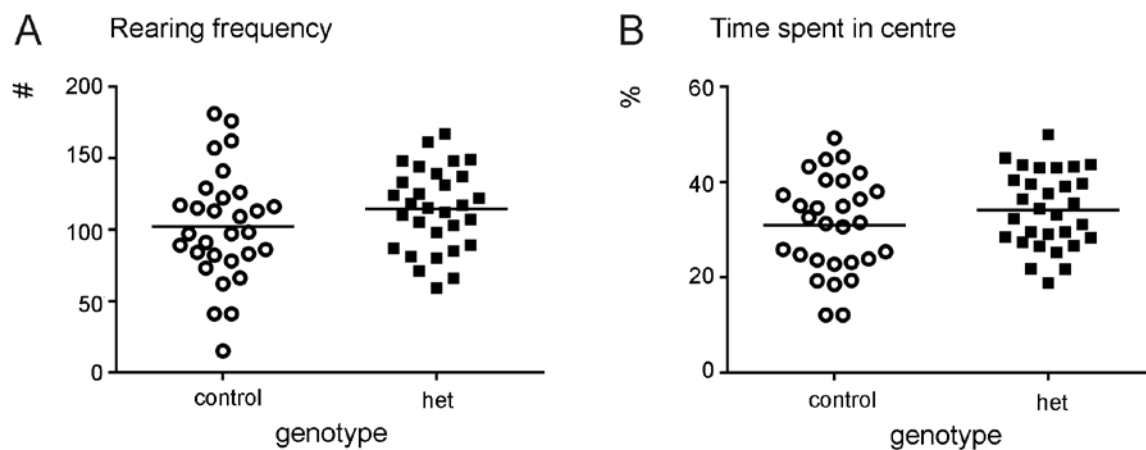
Transgenic embryos were staged according to Theiler (Theiler stages 19 and 20) (Theiler 1989). Blue color refers to regions expressing beta-galactosidase. Note the difference in signal intensity (compare left column, protective allele carriers, with right column, risk allele carriers) (bar represents 2,5 mm).

Figure S3: Volume estimations of the beta-galactosidase signal according to Cavalieri.



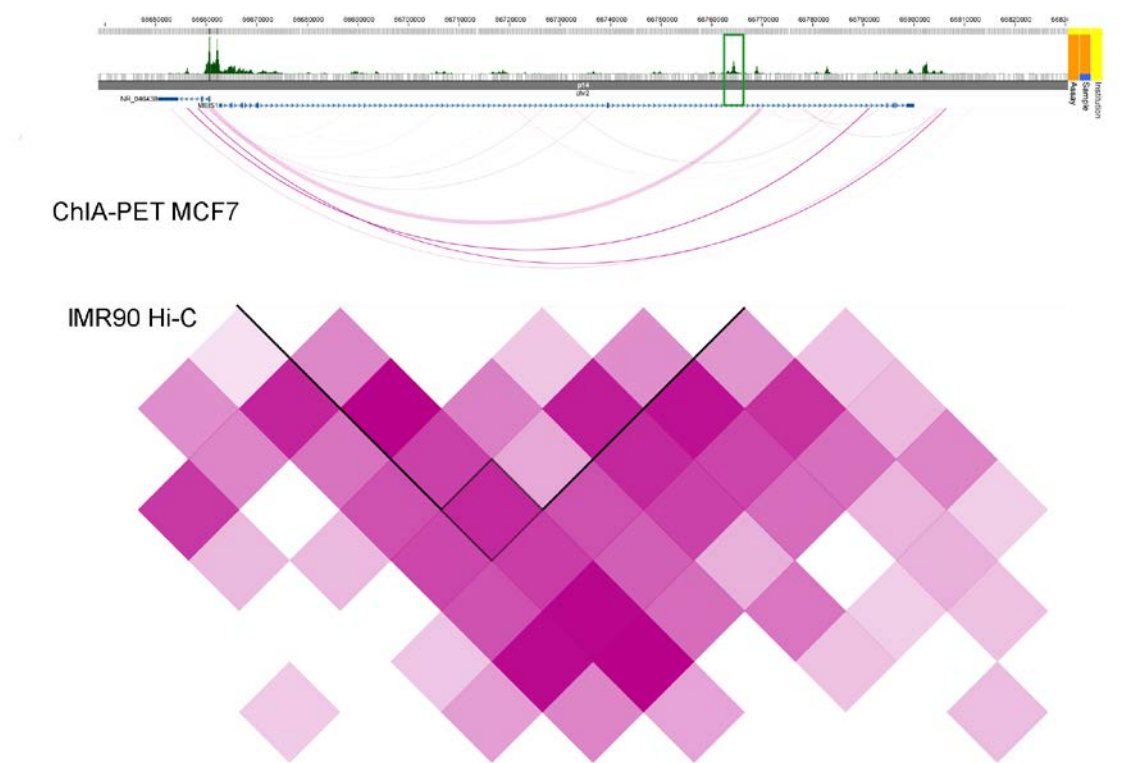
Analysis of transient transgenic mouse embryos via stereomicroscopy. The ganglionic eminences are split in lateral, medial and caudal. Asterisks refer to a significant difference in volume (Wilcoxon rank sum test, $P \leq 0.05$).

