Targeted Resequencing and Systematic In Vivo Functional Testing Identifies Rare Variants in MEIS1 as Significant Contributors to Restless Legs Syndrome

Eva C. Schulte,1,2,22 Maria Kousi,3,22 Perciliz L. Tan,3 Erik Tilch,2,4 Franziska Knauf,2 Peter Lichtner,2,4 Claudia Trenkwalder,5,6 Birgit Högl,7 Birgit Frauscher,7 Klaus Berger,8 Ingo Fietze,9 Magdalona Hornyak,1,10,11 Wolfgang H. Oertel,12 Cornelius G. Bachmann,13,14 Alexander Zimprich,15 Annette Peters,16 Christian Gieger,17 Thomas Meitinger,2,4,18 Bertram Müller-Myhsok,18,19,20 Nicholas Katsanis,3,23 and Juliane Winkelmann1,2,4,18,21,23,*

Restless legs syndrome (RLS) is a common neurologic condition characterized by nocturnal dysesthesias and an urge to move, affecting the legs. RLS is a complex trait, for which genome-wide association studies (GWASs) have identified common susceptibility alleles of modest (OR 1.2–1.7) risk at six genomic loci. Among these, variants in MEIS1 have emerged as the largest risk factors for RLS, suggesting that perturbations in this transcription factor might be causally related to RLS susceptibility. To establish this causality, direction of effect, and total genetic burden of MEIS1, we interrogated 188 case subjects and 182 control subjects for rare alleles not captured by previous GWASs, followed by genotyping of ~3,000 case subjects and 3,000 control subjects, and concluded with systematic functional annotation of all discovered variants using a previously established in vivo model of neurogenesis. We observed a significant excess of rare MEIS1 variants in individuals with RLS. Subsequent assessment of all nonsynonymous variants by in vivo complementation revealed an excess of loss-of-function alleles in individuals with RLS. Strikingly, these alleles compromised the function of the canonical MEIS1 splice isoform but were irrelevant to an isoform known to utilize an alternative 3′ sequence. Our data link MEIS1 loss of function to the etiology of RLS, highlight how combined sequencing and systematic functional annotation of rare variation at GWAS loci can detect risk burden, and offer a plausible explanation for the specificity of phenotypic expressivity of loss-of-function alleles at a locus broadly necessary for neurogenesis and neurodevelopment.

Introduction

Restless legs syndrome (RLS [MIM 102300]) is a common neurologic condition with an age-dependent prevalence of up to 10% in Europe and North America.1 It is characterized by an irresistible urge to move the legs accompanied by disagreeable, often painful, sensations in the lower limbs at night. Moving the affected legs or walking leads to prompt but only temporary relief.1 As a consequence, individuals suffer from persistent insomnia, leading to an impairment of quality of life and mental health. RLS is a highly familial trait but genetically complex, with estimates of narrow-sense heredity between 54% and 69% as derived from twin studies.2,3

Genome-wide association studies (GWASs) in large RLS case/control samples have identified common susceptibility alleles at six loci that together explain about 7% of the heritability.4 Among many models that can explain some of the “missing heritability,”5 we considered the possibility that a collection of rare variants of strong effect, which cannot be identified by means of GWASs,6,7 might be a contributory factor. Although a single potentially causal rare variant has been described in MEIS1 (MIM 601739),8,9 to date no variants of strong effect have been established. Nonetheless, for some other complex genetic diseases, such as diabetes or chronic inflammatory bowel disease, rare variants have recently been identified within known GWAS loci,10,11 supporting the concept of allelic series in complex genetic disorders. We therefore sought to assess the potential contribution of rare variants to disease burden both by using standard statistical methods and by assessing the incidence and contribution of...
functionally annotated variants relevant to MEIS1 biology (Figure S1 available online). For this purpose, we exploited two major resources: a well-phenotyped, ethnically homogeneous RLS cohort and an experimentally tractable method to assay MEIS1 functionality grounded on previously defined in vivo observations on the roles of this protein in neurogenesis, wherein suppression of meis1 in zebrafish embryos led to a quantitative reduction of the optic tectum, a major site of neurogenesis in the developing brain, and malformation of rhombomeres 3 and 5, which represent early hindbrain structures shown previously to be defective in the absence of meis1.12,13 Our data showed a significant enrichment of rare variants across both MEIS1 and also all seven RLS-associated genes. Based upon population statistics alone, only one single low-frequency variant in the 3’ UTR (rs11693221) showed significant association with RLS. However, the combinatorial usage of functional annotation and statistical analyses highlighted a major contribution of loss-of-function variants in MEIS1 and suggested that rare alleles in this locus pose significant RLS risk to individuals.

