Network: Systematic Gene Identification and Functional Analyses in Common CNS Disorders

Project: Identification and Functional Analysis of Genes in Human Idiopathic Epilepsy

Ortrud Steinlein – Ludwig-Maximilians University, School of Medicine, München - ortrud.steinlein@med.uni-muenchen.de

Introduction

Idiopathic epilepsies are those epilepsies in which a symptomatic background is neither detected nor suspected, but a genetic etiology is likely or proven. Most idiopathic epilepsies are caused by ion channel mutations, implicating an etiology based on imbalances in synaptic transmission or neuronal excitability [1]. However, since 2002, three genes have beeen identified which do not code for ion channels, but their mutations also cause idiopathic epilepsies. Among them, *LGI1* was the first identified gene in this group, its mutations cause a rare epilepsy syndrome called autosomal dominant lateral temporal lobe epilepsy (ADLTE) [2]. Mutations in *Mass1* were identified in a mouse model of audiogenic epilepsy[3]. Last year, *EFHC1* was found to be associated with human juvenile myoclonic epilepsy (JME), a subtype of idiopathic generalised epilepsies(IGE)[4].

The etiology of the epilepsies caused by these three non-channel genes is still unclear. Our group has been concentrating on studying the function of the two human genes *LGI1* and *EFHC1* using biochemical methods and zebrafish model, as well as trying to identify further non-channel epilepsy genes by looking for candidate genes with similar structure and same evolutionary origin.

Results

Phylogeny of the vertebrate LGI gene family

Our earlier work showed that *LGI1* belongs to a subfamily of leucine-rich repeat genes comprised of four members (*LGI1-LGI4*) in mammals[5]. In this study, both comparative developmental as well as molecular evolutionary methods were applied to investigate the evolution of the *LGI* gene family and, subsequently, of the functional importance of its different gene members.

We identified *LGI* genes with BLAST searches and RT-PCR experiments in chimpanzee, mouse, rat, chick, zebrafish and puffer fish genomes. No *LGI* ortholog was found in the invertebrate genomes of nematodes, the fruitfly and the ascidian, *Ciona intestinalis.* Our data suggest that the *LGI* gene family originated in the evolutionary lineage leading to the vertebrates.The phylogeny of the *LGI* gene family (Fig1) indicates an origin of the gene family through two rounds of gene or genome duplications. In this scenario, each of the two gene pairs *LGI1*/*LGI4* and *LGI2*/*LGI3* had one ancestral precursor gene. These two ancestral genes themselves may have arisen from a common "proto-*LGI*" gene. *LGI4* appears to be absent from zebrafish and puffer fish. The most probable scenario is that *LGI4* was lost in the lineage leading to the ray-finned fish. Alternatively, *LGI4* may have originated from a duplication of *LGI1* in the lineage leading to the sarcopterygians (lobe-finned fish), and also the mammals.

Expression pattern of the zebrafish *lgi* genes

We examined the embryonic expression patterns of all zebrafish *lgi* genes by whole-mount *in situ* hybridisation. Expression of *lgi1a* is first evident in the ventral diencephalon and at 24 hours post fertilisation (hpf) strong expression is observed in the developing eyes, in the ventral midbrain and hindbrain, and in the peripheral spinal cord (Fig.2*A-D*). By 48 hpf *lgi1a* is strongly expressed in the retinal ganglion cell layer, the diencephalon and along the ventral aspect of the hindbrain (Fig. 2*E-H*). Notably, all *lgi1a* expression domains are in neural tissues. *lgi1b* is expressed at 24 hpf in presumptive telencephalic and diencephalic bands and cranial paraxial mesenchyme. At 48 hpf, *lgi1b* transcripts are

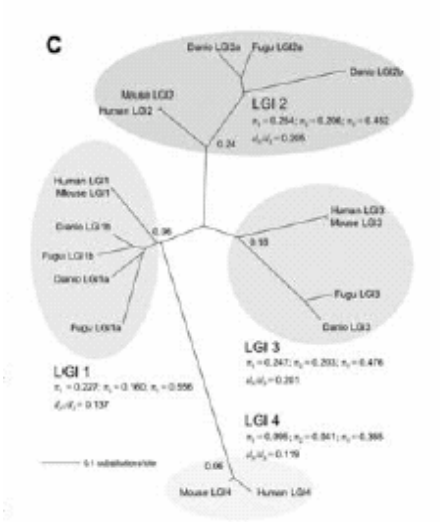

Fig 1: *Unrooted phylogeny of the LGI sub-families. Branch lengths are drawn in proportion to the expected number of nucleotide substitutions per codon.*