Figure S4: Missing difference in hyperactivity due to fear related behavior.



Rearing frequency (A) and the percentage of time spend in the aversive centre of the test arena (B), indicating anxiety related behavior, in a 20 min Open Field test for spontaneous locomotor activity in a novel enrichment. No differences could be detected between the genotypes (n.s.; n = 29-30).

Figure S5: Pol II ChIA-PET and Hi-C experiments in two human cell lines focusing on the RLS-associated intronic region.



Analysis of two human cell lines (human fetal lung cells, IMR90; breast cancer cells, MCF7) illustrates interaction (cf. three curved lines) between the 32 kb RLS-associated block and the promoter of *MEIS1*. Chromatin Interaction Analysis by Paired-End Tag Sequencing (Pol II ChIA-PET) and chromosome conformation capture (3C) coupled with next-generation sequencing (Hi-C) (Dixon et al. 2012; Li et al. 2012) identified interaction of the intronic RLS-associated 32 kb LD block with the regulatory landscape of *MEIS1*. Publicly available data illustrated with the human epigenomic browser of the Washington University School of Medicine, St.Louis, USA, (<http://epigenomegateway.wustl.edu/info/>) (Zhou et al. 2013). The triangle shapes in the Hi-C track depict chromatin domains in IMR90 cells and the arcs in the ChIA-PET tracks indicate similar domain structures in MCF7 cells.

Table S1: Results of the high resolution melting curve analysis (“SNP discovery panel”) of the RLS-associated 32 kb LD block (NCBI36/hg18).

Variant	HCNR element	Genome position	Minor allele	Major allele	MAF	Selection for follow-up genotyping
rs147475118	602	66602604	A	C	0.005	
rs114033747	602	66602608	A	C	0.005	
rs4316926	602	66602612	G	A	0.125	tagged by rs4430933 and rs4433986 with $r^2 = 1.0$
rs62144055	602	66602814	C	A	0.11	tagged by rs11688599 and rs6742861 with $r^2 = 1.0$
rs74559538	602	66603066	C	G	0.01	
rs4430933	602	66603114	A	G	0.13	genotyped
rs6724747	602	66603332	A	G	0.11	genotyped
ss947429183	602	66603477	A	G	0.003	
rs4544423	602	66603521	T	G	0.025	genotyped
rs6742861	602	66603656	C	T	0.315	genotyped
rs6728018	602	66603769	T	C	0.11	genotyped
rs73937943	602	66603788	A	C	0.003	
rs73937944	602	66603806	T	C	0.003	
rs113851554	602	66604068	T	G	0.42	genotyped
rs59088239	602	66604108	G	A	0.03	
rs34545468		66604989	A	G	0.003	
ss947429184		66605464	T	A	0.003	
ss947429185		66605609	G	C	0.003	
rs4233937		66605755	A	G	0.138	genotyped
rs145264250		66605850	DEL	AAAC	0.138	genotyped
rs62144057		66606060	G	C	0.109	tagged by rs11688599 and rs11678354 with $r^2 = 1.0$
rs114947399		66606301	C	G	0.008	
ss947429186	606	66606746	A	G	0.016	
rs72822902	606	66607381	A	G	0.005	
rs4480973	606	66607557	A	C	0.088	genotyped
ss947429187	606	66607732	G	A	0.003	
ss947429188	606	66607735	T	C	0.005	
ss947429189	606	66607870	T	C	0.005	
ss947429190		66608378	C	T	0.003	
rs72824804		66608457	G	A	0.064	tagged by rs6710341 with $r^2 = 0.941$
rs147083355		66608510	C	T	0.003	
rs56383906		66608878	C	T	0.011	
rs79464024		66609106	T	C	0.027	
ss947429191		66609324	A	G	0.003	
rs11687267		66609453	A	G	0.109	tagged by rs11688599 and rs11678354 with $r^2 = 1.0$
rs12465087		66609493	T	A	0.019	
rs13003040		66609774	G	A	0.13	tagged by rs430933 and rs4316926 with $r^2 = 1.0$
rs11677371		66609917	A	G	0.13	tagged by rs430933 and rs4316926 with $r^2 = 1.0$
rs11688599		66610054	T	C	0.13	genotyped
rs11694675		66610140	A	T	0.109	tagged by rs11688599 and rs11678354 with $r^2 = 0.927$
ss947429192		66610215	C	G	0.003	
rs3891585		66610480	A	G	0.13	tagged by rs430933 and rs4316926 with $r^2 = 1.0$
rs55742206		66610571	A	G	0.109	tagged by rs11688599 and