Materials and Methods

Participants

Both case and control populations were entirely of German and Austrian descent. In the case subjects, diagnosis was based on the diagnostic criteria of the International RLS Study Group1 as assessed in a personal interview conducted by an RLS expert. In keeping with the previous GWASs,4,14 we excluded individuals with secondary RLS due to uremia, dialysis, or anemia resulting from iron deficiency. The presence of secondary RLS was determined by clinical interview, physical and neurological examination, blood chemistry, and nerve conduction studies whenever deemed clinically necessary. Participants’ written informed consent was obtained prior to the initiation of the study. The institutional review boards of the contributing authors approved the study. The primary review board was located in Munich at the Bayerische Ärztekammer and Technische Universität München.

Genotyping by High-Resolution Melting Curve Analysis

In a first step, we used Idaho LightScanner high-resolution melting curve analysis (Biofire) to screen the coding regions and exon/intron boundaries of PTPRD (MIM 601598), BC034767, TOX3 (MIM 611416), BTBD9 (MIM 611237), MEIS1, and MAP2K5 (MIM 6002520) for variants. Due to the high GC content, the coding regions ± 10 bp of SKOR1 (MIM 611273) could not be subjected to LightScanner analysis and were Sanger sequenced instead. Included in the screening were 188 German RLS-affected case subjects and 182 general population control subjects belonging to the KORA cohort13 based in Southern Germany. Where possible, the 188 case subjects used were half homozygous and half heterozygous for the published risk alleles.5,14-16 The same set of control subjects was used for all screening experiments. MEIS1, TOX3, and BC034767 variants identified in the 188 case subjects have already been published.4,8 In the case of an altered melting pattern suggestive of variants, Sanger sequencing ensued to identify the underlying variant. The same method was used to screen the coding regions of MEIS1 isoforms 1 and 2 ± 10 bp as well as the 5’ and 3’ UTRs in 3,760 RLS case subjects of German and Austrian descent (62.2 ± 12.8 years; 30.8% male) and 3,542 general-population control subjects (55.1 ± 13.8 years; 40.1% male) belonging to the S4 and F4 surveys of KORA.17 For the discovery sample, group comparisons between case and control subjects were performed in R 18 for each gene and each type of variant separately, with and without the common risk allele genotype as covariate, using logistic regression (logreg) of the phenotype on aggregate minor allele counts of variants of each type.19 To account for a possible bias introduced by the comparison of different sets of risk-genotype-selected cases to a constant set of unselected controls, we performed a meta-analysis using meta to evaluate the contribution of rare coding variants across all seven genes.

Empirical p values were calculated with 1,000 permutations of the phenotype and assessing the ratio of p values equal to or smaller than the p value belonging to the original phenotype. For the large-scale screening of MEIS1, both logreg of the phenotype on number of variants as well as sequence kernel association tests (SKAT)20 with and without Madsen-Browning weights21 were performed. Empirical p values are based on 10,000 permutations of the phenotype, calculating the ratio of test statistics equal to or larger than the test statistic of the original phenotype.

