

# **Structural and functional promoter analyses of epilepsy-relevant genes as basis for targeted expression interference**

Doctoral thesis

to obtain a doctorate (PhD)

from the Faculty of Medicine

of the University of Bonn

**Despina Tsortouktzidis**

from Caracas, Venezuela

2023

Written with authorization of  
the Faculty of Medicine of the University of Bonn

First reviewer: Prof. Dr. Albert Becker

Second reviewer: Prof. Dr. Yvonne Weber

Day of oral examination: 31.05.2023

From the Institute of Neuropathology

Director: Prof. Dr. med. Torsten Pietsch

## Table of Contents

<b>List of abbreviations</b>	<b>4</b>
<b>1. Abstract</b>	<b>5</b>
<b>2. Introduction and aims</b>	<b>6</b>
2.1 Genetic epilepsies	6
2.2 Genetic aspects of acquired epilepsies	7
2.3 Genome-wide association studies (GWAS) and regulatory single nucleotide polymorphisms	7
2.4 Study and modulation of the non-coding genome by CRISPR	9
2.5 Aims	9
2.6 References	10
<b>3. Publications</b>	<b>14</b>
3.1 Publication 1: Gene expression analysis in epileptic hippocampi reveals a promoter haplotype conferring reduced aldehyde dehydrogenase 5a1 expression and responsiveness	14
3.2 Publication 2: SCN1A overexpression, associated with a genomic region marked by a risk variant for a common epilepsy, raises seizure susceptibility	21
3.3 Publication 3: A Versatile Clustered Regularly Interspaced Palindromic Repeats Toolbox to Study Neurological Ca <sup>v</sup> 3.2 Channelopathies by Promoter-Mediated Transcription Control	43
<b>4. Discussion</b>	<b>53</b>
4.1 Allelic variants of rSNP confer differential gene expression in epileptic hippocampi	53
4.2 Mechanisms that mediate transcriptional differences associated with SNPs	54
4.3 CRISPRa/i for identifying, validating, and functionally assessing regulatory elements of the genome	55
4.4 References	57
<b>5. Acknowledgements</b>	<b>60</b>

**List of abbreviations**

ALDH5A1	Aldehyde Dehydrogenase 5 Family Member A1
CNS	Central nervous system
CRISPR	Clustered regularly interspaced short palindromic repeats
dCas9	Dead Cas9
EMSA	Electrophoretic mobility shift assay
FS	Febrile seizures
GWAS	Genome-wide association studies
ILAE	International League Against Epilepsy
KRAB	Krüppel-associated box
MTF1	Metal regulatory transcription factor 1
MTLE	Mesial temporal lobe epilepsy
MTLEHS	Mesial temporal lobe epilepsy with hippocampal sclerosis
rSNPs	Regulatory single-nucleotide polymorphism
SCN1A	Sodium Voltage-Gated Channel Alpha Subunit 1
SE	Status epilepticus
SNP	Single-nucleotide polymorphism
SSADHD	Succinic semialdehyde dehydrogenase deficiency
TLE	Temporal lobe epilepsy
VPR	VP64-p65-Rta

## 1. Abstract

Background: Epilepsies are a group of brain disorders characterized by recurrent seizures. A fraction of those has a well defined genetic aetiology including monogenic epilepsies, and others are the result of acquired factors such as head trauma. Interestingly, the fact that (i) monogenic epilepsies can present with a wide variety of clinical manifestations and that (ii) the risk of epilepsy increases after an acquired factor in individuals with a familial history of epilepsy strongly suggests a genetic susceptibility. GWAS showed that most trait-associated SNPs lie in non-coding genomic regions. Thus, they may operate by altering gene expression. Functional characterization of non-coding SNPs is challenging. Hippocampal biopsies from patients, in combination with *in vitro* reporter assays are of great value to determine the functionality of regulatory SNPs in the context of epilepsy. A new technology based on CRISPR can manipulate gene expression by targeting non-coding regulatory elements and be used therefore to identify, validate, and functionally elucidate vulnerability loci in the genome.

Aims: (1) Investigate if candidate SNPs in non-coding genomic regions of genes previously associated with epilepsy alter gene expression in hippocampi of temporal lobe epilepsy (TLE) patients indicating genetic susceptibility, (2) reveal the molecular promoter control mechanism that confers distinct allelic expression and (3) engineer a CRISPR toolbox to manipulate the promoter activity of the epileptogenesis relevant calcium channel *cacna1h*, as a way to further understand the genomic architecture and its importance on gene expression.

Results and conclusion: The first publication demonstrated that promoter haplotypes of *ALDH5A1* confer differential gene expression in the hippocampi of TLE patients and that the promoter might be under regulation of the transcription factor Egr1. The second publication demonstrated that a GWAS-derived SNP in the *SCN1A* gene, confers differential expression in hippocampi of TLE patients and that the SNP lies in an intronic region that contains promoter-like activity. In the third publication, we showed that CRISPR modulates the promoter activity of *cacna1h* in a very specific manner, resulting in changes in gene expression. This work contributes to a better understanding of the importance of genetic regulatory regions in epilepsy and encourages the use of CRISPR methods in the field.

## 2. Introduction and aims

Epilepsies are a group of brain disorders characterized by recurrent epileptic seizures (The International League Against Epilepsy Consortium on Complex Epilepsies, 2018). They affect 50-65 million people, which represents a prevalence of 1 % worldwide (Cavalleri et al., 2007; Fernández-Marmiesse et al., 2019). Five etiological groups are currently recognized by the International League Against Epilepsy (ILAE): genetic, structural, metabolic, immune, and infectious (Scheffer et al., 2017). Of note, more than one etiological group maybe the involved in a single diagnosis. In this work, we will refer to “genetic epilepsies” as the ones known to have a genetic etiology and to “acquired or symptomatic epilepsies” as the ones that result from acquired factors.

### 2.1 Genetic epilepsies

According to the ILAE, genetic epilepsies are the consequence of a known or presumed mutation where seizures are the main symptom of the disorder (Scheffer et al., 2017). This group comprises monogenic epilepsies that include familial epilepsies and epilepsies caused by *de novo* mutations, and complex genetic epilepsies that are not inherited in a monogenic manner (Helbig and Lowenstein, 2013).

Of epilepsies suspected to have a genetic etiology only 1 to 2 % appear to be monogenic (Weber and Lerche, 2008). Well described examples are the syndrome of Benign Familial Neonatal Epilepsy where mutations in the potassium channels (*KCNQ2* or *KCNQ3*) are common (Scheffer et al., 2017), and the autosomal dominant nocturnal frontal lobe epilepsy where genes that encode subunits of the neuronal acetylcholine receptor (*CHRNA4* and *CHRNA2*) present with rare variants (Weber and Lerche, 2008).

A group of severe epilepsies named Developmental and epileptic encephalopathies (DEEs) is often caused by *de novo* genetic variants (Campbell et al., 2022). Although DEEs are primarily monogenic, they present with high genetic heterogeneity where genetic variants can show incomplete penetrance (Campbell et al., 2022; Scheffer et al., 2017). Intriguingly, DEE can also present with differences in clinical manifestations (Campbell et al., 2022). One of the best examples where monogenic epilepsies may cause a wide range of clinical outcomes is Dravet syndrome where more than 80 % of patients present with pathogenic mutations in *SCN1A* and the spectrum might range from mild to more severe forms of epilepsies (Scheffer et al., 2017). The heterogeneity seen in genetic

epilepsies and the low proportion of monogenic epilepsies suggest that most epilepsy syndromes are genetically complex where multiple genes and genetic variants contribute to their etiology (Ottman, 2005).

## 2.2 Genetic aspects of acquired epilepsies

Besides epilepsies caused by rare genetic variants, there is another group of epilepsies referred to as “symptomatic epilepsies” which result from acquired factors (Berkovic et al., 2006). Acquired causes are those that occur after a brain insult such as head trauma, stroke, infection, tumour, and perinatal injuries (Berkovic et al., 2006; Perucca and Scheffer, 2021). Interestingly, epidemiological studies have shown that only a subset of individuals develop epilepsy after a brain insult e.g., 2-5 % after brain trauma and 2-4 % after stroke (Perucca and Scheffer, 2021). Although, the risk of developing epilepsy after an acquired etiology mainly correlates to the severity of the injury, age and other factors, genetic contribution has been shown to play a role (Perucca and Scheffer, 2021). In a study, Eriksson et al. found that the risk of developing late post-stroke seizures was increased in patients that had a familial history of epilepsy in comparison to those without (Eriksson et al., 2019). Similarly, the risk of epilepsy after head trauma was found to be increased in patients that present with a familial history of epilepsy (Christensen et al., 2009).

The wide range of clinical manifestations seen in epilepsies of known or suspected genetic etiology as well as the increased risk of developing epilepsy after a brain insult when there is a familial history of epilepsy support the idea that there might be a genetic vulnerability that contributes to the epilepsy-phenotype and outcome. In this context, it has been hypothesized that common genetic variants that usually are of small effects per se in combination with either rare variants (in monogenic epilepsies) or brain insults (in acquired forms of epilepsy) might account for the heterogeneity seen in the clinical outcome (Campbell et al., 2022).

## 2.3 Genome-wide association studies (GWAS) and regulatory single nucleotide polymorphisms

GWAS seeks for an association between a trait and common variants in a population. The most common genetic variation is the single nucleotide polymorphism (SNP) occurring roughly every 500-1000 bp in the genome (Smigielski et al., 2000). A SNP consists in a

germline single nucleotide substitution that occurs at a specific region in the genome, and it is found in at least 1 % of the human population (Wang et al., 1998). SNPs can be classified according to their location in the genome and their functional consequences. SNPs present in protein-coding genomic regions are: (i) synonymous, where the nucleotide change does not lead to amino acid alteration, (ii) missense variant which results in amino acid substitution, and (iii) nonsense variants where the amino acid change results in a stop codon and therefore the production of a truncated protein. On the other hand, SNPs present in non-coding regulatory regions of the genome (e.g. promoter, enhancers) are named regulatory SNPs (rSNPs) and have functional implications by altering gene expression (Ramírez-Bello et al., 2013).

An important discovery in the past few decades is that the non-coding is considerably larger than the protein-coding genome and it is involved in gene regulation (Gloss and Dinger, 2018). Interestingly, GWAS have revealed that most risk-associated SNPs are in non-coding genomic regions (Tak and Farnham, 2015). These findings led to the hypothesis that variations in gene expression over than expression of certain genes play a major role in the differences in human phenotypes, raising the importance of rSNPs (Buckland, 2006). rSNPs might modify gene function by affecting gene expression and its spatial and temporal regulation (Prokunina and Alarcón-Riquelme, 2004). It has been suggested that a large proportion of rSNPs lie within gene promoters (Buckland, 2006) and that one of the mechanisms by which rSNPs may influence gene expression is by modifying, destroying, or creating binding sites for transcription factors that regulate gene expression (Prokunina and Alarcón-Riquelme, 2004).

In comparison to rare variants of large effect e.g. mutations in cases of monogenic epilepsies, common SNP variants have a small size functional effect but may confer susceptibility and/or be involved in the variability of the epilepsy phenotype. The potential role of rSNPs in genetic complex diseases and the contribution that they might confer in phenotypic variability highlights the importance of understanding better the genomic architecture in the context of gene expression regulation.

Identifying functional consequences of rSNPs is challenging per se, especially in neurosciences where the lack of availability of human brain tissue constitutes a clear limitation. Several *in vitro* assays are currently used for investigating the effect of allelic variants on gene expression. Although very valuable, those methods e.g. reporter assays



lack the genomic context in the experimental setup. Since most trait-associated SNPs lie in non-protein coding regions of the genome there is a current need to utilize methods that seek for identification and functional validation of regulatory elements. A new state-of-the-art technique based on the artificial manipulation of gene transcription might contribute considerably to the identification of genomic-relevant loci for gene regulation.

#### 2.4 Study and modulation of the non-coding genome by CRISPR

The clustered regularly interspaced short palindromic repeats (CRISPR) system allows gene sequence editing, epigenetic manipulation, and expression modulation of target genes. It consists of two components: the nuclease protein Cas9 and a single guide RNA (sgRNA) that directs the nuclease activity to its complementary DNA sequence in the genome (Kampmann, 2018). A modification of the conventional Cas9 into a nuclease-dead mutant Cas9 (dCas9) lacks the ability to cut the DNA but preserves the sgRNA-directed binding of the specific DNA sequence. Fusing a dCas9 with different effectors leads to a wide variety of functions. The CRISPR interference (CRISPRi) approach fuses dCas9 with the Krüppel-associated box (KRAB) and results in transcription inhibition of the target gene (Gilbert et al., 2013). Transcriptional activation is also possible through CRISPR activation (CRISPRa) where dCas9 is fused to different activator domains e.g. VPR (VP64-p65-Rta) leading to the activation of gene expression (Chavez et al., 2015). To modulate gene expression through CRISPRa or CRISPRi the sgRNAs have to be designed to target regulatory-relevant elements such as gene promoters and enhancers. Recent studies have taken advantage of this method to identify, validate and functionally assess regulatory regions of the genome (Fulco et al., 2016; Zhou et al., 2022). More specifically, using sgRNA molecules targeting risk loci in the genome and assessing gene expression allows the identification of these important non-coding elements.

#### 2.5 Aims

Considering that SNPs distribute throughout the genome and are commonly found in the human population it is possible that (i) the same genes that are substrates for severe forms of monogenic epilepsies may also be subjected to the effect of rSNPs and (ii) rSNPs in such genes might confer susceptibility in more common forms of epilepsy including TLE (where mutations are not frequently the main cause of the disease). An example is the gene *ALDH5A1*, where mutations result in succinic semialdehyde dehydrogenase

deficiency (SSADH) causing epilepsy in about half of the affected individuals. Also, a rSNP in the same gene has been found to be overrepresented in a TLE cohort (Pernhorst et al., 2011). Similarly, mutations in coding regions of the *SCN1A* gene cause Dravet syndrome (Brunklaus and Zuberi, 2014), whereas in the same gene, an intronic SNP was found to be associated with „mesial temporal lobe epilepsy with hippocampal sclerosis (MTLEHS) and a history of febrile seizures (FS)” (Kasperavičiute et al., 2013).

These observations highlight the importance of studying the non-coding genome and its implication in transcriptional regulation in epilepsy. In line with this statement, animal models of TLE have shown dynamic changes in the expression of transcription factors soon after SE (Beer et al., 1998), which in turn may alter the expression of ion channels. A cascade of promoter regulation has been well described for the T-Type calcium channel Cav3.2 (*Cacna1h*) in an animal model of TLE (Becker et al., 2008; van Loo et al., 2015). Importantly, this gene has been involved in both genetic as well as acquired epilepsies. The overall goal of this study is to improve the understanding of how regulatory regions of the genome may influence the function of epileptic-associated genes and to explore the applicability of new technologies based on the CRISPR approach as a new way to elucidate these mechanisms functionally.

Based on the importance that genetic motifs in non-coding regulatory regions may exert for transcriptional regulation in the context of epilepsy, this work has the individual aims to (1) investigate whether candidate rSNPs in genes previously associated with genetic epilepsies (*ALDH5A1* and *SCN1A*) affect gene expression in hippocampi of TLE patients, potentially indicating genetic susceptibility, (2) identify key transcription factors that influence distinct allelic expression (stratified by genotype) relevant for epilepsy and (3) engineer a CRISPR toolbox to manipulate the promoter activity of the epileptogenesis relevant calcium channel *cacna1h*, as a way to further understand the genomic architecture and its importance on gene expression.