detected in the optic tectum, the cerebellum and in the zone of migrating neurons that proliferated in the rhombic lip. Expression is further observed in the dorsal thalamus and in the retinal ganglion cell layers (Fig. 2*J-L*). Overall, *lgi1a* expression is predominant in ventral parts of the mid- and hindbrain, while *lgi1b* is more dorsally restricted in this region. *In situ*-PCRs on adult transversal brain sections (Fig.2*M,N*) show that *lgi1a* and *lgi1b* are expressed in the outer layer of the periventricular grey zone of the optic tectum, an area rich in tectal neurons. *lgi1b,* in addition, is strongly expressed in the cerebellum. Both genes colocalize with nuclear areas of ganglion cells. At this level of resolution we could not detect expression in adult brain glial cells.

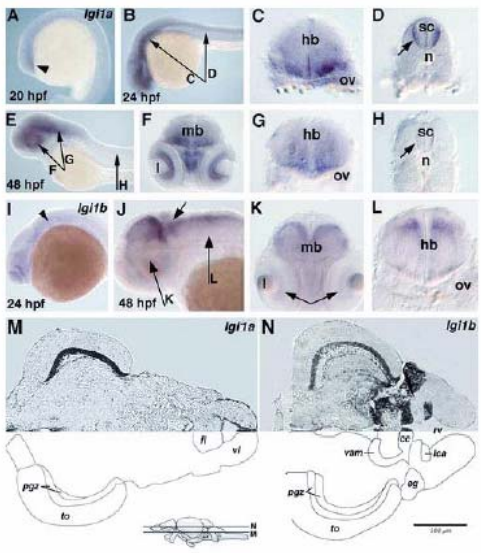

Fig 2: *Expression of lgi1 paralogs.*

NGFN Nationales Genomforschungsnetz

GEFÖRDERT VOM Bundesministerium für Bildung und Forschung

In contrast, expression of both *lgi2* paralogs is generally restricted to a few cells of putative (neuro-) ectodermal origin during embryogenesis. Both genes are expressed in trigeminal ganglion cells and in a few cells in the posterior head (Fig. 3*A-C*). More prominently, *lgi2a* is transiently detectable in dorsal spinal cord neurons. Finally, *lgi3* is expressed in cranial mesodermal cells and in a few cells on each side of the otic vesicle (Fig. 3*D* and not shown). *lgi3* appears to be co-expressed with *lgi1a* in the peripheral spinal cord in one- and two-day old embryos, and is detected in a reiterated symmetrical pattern of cells in the ventral hindbrain (Fig. 3*E-F*).

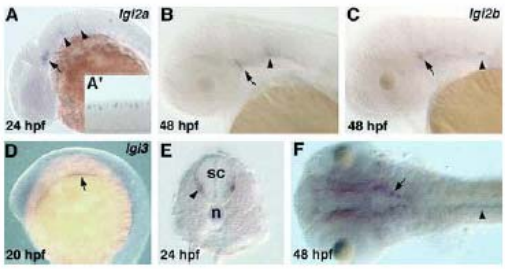

Fig 3: *Expression of lgi2 and lgi3 genes.*

Evolution of the vertebrate LGI gene family

Vertebrate *LGI* genes from mammals and teleosts were tested for signs of natural positive or negative selection in coding regions. Interestingly, *LGI1* and *LGI4* orthologs show evidence for strong negative natural selection (purifying selection), while the remaining groups of *LGI* orthologs exhibited rather moderate signs of negative selection pressure. Purifying selection is the form of natural selection that acts to eliminate selectively deleterious replacement mutations. Our data support the assumption that purifying selection is indicative of a gene that is more likely to cause human disease and points out the possible significance of *LGI4* for human neurological diseases. Interestingly, six out of ten amino acids mutated in human ADLTE exhibit high selection pressures, an observation which is in agreement with the role these mutations are assumed to play in the pathogenesis of this rare epilepsy.

Identification of EFHC2, the paralog of EFHC1

We identified and cloned a human paralog of *EFHC1*, which we named *EFHC2*. *EFHC2* is localised on chromosome Xp11.3 and is the only *EFHC1* paralog in human genome. Their protein products share an overall similarity of 41,6% at the amino acid level. The lacking of conservation of exon-intron boundaries between *EFHC1* and *EFHC2* argues that the assumed duplication event creating the two *EFHC* genes already happened before the vertebrate lineage. It's shown that massive intron gain and loss might accompany major evolutionary transitions, whereas little if any intron gains and limited amount of intron loss occur in the vertebrates. The ancient origin of both genes is also supported by the fact that homologs (i.e. Rib72) of *EFHC1*/*EFHC2* are found in species evolutionary as far away as *Chlamydomonas*.