Genotyping by Mass Spectrometry

Genotyping was carried out on the MassARRAY system by MALDI-TOF mass spectrometry with iPLEX Gold chemistry (Sequenom). Primers were designed with AssayDesign v.3.1.2.2 with iPLEX Gold default parameters. No assay could be designed for seven variants, largely those located in the extremely GC-rich gene SKOR1. Further, three assays failed two or more times and were, therefore, not pursued further. Automated genotype calling was carried out with SpectroTYPER v.3.4. Genotype clustering was visually checked by an experienced evaluator. SNPs with a call rate <90% were excluded. The genotyping sample consisted of 3,262 case subjects of German and Austrian descent (65.3 ± 11.3 years; 29.3% male) and 2,944 general population control subjects (56.1 ± 13.3 years; 48.7% male) from the KORA F4 survey.17 For the most part, case and control subjects used in both genotyping approaches were drawn from the same samples. Both logreg of the phenotype on number of variants as well as SKATs20 with and without Madsen-Browning weights21 were performed. Empirical p values are based on 200 permutations of the phenotype, calculating the ratio of test statistics larger than the test statistic of the original phenotype.

In Vivo Complementation in Zebrafish Embryos, In Situ Hybridization, and Whole-Mount Immunostaining

Splice-blocking morpholinos (MOs) against meis1 and map2k5 were designed and obtained from Gene Tools. We injected 1 nl of diluted MO (4 ng for meis1_MO1, 3 ng for meis1_MO2, and 5 ng for map2k5) and/or RNA (75 pg for meis1, 100 pg for map2k5) into wild-type zebrafish embryos at the 1- to 2-cell stage. To evaluate hindbrain organization, injected embryos were raised until the 10- to 13-somite stage, corresponding to 14 to 16 hr post-fertilization (hpf); they were then dechorionated and fixed in 4% paraformaldehyde (PFA) overnight. Fixed embryos were transferred to 100% methanol at –20°C for at least 2 hr and were then processed after standard protocols22 using a digoxigenin-labeled antisense riboprobe against knox20. For analysis of the
optic tectum, injected embryos were fixed overnight at 72 hpf in 4% PFA and stored in 100% methanol at −20°C. For acetylated tubulin staining, embryos were fixed in Dent’s fixative (80% methanol, 20% DMSO) overnight at 4°C. The embryos were permeabilized with proteinase K followed by postfixation with 4% PFA. PFA-fixed embryos were washed first in PBS and subsequently in IF buffer (0.1% Tween-20, 1% BSA in PBS) for 10 min at room temperature. The embryos were incubated in blocking buffer (10% FBS, 1% BSA in PBS) for 1 h at room temperature. After two washes in IF buffer for 10 min each, embryos were incubated in the primary antibody (anti-acetylated tubulin [T7451, mouse, Sigma-Aldrich], 1:1,000) in blocking solution, overnight at 4°C. After two additional washes in IF buffer for 10 min each, embryos were incubated in the secondary antibody solution (Alexa Fluor goat anti-mouse IgG [A21207, Invitrogen], 1:1,000) in blocking solution, for 1 h at room temperature.

For RNA rescue and overexpression experiments, the wild-type mRNAs of isoforms 1 (RefSeq accession number NM_002398.2/ENST00000272369) and 2 (no RefSeq ID/ENST00000381518) of MEIS1 as well as the canonical isoform of MAP2KS (RefSeq NM_145160.2/ENST00000178640) were cloned into the pCS2+ vector and transcribed in vitro using the SP6 Message Machine kit (Ambion). All variants identified in isoform 1 as well as all additional variants only coding in isoform 2 plus two functional null variants from isoform 1 (p.Ser204Thr [c.610T>G] and p.Arg272His [c.815G>A]; both RefSeq NM_002398.2/ENST00000272369) were introduced with Phusion high-fidelity DNA polymerase (New England Biolabs) and custom-designed primers. Additionally, a non-occurring occurring homeobox domain-dead variant of MEIS1 (p.Arg276Ala + p.Asn325Ala [c.826T>TGC; 0.973_974delinsGC]; RefSeq NM_002398.2/ENST00000272369) was created. All the experiments were repeated in triplicate and significance of the morphant phenotype was judged with Student’s t test.