## 2.6 References

Becker AJ, Pitsch J, Sochivko D, Opitz T, Staniek M, Chen CC, Campbell KP, Schoch S, Yaari Y, Beck H. Transcriptional upregulation of Cav3.2 mediates epileptogenesis in the pilocarpine model of epilepsy. *J Neurosci* 2008; 28: 13341-13353

Beer J, Mielke K, Zipp M, Zimmermann M, Herdegen T. Expression of c-jun, junB, c-fos, fra-1 and fra-2 mRNA in the rat brain following seizure activity and axotomy. *Brain Res* 1998; 794: 255-266

Berkovic SF, Mulley JC, Scheffer IE, Petrou S. Human epilepsies: interaction of genetic and acquired factors. *Trends Neurosci* 2006; 29: 391-397

Brunklaus A, Zuberi SM. Dravet syndrome--from epileptic encephalopathy to channelopathy. *Epilepsia* 2014; 55: 979-984

Buckland PR. The importance and identification of regulatory polymorphisms and their mechanisms of action. *Biochim Biophys Acta* 2006; 1762: 17-28

Campbell C, Leu C, Feng YCA, Wolking S, Moreau C, Ellis C, Ganesan S, Martins H, Oliver K, Boothman I, Benson K, Molloy A, Brody L, Epi4K Collaborative, Genomics England Research Consortium, Michaud JL, Hamdan FF, Minassian BA, Lerche H, Scheffer IE, Sisodiya S, Girard S, Cosette P, Delanty N, Lal D, Cavalleri GL, Epi25 Collaborative. The role of common genetic variation in presumed monogenic epilepsies. *EBioMedicine* 2022; 81: 104098

Cavalleri GL, Weale ME, Shianna K v., Singh R, Lynch JM, Grinton B, Szoeker C, Murphy K, Kinirons P, O'Rourke D, Ge D, Depondt C, Claeys KG, Pandolfo M, Gumbs C, Walley N, McNamara J, Mulley JC, Linney KN, Sheffield LJ, Radtke RA, Tate SK, Chisoe SL, Gibson RA, Hosford D, Stanton A, Graves TD, Hanna MG, Eriksson K, Kantanen AM, Kalviainen R, O'Brien TJ, Sander JW, Duncan JS, Scheffer IE, Berkovic SF, Wood NW, Doherty CP, Delanty N, Sisodiya SM, Goldstein DB. Multicentre search for genetic susceptibility loci in sporadic epilepsy syndrome and seizure types: a case-control study. *Lancet Neurol* 2007; 6: 970-980

Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, P R Iyer E, Lin S, Kiani S, Guzman CD, Wiegand DJ, Ter-Ovanesyan D, Braff JL, Davidsohn N, Housden BE, Perrimon N, Weiss R, Aach J, Collins JJ, Church GM. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* 2015; 12: 326-328

Christensen J, Pedersen MG, Pedersen CB, Sidenius P, Olsen J, Vestergaard M. Long-term risk of epilepsy after traumatic brain injury in children and young adults: a population-based cohort study. *Lancet* 2009; 373: 1105-1110

Eriksson H, Wirdefeldt K, Asberg S, Zelano J. Family history increases the risk of late seizures after stroke. *Neurology* 2019; 93: E1964-E1970

Fernández-Marmiesse A, Roca I, Díaz-Flores F, Cantarín V, Pérez-Poyato MS, Fontalba A, Laranjeira F, Quintans S, Moldovan O, Felgueroso B, Rodríguez-Pedreira M, Simón R, Camacho A, Quijada P, Ibanez-Mico S, Domingno MR, Benito C, Calvo R, Pérez-Cejas A, Carrasco ML, Ramos F, Couce ML, Ruiz-Falcó ML, Gutierrez-Solana L, Martínez-Atienza M. Rare Variants in 48 Genes Account for 42% of Cases of Epilepsy With or

Without Neurodevelopmental Delay in 246 Pediatric Patients. *Front Neurosci* 2019; 13: 1135

Fulco CP, Munschauer M, Anyoha R, Munson G, Grossman SR, Perez EM, Kane M, Cleary B, Lander ES, Engreitz JM. Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* 2016; 354: 769-773

Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi LS. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 2013; 154: 442-451

Gloss BS, Dinger ME. Realizing the significance of noncoding functionality in clinical genomics. *Exp Mol Med* 2018; 50: 1-8

Helbig I, Lowenstein DH. Genetics of the epilepsies: where are we and where are we going? *Curr Opin Neurol* 2013; 26: 179-185

Kampmann M. CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chem Biol* 2018; 13: 406-416

Kasperavičiute D, Catarino CB, Matarin M, Leu C, Novy J, Tostevin A, Leal B, Hessel EVS, Hallmann K, Hildebrand MS, Dahl HHM, Ryten M, Trabzuni D, Ramasamy A, Alhusaini S, Doherty CP, Dorn T, Hansen J, Krämer G, Steinhoff BJ, Zumsteg D, Duncan S, Kälviäinen RK, Eriksson KJ, Kantanen AM, Pandolfo M, Gruber-Sedlmayr U, Schlachter K, Reinthaler EM, Stogmann E, Zimprich F, Théâtre E, Smith C, O'Brien TJ, Meng Tan K, Petrovski S, Robbiano A, Paravidino R, Zara F, Striano P, Sperling MR, Buono RJ, Hakonarson H, Chaves J, Costa PP, Silva BM, da Silva AM, de Graan PNE, Koeleman BPC, Becker A, Schoch S, von Lehe M, Reif PS, Rosenow F, Becker F, Weber Y, Lerche H, Rössler K, Buchfelder M, Hamer HM, Kobow K, Coras R, Blumcke I, Scheffer IE, Berkovic SF, Weale ME, UK Brain Expression Consortium, Delanty N, Depondt C, Cavalleri GL, Kunz WS, Sisodiya SM. Epilepsy, hippocampal sclerosis and febrile seizures linked by common genetic variation around SCN1A. *Brain* 2013; 136: 3140-3150

Ottman R. Analysis of genetically complex epilepsies. *Epilepsia* 2005; 46: 7-14

Pernhorst K, Raabe A, Niehusmann P, van Loo KMJ, Grote A, Hoffmann P, Cichon S, Sander T, Schoch S, Becker AJ. Promoter variants determine  $\gamma$ -aminobutyric acid homeostasis-related gene transcription in human epileptic hippocampi. *J Neuropathol Exp Neurol* 2011; 70: 1080-1088

Perucca P, Scheffer IE. Genetic Contributions to Acquired Epilepsies. *Epilepsy Curr* 2021; 21: 5-13

Prokunina L, Alarcón-Riquelme ME. Regulatory SNPs in complex diseases: their identification and functional validation. *Expert Rev Mol Med* 2004; 6: 1-15

Ramírez-Bello J, Vargas-Alarcón G, Tovilla-Zárate C, Fragoso JM. Single nucleotide polymorphisms (SNPs): functional implications of regulatory-SNP (rSNP) and structural RNA (srSNPs) in complex diseases. *Gac Med Mex* 2013; 149: 220-228

Scheffer IE, Berkovic S, Capovilla G, Connolly MB, French J, Guilhoto L, Hirsch E, Jain S, Mathern GW, Moshé SL, Nordli DR, Perucca E, Tomson T, Wiebe S, Zhang YH, Zuberi SM. ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia* 2017; 58: 512-521

Smigielski EM, Sirotkin K, Ward M, Sherry ST. DbSNP: a database of single nucleotide polymorphisms. *Nucleic Acids Res* 2000; 28: 352-355

Tak YG, Farnham PJ. Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin* 2015; 8: 57

The International League Against Epilepsy Consortium on Complex Epilepsies. Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nat Commun* 2018; 9: 5269

van Loo KMJ, Schaub C, Pitsch J, Kulbida R, Opitz T, Ekstein D, Dalal A, Urbach H, Beck H, Yaari Y, Schoch S, Becker AJ. Zinc regulates a key transcriptional pathway for epileptogenesis via metal-regulatory transcription factor 1. *Nat Commun* 2015; 6: 8688

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 1998; 280: 1077-1082

Weber YG, Lerche H. Genetic mechanisms in idiopathic epilepsies. *Dev Med Child Neurol* 2008; 50: 648-654

Zhou T, Zhu X, Ye Z, Wang YF, Yao C, Xu N, Zhou M, Ma J, Qin Y, Shen Y, Tang Y, Yin Z, Xu H, Zhang Y, Zang X, Ding H, Yang W, Guo Y, Harley JB, Namjou B, Kaufman KM, Kottyan LC, Weirauch MT, Hou G, Shen N. Lupus enhancer risk variant causes dysregulation of IRF8 through cooperative lncRNA and DNA methylation machinery. *Nat Commun* 2022; 13: 1855

### 3. Publications





3.1 Publication 1: Gene expression analysis in epileptic hippocampi reveals a promoter haplotype conferring reduced aldehyde dehydrogenase 5a1 expression and responsiveness

**Tsourtoukzidis D**, Schulz H, Hamed M, Vatter H, Surges R, Schoch S, Sander T, Becker AJ, van Loo KMJ. Gene expression analysis in epileptic hippocampi reveals a promoter haplotype conferring reduced aldehyde dehydrogenase 5a1 expression and responsiveness. **Epilepsia**. 2021 Jan;62(1):e29-e34. doi: 10.1111/epi.16789. Epub 2020 Dec 15. PMID: 33319393.

## BRIEF COMMUNICATION

Epilepsia®

# Gene expression analysis in epileptic hippocampi reveals a promoter haplotype conferring reduced aldehyde dehydrogenase 5a1 expression and responsiveness

Despina Tsortouktzidis<sup>1</sup> | Herbert Schulz<sup>2</sup> | Motaz Hamed<sup>3</sup> | Hartmut Vatter<sup>3</sup> |  
 Rainer Surges<sup>4</sup>  | Susanne Schoch<sup>1</sup> | Thomas Sander<sup>2</sup>  | Albert J. Becker<sup>1</sup>  |  
 Karen M. J. van Loo<sup>1,5</sup> 

<sup>1</sup>Section for Translational Epilepsy Research, Department of Neuropathology, University of Bonn Medical Center, Bonn, Germany

<sup>2</sup>Cologne Center of Genomics, University of Cologne, Germany

<sup>3</sup>Clinic for Neurosurgery, University of Bonn Medical Center, Bonn, Germany

<sup>4</sup>Department of Epileptology, University of Bonn Medical Center, Bonn, Germany

<sup>5</sup>Department of Epileptology, Neurology, RWTH Aachen University, Aachen, Germany

## \*Correspondence

Karen M. J. van Loo, Department of Epileptology, Neurology, RWTH Aachen University, Pauwelsstrasse 30, 52074 Aachen, Germany.  
 Email: kvanloo@ukaachen.de

## Funding information

BONFOR; Deutsche Forschungsgemeinschaft, Grant/Award Number: FOR 2715, SCHO 820 5-2, SCHO 820/4-1, SCHO 820/5-2, SCHO 820/6-1, SCHO 820/7-2 and SFB 1089

## Abstract

Increasing evidence indicates the pathogenetic relevance of regulatory genomic motifs for variability in the manifestation of brain disorders. In this context, *cis*-regulatory effects of single nucleotide polymorphisms (SNPs) on gene expression can contribute to changing transcript levels of excitability-relevant molecules and episodic seizure manifestation in epilepsy. Biopsy specimens of patients undergoing epilepsy surgery for seizure relief provide unique insights into the impact of promoter SNPs on corresponding mRNA expression. Here, we have scrutinized whether two linked regulatory SNPs (rs2744575; 4779C > G and rs4646830; 4854C > G) located in the aldehyde dehydrogenase 5a1 (succinic semialdehyde dehydrogenase; *ALDH5A1*) gene promoter are associated with expression of corresponding mRNAs in epileptic hippocampi ( $n = 43$ ). The minor *ALDH5A1*-GG haplotype associates with significantly lower *ALDH5A1* transcript abundance. Complementary *in vitro* analyses in neural cell cultures confirm this difference and further reveal a significantly constricted range for the minor *ALDH5A1* haplotype of promoter activity regulation through the key epileptogenesis transcription factor Egr1 (early growth response 1). The present data suggest systematic analyses in human hippocampal tissue as a useful approach to unravel the impact of epilepsy candidate SNPs on associated gene expression. Aberrant *ALDH5A1* promoter regulation in functional terms can contribute to impaired  $\gamma$ -aminobutyric acid homeostasis and thereby network excitability and seizure propensity.

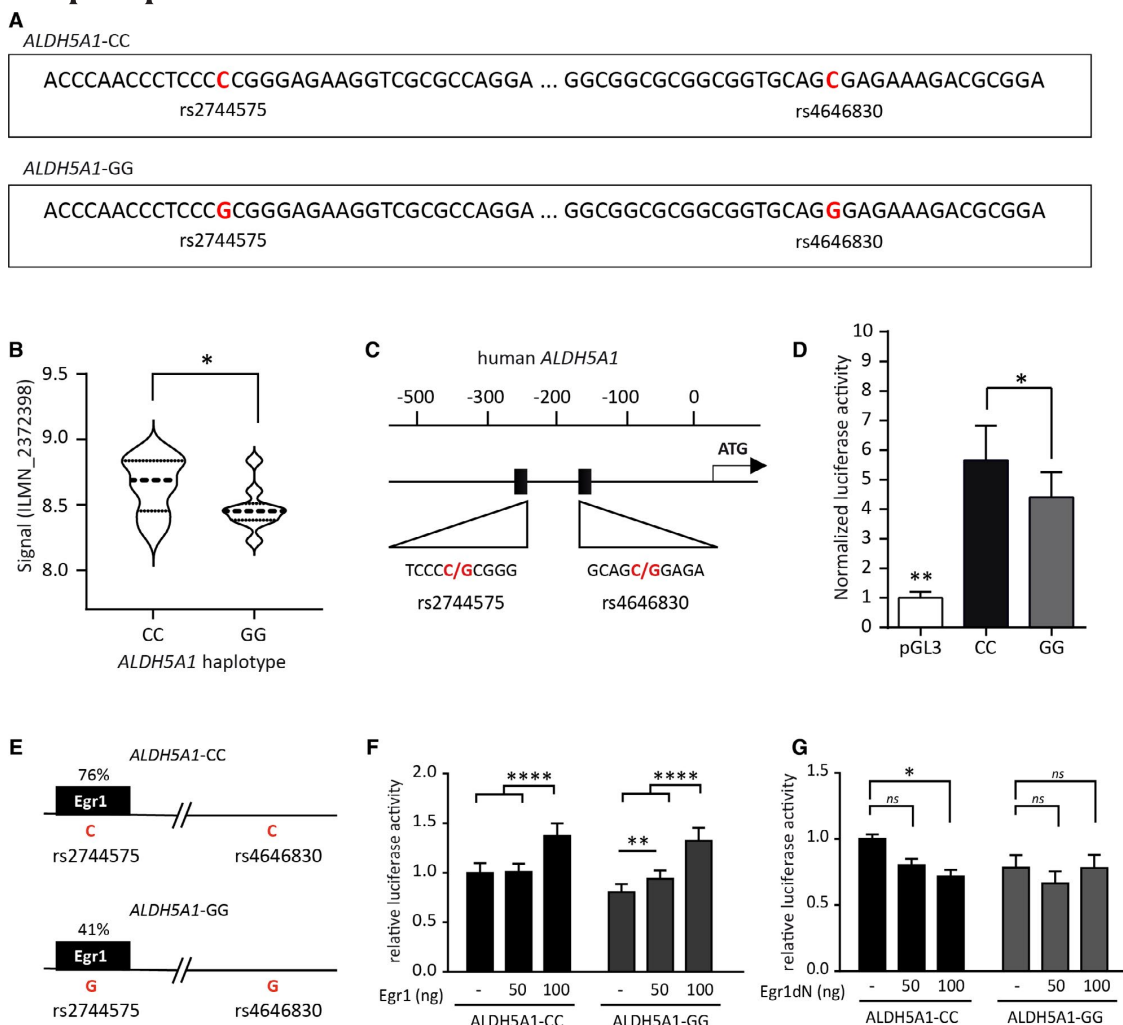
## KEY WORDS

*ALDH5A1*, early growth response 1, rSNPs, temporal lobe epilepsy

Albert J. Becker and Karen M. J. van Loo contributed equally to the work.