Analysis of EFHC2 as a candidate suscepbility gene for IGE

To test whether genetic variation of the *EFHC2* gene region confers susceptibility to JME and other common IGE syndromes, we carried out an association study of six single nucleotide polymorphisms of the *EFHC2* gene in 654 German IGE patients and 662 population controls. Because of its X-chromosomal localisation, male and female patients were analysed seperately. None of the genotype distributions in female controls deviated significantly from that expected by Hardy-Weinberg equilibrium ($P > 0.15$). The scoring rate was >99% for the study samples. Genotyping reliability was ensured on duplicated samples (6.5%) and consistent Mendelian inheritance was demonstrated in eight control families. The LD pattern between adjacent SNPs in male controls showed week LD between SNP rs5906926 in the 5´-region of the *EFHC2* gene and SNP rs1885293 in intron 1 ($D´= 0.08$), whereas strong LD between the five intragenic SNPs ($D´ > 0.95$) indicates that these SNPs are located within a single haplotype block. Haplotype analysis of the five intragenic SNPs revealed four common haplotypes (> 3%) in the controls, accounting for > 93% of all haplotypes in male controls.

In males, the frequency of the T allele (Y430) of the amino acid substitution polymorphism (S430Y; rs2208592) was significantly increased in the JME patients (14.4%) compared to the controls (7.2%; $\chi^2 = 4.705$, df = 1, $P = 0.030$; OR = 2.17; 95-CI: 1.06 – 4.43). This allelic association was even stronger for 81 males with "classical" JME (JME without absence seizures) ($\chi^2 = 6.06$, df = 1, $P = 0.014$; OR = 2.46; 95-CI: 1.18 – 5.13). No allelic association of the Y430 allele was observed for the entire group of IGE males ($\chi^2 = 0.305$, df = 1, $P = 0.58$; OR = 1.19; 95-CI: 0.64 – 2.20). None of the allele frequencies of the other SNPs differed significantly between the IGE groups when compared with the controls ($P > 0.18$). Likewise, no significant difference in the distribution of the common intragenic five-SNP haplotypes was found in the IGE groups compared with the controls ($P > 0.18$). The observed allelic association between JME and Y430 was no longer significant after Bonferroni adjustment for multiple comparisons.

In females, there was no significant difference in the allele frequencies of any SNP between the IGE groups when compared with the controls ($P > 0.21$). Likewise, a similar distribution of the common intragenic five-SNP haplotypes was found in the IGE groups as in the controls ($P > 0.82$).

An association with the gonosomal gene *EFHC2* would be in accordance with the observed preponderance of maternal inheritance in JME. Independent replication studies are needed to further analyse the tentative association described here.

Outlook

The expression analysis of LGI genes in zebrafish provides the basis for the planned morpholino knock-down experiments, with which the question of LGI function can be adressed directly in a simplified although still considerably conserved vertebrate nervous system. Besides the evolutionary and expression pattern analysis of the LGI gene family, we are also studying the function of LGI1 protein with biochemical methods. Using yeast two-hybrid screenings, two candidate proteins interacting with LGI1 protein have been identified. We are currently verifying them with immuno-precipitation assays. Also in this sense, zebrafish offers an attactive model system to test the biochemical results *in vivo*. The EFHC genes are being studied in zebrafish model as well. As a summary, by combining molecular genetical, biochemical and zebrafish model mehtods, we hope to get further insight into the function of the non-channel epilepsy genes and to understand the etiology of the epilepsies caused by their mutations.

Lit.: ***1.*** *Steinlein OK. Genetic mechanisms that underlie epilepsy. Nat. Rev. Neurosci. 5:400-8.* ***2.*** *Kalachikov S et al. Mutations in LGI1 cause autosomal-dominant partial epilepsy with auditory features.Nat Genet. 2002 30:* 335-41. ***3.*** *Skradski SL et al. A novel gene causing a mendelian audiogenic mouse epilepsy.Neuron. 2001 31:537-44.* ***4.*** *Suzuki T et al. Mutations in EFHC1 cause juvenile myoclonic epilepsy.Nat Genet. 2004 36:842-9.* ***5.*** *Gu W et al. The LGI1 gene involved in lateral temporal lobe epilepsy belongs to a new subfamily of leucine-rich repeat proteins. FEBS Lett. 2002 519:71-6.*