Results

Variant Screening of Seven RLS GWAS Candidate Genes

To assess low-frequency (1% < minor allele frequency [MAF] < 5%) coding variation at the known RLS susceptibility loci encompassing PTPRD, TOX3, BTB9, MEIS1, MAP2KS, SKORI, and the noncoding RNA BC034767,4,14–16 we screened the coding regions and exon-intron boundaries (±10 bp) in 188 German individuals with RLS and 182 control subjects belonging to the KORA population cohort13 based in Southern Germany. A total of 49 variants with MAF < 5% were identified in case and control subjects together (Table S1). When collapsed across all seven genes, rare and low-frequency nonsynonymous variants showed a trend toward being more frequent in case than in control subjects (39 in case subjects versus 24 in control subjects; p = 0.103; logreg meta-analysis; odds ratio [OR] = 1.40). Nonsynonymous and synonymous variants combined, however, showed a stronger enrichment in individuals with RLS (77 case subjects versus 46 control subjects; p = 0.023, logreg meta-analysis; OR = 1.51). Addition of the common risk allele genotype4,14–16 into the analysis as a covariate decreased the permitted p value (nonsynonymous only: p = 0.079; all variants: p = 0.008; n = 1,000 permutations), suggesting that the enrichment of rare variants across all loci is independent of the common risk allele genotype. However, the degree of interdependence between common and rare variants differed between genes (Table S2).

Within MEIS1, synonymous or nonsynonymous variants with a MAF below 5% were present in nine case subjects but only one control subject (p = 0.021) although no marked difference in the amount of variation was obvious in any of the other genes (Table S1). Here, the p value increased modestly after addition of the common risk variant (rs2300478) as covariate (p = 0.080: 1,000 permutations), indicating some interdependence of rare and common alleles at the MEIS1 locus. Variants did not seem to cluster within specific regions of the examined genes (Figure S2).

Genotyping of Identified Low-Frequency and Rare Variants

To assess a possible association with the RLS phenotype, we next genotyped 39 of the 49 identified variants in 3,262 German and Austrian RLS case subjects (65.3 ± 11.3 years; 29.3% male) and 2,944 KORA control subjects (56.1 ± 13.3 years; 48.7% male). Of the 49 variants, 10 could not be included for technical reasons (see Material and Methods section above). Although variants (either altogether or only nonsynonymous ones) with MAF < 5% and MAF < 1% were not significantly enriched in RLS, we observed a distinct excess of very rare variants with a KORA-derived MAF < 0.1% (total: 57 case versus 16 control subjects, p = 4.99 × 10−4; OR = 2.50; nonsynonymous only: 23 case versus 5 control subjects, p = 0.0019; OR = 3.92; logreg; Figure 1).

We then went on to assess whether more sophisticated tests used to analyze rare variant associations of bidirectional effects (SKAT)20 and differing allele frequency (SKAT with Madsen-Browning weights)21 would change the association signal. However, although frequency-weighted results were still superior to unweighted results in most cases, overall SKAT analysis led to an increase in association p values both for the joined analysis across all loci as well as for MEIS1 alone (Table S2).

We next asked whether this signal is distributed across all seven tested genes, or whether a specific subset of tested loci cause the apparent enrichment of rare nonsynonymous variation in individuals with RLS. A significant association with RLS was observed for MEIS1; the aggregate risk conferred by the variants showed a large increase as we transitioned from rare (MAF < 1%) to very rare (MAF < 0.1%) (Figure 1). More specifically, rare variants with MAF < 1% (total: 116 case subjects versus 67 control subjects, p = 0.0064, OR = 1.51; nonsynonymous only: 39 versus 14, p = 0.0024, OR = 2.46) and MAF < 0.1% (both total and nonsynonymous only: 9 versus 1: p = 0.014; OR = 8.14; Figure 1) were seen more frequently in cases than in controls. SKAT (all: p = 0.049;
nonsynonymous only: p = 0.009) and Madsen-Browning-weighted SKAT (all: p = 0.019; nonsynonymous only: p = 0.029; 10,000 permutations each) substantiated this finding (Table S2). No low-frequency coding variants with MAF between 1% and 5% were found within the coding regions of MEIS1.