This is an open access article under the terms of the Creative Commons Attribution NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2020 The Authors. *Epilepsia* published by Wiley Periodicals LLC on behalf of International League Against Epilepsy



**FIGURE 1** *ALDH5A1* promoter haplotypes correlate with *ALDH5A1* expression levels in hippocampi of temporal lobe epilepsy (TLE) patients and in vitro, and are potentially regulated by Egr1. (A) Schematic representation of *ALDH5A1* promoter haplotypes. The *ALDH5A1*-CC haplotype represents the rs2744575/C and rs4646830/C genotype, whereas *ALDH5A1*-GG refers to rs2744575/G and rs4646830/G (single nucleotide polymorphism [SNP] positions are indicated in red). (B) Violin plots show the expression levels of *ALDH5A1* in hippocampi of TLE patients stratified by haplotype. Patients homozygous for *ALDH5A1*-CC haplotype have higher *ALDH5A1* expression levels than patients homozygous for the *ALDH5A1*-GG haplotype (*ALDH5A1*-CC:  $n = 33$ ,  $8.652 \pm 0.03716$ ; *ALDH5A1*-GG:  $n = 10$ ,  $8.467 \pm 0.05144$ ;  $t$ -test,  $p = .0160$ ). Dashed and dotted lines in the violin plot show median and quartiles (median: *ALDH5A1*-CC = 8.690, *ALDH5A1*-GG = 8.452). (C) Schematic representation of the location of the *ALDH5A1* promoter SNPs relative to the ATG. The SNPs rs2744575 and rs4646830 are located 250 bp and 175 bp upstream of the start-ATG, respectively. (D) Luciferase activity of NG108-15 cells transfected with *ALDH5A1*-CC, *ALDH5A1*-GG, and the pGL3 empty vector. Both haplotypes have higher affinity than cells transfected with the pGL3 vector, and the *ALDH5A1*-CC haplotype has a stronger promoter activity compared to the *ALDH5A1*-GG haplotype ( $N = 9$ ,  $n = 3$ , data normalized to pGL3 and represented as mean  $\pm$  SEM, pGL3  $1.00 \pm 0.2024$ , *ALDH5A1*-CC  $5.657 \pm 1.170$ , and *ALDH5A1*-GG  $4.398 \pm 0.8484$ , one-way analysis of variance [ANOVA] of repeated measures, Tukey multiple comparisons test,  $*p \leq .05$ ,  $**p \leq .01$ ). (E) Schematic representation of Egr1 binding affinity to allelic variants of SNP rs2744575 shows higher affinity for variant rs2744575/C. (F) Relative luciferase activity of the *ALDH5A1* promoter haplotypes after overexpression with Egr1. Both haplotypes show increased promoter activity when exposed to 100 ng of Egr1 ( $N = 5$ ,  $n = 3$ , data normalized to *ALDH5A1*-CC basal and represented as mean  $\pm$  SEM, two-way ANOVA, Sidak multiple comparisons test,  $**p \leq .01$ ,  $****p \leq .0001$ ). (G) Luciferase activity of NG108-15 cells transfected with the *ALDH5A1*-CC and *ALDH5A1*-GG promoter haplotypes and Egr1dN. The *ALDH5A1*-CC haplotype shows a decreased luciferase activity after exposure to Egr1dN, whereas the *ALDH5A1*-GG haplotype remains constant (data normalized to *ALDH5A1*-CC basal and represented as mean  $\pm$  SEM, two-way ANOVA, Dunnett multiple comparisons test;  $p = .0497$ ,  $N = 4$ ,  $n = 3$ ). ns, not significant



## 1 | INTRODUCTION

Transient neuronal network hyperexcitability manifested by seizures represents the shared leading symptom of both acquired focal and genetic generalized epilepsies. Genome-wide association studies have identified a varying number of susceptibility loci for different forms of epilepsies.<sup>1–3</sup> Many associated single nucleotide polymorphisms (SNPs) are located in noncoding genomic regions. Functionally, their effects often stay unresolved.<sup>4</sup> However, SNP variants located in promoter regions can differentially impact the dynamic expression of corresponding genes, which can have relevance in epilepsy syndromes with very distinct genetic components. These can include (1) acquired epilepsies, including temporal lobe epilepsy (TLE) with only minor genetic contribution; (2) idiopathic generalized epilepsies with a complex pattern of inheritance suggesting an interaction of several susceptibility genes<sup>5</sup>; (3) monogenic epileptic encephalopathies; and (4) rare genetic disorders affecting aspects of neurotransmission. The latter includes deficiency in the succinic semialdehyde dehydrogenase (SSADH) resulting in impaired  $\gamma$ -aminobutyric acid (GABA) degradation.<sup>6–9</sup> This epilepsy syndrome shows differences in the clinical manifestation albeit the presence of the same mutational event. Here, we have analyzed two linked regulatory (r)SNPs (rs2744575; 4779C > G and rs4646830; 4854C > G) located in the promoter of *ALDH5A1*, the gene mutated in SSADH, by using a unique tissue repository consisting of hippocampal biopsies of patients with chronic epilepsies undergoing neurosurgery for seizure relief and complemented the approach by in vitro promoter analyses.

## 2 | MATERIALS AND METHODS

### 2.1 | TLE patients, SNP genotyping, and expression analysis

Biopsies of hippocampal tissue from pharmacoresistant TLE patients with hippocampal sclerosis (HS) were included in this study ( $n = 74$ ). Patients underwent surgical treatment in the Epilepsy Surgery Program at the University of Bonn Medical Center. SNP genotyping using Human660W SNP array (Illumina) and gene expression using the HumanHT-12 v3 Expression BeadChip (Illumina) were performed and analyzed as before.<sup>2</sup>

### 2.2 | Plasmids

*ALDH5A1*-GG was made by amplifying a fragment of 566 bp (chr6:24494417–24494982, GRCh38/hg38) located upstream of the *ALDH5A1* start-ATG (Figure 1C) and cloning it into

the pGL3-basic vector (Promega). Subsequent mutagenesis using QuikChange Site-Directed Mutagenesis (Agilent) resulted in *ALDH5A1*-CC.

### 2.3 | Cell cultures, transient transfection, and luciferase reporter assays

NG108-15 cells were maintained and transfected as described previously,<sup>10</sup> using 100 ng of luciferase reporter plasmid (pGL3, *ALDH5A1*-CC, or *ALDH5A1*-GG), 25 ng of Renilla luciferase vector (Promega), and early growth response 1 (Egr1)/Egr1dN as indicated in the single experiments. Luciferase assays were performed using the Dual Luciferase Reporter Assay System (Promega) according to the manufacturer's specifications. *Renilla* and firefly luciferase activities were determined using the Glomax Luminometer (Promega).

### 2.4 | Statistical analyses

Statistical analyses were performed with GraphPad Prism software. Student *t*-test, one-way analysis of variance (ANOVA), and two-way ANOVA followed by multiple comparison tests were used to evaluate the statistical significance of the results. All in vitro experiments were independently repeated at least four times ( $N$ ), each with three technical replicates ( $n$ ). Values were considered significant at  $p < .05$ . Results are plotted as mean  $\pm$  SEM.

## 3 | RESULTS

We found the two SNPs present in the *ALDH5A1* promoter region to be associated with differential *ALDH5A1* expression in the present TLE cohort. Patients with the *ALDH5A1*-CC haplotype (rs2744575; 4779C and rs4646830; 4854C; Figure 1A) showed an augmented hippocampal *ALDH5A1* expression compared to patients with the *ALDH5A1*-GG haplotype (rs2744575; 4779G and rs4646830; 4854G; Figure 1A,B), indicating that the *ALDH5A1* promoter SNPs could function as regulatory (r)SNPs in TLE pathogenesis.

In line with the data available in the linkage disequilibrium database (LDlink database; <https://ldlink.nci.nih.gov/>), we observed for both rSNPs identical genotype and allele frequencies; the homozygous CC variant was observed in 33 (44.6%) patients, the heterozygous CG variant in 31 (41.9%) patients, and the homozygous GG variant in 10 (13.5%) patients of our HS cohort, resulting in an allele frequency of 65.5% and 34.5% for the C-allele and G-allele, respectively. No differences in genotype and allele distribution were observed between our HS cohort and a control group ( $p = .80$ ).<sup>11</sup> Because the two rSNPs were in complete linkage

disequilibrium, we now refer to the two SNPs as one haplotype, consisting of two variants, that is, the *ALDH5A1*-CC haplotype block (both SNPs harbor the C-allele) and the *ALDH5A1*-GG haplotype block (both SNPs harbor the G-allele; Figure 1A).

To examine whether the two *ALDH5A1* haplotypes can lead to differential *ALDH5A1* promoter activity, we tested them in neuronal NG108-15 cells. Using a luciferase reporter assay with *ALDH5A1* reporter fragments harboring both rSNPs (Figure 1C), we found that both haplotypes had a stronger luciferase activity compared to the pGL3 empty control vector, indicating that the 566-bp *ALDH5A1* promoter fragment harbors promoter activity. Furthermore, we observed the *ALDH5A1*-CC haplotype to have a stronger basal promoter activity compared to the *ALDH5A1*-GG haplotype (Figure 1D).

To unravel the underlying molecular mechanisms of the observed differential expression of the two haplotype blocks, we compared the *ALDH5A1*-CC and *ALDH5A1*-GG haplotypes for differential transcription factor (TF) binding affinities. Bioinformatic analysis using GWAS4D revealed the strongest difference in TF binding affinity for the allelic variants of SNP rs2744575 (effect  $p$ -value =  $9.6252E-04$ ). In line with GWAS4D, another bioinformatics tool, EPOSSUM2/Jaspar2018, predicted a higher binding affinity for EGR1 to allelic variant rs2744575-C in comparison to rs2744575-G (76.3% and 40.7%, respectively; Figure 1E). A key role of Egr1 in epileptogenic processes as well as in absence epilepsy has been described before.<sup>10,12</sup>

To determine whether Egr1 differentially regulates the *ALDH5A1*-CC and *ALDH5A1*-GG haplotypes, we transfected an expression vector for Egr1<sup>10</sup> into NG108-15 cells and determined the luciferase activity of the two haplotype blocks. Applying increasing concentrations of Egr1 resulted in significant *ALDH5A1* promoter activation for both haplotypes (Figure 1F), indicating that the *ALDH5A1* promoter is under the control of Egr1. Intriguingly, stepwise increased concentrations of Egr1 exposure were reflected by a significant increase of promoter activity only in the case of the *ALDH5A1*-GG variant, whereas for the *ALDH5A1*-CC haplotype only the maximal Egr1 concentration exposure resulted in significant activation of the promoter (Figure 1F). We extrapolated on the regulation of the *ALDH5A1* promoter through Egr1 by applying a dominant-negative variant of Egr1 (Egr1dN).<sup>12</sup> This modified TF contains the DNA-binding domain, but lacks the activating domain. Upon binding of Egr1dN to the DNA, endogenous Egr1 can no longer bind to the DNA. By this approach, we observed a reduction of promoter activity only for the *ALDH5A1*-CC haplotype at the highest concentrations of Egr1dN, whereas exposure of the *ALDH5A1*-GG haplotype promoter to increasing concentrations of Egr1dN did not exert significant effects with

respect to activity as determined by luciferase reporter assays (Figure 1G).

## 4 | DISCUSSION

Here, we have detected a minor haplotype in the *ALDH5A1* promoter that is associated with reduced expression of the corresponding mRNA in human epileptic hippocampal tissue. In contrast to a previous study on another regulatory motif in the context of *ALDH5A1* mRNA expression,<sup>13</sup> we have now confined the present analysis to hippocampal biopsies with the damage pattern of HS, where all patients had a clinical history of epileptogenesis. Furthermore, we now report on rSNPs located in the functional core promoter of the *ALDH5A1* gene.<sup>14</sup>

Intriguingly, the present haplotype consists of two expression quantitative trait locus SNPs. Of note, there was no significant abundance of the minor haplotype with functional impairment of gene expression in the TLE patient cohort. *ALDH5A1* encodes a protein with critical importance for GABA recycling and thus inhibition. Functionally, impairment of GABA homeostasis has been shown to be a key pathogenetic aspect of TLE.<sup>15,16</sup> Egr1 represents a key TF in epileptogenesis.<sup>12</sup> The present in vitro data suggest that under basal conditions, a higher amount of Egr1 is already bound to the Egr1 binding sites in the *ALDH5A1*-CC variant. This is reflected by a stronger activity of the *ALDH5A1*-CC promoter under basal conditions and lack of activation by only slightly increased concentrations of Egr1 (50 ng). In contrast, the *ALDH5A1*-GG variant has a lower binding affinity for Egr1, that is, 41% for the *ALDH5A1*-GG variant compared to 76% for the *ALDH5A1*-CC variant, and can become activated upon slight Egr1 (50 ng) increases. However, after exposure to maximal concentrations of Egr1, reflecting, for example, post-status epilepticus conditions, the *ALDH5A1* promoter activation status is indistinguishable for both haplotype variants. Thus, under stress conditions, both haplotypes can confer virtually indistinguishably strong promoter activation putatively leading to the production of high amounts of the corresponding *ALDH5A1* protein. However, under basal conditions, the dominant-negative variant for Egr1 can only reduce the *ALDH5A1*-CC promoter activity, which means that (1) under normal in vitro conditions, the *ALDH5A1*-GG variant already lacks binding of Egr1; and (2) the "regulatory range" of varying Egr1 concentrations is larger for the *ALDH5A1*-CC variant. As a functional consequence, the *ALDH5A1*-GG variant may contribute to functionally impaired GABA recycling in TLE.

Furthermore, mutations of *ALDH5A1* have been related to several forms of epilepsy.<sup>9,11,17</sup> However, the clinical manifestation of individual *ALDH5A1* mutations has been

extremely diverse, and only approximately 50% of patients with SSADH deficiency suffer from seizures.<sup>18</sup> Missense mutations considered to be causative of SSADH deficiency have been demonstrated to reduce the SSADH activity to less than 5% of the normal activity in *in vitro* expression systems.<sup>18</sup> Therefore, differences in the activity and the regulatory range of the *ALDH5A1* promoter dependent on the presence of distinct rSNPs in individual patients may either confer a higher propensity for the availability of the mutated *ALDH5A1* protein variant in neurons of patients with SSADH deficiency or induce compensatory effects.

The present data may be paradigmatic for the relevance of distinct SNPs in the regulation of neurotransmitter homeostasis-related gene promoters in human epilepsies, an aspect that appears important to be considered for an improved understanding of the phenotypic diversity of epilepsies.

## ETHICAL PUBLICATION STATEMENT

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

## ACKNOWLEDGMENTS

Our work is supported by Deutsche Forschungsgemeinschaft (SFB 1089 to A.J.B., S.S., K.M.J.v.L.; FOR 2715 to A.J.B.; SCHO 820/7-2, SCHO 820/5-2, SCHO 820/6-1, SCHO 820/4-1, SCHO 820 5-2 to S.S.) and BONFOR. We thank Sabine Optiz for the technical assistance.

## CONFLICT OF INTEREST

None of the authors has any conflict of interest to disclose.

## AUTHOR CONTRIBUTIONS

Despina Tsortouktzidis and Karen M. J. van Loo carried out the molecular biological experiments. Herbert Schulz and Thomas Sander provided genetic analyses. Hartmut Vatter, Motaz Hamed, and Rainer Surges provided key clinical data and revised the manuscript. Susanne Schoch contributed to writing key sections of the manuscript. Karen M. J. van Loo and Albert J. Becker planned and supervised the work, and contributed to writing key sections of the manuscript.