						rs11678354 with r2 = 1.0
ss947429193		66610676	T	G	0.003	
rs62145760		66610954	T	C	0.109	tagged by rs11688599 and rs11678354 with r2 = 1.0
rs182588061		66611213	T	G	0.005	
rs6710341		66611926	A	G	0.197	genotyped
rs12621948	612	66612416	G	C	0.1	genotyped
ss947429194	612	66612421	G	C	0.003	
rs115566126	612	66612680	C	T	0.01	
ss947429195	612	66612744	C	A	0.005	
rs11693646	612	66613347	A	G	0.11	genotyped
rs11678354	612	66613583	A	T	0.136	genotyped
rs4433986		66613815	A	G	0.136	genotyped
ss947429196		66613849	A	G	0.005	
rs55855986		66613965	DEL	T	0.13	assay design not possible, no tagging proxies available
ss947429197		66614166	A	T	0.003	
ss947429198		66614210	A	G	0.005	
rs4316931		66614409	T	G	0.088	genotyped
rs142456824		66614805	AT	DEL	0.106	genotyped
rs191999528		66614962	A	G	0.003	
ss947429199		66615029	C	T	0.003	
ss947429200		66615134	T	C	0.005	
rs7600007		66615159	G	C	0.003	
rs9789607		66615652	T	A	0.106	genotyped
ss947429201		66615877	G	C	0.003	
rs7603415		66615880	G	C	0.019	tagged by rs2216125 and rs2216120 with r2 = 0.915
rs7603236		66615915	C	T	0.125	genotyped
rs150221482		66615978	TATA	DEL	0.021	
rs5006732		66615984	T	C	0.021	
rs74514787		66616019	G	A	0.003	
rs150277491		66616114	T	C	0.035	
ss947429203		66616178	T	C	0.003	
rs189634091	616	66616545	T	C	0.02	
rs147099389	616	66616717	T	G	0.003	
ss947429205		66617103	A	G	0.005	
ss947429206		66617173	A	G	0.003	
rs12469063	617	66617812	G	A	0.25	genotyped
ss947429207	617	66617888	T	C	0.003	
rs141399173	617	66617969	T	C	0.003	
ss947429208	617	66618106	G	A	0.003	
rs11681729	617	66618204	G	A	0.11	genotyped
rs139989120		66618445	A	G	0.011	
rs10865355		66618501	A	G	0.128	genotyped
rs9789535		66619082	A	G	0.144	genotyped
ss947429209		66619204	T	G	0.003	
ss947429210		66619219	G	C	0.003	
rs186720691		66619330	G	A	0.003	
ss947429221		66619520	DEL	T	0.021	
rs11678321		66619522	T	G	0.021	
ss947429211		66620016	G	A	0.003	
rs2216125		66620287	A	G	0.016	genotyped
rs2192960		66620468	C	T	0.271	tagged by rs12469063 with r2 = 0.823
rs189306611		66620521	A	G	0.011	
rs7563429		66621051	A	T	0.016	
ss947429202		66621636	C	G	0.003	

rs13000004		66621734	T	C	0.258	tagged by rs12472245 with $r^2 = 0.810$
rs17625694		66622087	G	A	0.237	genotyped
ss947429213	622	66622320	G	A	0.003	
ss947429214	622	66622328	A	G	0.003	
rs17625724	622	66622479	A	G	0.11	genotyped
ss947429215	622	66622598	C	T	0.003	
rs17625742	622	66622658	T	C	0.125	genotyped
ss947429222	622	66623108	DEL	T	0.019	
rs11678796	622	66623539	A	G	0.11	genotyped
rs12472245	622	66624057	A	T	0.095	genotyped
rs997153	622	66624551	T	C	0.125	genotyped
rs10194077	622	66625367	C	T	0.135	genotyped
rs11897119	622	66625504	C	T	0.135	tagged by rs10194077 and rs17625694 with $r^2 = 1.0$
ss947429216		66626044	A	C	0.003	
rs13029520		66626466	T	C	0.098	genotyped
rs139413629		66626528	C	T	0.003	
rs75313256		66626823	A	G	0.016	
rs10188003		66626973	T	C	n.d.	genotyped
rs367929915	628	66628521	G	A	0.005	
ss947429217	629	66629795	G	A	0.01	
rs2216120	629	66630010	C	T	0.02	genotyped
rs13019989	629	66630028	A	G	0.08	genotyped
rs141179647	629	66630214	T	C	0.003	
rs2300477	629	66630233	G	A	0.195	genotyped
rs2192954		66631025	G	A	0.41	genotyped
rs62145814	631	66631567	T	C	0.308	genotyped
ss947429259	631	66631639	T	C	n.d.	genotyped
rs189356944	631	66631764	G	A	0.01	
rs192448151	631	66631875	C	T	0.003	
rs149534896		66632176	C	T	0.005	
rs76977770		66632785	C	T	0.003	
rs3772026		66633370	T	C	0.13	tagged by rs10194077 and rs17625694 with $r^2 = 0.933$
rs3755517		66633878	C	T	0.021	genotyped
ss947429260		66634004	T	C	0.005	
rs75226702		66634198	T	C	0.032	
ss947429218		66634481	A	G	0.003	
ss947429219		66634815	C	T	0.003	
rs76205499		66634930	T	A	0.003	
rs2300478		66634957	T	G	0.25	genotyped
rs2300480		66635212	C	T	0.093	genotyped

Results of the high resolution melting curve analysis (“SNP discovery panel”) of the RLS-associated 32 kb LD block in 188 unrelated RLS-affected individuals (94 patients homozygote for the risk allele of rs12469063 (G/G), 94 patients heterozygote (G/A)). MAF was calculated in the respective sample, n.d. – not determined, refers to common variants for which exact MAF calculation was not possible due to overlapping melting curve profiles. Variants are annotated based on dbSNP 138 with novel variants identified in this study with their new ss-id. Positions refer to genome build NCBI36/hg18. Selection for genotyping indicates common variants selected for genotyping in additional samples and with additional

information if SNP was genotyped directly or tagged. Bold highlighting indicates SNPs, which did not pass quality control and were not included in the final analysis.