After exclusion of MEIS1, logreg (all variants: p = 0.009; nonsynonymous only: p = 0.024; 2,000 permutations) for variants with MAF < 0.1% and SKAT using Madsen-Browning weights (all variants: p = 0.019; nonsynonymous only: p = 0.019; 100 permutations) showed a nominally significant enrichment across all other six genes. In PTPRD alone, rare variants of all classes were also encountered more frequently in individuals with RLS (all variants with MAF < 0.1%: p = 9.99 × 10^{-4}; logreg; all variants: p = 0.029, SKAT with MB) whereas this enrichment was not observed for nonsynonymous variants only. Of note, several individual rare variants in MEIS1 (p.Arg272His [c.815G>C] and p.Met453Thr [c.1359T>C] [ENST00000381518]), TOX3 (p.Ala233Ala [c.699T>C]; RefSeq NM_001080430.2/ENST00000219746), and PTPRD (p.453Thr [c.1359T>C] and p.Pro278Pro [c.834T>G]; both RefSeq NM_002839.3/ENST00000381196) were associated nominally with RLS in the large case/control sample; however, associations did not withstand correction for multiple testing (Table S1).

Assessment of Rare Variation in MEIS1
The excess of low-frequency and rare variants at RLS-associated GWAS loci was most pronounced for MEIS1. Therefore, we sought to expand our analysis to a more comprehensive investigation of genetic variation with regard to frequency and location within MEIS1 by screening the coding regions ±10 bp as well as the 5’ and 3’ UTRs for variants with MAF < 5% in 3,760 German and Austrian RLS case subjects (62.2 ± 12.8 years; 30.8% male) and 3,542 KORA control subjects (55.1 ± 13.8 years; 40.1% male). We identified a total of 75 such variants (Tables S3 and S4); 28 of these lie in either the canonical isoform 1 (ENST000000272369) or a longer isoform 2 (ENST00000398506) of MEIS1 that encodes an alternate start site and an alternate C terminus. All synonymous and nonsynonymous variants identified in the initial screening of 188 case subjects and 182 control subjects were observed again.

Overall, we observed an excess of variants with MAF < 5% across all examined regions of MEIS1 (1,383 variant counts in case subjects versus 606 in control subjects, p = 1.04 × 10^{-11}; not permuted), which was driven primarily by a low-frequency variant, rs11693221, in the 5’ UTR (MAF cases = 13.55%; MAF controls = 3.58%; p = 1.27 × 10^{-11}; OR = 4.42 [95% CI: 3.83–5.11]); not permuted; Figure 2, Tables S3 and S4). After exclusion of this variant, the remainder of individuals with low-frequency and rare variants across all regions of MEIS1 was similar in case and control subjects (432 versus 396; p = 0.68). However, stratification of variants according to their localization showed an excess of rare variants with MAF < 5% in the 5’ UTR (16 case subjects versus 2 control subjects; logreg: p = 0.001, OR = 7.56; SKAT: p = 0.01; SKAT with MB: p = 0.006) and among nonsynonymous variants in isoform 2 (34 case subjects versus 15 control subjects; logreg: p = 0.007, OR = 2.31; SKAT: p = 0.004; SKAT with MB: p = 0.0005) (Figure 2, Table S3).