## ORCID

Rainer Surges  <https://orcid.org/0000-0002-3177-8582>  
 Thomas Sander  <https://orcid.org/0000-0002-4011-1798>  
 Albert J. Becker  <https://orcid.org/0000-0003-2661-3705>  
 Karen M. J. van Loo  <https://orcid.org/0000-0003-3074-5612>

## REFERENCES

1. International League Against Epilepsy Consortium on Complex Epilepsies. Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nat Commun.* 2018;9:5269.
2. Schulz H, Ruppert AK, Herms S, Wolf C, Mirza-Schreiber N, Stegle O, et al. Genome-wide mapping of genetic determinants influencing DNA methylation and gene expression in human hippocampus. *Nat Commun.* 2017;8:1511.
3. Srivastava PK, Bagnati M, Delahaye-Duriez A, Ko JH, Rotival M, Langley SR, et al. Genome-wide analysis of differential RNA editing in epilepsy. *Genome Res.* 2017;27:440–50.
4. Barr CL, Misener VL. Decoding the non-coding genome: elucidating genetic risk outside the coding genome. *Genes Brain Behav.* 2016;15:187–204.
5. Greenberg DA, Durner M, Delgado-Escueta AV. Evidence for multiple gene loci in the expression of the common generalized epilepsies. *Neurology.* 1992;42:56–62.
6. McTague A, Howell KB, Cross JH, Kurian MA, Scheffer IE. The genetic landscape of the epileptic encephalopathies of infancy and childhood. *Lancet Neurol.* 2016;15:304–16.
7. Allen NM, Weckhuysen S, Gorman K, King MD, Lerche H. Genetic potassium channel-associated epilepsies: clinical review of the Kv family. *Eur J Paediatr Neurol.* 2020;24:105–16.
8. Pop A, Smith DEC, Kirby T, Walters D, Gibson KM, Mahmoudi S, et al. Functional analysis of thirty-four suspected pathogenic missense variants in *ALDH5A1* gene associated with succinic semialdehyde dehydrogenase deficiency. *Mol Genet Metab.* 2020;130:172–8.
9. DiBacco ML, Pop A, Salomons GS, Hanson E, Roulet J-B, Gibson KM, et al. Novel *ALDH5A1* variants and genotype: phenotype correlation in SSADH deficiency. *Neurology.* 2020;95(19):e2675–82.
10. Van Loo KMJ, Schaub C, Pernhorst K, Yaari Y, Beck H, Schoch S, et al. Transcriptional regulation of T-type calcium channel *CaV3.2*: bi-directionality by early growth response 1 (*Egr1*) and repressor element 1 (*RE-1*) protein-silencing transcription factor (*REST*). *J Biol Chem.* 2012;287:15489–501.
11. Lorenz S, Heils A, Taylor KP, Gehrmann A, Muhle H, Gresch M, et al. Candidate gene analysis of the succinic semialdehyde dehydrogenase gene (*ALDH5A1*) in patients with idiopathic generalized epilepsy and photosensitivity. *Neurosci Lett.* 2006;397:234–9.
12. Van Loo KMJ, Rummel CK, Pitsch J, Müller JA, Bikbaev AF, Martinez-Chavez E, et al. Calcium channel subunit  $\alpha 2\delta 4$  is regulated by early growth response 1 and facilitates epileptogenesis. *J Neurosci.* 2019;39:3175–87.
13. Pernhorst K, Raabe A, Niehusmann P, Van Loo KMJ, Grote A, Hoffmann P, et al. Promoter variants determine  $\gamma$ -aminobutyric acid homeostasis-related gene transcription in human epileptic hippocampi. *J Neuropathol Exp Neurol.* 2011;70:1080–8.
14. Blasi P, Boyd PP, Ledda M, Novelletto A, Gibson KM, Jakobs C, et al. Structure of human succinic semialdehyde dehydrogenase gene: identification of promoter region and alternatively processed isoforms. *Mol Genet Metab.* 2002;76:348–62.
15. Ferando I, Mody I. GABAA receptor modulation by neurosteroids in models of temporal lobe epilepsies. *Epilepsia.* 2012;53(Suppl 9):89–101.
16. Dengler CG, Coulter DA. Normal and epilepsy-associated pathologic function of the dentate gyrus. *Prog Brain Res.* 2016;226:155–78.

17. Malaspina P, Picklo MJ, Jakobs C, Snead OC, Gibson KM. Comparative genomics of aldehyde dehydrogenase 5a1 (succinate semialdehyde dehydrogenase) and accumulation of gamma-hydroxybutyrate associated with its deficiency. *Hum Genomics*. 2009;3:106–20.
18. Akaboshi S, Hogema BM, Novelletto A, Malaspina P, Salomons GS, Maropoulos GD, et al. Mutational spectrum of the succinate semialdehyde dehydrogenase (ALDH5A1) gene and functional analysis of 27 novel disease-causing mutations in patients with SSADH deficiency. *Hum Mutat*. 2003;22:442–50.

**How to cite this article:** Tsortouktzidis D, Schulz H, Hamed M, et al. Gene expression analysis in epileptic hippocampi reveals a promoter haplotype conferring reduced aldehyde dehydrogenase 5a1 expression and responsiveness. *Epilepsia*. 2021;62:e29–e34. <https://doi.org/10.1111/epi.16789>

### 3.2 Publication 2: SCN1A overexpression, associated with a genomic region marked by a risk variant for a common epilepsy, raises seizure susceptibility

Silvennoinen K, Gawel K, **Tsortouktzidis D**, Pitsch J, Alhusaini S, van Loo KMJ, Picardo R, Michalak Z, Pagni S, Martins Custodio H, Mills J, Whelan CD, de Zubicaray GI, McMahon KL, van der Ent W, Kirstein-Smardzewska KJ, Tiraboschi E, Mudge JM, Frankish A, Thom M, Wright MJ, Thompson PM, Schoch S, Becker AJ, Esguerra CV, Sisodiya SM. SCN1A overexpression, associated with a genomic region marked by a risk variant for a common epilepsy, raises seizure susceptibility. **Acta Neuropathol.** 2022 Jul;144(1):107-127. doi: 10.1007/s00401-022-02429-0. Epub 2022 May 12. PMID: 35551471; PMCID: PMC9217876.



## SCN1A overexpression, associated with a genomic region marked by a risk variant for a common epilepsy, raises seizure susceptibility

Katri Silvennoinen<sup>1,2</sup> · Kinga Gawel<sup>3,4</sup> · Despina Tsortouktzidis<sup>5,6</sup> · Julika Pitsch<sup>5,6</sup> · Saud Alhusaini<sup>7,8</sup> · Karen M. J. van Loo<sup>5,9</sup> · Richard Picardo<sup>10</sup> · Zuzanna Michalak<sup>10</sup> · Susanna Pagni<sup>1,2</sup> · Helena Martins Custodio<sup>1,2</sup> · James Mills<sup>1,2</sup> · Christopher D. Whelan<sup>7,11</sup> · Greig I. de Zubicaray<sup>12</sup> · Katie L. McMahon<sup>13</sup> · Wietske van der Ent<sup>3</sup> · Karolina J. Kirstein-Smardzewska<sup>3</sup> · Ettore Tiraboschi<sup>3</sup> · Jonathan M. Mudge<sup>14</sup> · Adam Frankish<sup>14</sup> · Maria Thom<sup>10</sup> · Margaret J. Wright<sup>15</sup> · Paul M. Thompson<sup>11</sup> · Susanne Schoch<sup>5,6</sup> · Albert J. Becker<sup>5</sup> · Camila V. Esguerra<sup>3</sup> · Sanjay M. Sisodiya<sup>1,2</sup>

Received: 16 March 2022 / Revised: 29 April 2022 / Accepted: 30 April 2022 / Published online: 12 May 2022  
© The Author(s) 2022

### Abstract

Mesial temporal lobe epilepsy with hippocampal sclerosis and a history of febrile seizures is associated with common variation at rs7587026, located in the promoter region of *SCN1A*. We sought to explore possible underlying mechanisms. *SCN1A* expression was analysed in hippocampal biopsy specimens of individuals with mesial temporal lobe epilepsy with hippocampal sclerosis who underwent surgical treatment, and hippocampal neuronal cell loss was quantitatively assessed using immunohistochemistry. In healthy individuals, hippocampal volume was measured using MRI. Analyses were performed stratified by rs7587026 type. To study the functional consequences of increased *SCN1A* expression, we generated, using transposon-mediated bacterial artificial chromosome transgenesis, a zebrafish line expressing exogenous *scn1a*, and performed EEG analysis on larval optic tecta at 4 day post-fertilization. Finally, we used an in vitro promoter analysis to study whether the genetic motif containing rs7587026 influences promoter activity. Hippocampal *SCN1A* expression differed by rs7587026 genotype (Kruskal–Wallis test  $P=0.004$ ). Individuals homozygous for the minor allele showed significantly increased expression compared to those homozygous for the major allele (Dunn's test  $P=0.003$ ), and to heterozygotes (Dunn's test  $P=0.035$ ). No statistically significant differences in hippocampal neuronal cell loss were observed between the three genotypes. Among 597 healthy participants, individuals homozygous for the minor allele at rs7587026 displayed significantly reduced mean hippocampal volume compared to major allele homozygotes (Cohen's  $D=-0.28$ ,  $P=0.02$ ), and to heterozygotes (Cohen's  $D=-0.36$ ,  $P=0.009$ ). Compared to wild type, *scn1lab*-overexpressing zebrafish larvae exhibited more frequent spontaneous seizures [one-way ANOVA  $F(4,54)=6.95$  ( $P<0.001$ )]. The number of EEG discharges correlated with the level of *scn1lab* overexpression [one-way ANOVA  $F(4,15)=10.75$  ( $P<0.001$ )]. Finally, we showed that a 50 bp promoter motif containing rs7587026 exerts a strong regulatory role on *SCN1A* expression, though we could not directly link this to rs7587026 itself. Our results develop the mechanistic link between rs7587026 and mesial temporal lobe epilepsy with hippocampal sclerosis and a history of febrile seizures. Furthermore, we propose that quantitative precision may be important when increasing *SCN1A* expression in current strategies aiming to treat seizures in conditions involving *SCN1A* haploinsufficiency, such as Dravet syndrome.

**Keywords** Febrile seizures · Hippocampal sclerosis · Zebrafish · MRI · Genotype

Katri Silvennoinen, Kinga Gawel and Despina Tsortouktzidis have contributed equally to this work. Albert J. Becker, Camila V. Esguerra and Sanjay M. Sisodiya have contributed equally to this work.

✉ Sanjay M. Sisodiya  
s.sisodiya@ucl.ac.uk

Extended author information available on the last page of the article

### Introduction

The epilepsies are a set of conditions defined by the occurrence of seizures, often associated with comorbidities, impaired quality of life and shortened lifespan [56]. In many epilepsies, there is evidence of risk susceptibility imparted by common genetic variation [25, 35], but the mechanistic

link between genetic variation and phenotype is typically obscure [25]. Increasingly, rare variants are being shown to underlie many of the rare epilepsies, especially the early onset conditions called the developmental and epileptic encephalopathies. *SCN1A* is amongst the best studied genes in epilepsy: rare pathogenic variants cause a variety of types of epilepsy, often characterised by seizures provoked by fever, as seen in Dravet syndrome [9]. Variants leading to Dravet syndrome cause loss of function, and are typically de novo in origin [26]. *SCN1A* variants may also cause familial epilepsies, such as genetic epilepsy with febrile seizures plus (GEFS+) [64, 65]. Moreover, we showed an association between common variation in *SCN1A* and the distinct epilepsy syndrome of mesial temporal lobe epilepsy with hippocampal sclerosis with history of febrile seizures (MTLEHS + FS) [29]. Despite its location overlapping with promoter regions of *SCN1A*, we were previously unable to demonstrate an association between the most strongly linked single nucleotide polymorphism (SNP; rs7587026) and expression of *SCN1A* in therapeutically resected human lateral temporal neocortex [29]. We hypothesized that minor allele homozygosity would influence *SCN1A* expression levels in the hippocampus with consequences linked to MTLEHS + FS.

Through a combination of expression data from affected human tissue, an MRI study of the impact of this variant on hippocampal size in healthy humans and functional data from an experimental model, we show that rs7587026 is associated with increased expression of *SCN1A* in human hippocampal tissue, which in turn reduces human hippocampal volume in healthy controls. Increased *Scn1a* expression lowers seizure threshold in an animal model. Finally, using in silico and in vitro promoter analysis, we demonstrate that rs7587026 lies in a highly constrained region and that the rs7587026-containing genomic motif exerts a strong regulatory role on *SCN1A* expression. We thus reveal a possible path between common genetic variation and a common human epilepsy (MTLEHS + FS).

## Materials and methods

### *SCN1A* expression and neuronal loss in hippocampal specimens from individuals with mesial temporal lobe epilepsy with hippocampal sclerosis

We analysed *SCN1A* mRNA expression in all available hippocampal biopsy specimens from 91 patients (44 female; aged 4–65 years) with pharmacoresistant mesial temporal lobe epilepsy with hippocampal sclerosis (MTLEHS) who had undergone surgical treatment at the University of Bonn Medical Centre between 1998 and 2008. Case ascertainment and clinical characteristics have been detailed previously [33,

43, 44]. All procedures were conducted in accordance with the Declaration of Helsinki. This study was approved by the Ethical Commission of University Hospital Bonn (222/16). Informed written consent was obtained from every patient. We adhered to the legal provisions governing the handling of personal data.

For all patients, genotyping was performed from surgical tissue for rs7587026. Typing was also performed for rs922224, another intronic SNP around *SCN1A* previously associated with MTLEHS + FS [29], unlinked to rs7587026 ( $r^2 < 0.2$ ), and rs6432860, an *SCN1A* SNP associated with febrile seizures (FS) only [15]. To examine possible allele-specific differences in expression levels, we compared *SCN1A* mRNA expression in hippocampal biopsy specimens by SNP types.

### Sample preparation and SNP genotyping analysis

Tissue processing and DNA isolation were carried out as described previously [49]. Genotyping of the SNPs was performed using TaqMan SNP™ Genotyping Assays (rs580041: C\_778399\_20; SNP rs922224: C\_8945633\_10; SNP rs7587026: C\_3041377\_10; Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocol on an ABI PRISM 9700HT sequence detection system (PE Applied Biosystems) [44]. The assignments of the alleles of the SNPs refer to the forward genome strand orientation relative to the NCBI reference genome build 36. Allelic discrimination was carried out using the SDS 2.2 software.

### RNA isolation and mRNA expression analysis

Total RNA for gene expression microarray analysis was isolated from human hippocampal tissue samples using All Prep DNA/RNA Mini Kit (Qiagen) according to the manufacturer's protocol. To synthesize cDNA from total RNA and in vitro transcription to biotin-labelled cRNA, Illumina Total Prep-96 RNA Amplification Kit (Life Technologies Corporation, Darmstadt) was used according to manufacturer's protocol. cRNA was then hybridized on Human HT-12v3 Expression Bead Chips using Illumina Direct Hybridization Assay Kit (Life Technologies Corporation, Darmstadt, Germany). The Illumina Bead Array Reader was applied for scanning and the data were analysed using Illumina's Genome Studio Gene Expression Module. Gene expression data were normalised using the Illumina Bead Studio software suite by means of quantile normalisation with background subtraction.