Table S2: Results of the case/control association study with the identified variants covering the RLS-associated 32 kb LD block.

Variant	HCNR element	Genome position	Risk allele	MAF (minor allele)	P _{nom}	P _{corr}
rs4430933	602	66603114	G	0.37 (A)	2.88E-05	0.00010
rs6724747	602	66603332	G	0.32 (A)	6.68E-04	0.00922
rs4544423	602	66603521	G	0.37 (T)	1.30E-06	0.00010
rs6742861	602	66603656	T	0.31 (C)	3.08E-04	0.00431
rs6728018	602	66603769	C	0.31 (T)	4.71E-04	0.00772
rs4480973	606	66607557	C	0.26 (A)	1.03E-01	0.63502
rs11688599		66610054		0.31 (T)	1.47E-04	0.00140
rs6710341		66611926		0.14 (G)	4.00E-01	0.99008
rs12621948	612	66612416	C	0.28 (G)	6.89E-02	0.50241
rs11693646	612	66613347	G	0.3 (A)	1.08E-05	0.00010
rs11678354	612	66613583	T	0.3 (A)	2.01E-05	0.00010
rs4433986		66613815		0.37 (A)	8.65E-06	0.00010
rs4316931		66614409		0.27 (T)	1.10E-01	0.65457
rs9789607		66615652		0.31 (T)	7.25E-05	0.00010
rs7603236		66615915		0.36 (C)	6.93E-08	0.00010
rs12469063	617	66617812	G	0.29 (G)	7.68E-18	0.00010
rs11681729	617	66618204	A	0.31 (G)	5.38E-05	0.00010
rs10865355		66618501		0.37 (A)	1.52E-05	0.00010
rs9789535		66619082		0.41 (A)	2.83E-04	0.00391
rs2216125		66620287		0.06 (A)	6.35E-02	0.47604
rs17625694		66622087		0.37 (G)	9.35E-06	0.00010
rs17625724	622	66622479	G	0.3 (A)	6.53E-05	0.00010
rs17625742	622	66622658	C	0.31 (T)	1.48E-05	0.00010
rs11678796	622	66623539	G	0.07 (A)	4.84E-01	0.99810
rs12472245	622	66624057	T	0.28 (A)	1.13E-01	0.66510
rs997153	622	66624551	C	0.37 (T)	3.38E-06	0.00010
rs10194077	622	66625367	T	0.37 (C)	1.53E-06	0.00010
rs13029520		66626466		0.37 (T)	2.30E-05	0.00010
rs10188003		66626973		0.37 (T)	1.99E-06	0.00010
rs2216120	629	66630010	T	0.06 (C)	2.54E-01	0.91680
rs13019989	629	66630028	G	0.27 (A)	8.05E-02	0.55443
rs2300477	629	66630233	A	0.44 (A)	1.21E-09	0.00010
rs2192954		66631025		0.48 (A)	1.61E-08	0.00010
rs62145814	631	66631567	C	0.31 (T)	5.11E-05	0.00010
ss947429259	631	66631639	C	0.31 (T)	4.96E-04	0.00792
rs3755517		66633878		0.07 (C)	2.64E-01	0.92562
rs2300478		66634957		0.3 (G)	4.24E-18	0.00010
rs2300480		66635212		0.28 (C)	9.68E-03	0.12390

Results of the case/control association study performed in 1,302 unrelated RLS-affected individuals from Germany and 1,259 age-, sex-, and ethnicity-matched population-based controls of the KORA cohort. Variants within the RLS-associated 32 kb LD block with a $MAF \geq 0.05$, identified in the “SNP discovery panel” of 188 unrelated RLS patients, were included in statistical analysis. P-values were calculated using logistic regression (P_{nom}) with gender as a covariate and corrected for multiple testing by permutation with $n = 10.000$ permutations (P_{corr}). Positions refer to genome build NCBI36/hg18.

Table S3: List of the 51 proteins identified by affinity chromatography/mass spectrometry as differentially bound by rs12469063.

The acquired spectra were loaded into the Progenesis LC-MS software (version 4.0, Nonlinear) for label free quantification based on peak intensities. Spectra were searched against the Ensembl mouse database (Release 62; 54576 sequences). Normalized abundances of all unique peptides were summed up and allocated to the respective protein. Fold changes were calculated separately for every independent experiments based on normalized intensities, comparing the protective (NR) with the risk (R) condition. Mean fold change was calculated and proteins with fold changes ≥ 1.5 were considered differentially binding.