Functional Analysis of Rare Nonsynonymous Variants in MEIS1 by In Vivo Complementation in Zebrafish Embryos
Resequencing of any locus is certain to reveal rare variants in both case and control subjects which, benefit of a means of preselecting for variants relevant to protein function, can dampen or extinguish bona fide association signals. We, therefore, considered a paradigm grounded on prior knowledge of MEIS1 to test whether each of the rare discovered variants in our study have an effect on function...
and to then use this information to assess the burden of deleterious genetic lesions in our case/control RLS study. During zebrafish development, suppression of meis1 has been shown to impact neurogenesis, a phenotype captured prominently by the quantitative reduction of the size of the optic tectum 72 hpf\(^{12}\) as well as the disruption of hindbrain patterning at 14 hpf.\(^{13}\) We first tested the ability of human MEIS1 mRNA to rescue the optic tectum size phenotype and to thus establish a baseline assay for the evaluation of the identified nonsynonymous variants. To this end, we designed two independent MOs that target different exon-intron splice junctions of the endogenous zebrafish meis1. Both MOs gave rise to the same phenotype, bolstering our confidence in the specificity of the assay: we observed a reduction of the size of the optic tectum by \(\sim 30\%\) when injecting 4 ng of meis1_MO1 targeting exon-intron boundary 2, and by \(\sim 20\%\) when injecting 3 ng of a previously characterized MO (meis1_MO2) against the acceptor site of exon \(2^{\text{c}}\) (Figures S3 and S4). In both cases, the phenotype was rescued significantly and reproducibly (\(p < 0.0001\); performed in triplicate, scored blind to injection cocktail) by coinjection with 75 pg of human capped MEIS1 mRNA (Figures S3 and S4). The optic tectum phenotype could be rescued by either the canonical MEIS1 isoform or by an isoform utilizing an alternative 3’ terminus. By contrast, injection of the “domain-dead” human mRNA (MEIS1_DD) bearing two variants engineered to ablate DNA binding ability (p.Arg276Ala+Asn325Ala) was indistinguishable from MO alone (\(p < 0.68\)). As a test of the relevance of the phenotype to RLS, MO suppression of map2k5, another GWAS-associated RLS gene, yielded a similar phenotype with regard to the size of the optic tectum, which was also specific as it could be rescued by coinjection with human MAP2K5 mRNA (Figure S5).

Given these data, we proceeded to perform in vivo complementation assays on all 13 nonsynonymous coding variants identified in isoform 1, as well as in the 4 nonsynonymous variants that lie in the unique sequence of isoform 2 (Figure 3, Table S4), wherein human mRNA bearing one test variant was coinjected with MO and compared to the rescue ability of WT mRNA (\(n = \) at least 51 embryos tested per injection, Figure 3 and Tables S5 and S6). We classified variants as benign (rescue indistinguishable from WT), hypomorphic (mutant rescue significantly worse than WT but better than MO alone), or functionally null (indistinguishable from MO alone). Among the 13 variants of isoform 1, we identified three benign, four hypomorphic, and six null variants (Figures 3 and 4). Overexpression of MEIS1 WT mRNA or mRNAs harboring each of the 13 variants had no effect on optic tectum size (Figure S6).

Layering these data over the incidence and distribution of these variants in our case/control data set, we found a significant excess of functionally null variants in individuals with RLS compared to control subjects (14 in case subjects versus 2 in control subjects; \(p = 0.0012\); OR = 7.48 [95% CI: 1.68–33.40]) (Figures 3 and 5) whereas hypomorphic (2 in case subjects versus 4 in control subjects; \(p > 0.05\)) and benign (2 in cases subject versus 4 in control subjects; \(p > 0.05\)) variants showed similar distributions in case and control subjects.

To corroborate these data, we designed a second assay at an earlier developmental time point, grounded on the known requirement of meis1 for the organization of the hindbrain; in triplicate experiments, \(\sim 30\%\) of the embryos suppressed for meis1 developed hindbrain defects that reproduced previously reported meis1 phenotypes.\(^{13}\) These consisted of significant widening of rhombomeres 3 and/or 5 (r3 and r5); shortening of the distance between r3 and r5; thinning of either one of those structures; or absence of r3 and/or r5 altogether (Figure 4). This phenotype was rescued by coinjection with 75 pg of WT human MEIS1 mRNA (\(p = 0.045\)). Assessment of six variants from our series (three variants scored as null and three variants scored as benign in the optic tectum assay) and blind triplicate scoring confirmed this result: p.Ser204Thr, p.His239Tyr (c.715C>T), and p.Arg272His were verified as null variants and p.Ala122Val (c.365C>T), p.Ser175Tyr (c.524G>A), and p.Met366Leu (c.1096A>G) were validated as benign (all RefSeq NM_002398.2/ENST00000272369).

By contrast, all four nonsynonymous variants exclusive to isoform 2 scored benign in our assay. Strikingly, two
Figure 3. Functional Assessment of Rare Nonsynonymous Variants in MEIS1 by In Vivo Complementation in Zebrafish Embryos

(A) Location and frequency of nonsynonymous MEIS1 variants examined in zebrafish. Variants found in case subjects are given above the gene, those found in control subjects below. The short, canonical isoform 1 of MEIS1 is given in dark gray (ENST00000272369); the additional amino acids in the longer isoform 2 in light gray (ENST00000398506).