### Analysis of severity of neuronal cell loss in different hippocampal subfields

To assess the possibility that the observed effect of rs7587026 on *SCN1A* expression might be compounded

by differences in severity of hippocampal sclerosis (HS), we used immunohistochemistry for NeuN (encoded by *RBFOX3*) to quantify neuronal loss. Among the 91 individuals included in the *SCN1A* expression analysis, suitable samples for immunohistochemistry were available for 73. Neuronal loss was categorised qualitatively as severe, moderate, mild, or no neuronal loss by experienced neuropathologists as described previously [4].

### Statistical analyses

Statistical analyses were performed using GraphPad Prism 9.00 (San Diego, CA, USA). Differences in *SCN1A* expression and neuronal loss by genotype were assessed using a Kruskal–Wallis test with a post hoc Dunn's test. The significance level was set at  $P < 0.05$ .

### Patterns of neuronal loss in surgical hippocampal samples from individuals with MTLEHS—replication

To replicate and quantify the analysis of neuronal loss, we employed postsurgical samples from therapeutic anterior temporal lobectomies from individuals with MTLEHS archived in the UCL Epilepsy Society Brain and Tissue Bank. The maximum available number of cases with sufficient tissue for diagnosing HS as per ILAE guidelines [5] and accessible genetic data on rs7587026 were chosen (96 in total). The study has ethical approval (UK National Research Ethics 17/SC/0573) and informed consent was obtained from all participants.

For visualisation of the extent of neuronal loss, two sections were immunostained per case, one with NeuN, and the other with MAP2. We compared the total percentage area labelled for each region of interest (ROI)—CA1, CA2, CA3, CA4—across the three genotypes, separately for NeuN and MAP2-stained slides.

### Immunohistochemistry

A representative formalin-fixed and paraffin-embedded tissue block was selected from the hippocampal resection which showed maximal representation of all subfields. Immunohistochemistry for NeuN and MAP2 was performed on 5  $\mu\text{m}$  thick tissue sections on the Bond-MAX Autostainer (Leica, UK). The NeuN antibody was mouse clone A60 (Chemicon/Millipore, UK, diluted 1:2000). The MAP2 antibody was mouse clone HM-2 (Sigma-Aldrich, UK, diluted 1:5000). The sections were dewaxed and rehydrated, and put through heat-induced antigen retrieval (citrate buffer pH 5.9–6.1, Epitope Retrieval Solution 1), at 100 °C for either 20 min (NeuN) or 30 min (MAP2). Hydrogen peroxide block was carried out for 20 min, followed by primary antibody incubation for 15 min. Slides were incubated first with

the blocking agent, post-primary horseradish peroxidase (HRP; Bond Polymer Refine Red Detection, Leica, UK), for 8 min, and then in polymer HRP (Bond Polymer Refine Red Detection, Leica, UK) for further 8 min. For both markers, immunocomplexes were then visualized with 3,3'-Diaminobenzidine (DAB), followed by 5 min incubation with DAB enhancer. Slides were counterstained with haematoxylin, dehydrated and coverslipped. Between each step, all sections were washed with phosphate-buffered saline (Fisher Scientific, Ltd., UK). Negative controls were run simultaneously but without primary antibodies. Slides were digitized using a whole slide scanner (Leica SCN400 scanner, Leica, UK) at  $\times 40$  magnification. Images were acquired using Leica SlidePath Digital Image Hub software (Leica, UK).

### Quantitative analysis

On each digitised slide, four ROIs were manually selected—CA1, CA2, CA3, and CA4. The percentage of light brown and dark brown staining within each selected region was then calculated by the Definiens Developer XD 64 Life software (Definiens AG Munich, Germany), giving the total percentage of staining for each ROI, with higher values implying greater preservation of the respective region. For a subset of 10 samples, the assessment was carried out by two separate researchers, with excellent internal consistency (Cronbach alpha coefficient  $\alpha = 0.994$  for NeuN and  $\alpha = 0.985$  for MAP2).

### Statistical analysis

Statistical analysis was carried out with SPSS (versions 24 for Windows and 25 for Mac, IBM Corp, Armonk, NY). GraphPad Prism 9.00 (San Diego, CA, USA) was used for the graphic representation of the data. In case of any missing values, the individual was excluded for the analysis of that ROI. Data were not normally distributed and Mann–Whitney  $U$  tests were used to compare staining between samples from individuals with rs7587026 genotypes AA and AC, and genotypes AA and CC. A two-tailed value of  $P \leq 0.05$  was considered significant.

### Hippocampal and amygdalar MRI volume measurements in healthy individuals by rs7587026 type

To determine whether rs7587026 minor allele homozygosity influences brain anatomy, we examined, using high-resolution structural brain MR imaging data, the volume of subcortical structures typically associated with MTLEHS (hippocampus, amygdala, and thalamus), and total intracranial volume (ICV) in 597 healthy individuals with available genotype data. All eligible participants were included from the Queensland Twin



Imaging (QTIM) study, a longitudinal study of healthy young twins and their siblings who underwent neuroimaging, genetic, and comprehensive cognitive assessments [66]. The QTIM study was approved by the Human Research Ethics Committees of the University of Queensland and the QIMR Berghofer Medical Research Institute. Informed consent was obtained from each participant. See Supplementary Methods (online resource) for full inclusion criteria. For the present analysis, one member of each set of monozygotic twins was randomly excluded.

### Genotyping

Genome-wide genotype data were collected on the Human610-Quad BeadChip (Illumina, Inc., San Diego, CA), and subjected to standard quality control measures used in large genome-wide association study (GWAS) analyses. Genotype for rs7587026 was extracted for each individual.

### Brain MRI acquisition and processing

High-resolution structural MRI scans were obtained on a single 4-Tesla scanner (Bruker Medspec) using the same imaging protocol. Three-dimensional T1-weighted images were acquired with an inversion recovery rapid gradient echo sequence (TI/TR/TE = 700/1500/3.35 ms; flip angle = 8°; slice thickness = 0.9 mm). MR Images were processed using FreeSurfer (<http://surf.nmr.mgh.harvard.edu/>) [11, 16–18]. Quality control of image segmentation was performed according to the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) Consortium procedures (<http://enigma.ini.usc.edu/protocols/imaging-protocols>) [52].

### Data analysis

Differences in mean hippocampal, amygdalar and other subcortical (thalamus, caudate nucleus, putamen, and globus pallidus) volume measures by genotype were examined using generalized linear regression models (covariates: ICV, age, gender, and zygosity). The mean was calculated from left and right subcortical structural volumes in mm<sup>3</sup>. Differences in participants' age by rs7587026 genotype were examined using an unpaired, two-tailed *t* test. Sex and zygosity differences between groups were tested using a chi-squared test. An alpha-level of 0.05 was used to determine statistical significance. All statistical analyses were performed using the R statistics package (<https://www.r-project.org/>) [55].

### Generation of *scn1lab*-overexpressing zebrafish larvae and experimental setup

To study the functional consequences of increased *SCN1A* expression, we generated, using transposon-mediated

bacterial artificial chromosome (BAC) transgenesis, a zebrafish line expressing exogenous *scn1a* (in zebrafish known as *scn1lab*) (*scn1lab*-OE). We studied the seizure propensity of the *scn1lab*-OE larvae using EEG, exposure to sodium channel blocking antiseizure medications, and exposure to hyperthermia.

### Ethical approval

All experiments were performed in compliance with the European Community Council Directive of November 2010 for Care and Use of Laboratory Animals (Directive 2010/63/EU), and the ARRIVE guidelines. The Norwegian Food Safety Authority via its experimental animal administration's supervisory and application system approved all animal experimentation (FOTS ID 15469 and 23935).

### Construct generation and zebrafish husbandry

The CH211-74H7 clone in a pTARBAC2.1 vector (Source BioScience CHORB736H0774Q) was modified to contain Tol2 elements flanking either side of *scn1lab*. For the over-expression construct, an mCherry reporter was linked to the last exon of *scn1lab* with a self-cleaving T2A sequence to separate the fluorescence reporter and *scn1a* proteins after translation. As a control construct, mCherry was placed upstream of the first exon of *scn1lab*, preventing transcription of *scn1lab*. BAC constructs (20 ng/μl) were injected together with capped tol2 transposase mRNA (50 ng/μl) into the cytoplasm of 1-cell stage fertilized embryos, at a volume of 1.5 nl.

Adult zebrafish of the AB strain were raised under standardized aquaculture conditions, in a 14/10 h light/dark cycle at 28.5 °C. Eggs from natural spawning of adult fish were collected and microinjected with the BAC construct. Next, injected eggs were transferred to petri dishes and raised until 4 day post-fertilization (dpf) in embryo medium (17 mM NaCl, 2 mM KCl, 1.8 mM Ca(NO<sub>3</sub>)<sub>2</sub>, 0.12 mM MgSO<sub>4</sub>, 1.5 mM HEPES buffer pH 7.1–7.3 and 0.6 μM methylene blue; 14/10 h dark/light cycle at 28.5 °C). 3 or 4 dpf larvae were screened using fluorescence stereomicroscopy. Only those larvae which expressed the BAC fluorescent reporter in the brain were used for experiments.

### Morphological assessment

4 dpf larvae were photographed using a Leica M205 FA stereomicroscope and assembled using Adobe Photoshop 2020. All pictures were taken at the same resolution.

### EEG recordings

We performed EEG analysis on larval optic tecta as described by Afrikanova et al. [2]. Epileptiform-like discharges were detected by inserting a glass electrode filled with artificial cerebrospinal fluid (124 mM NaCl, 2 mM KCl, 2 mM MgSO<sub>4</sub>, 2 mM CaCl<sub>2</sub>, 1.25 mM KH<sub>2</sub>PO<sub>4</sub>, 26 mM NaHCO<sub>3</sub>, 10 mM glucose) into the optic tectum of individual 4 dpf zebrafish larvae for 20 min (MultiClamp 700B amplifier, Digidata 1550 digitizer, Axon instruments, USA). The larvae were restrained with the aid of a thin layer of 2% low melting point agarose. The Clampfit version 10.6.2 software (Molecular Devices Corporation, USA) was used for processing the EEG recordings. The data were analysed manually by a highly trained observer, blind to treatment group.

### Drugs

We tested the response of 3 dpf *scn1lab*-OE larvae to three different sodium channels blockers, namely, phenytoin (100 μM; Sigma Aldrich), oxcarbazepine (170 μM; Sigma Aldrich) and valproic acid (100 μM; Sanofi Aventis). All drugs were dissolved in dimethyl sulfoxide (DMSO) at a final concentration of 0.5% v/v DMSO (Sigma Aldrich). The appropriate vehicle control was prepared by dissolving DMSO in zebrafish medium. 3-day-old larvae were screened under fluorescence microscope, and subsequently treated for 20 h in respective drugs. Next, EEG recordings were conducted as detailed above.

### qRT-PCR

To determine whether the level of *scn1lab* overexpression correlates with the number of seizures in *scn1lab*-OE, following EEG recording, each single larva was collected for qRT-PCR analysis. Larvae were transferred to Eppendorf tubes filled with 200 μl Trizol. mRNA was purified from single larvae as described by Dupret et al. [12] Next, cDNA was synthesized using SuperScript™ IV First-Strand Synthesis System (Invitrogen) and amplified using PowerUp™ SYBR™ Green Master Mix (Applied Biosystems) according to manufacturer's instructions. Relative enrichment was computed according to the 2<sup>-ΔΔt</sup> method [38]. Expression levels were normalized against glyceraldehyde 3-phosphate dehydrogenase (*gapdh*), which at the stage of larval development used in this study (i.e., 4 dpf), has been shown to be stably expressed [6, 42].

### Hyperthermia-induced abnormalities

To explore possible effects of a rapid ambient temperature increase on seizure propensity in relation to *scn1lab*

overexpression, 4-day-old *scn1lab*-OE or control larvae were transferred to 50 ml falcon tubes filled with 10 ml of preheated (33, 35 or 37 °C) medium. Next, tubes were bathed in water bath (33, 35 or 37 °C) for 5 min. Subsequently, larvae were monitored for 5 min under a microscope for occurrence of convulsion-like behavior, including repetitive pectoral fin fluttering, lying on one side (loss of posture), tail wagging, or myoclonus-like jerks by a highly experienced researcher, blinded to the group of animals being scored. Larvae were scored for heat-induced convulsive-like phenotype.

### Statistical analysis

For statistical purposes, GraphPad Prism 8.00 (San Diego, CA, USA) was used. For comparisons, one-way or two-way ANOVA with Tukey's or Sidak's multiple comparisons test was used as appropriate. All data are shown as mean ± standard error of the mean (SEM).

### Bioinformatics

We examined the genomic sequence context of rs7587026 using GENCODE [20] release v28. As the GENCODE project is incomplete, we also reviewed transcriptomics data not yet incorporated into the annotation catalogue. These data sets included RNAseq read coverage graphs (from HPA [28, 60]), RNAseq model collections (including FANTOM [24], GTEx [39] and PLAR [23]) and long-read RNA libraries (from GENCODE [34, 58]). We then employed the ReMap resource [8], which seeks to characterise regulatory elements via the large-scale curation and integration of publicly available ChIP-seq data sets (see Supplementary Material in online resource for more details).

### Human sequence constraint

The map of sequence constraint for the human genome created by di Iulio et al. in 2018 was used to identify sequences that are rarely mutated in healthy individuals, intolerant to genetic variation and thus more likely to be functionally relevant [27]. The map, which was produced using whole-genome sequencing (WGS) data from 11,257 individuals, assigns a context-dependent tolerance score (CDTS) to each 10 bp long bin of the genome, indicating the likelihood of variation: the lower the score, the less frequently the bin is affected by variation, and the more mutation intolerant the bin is [27].

### Linkage disequilibrium (LD) blocks

The BigLD function from the gpart R package was used to detect the LD block structure of the *SCN1A* locus. Big-LD is a block-partitioning algorithm that estimates the LD block

organisation of the genetic region of interest using the interval graph modelling of LD bins (clusters of strong pairwise LD SNPs) [31]. The BigLD algorithm employs an agglomerative strategy, which involves detecting small communities of SNPs in strong LD and then merging them if they share SNPs in high LD, ultimately resulting in larger LD blocks than other alternative methods. To account for this feature, the LD block structure of the *SCN1A* locus was also estimated using the command line tool PLINK 1.9 [7, 46].

### Transcription start sites (TSSs)

The functional annotation of the mammalian genome (FANTOM) 5 data repository was used. The FANTOM5 project performed cap analysis gene expression (CAGE) across 975 samples, including human primary cells, tissue samples and cancer cell lines and mapped transcription start sites (TSSs) throughout the genome and their differential usage [1, 19]. Using the TSS peaks identified by the FANTOM5 project, the genetic location of TSSs in the *SCN1A* locus was explored. The FANTOM5 project also measured the relative activity of each TSS as normalized tags per million (TPM), calculated using the relative log expression (RLE) method in edgeR [1, 19, 47]. The relative activity of the TSSs in the *SCN1A* locus was compared across hippocampus and cortex samples.

### In vitro regulatory analysis

To investigate whether the *SCN1A* region comprising the SNP rs7587026 area possesses regulatory activity can confer promoter activity and, if so, whether this promoter activity is differentially affected by genotype, we cloned a 50 bp fragment surrounding the SNP into a luciferase reporter vector. This fragment was chosen as according to the UCSC Genome Browser, it has a high conservation score among 100 vertebrates [30]. In our experimental experience, this feature in particular provides a strong argument for biologically relevant genomic motifs [40, 63]. The conservation score broke down markedly up- and downstream of this 50 bp fragment [30]. Furthermore, two large repeats flanking the SNP region prevented extension of the luciferase fragment: LIM4, a long interspersed nuclear element (LINE) located at 300 bp from the SNP on the 5' side, and ERVL-MaLR, a long terminal repeat (LTR) located at 90 bp from the SNP on the 3'.