^a number of peptides assigned to the respective protein, ^b the number of unique peptides used for quantification of the assigned protein, ^c Mascot score: number reflecting the combined mascot scores of all observed spectra matching to amino acid sequences within the respective protein. A higher score indicates a more confident match, ^d DNA binding: “yes” indicates, if the identified protein binds to the DNA in a sequence-specific manner based on GeneRanker software (Genomatix software suite) (Berriz et al. 2003), which uses annotation data from various sources, like Gene Ontology or Genomatix proprietary annotation, ^e Consistent allele-specific binding in three biological replicates: “yes” indicates, if the protein was identified consistently to bind preferentially the protective or the risk allele, in all three independent biological experiments, ^{f-h} Fold change (R/NR) 1st/2nd/3rd: the ratio R/NR of normalized abundances in the 1st/2nd/3rd biological replicate, respectively, ⁱ the mean of the ratios R/NR of normalized abundances for all three biological replicates, ^{j-k} the respective allele (NR or R) with the highest mean normalized abundance, ^l detailed protein description including abbreviation according to Mouse genome Informatics (<http://www.informatics.jax.org/>), ^{m-o} raw normalized abundances for the risk allele in all three independent experiments, ^{p-r} raw normalized abundances for the protective (non-risk) allele in all three independent experiments

[Please refer to separate excel file in the online Supplemental Material.]

Table S4: Heat map from unpaired two class analysis (SAM) of genes regulated between *Meis1*^{tm1Mtor/wt} and wildtype forebrain tissue at E12.5 expanded to the level of individual genes.

Fold changes were calculated as ratios of signal intensities of individual mutant embryos and the arithmetic mean signal of the respective wildtype control group. Fold changes are given as log2 values and are color encoded according to the scale bar at the bottom of the heat map.

Blue is down- and yellow is up-regulated in mutant embryos. The corresponding Illumina IDs, official gene symbols and gene names are given to the right of the heat map. Genes (rows in heat map) are clustered according to similarities in expression profiles between individual mutant embryos (columns in heat map). Letters on the right of the figure refer to groups of genes with similar expression patterns and are identical to those in Fig. 5.

[Please refer to separate excel file in the online Supplemental Material.]

Table S5: Results of calorimetry trial: oxygen consumption and mean distance travelled.

	female		male		linear model	linear model	linear model	linear model
	control	mutant	control	mutant	sex	genotype	body mass	sex:genotype
	n = 11	n = 9	n = 11	n = 9				
	mean ± sd	mean ± sd	mean ± sd	mean ± sd	P-value	P-value	P-value	P-value
avg. mass	22.6±1.5	20.9±1.7	29.8±3.1	27.7±2.5	< 0.001	0.002	NA	0.719
food intake [g]	1.9±0.6	2.2±0.5	2.2±0.9	2.5±1	0.75	< 0.001	< 0.001	0.215
avg. VO2 [ml/(h animal)]	82.24±5.41	82.08±7.48	92.98±9.25	94.83±8.71	0.003	< 0.001	< 0.001	0.406
min. VO2 [ml/(h animal)]	57.8±5.17	57.13±6.64	66.53±10.95	66.8±6.84	< 0.001	0.001	< 0.001	0.534
max. VO2 [ml/(h animal)]	111.27±7.32	116.4±11.39	128.8±13.59	129.47±15.11	0.023	< 0.001	< 0.001	0.534
avg. RER ^a [VCO2/VO2]	0.85±0.003	0.85±0.03	0.850±0.04	0.86±0.03	0.73	0.426	NA	0.975
avg. distance [cm]	1101±271	1192±296	815±154	945±123	< 0.001	0.06	NA	0.734
avg. rearing [counts]	143±52	148±53	98±30	114±20	0.001	0.319	NA	0.476

^a RER = respiratory exchange ratio

Table S6: PreSTIGE-analysis focusing on the RLS-associated LD block suggests solely *MEIS1* as regulated gene.

Genome	Chr	Start (position)	Midpoint (position)	End (position)	HCNR	Predicted enhancer	Predicted Genes (Cell lines) ^a
hg18	2	66601963	66603150	66604337	602	chr2:66602948- 66605355	<i>MEIS1</i> (NHLF)
hg18	2	66606673	66607387	66608102	606	chr2:66606229- 66607067	<i>MEIS1</i> (NPC)
hg18	2	66612005	66612805	66613605	612	NA	NA
hg18	2	66617411	66617882	66618354	617	chr2:66617222- 66618938	<i>MEIS1</i> (NHLF, NPC)
hg18	2	66622163	66623916	66625669	622	NA	NA
hg18	2	66628857	66629739	66630622	629	NA	NA
hg18	2	66631456	66631695	66631935	631	NA	NA

Only *MEIS1* is predicted to be regulated by the RLS-associated 32 kB LD block (indicated in **BOLD**).

^a NHLF, NPC = cell lines in which *MEIS1* was predicted to be regulated by the RLS-associated LD block

Table S7: Oligonucleotide sequences used for "high resolution melting curve analysis" amplicons

Amplicons are listed for highly conserved non-coding regions (HCNR) and the remaining areas (non-highly conserved non-coding regions (NHCNR) separately.

[Please refer to separate excel file in the online Supplemental Material.]

Table S8: Oligonucleotide sequences used in the “Sequenom iPLEX Gold” genotyping

For each genotyped SNP oligonucleotide sequences of the respective forward-, reverse- and extension-primer are given.

[Please refer to separate excel file in the online Supplemental Material.]