(legend continued on next page)
variants found to be functional nulls in the isoform 1 complementation assay (p.Ser204Thr and p.Arg272His) were able to fully rescue the tectum size phenotype in isoform 2 (Figures 5 and S6), suggesting that the contribution of rare variants to RLS is mediated specifically by reduced activity of MEIS1 isoform 1-encoded protein.

**Discussion**

Previous GWASs have established the genomic locus encompassing MEIS1 as the most significant susceptibility region for RLS. The most likely candidate gene in this region is MEIS1, a TALE homeobox transcription factor known to be involved in specifying spinal motor neuron pool identity and connectivity as well as proximo-distal limb patterning and expressed in forebrain neurons and astrocytes during embryonic development. A common RLS-linked intronic variant in MEIS1 was also shown to induce differential forebrain enhancer activity during development. Additional studies in the context of RLS have suggested a link between MEIS1 and iron metabolism in the central nervous system.

The excess of rare alleles of functional effect in RLS case subjects compared to control subjects shown here substantiates MEIS1 as the causal genetic factor underlying the observed associations. Moreover, it implicates loss of function as the underlying mechanism, at least with regard to rare variants. We also observed a new association with a low-frequency variant (rs11693221) located in the 3’ UTR of the ENSEMBL-derived canonical isoform 1 of MEIS1 (ENST00000272369) that represents the largest single-allele genetic risk factor for RLS identified to date. In the 1000 Genomes data, the linkage disequilibrium (LD, r²) between rs11693221 and the most significantly associated common variant within MEIS1 (rs12469063) is low (r² = 0.080). In the same data set, rs11693221 is in high LD (r² > 0.8) solely with a single low-frequency variant (rs113851554, r² = 0.83). rs113851554 is located within a highly conserved noncoding region in intron 8 and has previously also been shown to be associated with RLS in a Canadian case/control sample. Follow-up analyses are needed to fully dissect all functional effects underlying the MEIS1 association signal.

Nonetheless, the finding of an associated low-frequency variant (rs11693221) in the 3’ UTR of isoform 1 in conjunction with the observed excess of rare variants in the 5’ UTR of MEIS1 in individuals with RLS implicates the UTRs in disease pathogenesis, potentially through regulating expression and/or mRNA stability. The excess of rare noncoding variants in the 100 bp surrounding the exons of nine genes associated with asthma and the fact that fine-mapping studies located about 22% of 36 GWAS association signals for celiac disease to either the 5’ or the 3’ noncoding regions (UTRs and several kilobases up- or downstream) could indicate that these regions are indeed important in the context of complex genetic diseases and might be overlooked by the current surge of whole-exome sequencing studies.

The functional experiments conducted in zebrafish allowed us to differentiate between potentially benign and pathological sequence variation and thereby increased both effect size and significance levels observed in burden testing. Moreover, the identified rare nonsynonymous MEIS1 variants, which showed an effect on optic tectum size in zebrafish embryos, were restricted to isoform 1 of MEIS1 (ENST00000272369). Previous studies have demonstrated that rare variants can exhibit isoform-specific effects on a given phenotype, such as in the case of DNAJB6 in limb-girdle muscular dystrophy. Because the pathophysiology of RLS is just beginning to be elucidated, it will be of importance to see when and where this isoform of MEIS1 is expressed as the temporal and spatial expression patterns of the different MEIS1 transcripts are currently unknown. Given our observations, we speculate that understanding the differential biological roles of isoform 1 will help dissect the subset of MEIS1 functions relevant to the etiopathology of RLS.