The fragment was cloned in the sense direction respective to the ATG-start codon; therefore, in this section, the major allele or wild type allele (C) is denoted as G and the minor allele (A) as T. Luciferase constructs are denoted as *SCN1A*-50 bp-rs7587026-G(WT)-Luciferase and

*SCN1A*-50 bp-rs7587026-T-Luciferase. Bioinformatic analyses had revealed a consensus binding motif for the transcription factor Sox2 which contained the SNP. Therefore, we generated two additional luciferase reporter constructs, one that comprised only the Sox2 binding site for each genotype (*SCN1A*-20 bp-rs7587026-G(WT)-Luciferase and *SCN1A*-20 bp-rs7587026-T-Luciferase) and one in which the Sox2 binding site had been replaced by a scrambled sequence.

### Ethical approval

Approval was obtained from the Ethical Commission of University Hospital Bonn (196/17).

### Cloning

Luciferase constructs containing a 50 bp fragment surrounding rs7587026 were generated by in-fusion cloning. Fragments were PCR amplified from human genomic DNA isolated from blood either homozygous for the major (wild type) allele G or the minor allele T (see primers in Supplementary Table 1, online resource). The PCR product was cloned in the *NheI/BglIII* sites of the pGL3-basic vector (Promega) generating the constructs *SCN1A*-50 bp-rs7587026-G(WT)-Luciferase and *SCN1A*-50 bp-rs7587026-T-Luciferase. To generate the constructs *SCN1A*-50 bp-Scramble3-Luciferase, *SCN1A*-20 bp-rs7587026-G(WT)-Luciferase and *SCN1A*-20 bp-rs7587026-T-Luciferase oligonucleotides (see oligonucleotide sequence in Supplementary Table 1, online resource) were first annealed and inserted in the *NheI/BglIII* sites of the pGL3-basic vector (Promega).

### Cell culture, transfection and luciferase assay

NS20Y cells (Sigma, 08062517) were cultured in DMEM (Sigma, D6546) supplemented with 10% (v/v) heat inactivated FBS, 2 mM L-Glutamine, 100 units/ml penicillin/streptomycin and kept at 37 °C and 5% CO<sub>2</sub>. Cells were plated in 24-well plates and after 24 h transfected using Lipofectamine (Invitrogen) following the manufacturer's instructions. Transfection was performed using 100 ng of the luciferase constructs and 25 ng of the Renilla luciferase control construct (Promega) and collected 48 h after transfection. Luciferase assay were performed as described before [59].

HEK293T cells were cultured in DMEM (Invitrogen) supplemented with 10% FCS and 1% penicillin–streptomycin and kept at 37 °C and 5% CO<sub>2</sub>. Cells were plated

in 10 cm<sup>2</sup> culture dishes and transfected using calcium phosphate with 10 µg of the constructs: pCMV-Sox2-T2A-GFP (Addgene plasmid #127537) or hypB-CAG-2A-eGFP (derived from a PiggyBac backbone) used as a control. Proteins were isolated 48 h after transfection.

#### Protein extraction and electrophoretic mobility shift assay (EMSA)

Nuclear protein extracts prepared either from mouse brain or from HEK293T cells transfected with pCMV-Sox2-T2A-GFP (Addgene#127537) was performed using the NE-PER kit (Thermo Fisher Scientific, #78833) following the manufacturer's instruction. Adult male mice (~50 days, > 20 g) were obtained from Charles River (C57Bl/6-N). Animals were decapitated under deep isoflurane (Forene) anaesthesia. Brains were prepared, the cerebellum was removed and the right hemisphere used for nuclear protein isolation.

The oligonucleotides for EMSA (Supplementary Table 2, online resource) were annealed in annealing buffer (10 mM Tris pH 7.5, 50 mM NaCl, 1 mM EDTA) by incubation for 2 min at 95 °C followed by 1 h of slowly cooling down. The EMSA reaction was performed using the commercially available EMSA kit (Invitrogen, #E33075) following the manufacturer's instructions. Briefly, 0.5 µg of annealed oligonucleotides and 5 µg of nuclear protein extract were mixed in 1X binding buffer and incubated at 30 °C for 30 min. Following incubation, samples were mixed with 6X EMSA loading solution. The reaction was separated in 6% non-denaturing polyacrylamide gels. Gels were stained with SYBR Green for 20 min and imaged.

#### Statistical analysis

Data for luciferase activity are shown as mean ± standard error of the mean (SEM). Activity in different conditions was compared using one-way ANOVA, with Tukey's test for multiple comparisons, or one-sample *t* test as appropriate.

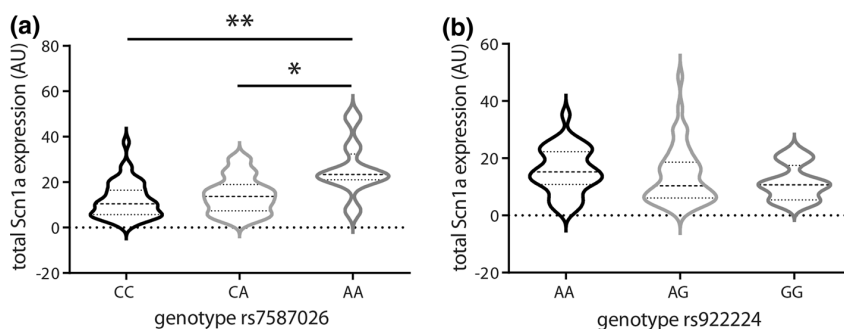
#### Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request and subject to local requirements for each source data set.

## Results

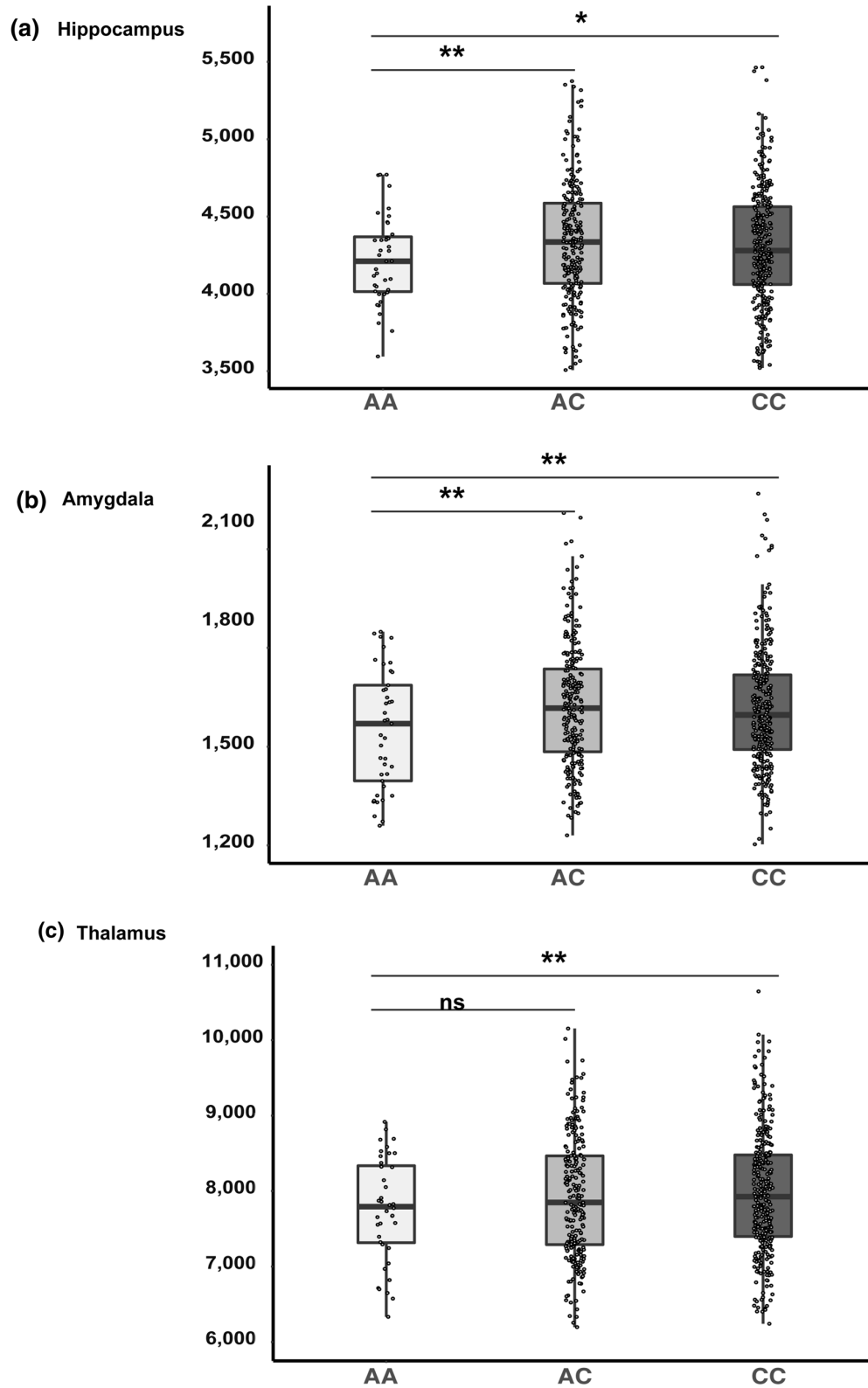
### *SCN1A* mRNA expression levels in hippocampi of patients with pharmacoresistant MTLEHS differ by genotype

Hippocampal *SCN1A* expression differed by rs7587026 genotype (Kruskal–Wallis test  $H = 10.92$ ,  $P = 0.004$ ). Individuals homozygous for the minor allele (AA;  $n = 8$ ) showed significantly increased expression compared to those homozygous for the major allele (CC;  $n = 43$ ; Dunn's test:  $P = 0.003$ ; Fig. 1a), and to heterozygotes (AC;  $n = 40$ ; Dunn's test:  $P = 0.035$ ). Neither rs922224 nor rs6432860 genotypes showed association with hippocampal *SCN1A* expression (Kruskal–Wallis test for rs922224:  $H = 2.98$ ,  $P = 0.226$ ; Fig. 1b; for rs6432860, see Supplementary Fig. 1, online resource).



**Fig. 1** rs7587026 genotype is correlated with *SCN1A* expression in individuals with MTLEHS, whereas rs922224 genotype is not. **a, b** Violin plots showing the expression levels of *SCN1A* in hippocampi of individuals with MTLEHS stratified by SNP genotype. Horizontal lines within the plots present median and quartiles. **a** Individuals with MTLEHS homozygous for the minor allele (rs7587026) have higher *SCN1A* expression levels compared to the other two genotypes

(CC:  $n = 43$ , CA:  $n = 40$ , AA:  $n = 8$ ; Kruskal–Wallis test:  $**P = 0.004$ , Dunn's test: CC vs. AA:  $**P = 0.003$ , CA vs. AA:  $*P = 0.035$ ). **b** No differences in *SCN1A* expression levels were observed by rs922224 genotype in MTLEHS (AA:  $n = 33$ , AG:  $n = 46$ , GG:  $n = 12$ ; Kruskal–Wallis test:  $P = 0.226$ ). See Supplementary Table 3 (online resource) for the expression level values



**Fig. 2** Subcortical volumes by genotype. The mean volumes of the **a** hippocampus, **b** amygdala and **c** thalamus are presented based on rs7587026 genotype in healthy young individuals. The mean volume was calculated from left and right hemisphere structural volumes in mm<sup>3</sup>. In each figure, the x-axis displays three groups of the QTIM sample based on rs7587026 genotype: minor allele homozygotes (AA;  $n=41$ ), minor allele heterozygotes (AC;  $n=242$ ), and major allele homozygotes (CC;  $n=314$ ). The y-axis displays the structural volume in mm<sup>3</sup>. Error bars represent standard error (SE) of the means. \* $P<0.05$ ; \*\* $P<0.01$ ; *ns* non-significant

### rs7587026 type is not associated with severity of hippocampal cell loss in MTLHS

No significant correlation was identified between allelic variants and degree of hippocampal cell loss in any analysed subregions (CA1–CA4; Supplementary Fig. 2, online resource) for the samples in which expression was studied (detailed in 1 above).

We examined this question in a second independent cohort also. The cases comprised nine rs7587026 minor allele homozygotes (AA), 35 heterozygotes (CA), and 52 major allele homozygotes (CC). For NeuN, median percentage staining ranged from 0.49 (AC in CA4) to 4.36 (AA in CA2). For MAP2, median percentage staining ranged from 31.1 (AA in CA1) to 87.2 (AA in CA2). No statistically significant differences in degree of staining by genotype were observed for any of the ROIs using either immunolabel (Supplementary Fig. 3 and Supplementary Tables 4 and 5, online resource).

### Hippocampal and amygdalar volumes differ by rs7587026 genotype as measured by MRI in healthy individuals

Among the 597 individuals, 236 were male and 361 female, with a mean age of 23.5 years (standard deviation  $\pm 3.1$ ; Supplementary Table 6, online resource). Hippocampal, amygdalar, and thalamic volumes are presented in Fig. 2 and Table 1. Minor allele homozygotes (AA;  $n=41$ ) displayed significantly reduced mean hippocampal volume compared to major allele homozygotes (CC;  $n=314$ ; Cohen's  $D=-0.28$ ,  $P=0.02$ ), and to heterozygotes (AC;  $n=242$ , Cohen's  $D=-0.36$ ,  $P=0.009$ ; Fig. 2a). AA homozygotes also displayed reduced mean amygdalar volume relative to CC homozygotes (Cohen's  $D=-0.35$ ,  $P=0.01$ ) and heterozygotes (Cohen's  $D=-0.39$ ,  $P=0.004$ ) (Fig. 2b). Similarly, AA homozygotes showed significantly reduced mean thalamic volume relative to CC homozygotes (Cohen's  $D=-0.29$ ,  $P=0.009$ ) and a trend of reduced thalamic volume when compared to heterozygotes (Cohen's  $D=-0.21$ ,  $P=0.07$ ; Fig. 2c). Other subcortical volumes, including the

caudate nucleus, putamen, and globus pallidus, did not differ by rs7587026 genotype (Table 1).

### scn1a-overexpressing zebrafish larvae exhibit spontaneous seizures and are more prone to heat-induced convulsions

Morphologically, *scn1lab*-overexpressing (hereafter, *scn1lab*-OE) larvae did not exhibit any overt malformations (Fig. 3a), though a subset were slightly hyperpigmented and occasionally lacked a swim bladder. Most (16/19) *scn1lab*-OE larvae exhibited spontaneous seizures in the form of high-voltage spikes, spike-wave complexes and polyspike-wave discharges, while only 2/8 control larvae displayed only a single seizure [one-way ANOVA  $F(4,54)=6.95$  ( $P<0.001$ ); Fig. 3b, c]. As predicted, oxcarbazepine ( $P<0.01$ ) and valproic acid ( $P<0.05$ ) led to a decreased number of EEG discharges in *scn1lab*-OE larvae. Interestingly, phenytoin did not exert antiseizure activity in *scn1lab*-OE larvae (Fig. 3b).

The number of EEG discharges in *scn1lab*-OE larvae differed by the level of *scn1lab* overexpression [one-way ANOVA  $F(4,15)=10.75$  ( $P<0.001$ ); Fig. 3d] and was remarkably dose-sensitive. Even a modest rise in *scn1lab* transcript levels of 11% was sufficient to increase seizures to between 3 and 5 per 20 min recording ( $P<0.05$ ), while 24% and 31% increases resulted in 6 to 8 ( $P<0.01$ ) and 9 to 12 seizures ( $P<0.001$ ) in *scn1lab*-OE larvae, respectively.