Supplemental Experimental Methods

Selection of variants for follow-up genotyping

42 variants were genotyped directly. For 16 variants, tagging SNPs with r^2 -values of 0.85 and above were contained in the 42 genotyped variants. One variant (rs55855986) could not be converted to a genotyping assay and had no proxy SNPs in the set of genotyped SNPs. Quality control criteria for inclusion into statistical analysis were a call rate $\geq 90\%$ for individuals, and call rate $\geq 95\%$ and P-value for Hardy-Weinberg-Equilibrium > 0.001 for SNPs.

HCNR selection criteria and generation of HCNR reporter vectors

For determination of HCNRs we used the VISTA Genome Browser v2.0 (<http://pipeline.lbl.gov/cgi-bin/gateway2>) (Frazer et al.2004). First, we looked for $\geq 70\%$ sequence identity over 100 bp between human (*March 2006; hg18*) and the respective species: mouse (*Jul. 2007; mm9*), chicken (*May 2006; galGal3*), frog (*Xenopus tropicalis v4.1*), and pufferfish (*Takifugu rubripes v4.0*). In case a region met this criterion in at least two species a HCNR was appointed. Distinct peaks were defined as one HCNR block, when the conservation in mouse was continuous and the corresponding peaks in the chicken conservation were closer than 700 bp. Thus, eight HCNRs were determined, which were named according to their genomic location (for example: HCNR 602 refers to the block beginning at genomic position ~66,602,000 on chromosome 2 etc.).

HCNR were cloned using Phusion polymerase (Finnzymes, Vantaa, Finland) and Qiagen Taq Polymerase to adapt the blunt-ended PCR product for the TOPO-TA-Cloning (pCR8GW/TOPO. Based on the allele-specific pCR8GW/TOPO vectors, all HCNRs were transferred via the Gateway-technology (Invitrogen) into the respective reporter vector and Sanger sequenced. For zebrafish, we used the zebrafish reporter detection vector (ZED; EGFP-gata2-promoter reporter vector flanked by Tol2 transposase recognition sites) (Bessa et al. 2009). This vector includes an intrinsic control with red fluorescent protein (dsRed; (Baird et al. 2000; Wall et al. 2000)). For mice, we used the beta-galactosidase reporter vector “Hsp68-lacZ” (Pennacchio et al. 2006). The following primers were used (in brackets length of PCR product): 602_F: 5'-AAACCTTCTAACACAGAATTTAGCTC and 602_R: 5'-TGCCACATTTGAATGCTACTTTAC (2375 bp), 606_F: 5'-

TGTATTCCCCTGCCTTGTG and 606_R: 5'-AAAGGCATGACTCTGATGAGG (1429 bp), 612_F: 5'-TGTGAAGTCTCTGTTTAAATAGGAAGG and 612_R: 5'-ATTTGATGGCAGGATTTTGG (1601 bp), 617_F: 5'-AATGCATAAAAAGTGGGCATT and 617_R: 5'-ACGCCATTTTGAATGAGTC (944 bp), 622_F: 5'-ACTGGCAACTTCTTTTAACTGC and 622_R: 5'-TTGCATGCCTGTTTATGAGC (3507 bp), 629_F: 5'-TCCTTTATAAGTTGACAATTTTATGC and 629_R: 5'-GCTCTCCGGCAGAGACTGT (1766 bp), 631_F: 5'-CCAGGCTGGTCTCTAACTCC and 631_R: 5'-TCTCCTCTTTTGCCTTTCTCC (652bp).

Genotyping of transient transgenic mice

Genotyping was performed using extra-embryonic tissue and the transgene-specific PCR was performed using two primer pairs:

5'-GGTGGCGCTGGATGGTAA, 5'-CGCCATTTGACCACTACC
and 5'-GATGTTCTGGAGCTCGGTA, 5'-GTGGATTGAAGCCCAGCTAA.

Histological methods

Templates for in vitro transcription of riboprobes: for *Skor1* a full-length cDNA clone from Imagenes: (IMAGp998I071138Q), for *Meis1* cDNA as described previously (Mercader et al. 2009) and cDNA fragments amplified by following primers: 5'-gacgttgggggtcattgaag-3' and 5'-tgaggagagcgtgcagacag-3' for *Ptprd*; 5'-tggtgccttggtgattggtg-3' and 5'-cgggtccacaggagtagacg-3' for *Btbd9*; 5'-ccaccgagacactcaggac-3' and 5'-aggccatgtatgtccgttc-3' for *Map2K5*; 5'-caaattccagtcgcagactca-3', 5'-gggctaacagtggccataga-3' for *Tox3* and 5'-tgtagtttgacgcggtgtgt-3' and 5'-cttgagggcagaagtggaag-3' for *Creb1*.

Immunohistochemistry on paraffin sections was performed according to standard procedures using polyclonal mouse anti-MEIS1 antibody (Abnova, 1:200 dilution), monoclonal rabbit anti-CREB antibody (Ab32515, 1:200 dilution) and donkey anti-mouse/anti-rabbit IgG Alexa Fluor 488/ Fluor 594, respectively (Invitrogen, 1:1,000 dilution). The specificity was verified by omission of the primary antibody and pre-absorption of the antibody with an excess of MEIS1 peptide (Abnova, 1:100 dilution). Histological images were taken with an Axioplan2 microscope or StemiSV6 using a AxioCam MRC camera and Axiovision 4.6 software (Zeiss, Germany) and processed with Adobe Photoshop and Illustrator CS3 software (Adobe Systems Inc., USA).