Recent studies have implicated allelic series of variants of different frequency and effect sizes at loci identified in the context of GWASs in complex genetic diseases. In single cases, individual rare variants were shown to be associated with the phenotype whereas in other cases it was the collective of rare variation either within a single gene or across a number of GWAS-identified loci. Our data substantiate this role of the whole of genetic variation, from common to low-frequency to rare variation, at a GWAS-identified locus in the genetic architecture of a complex genetic disease. Interestingly, the addition of synonymous variants to the burden analyses yielded a more significant enrichment of rare variants in many situations. Whether this is due to increased power, fine-scale population substructure, artificial signal amplification driven by high LD between the synonymous and the nonsynonymous or causal variants, or a true causal contribution of rare synonymous
variation as was recently reported for sporadic Alzheimer disease cannot be ascertained within the bounds of this study. We also note that we were unable to establish a single rare variant of large effect involved in RLS at the examined loci, possibly due to a lack of power in sight of the large amount of background rare genetic variation and the excess of singletons known to exist in the human genome. However, systematic functional annotation of such singletons has improved our interpretative ability and has suggested that, in addition to the risk conferred by common and low-frequency alleles, rare variants contribute significantly to the genetic burden in RLS. Our data are consistent with previous observations wherein the rarer a genetic variant, the more likely it is to harbor a functional effect and extends these to a disease context. It is particularly notable that, subsequent to functional tagging of alleles, we observed association with RLS only for functionally null variants, but not for hypomorphs, potentially intimating a threshold effect on total MEIS1 function necessary to drive pathology. Nonetheless, although our positive and negative controls for the in vivo complementation assay support previously reported high specificity and sensitivity for the approach, and despite the fact that we achieved full concordance of allele effect tagging by two independent in vivo complementation assays, it will be important to validate our observations further in an independent model system; the evaluation of the two functionally null variants MEIS1 p.Arg272His and p.Gln353His (c.1059G>C) (both RefSeq NM_002398.2/ENST00000272369) found primarily or exclusively in case subjects, located in the homeobox and transcription activation domains, respectively, in animal models could prove worthwhile to establish their relevance to the RLS phenotype and to further explore the pathophysiology of the disorder.

Figure 4. Functional Assessment of Null and Benign MEIS1 Variants by In Vivo Complementation in Zebrafish Embryos and Evaluation of Hindbrain Patterning

(A–E) At 14–16 hpf, developing zebrafish embryos were evaluated for the integrity of rhombomeres 3 and 5 (r3 and r5) by in situ hybridization with a riboprobe against krox20. Upon disruption of meis1, we observed rhombomeric defects that involved widening of the evaluated structures (B and D) or shortening of the distance between r3 and r5 (D), as well as thinning or absence of the evaluated structures.

(F) Quantification showing that the aberrant phenotypes were especially pronounced in the morphant embryos and embryos coinjected with MO+null mRNA (n ≥ 26 embryos per genotype). Abbreviations are as follows: MO, morpholino; WT, wild-type.
Supplemental Data

Supplemental Data include six figures and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2014.06.005.

Acknowledgments

We are gratefully indebted to Katja Junghans, Susanne Lindhof, Jelena Golic, Sybille Frischholz, and Regina Feldmann at the Institut für Humangenetik, Helmholtz Zentrum München (Munich, Germany) for their expert technical assistance in performing Sequenom genotyping and Light Scanner analyses. Moreover, we thank Rene Rezsohazy at the Université de Louvain (Louvain, Belgium) for his advice in the construction of the domain-dead construct of MEIS1. This study was funded by in-house institutional funding from Technische Universität München and Helmholtz Zentrum München, a grant entitled “Functional Analysis of Rare Variants in Restless Legs Syndrome” from the Else Kröner-Fresenius-Stiftung (2013_A124), by seed funding from the Center for Human Disease Modeling, Duke University, and by P50 MH094268 to N.K. Recruitment of case and control cohorts was supported by institutional (Helmholtz Zentrum München) and government funding from the German Bundesministerium für Bildung und Forschung (03.2007-02.2011 FKZ 01ET0713). W.H.O. is a Senior Research Professor of the Charitable Hertie Foundation, Frankfurt/Main, Germany. N.K. is a Distinguished Brumley Professor.

Received: December 9, 2013
Accepted: June 10, 2014
Published: July 3, 2014

Web Resources

The URLs for data presented herein are as follows:
1000 Genomes, http://browser.1000genomes.org

Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/

Accession Numbers

The dbSNP accession numbers of the variants reported in this paper are available in Table S7.

References