To explore possible effects of a rapid ambient temperature increase on seizure propensity in relation to *scn1lab* overexpression, we bathed control and *scn1lab*-OE larvae (normally reared at 28.5 °C) in embryo medium pre-heated to 33, 35 or 37 °C for 5 min. We then assessed larvae for convulsion-like behavior (i.e., loss of posture, excess tail beating, pectoral fin fluttering). Here, two-way ANOVA with Sidak's multiple comparisons test revealed a difference between tested groups of animals [group of animals  $F(1,6)=88.33$ ,  $P<0.001$ ; temperature  $F(2,6)=39.41$ ,  $P<0.001$ ; group of animals  $\times$  temperature interaction  $F(2,6)=8.87$ ,  $P<0.05$ ; Fig. 3e]. Statistically significant differences in the number of fish showing convulsive-like behaviors were observed between control and *scn1lab*-OE larvae on exposure to 35 °C ( $P<0.01$ ) and 37 °C ( $P<0.001$ ), but not 33 °C (Fig. 3e).

### rs7587026: bioinformatic analysis

rs7587026 does not overlap with any exonic sequences found in GENCODE [20] release v28, being instead found within intronic sequence of the *SCN1A* 5' untranslated region (as well as intronic sequence of lncRNA ENSG00000236107 on

**Table 1** Mean volumes of subcortical structures in healthy participants according to rs7587026 type

	AA Mean (SD)	AC Mean (SD)	CC Mean (SD)	AA vs. AC		AA vs. CC	
				Cohen's D	P value	Cohen's D	P value
Hippocampus	4198 (308)	4333 (389)	4307 (402)	− 0.36	0.009	− 0.28	0.02
Amygdala	1552 (180)	1626 (186)	1616 (187)	− 0.39	0.004	− 0.35	0.01
Thalamus	7756 (683)	7925 (805)	7985 (834)	− 0.21	0.07	− 0.29	0.009
Caudate nucleus	3790 (499)	3866 (505)	3915 (494)	− 0.15	0.24	− 0.25	0.07
Putamen	5695 (629)	5716 (663)	5740 (684)	− 0.03	0.78	− 0.07	0.60
Globus Pallidum	1399 (174)	1433 (192)	1445 (203)	− 0.17	0.22	− 0.22	0.09

Volumes are expressed in mm<sup>3</sup>

the opposite strand). Our extensive investigations of other data sets including RNAseq read coverage graphs (from HPA [28, 60]), RNAseq model collections (including FANTOM, [24] GTEX [39] and PLAR [23]) and long-read RNA libraries (from GENCODE [34, 58]) did not suggest that rs7587026 is exonic. Using the ReMap resource [8], in silico analysis showed that the variant falls within DNA binding regions of the transcription factor SOX2 but also of other molecules including AR proteins (see also Supplementary Material in online resource).

According to both the BigLD and PLINK estimates of LD blocks, rs7587026 is located between the P1b and P1c *SCN1A* promoters, in the same LD block as P1b (Fig. 4). In terms of the level of sequence constraint, rs7587026 falls within a broader genetic region characterised by a negative CDTS score, indicating a highly constrained region of the human genome, that is infrequently mutated in healthy individuals (Fig. 4). In the di Iulio et al. map [27], the CDTS score for the 10 bp bin upstream of rs7587026 (chr2:166,122,230–166,122,240) was −2.3445, and the score for the 10 bp bin downstream of rs7587026 (chr2:166,122,240–166,122,250) was −2.3232.

The FANTOM5 CAGE data set identified seven transcription start sites (TSS) located upstream of the *SCN1A* gene body: three mapped to the P1a promoter, three to the P1b promoter, and one closer to rs7587026 (distance: 981 bp) (Fig. 5). The six TSSs falling in the *SCN1A* promoters were recognised as *SCN1A*-related TSSs by the FANTOM5 project. The TSS located 981 bp from rs7587026 was not identified as a TSS for *SCN1A* or any other gene.

Considering the relative activity of the TSSs across the hippocampus and brain cortex, the TSS closest to rs7587026 was the least active in both tissues, accounting for a mean of 0.32 TPM in the hippocampus and 0.26 in the cortical samples, indicating that for every 1,000,000 CAGE tags in the CAGE library, an average of 0.32 in hippocampus and 0.26 in the cortex originated from this TSS. In both hippocampus and cortex, TSS:166128014, in P1b, was the most active,

accounting for a mean of 37.61 TPM in the hippocampal samples and 81.11 in the cortical samples. The second most active TSS was TSS:166149160, which accounted for a mean of 7.48 TPM in hippocampus and 79.04 in the cortex (Fig. 6).

### rs7587026-containing genomic motif regulates *SCN1A* expression

After transfection of NS20Y cells with the luciferase constructs (Fig. 7a), we observed a dramatic increase of luciferase activity in the 50 bp fragments for both genotypes (G-major allele and T-minor allele) when compared to the pGL3 basic empty vector (Fig. 7b), showing that this fragment indeed contains transcription factor binding sites and can fundamentally activate gene transcription. However, we did not detect a difference in promoter activity between the two genotypes.

Luciferase activity measurements of NS20Y cells transfected with the three additional reporter constructs showed that the promoter activity was strongly reduced if the Sox2 binding site was destroyed (Scr3, Fig. 7b), and even completely abolished if the flanking sequences were deleted (20 bp fragment; Fig. 7b), suggesting that Sox2 plays a role in mediating transcriptional activation, but is not solely sufficient.

To further examine if Sox2 on its own has the potential to activate the transcriptional activity of the 50 bp genomic *SCN1A* fragment, we co-transfected the reporter plasmids under transcriptional control of the 50 bp fragments (G-major allele and T-minor allele) together with an expression plasmid for Sox2 into NS20Y cells. We found that Sox2 increased luciferase activity when compared to the untreated condition (Fig. 7c; fold increase of  $4.255 \pm 0.5385$ ,  $P = 0.0042$  and  $5.233 \pm 0.7104$ ,  $P = 0.0052$  for *SCN1A*-50 bp-rs7587026-G(WT)-Luciferase and *SCN1A*-50 bp-rs7587026-T-Luciferase, respectively,  $n = 4$ ,

one sample *t* test). We did not observe a significant difference in activation between the two genotypes tested. Next, we investigated if Sox2 directly binds to the genomic *SCN1A* fragments and if there is a difference in binding efficiency between the two genotypes. To this end, we performed electrophoretic mobility shift assays (EMSA) with the 50 bp genomic *SCN1A* fragments and nuclear protein extracts derived from HEK293T cells transfected with CAG-eGFP (control) or pCMV-Sox2-T2A-GFP. Interestingly, we observed three bands corresponding to DNA–protein complexes between the 50 bp fragments of both genotypes (G-major allele and T-minor allele), which could only be detected if Sox2 was expressed (Fig. 7d). One of these DNA–protein complexes remained even if the site surrounding the SNP was mutated (scrambled condition, scr), indicating that Sox2 can also bind to the flanking sequence. When the 20 bp genomic *SCN1A* fragments were subjected to an EMSA with nuclear extract from HEK293 cells expressing Sox2, only one band remained (Fig. 7e). Together, these results suggest that the sequence surrounding the SNP can bind to Sox2, as suggested by bioinformatic data given above; there does not appear to be a significant difference in binding efficiency between the genotypes. However, our findings also indicate that Sox2 can also bind to sites outside of this region. Finally, we examined if the genomic 50 bp *SCN1A* fragment containing the SNP binds to proteins present in nuclear extracts prepared from mouse brain (Fig. 7f). EMSA showed that multiple defined complexes formed for both genotypes. The pattern for the scrambled oligonucleotide strongly differed from the observed native band patterns, suggesting that the region around the SNP defines which transcription factors bind to this sequence. Additional bioinformatic analyses using the JASPAR database with a stringent cut off of 90% predicted that additional transcription factors (Supplementary Table 8, online resource) could bind to the 50 bp-*SCN1A* fragment both within the area surrounding the SNP (e.g., *FOXLI* and *RFX7*) as well as the flanking sequences (e.g., *GATA2*).

## Discussion

MTLEHS is a drug-resistant epilepsy syndrome of unknown—and likely multifactorial—causation [51, 53, 57]. A significant proportion of patients with MTLEHS have a history of FS [21, 45]. We previously reported a genetic association between MTLEHS + FS and rs7587026 [29]. The current findings suggest that the association is mediated by the consequences of regionally increased *SCN1A* expression related to the risk variant.

Between incident FS and onset of habitual seizures in MTLEHS + FS, there is a ‘latent period’, during which

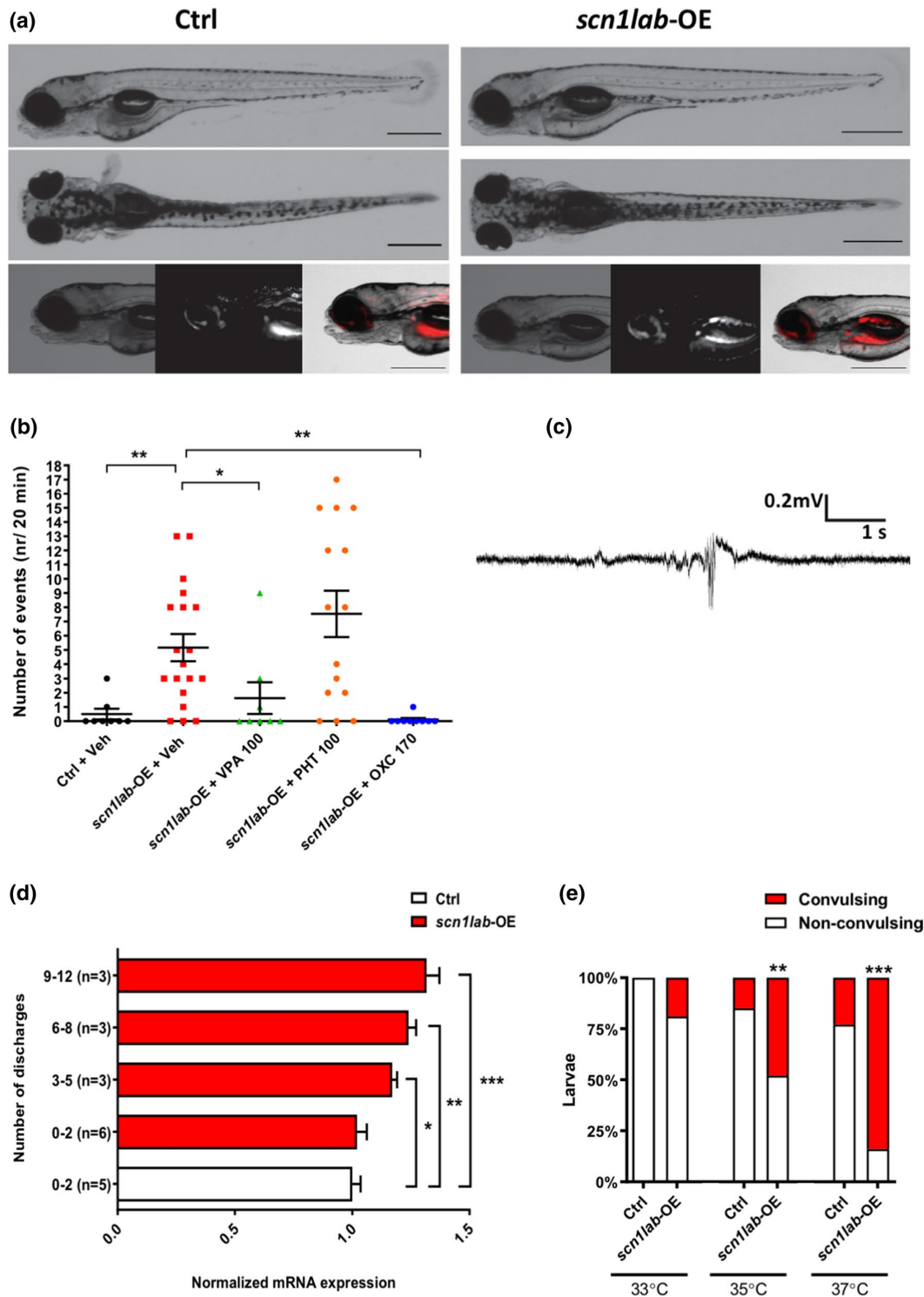
epileptogenesis is believed to occur [21, 62], as supported by data from animal models [36]. The period represents a potential window of opportunity to interfere with processes of epileptogenesis and neurodegeneration [41, 54]; it is a major area of research interest in developing anti-epileptogenesis strategies [41]. In humans, intervention would require both identification of which children are at risk of developing MTLEHS after FS and an effective anti-epileptogenic therapeutic strategy based on a mechanistic understanding.

The possibility that, in people who go on to develop MTLEHS after FS, the hippocampus may have been ‘vulnerable’ to FS/febrile status-induced damage has long been debated [3, 50]. If correct, demonstration of pre-FS hippocampal abnormalities might identify children at risk of developing MTLEHS after FS. There are no published MRI studies showing underlying pre-FS structural brain abnormality in children who have FS and then go on to have MTLEHS. Acute MRI post-febrile seizures, particularly post-febrile status, has shown hippocampal swelling and T2 hyperintensity that may in some cases proceed to the appearances of HS [37]. Underlying hippocampal ‘vulnerability’ is supported by the FEBSTAT study: compared to children with simple FS, children with febrile status showed, at baseline after febrile seizures, reduced hippocampal volumes and reversed right/left hippocampal volume ratios, even in the absence of signs of acute hippocampal damage [37]. However, as measuring pre-FS hippocampal volumes in healthy children is not feasible, an alternative strategy is needed to identify children who may be at risk following FS.

Genetics offers one possible route to identifying at risk individuals. FS are amongst the most heritable type of seizure [32]. Feenstra et al. identified a total of six SNPs associated with FS after measles, mumps, and rubella (MMR) vaccination (*IFI44L* rs273259; *CD46* rs1318653), or non-MMR linked FS in general (*SCN1A* rs6432860; *SCN2A* rs3769955; *ANO3* rs114444506; 12q21.33 rs1110546), but they also provided evidence that the six SNPs were not associated with post-FS epilepsy [15]. We previously showed that common variation in *SCN1A* at rs7587026 is associated specifically with MTLEHS + FS [29]. rs7587026 was not associated with FS in general—either in our study, or in the Feenstra study [15]—suggesting that there are genetically distinct susceptibilities to FS in general and to FS that are specifically associated with subsequent epilepsy.

It is not feasible to conduct large-scale MRI studies in healthy children to directly address the hypothesis that in children who experience FS and go on to develop MTLEHS, the hippocampus is structurally vulnerable to FS-induced injury. In adults with epilepsies, secondary processes, such as seizure-related neurodegeneration, may be confounding factors; for example subcortical volume differences seen on





MRI in mesial temporal lobe epilepsy are correlated with disease duration [61]. Therefore, to determine whether rs7587026, with its demonstrated consequences on *SCN1A*

expression in disease, affects brain structure, we examined healthy humans using the QTIM cohort. We show that rs7587026 is robustly associated with smaller hippocampal

**Fig. 3** Overexpression of *scn1a* in zebrafish increases seizure susceptibility and temperature sensitivity. **a** Representative images of control (Ctrl; left panel) and *scn1lab*-OE larvae (right panel). Red colour depicts mCherry fluorescence. **b** Number of EEG discharges in control and *scn1lab*-OE larvae. Horizontal bars represent mean  $\pm$  SEM. Sample sizes: Ctrl,  $n=8$ ; *scn1lab*-OE+Veh,  $n=19$ ; *scn1lab*-OE+valproic acid (VPA) (100  $\mu$ M),  $n=8$ ; *scn1lab*-OE+phenytoin (PHT) (100  $\mu$ M),  $n=15$ ; *scn1lab*-OE+oxcarbazepine (OXC) (170  $\mu$ M),  $n=9$ .  $**P<0.01$ ,  $*P<0.05$  (One-way ANOVA with Tukey's *post-hoc*). Abbreviations: VPA-valproic acid, PHT-phenytoin, OXC- oxcarbazepine. **c** An example of EEG recordings obtained from the optic tecta of 4 dpf *scn1lab*-OE larvae. **d** Number of discharges recorded in *scn1lab*-OE larvae relative to *scn1lab* mRNA levels. Vertical bars represent mean  $\pm$  SEM. **e** Percentage of larvae developing heat-induced seizures. Sample sizes: Ctrl,  $n=18$ ,  $n=20$  or  $n=22$  (for 33, 35 or 37  $^{\circ}$ C, respectively); *scn1lab*-OE,  $n=16$ ,  $n=19$  or  $n=18$  (for 33, 35 or 37  $^{\circ}$ C, respectively).  $***P<0.001$ ,  $**P<0.01$  (Two-way ANOVA with Sidak's multiple comparisons test)

volume, in support of the idea that individuals at risk of MTLEHS after FS may have hippocampi with pre-existing vulnerability [3, 50]. Amygdala and thalamic volumes are also reduced in healthy minor allele homozygotes, suggesting that structures other than the hippocampus may also have pre-existing vulnerability upon which disease may act. We note that age does not have a differential effect on these subcortical volumes by minor allele homozygosity (Supplementary Fig. 4 and Supplementary Table 7, online resource).