Electrophoretic mobility shift assay (EMSA) and overexpression of CREB1 in 293T cells

5'-terminus Cy5-labeled oligonucleotides were annealed with the respective unlabeled complementary strand and subsequently purified using a native 12% polyacrylamide gel: 5'-GCTTCCAGCTGTGGCAGGCATGATGCAGTGAATTGCTTTT-3' (protective allele, underlined) and 5'-GCTTCCAGCTGTGGCAGGCGTGATGCAGTGAATTGCTTTT-3' (risk allele, underlined) ,

5' GATTCCAGACACAAACCCCAAGATGCAGTGAATTGCTTTT-3' (oligonucleotide with predicted CREB1 binding site mutated, underlined), CREB consensus (AGAGATTGCCTGACGTCAGAGAGCTAG) and mutant (AGAGATTGCCTGTGGTCAGAGAGCTAG) oligonucleotide (Santa Cruz, sc-2504 and sc-2517, respectively).

5 µg of nuclear protein extract were incubated with 1x EMSA binding buffer (3% glycerol, 0.75 mM MgCl₂, 0.375 mM EDTA, 0.375 mM DTT, 37.5 mM NaCl, 7.5 mM TrisHCl pH 7.5) and 90 ng/µl poly-(dIdC) for 10 min on ice, prior to the addition of the labeled oligonucleotide probe. For competition the unlabeled probe was added in increasing amounts (33x, 66x, 100x nM) together with the labeled probe. After adding the oligonucleotide probe the reaction was incubated for additional 20 min on ice and in the dark. DNA-protein complexes were resolved by electrophoresis for 4 hours on a 5.3% native polyacrylamide gel in 0.5% TBE buffer, at 4 C and in the dark. The fluorescence of the Cy5-labeled oligonucleotides was detected with the Typhoon Trio+ imager (GE Healthcare).

For the supershift assays, 1 µg of rabbit anti-CREB (Abcam, ab32515) or 1 µg control antibody (rabbit IgG, same isotype, Santa Cruz) was added to the reaction before adding the oligonucleotide probe and incubated for 30 min on ice.

293T cells were transiently transfected with the CMV-*Creb1* overexpression vector (pcDNA3.1⁺, cDNA of *Creb1* cloned in NheI and NotI sites), using standard methods (Lipofectamine 2000 Invitrogen). Nuclear extract was obtained using the NE-PER Extraction Reagents (Thermo Fisher Scientific).

RNA isolation and transcriptome analysis

Each embryo was genotyped for *Meis1* alleles (Azcoitia et al. 2005) and gender, and stored in liquid nitrogen. Forebrains of E12.5 mouse embryos were carefully dissected in order to isolate *Meis1* expressing areas of the MGE/LGE. Total RNA of eleven samples (males: *Meis1*^{tm1Mtor/-} n = 3, wt n = 3; females: *Meis1*^{tm1Mtor/-} n = 3, wt n = 2) was extracted according

to a standard protocol (RNeasy, QiagenTM) and RNA quality was examined using an Agilent Bioanalyzer 2100. Using the Illumina TotalPrep RNA Amplification kit (Ambion), 500 ng of high quality total RNA was amplified and hybridized to MouseRef-8v2 Expression BeadChips (Illumina, San Diego, CA, USA). After 16 hours hybridization, staining and scanning, data were normalized applying cubic spline normalization and background subtraction (GenomeStudioV2011.1, Illumina). Statistical analyses were performed using SAM (Significant Analysis of Microarrays) included in the TM4 software package (Saeed et al. 2003; Horsch et al. 2008). To estimate the false discovery rate, nonsense genes were identified by calculating 1000 random permutations of measurements. The selection of the top differentially expressed genes between heterozygous mutant and wildtype embryos with reproducible up- or down-regulation for the one class and two-class SAM analysis was based on an FDR < 10% in combination with a fold change above 1.5 fold. Hierarchical cluster analysis (HCL) (Eisen et al. 1998) was employed for identification of similar expression patterns applying the average-linkage-method as distance and the Euclidean Distance as distance-metric. Overrepresented functional annotations within the data sets were provided as GO terms of the category biological processes (Ingenuity Pathway Analysis, IPA). The complete array data set is available from the GEO database under accession number GSE44592.

Behavioral and metabolic phenotyping of *Meis1*^{tm1Mtor/-} mice

For the evaluation of energy balance, single mice (9-11 per sex and genotype) were kept in respirometry cages (Phenomaster System, TSE Systems, Germany) (Gailus-Durner et al. 2009). The indirect calorimetry measurement started at 1 pm (CET) after about two hours adaptation to the cages and continued until 10 am the next morning. The set-up allowed the analysis of gas samples of individual mice every 20 minutes resulting in 63 readings per individual and trial. Mean oxygen consumption (ml/h) and respiratory exchange ratio (VCO₂/VO₂) were calculated from gas exchange data. Minimum and maximum rates were estimated from the lowest and the highest VO₂ detected during the trial. Two infrared light beam frames allowed the monitoring of physical activity (lower frame: distance travelled per 20 minutes, upper frame: number of rearings per 20 minutes). Food hoppers were attached to electronic scales to determine food consumption. Two way ANOVA was calculated for genotype, sex and genotype:sex effects on respiratory exchange ratio, distance travelled and number of rearings. Oxygen consumption and food intake were analyzed by linear regression

models including the factors genotype, sex, genotype:sex effects, and body mass as covariate (TIBCO Spotfire S+ 8.1 for Windows).

The Open Field test was assessed according to the standardized phenotyping screens developed by the Eumorphia partners (Mandillo et al. 2008), available as EMPreSSslim protocols (see www.eumodic.org). The Open Field apparatus consisted of a transparent and infrared light permeable acrylic test arena with a smooth floor (internal measurements: 45.5 x 45.5 x 39.5 cm). Illumination levels were set at approximately 150 lux in the corners and 200 lux in the middle of the test arena. Data were recorded and analyzed using the ActiMot system (TSE, Bad Homburg, Germany).

Poll II ChIA-PET and Hi-C data and “PreSTIGE”-software analysis

For all seven enhancers of the 32 kb RLS-associated region, interactions were analyzed based on predictions from “PreSTIGE” software using thirteen human cell lines, as available at <http://genetics.case.edu/prestige/> (Corradin et al. 2014). Poll II ChIA-PET (Fullwood and Ruan 2009; Li et al. 2012) and Hi-C (Belton et al. 2012; Dixon et al. 2012) data from the human epigenomic browser of the Washington University School of Medicine, St. Louis, USA, (<http://epigenomegateway.wustl.edu/info/>) (Zhou et al. 2013) were analyzed with respect to the 32 kb RLS-associated LD block.

References

- Azcoitia V, Aracil M, Martinez AC, Torres M. 2005. The homeodomain protein Meis1 is essential for definitive hematopoiesis and vascular patterning in the mouse embryo. *Dev Biol* **280**(2): 307-320.
- Baird GS, Zacharias DA, Tsien RY. 2000. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc Natl Acad Sci U S A* **97**(22): 11984-11989.
- Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. 2012. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* **58**(3): 268-276.
- Berriz GF, King OD, Bryant B, Sander C, Roth FP. 2003. Characterizing gene sets with FuncAssociate. *Bioinformatics* **19**(18): 2502-2504.
- Bessa J, Tena JJ, de la Calle-Mustienes E, Fernandez-Minan A, Naranjo S, Fernandez A, Montoliu L, Akalin A, Lenhard B, Casares F et al. 2009. Zebrafish enhancer detection (ZED) vector: a new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish. *Dev Dyn* **38**(9): 2409-2417.
- Corradin O, Saiakhova A, Akhtar-Zaidi B, Myeroff L, Willis J, Cowper-Sal Lari R, Lupien M, Markowitz S, Scacheri PC. 2014. Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. *Genome Res* **24**(1): 1-13.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**(7398): 376-380.
- Eisen MB, Spellman PT, Brown PO, Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* **95**(25): 14863-14868.
- Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for comparative genomics. *Nucleic Acids Res* **32**: W273-279.
- Fullwood MJ, Ruan Y. 2009. ChIP-based methods for the identification of long-range chromatin interactions. *J Cell Biochem* **107**(1): 30-39.
- Gailus-Durner V, Fuchs H, Adler T, Aguilar Pimentel A, Becker L, Bolle I, Calzada-Wack J, Dalke C, Ehrhardt N, Ferwagner B et al. 2009. Systemic first-line phenotyping. *Methods Mol Biol* **530**: 463-509.
- Horsch M, Schadler S, Gailus-Durner V, Fuchs H, Meyer H, de Angelis MH, Beckers J. 2008. Systematic gene expression profiling of mouse model series reveals coexpressed genes. *Proteomics* **8**(6): 1248-1256.
- Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, Poh HM, Goh Y, Lim J, Zhang J et al. 2012. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**(1-2): 84-98.
- Mandillo S, Tucci V, Holter SM, Meziane H, Banchaabouchi MA, Kallnik M, Lad HV, Nolan PM, Ouagazzal AM, Coghill EL et al. 2008. Reliability, robustness, and reproducibility in mouse behavioral phenotyping: a cross-laboratory study. *Physiol Genomics* **34**(3): 243-255.
- Mercader N, Selleri L, Criado LM, Pallares P, Parras C, Cleary ML, Torres M. 2009. Ectopic Meis1 expression in the mouse limb bud alters P-D patterning in a Pbx1-independent manner. *Int J Dev Biol* **53**(8-10): 1483-1494.
- Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M et al. 2003. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34**(2): 374-378.

- Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD et al. 2006. In vivo enhancer analysis of human conserved non-coding sequences. *Nature* **444**(7118): 499-502.
- Wall MA, Socolich M, Ranganathan R. 2000. The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. *Nat Struct Mol Biol* **7**(12): 1133-1138.
- Zhou X, Lowdon RF, Li D, Lawson HA, Madden PA, Costello JF, Wang T. 2013. Exploring long-range genome interactions using the WashU Epigenome Browser. *Nat Methods* **10**(5): 375-376.