We show that homozygosity for the risk variant has further functional consequences that may underpin disease susceptibility. In the human hippocampus, rs7587026 is associated with increased *SCN1A* expression (which is approximately doubled). The effect is specific for this SNP, and is not seen with unlinked common variants in *SCN1A*, including the *SCN1A* SNP (rs6432860) associated with FS in general identified by Feenstra et al. [15]. Addressing the possibility that this observed consequence is an artefact of HS severity affecting the number of cells available to express *SCN1A*, we show that rs7587026 is not associated with severity of HS/cell loss in resected sclerosed hippocampi from people with MTLEHS.

An alternative to the idea that FS leading to MTLEHS are a marker of underlying hippocampal vulnerability is that they might directly cause hippocampal injury and epileptogenesis. In a murine model for pleiotropic *SCN1A*-related epilepsy (*Scn1a*<sup>RH $\pm$</sup> ), both prolonged [13] and short [48] exposure to hyperthermia-induced seizures precipitated development of a more severe phenotype than in those without exposure, or in exposed wildtype mice [48], suggesting that febrile seizures and *Scn1a* haploinsufficiency interact to produce a severe phenotype [48]. MTLEHS + FS involves an analogous possible precipitating event. As opposed to haploinsufficiency, however, we show that there is an association between rs7587026 genotype and *SCN1A* expression levels. Contrasting under- and over-expression directly is likely too simplistic and further understanding of regional

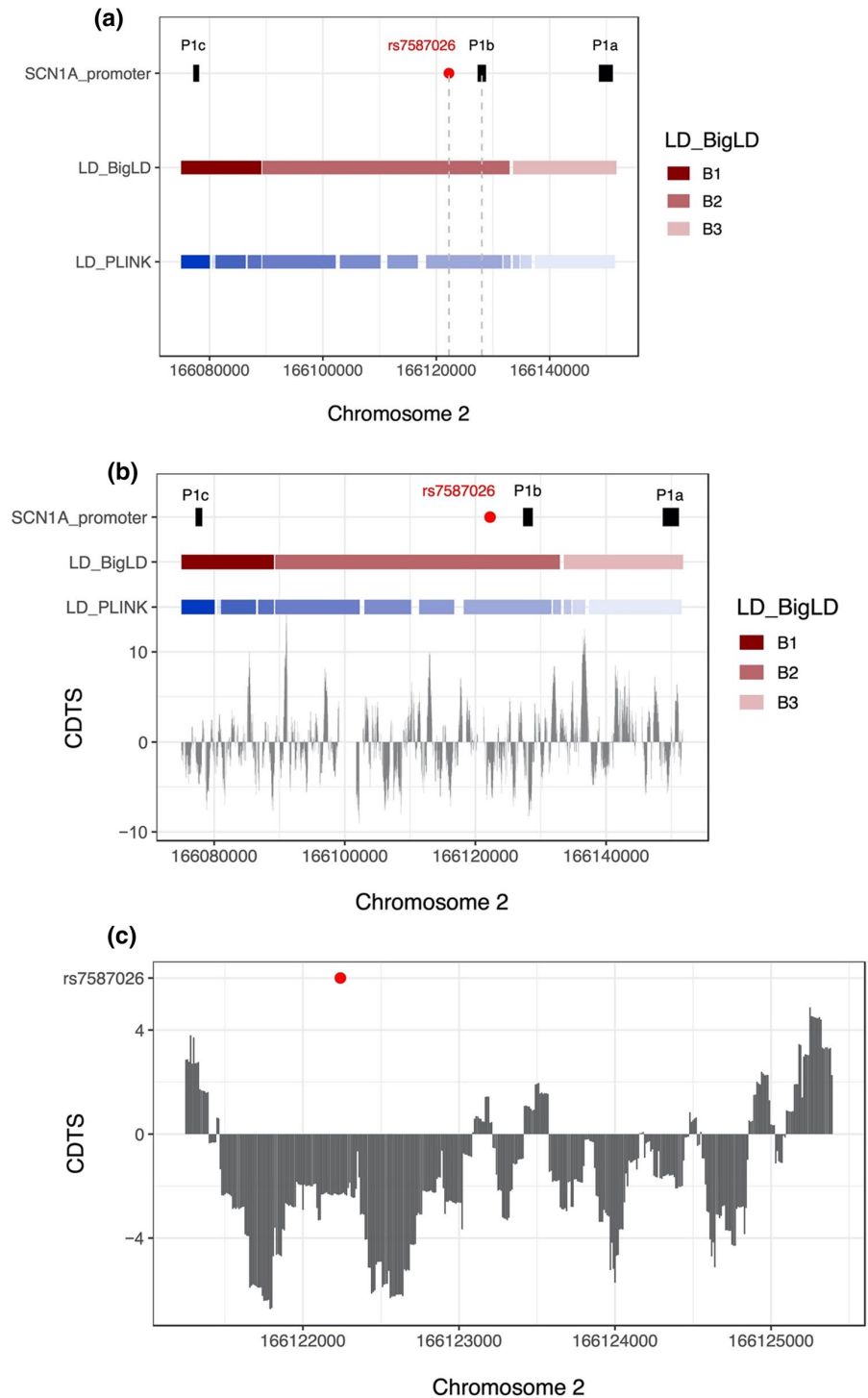
and temporal patterns of expression, at cellular levels during development, and how these might affect neurophysiological properties at a network level, is required. For example, common genetic variants associated with brain surface area identified in a GWAS meta-analysis were enriched for regulatory elements active in the mid-foetal period [22]. In any case, based on the localised overexpression of *SCN1A* in the epileptogenic foci in adult brain, and smaller hippocampal volumes in healthy adult brain, we hypothesised that overexpression of *SCN1A* is causally related to increased seizure susceptibility.

To test this hypothesis, we generated a new zebrafish *scn1lab* overexpression model. Most (84%) *scn1lab*-OE larvae exhibited spontaneous seizures as measured by EEG. Moreover, *scn1lab*-OE zebrafish larvae had heightened incidence of temperature-induced convulsions. As expected, sodium channel blockers oxcarbazepine and valproic acid successfully repressed electric seizure activity; that the sodium-channel blocker phenytoin did not may be a species-dependent phenomenon [2], perhaps due to lack of absorption and/or rapid metabolism and excretion in larval zebrafish. Overall, our results provide functional evidence for a thermosensitive seizure-promoting effect of *scn1a* overexpression.

We were unable to determine a definite route for the association of rs7587026 on *SCN1A* transcription, leading us to seek alternative explanations for its prospective functional relevance. Our in silico analysis did not suggest a protein-coding function for the SNP. According to the most up to date *SCN1A* annotation, rs7587026 is located in an intronic region between the P1b and P1c *SCN1A* promoters. Using the map of sequence constraint as an indicator of the likelihood of functionality, rs7587026 lies within a highly constrained region of the human genome, suggesting that this region might be functionally relevant. The FANTOM5 CAGE data set also identified, close to rs7587026, a TSS peak not identified as a TTT for *SCN1A* or any other gene. While the lack of RNAseq supported introns in the region suggests that transcription here does not generate an alternative transcript of *SCN1A*, this signal could indicate an enhancer. Together, these observations led us to hypothesise that rs7587026 falls within a regulatory element that may operate in brain development; such epigenetic regulation may also explain regional specificity of the association between rs7587026 genotype and *SCN1A* expression [20].

To test this hypothesis, we performed several complementary in vitro analyses. We confirmed that a short genomic fragment in the *SCN1A* gene surrounding rs7587026 has intriguingly strong transcriptional activity and is able to bind, as predicted by the bioinformatic analysis (see Supplementary Material in online resource), to the transcription factor SOX2; the short genomic motif, furthermore, can be

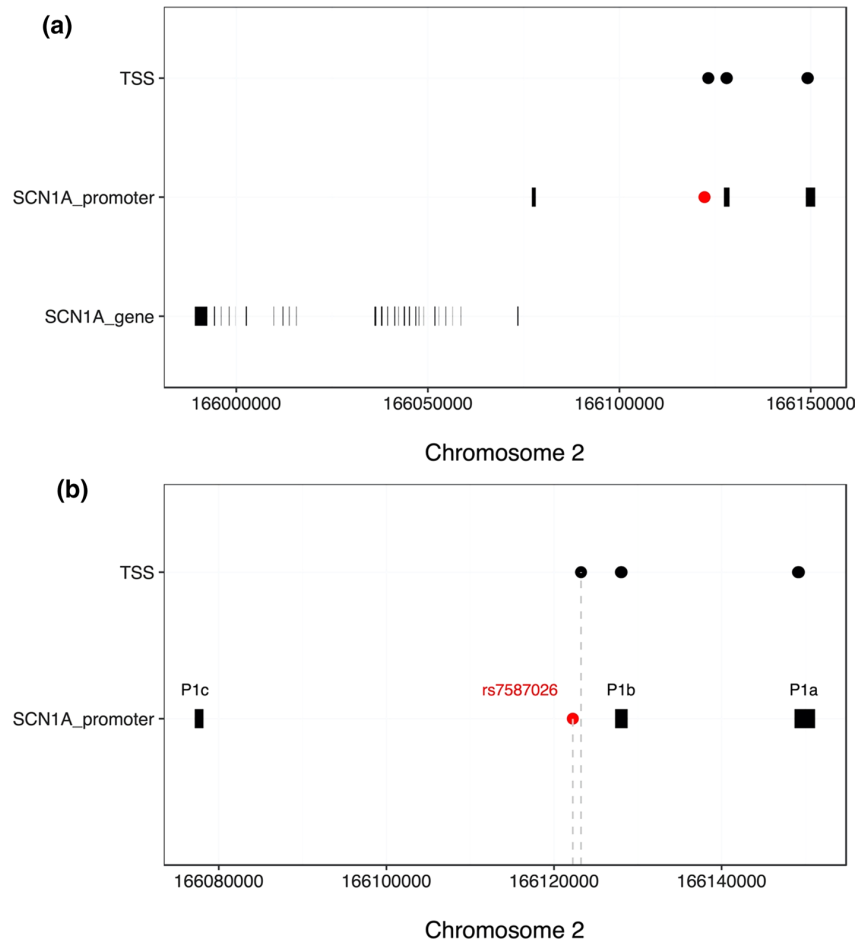
**Fig. 4** **a** Chromosomal position of rs7587026 with respect to the three known *SCN1A* promoter regions. rs7587026 falls within the same LD block as the P1b promoter. X-axis: position on chromosome 2, Y-axis: elements that were considered (*SCN1A* promoter regions, LD blocks defined using the BigLD function and PLINK). **b** The top half of the figure shows the genetic position of the three *SCN1A* promoter regions, rs7587026, as well as a representation of the LD blocks that span the locus; the bottom half shows the sequence constraint level. **c** Level of sequence constraint of the genetic region surrounding rs7587026. Sequence constraint is expressed as context-dependent tolerance score (CDTS), indicating the likelihood of variation. Negative CDTS scores indicate highly constrained regions, infrequently mutated in healthy individuals and more likely to be functionally relevant



bound in vivo by a striking multitude of transcription factors. These data underline the biological complexity and relevance of transcriptional regulation involving rs7587026. In this simplified in vitro system, our analyses did not reveal a

difference in the promoter activity or SOX2 binding affinity between the two rs7587026 genotypes. Several reasons can account for this discrepancy between the in vitro data and our complementary experiments on rs7587026 effects

**Fig. 5** **a** Transcription start sites (TSSs) identified by the FANTOM5 project, located upstream of the *SCN1A* gene body. Six TSSs fell within two of the three *SCN1A* promoters and one TSS was located 981 bp away from rs7587026. **b** X-axis: genetic positions on chromosome 2, Y-axis: genetic elements that were considered

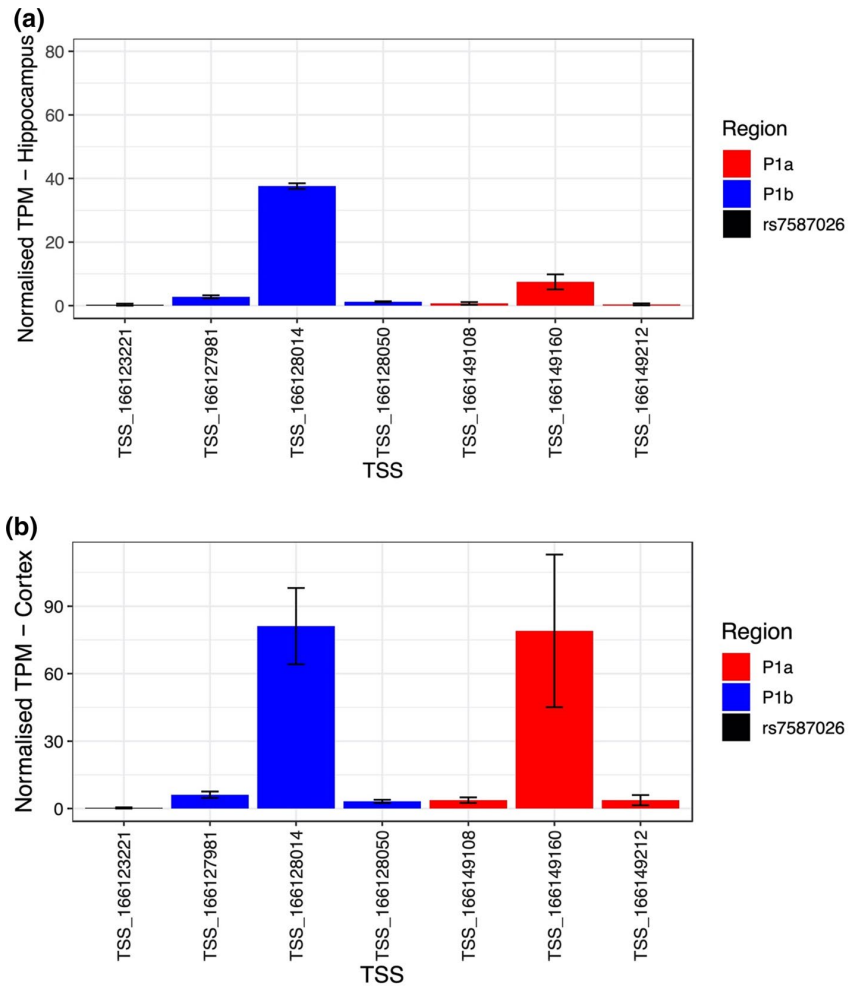


on *SCN1A* transcript abundance under in vivo conditions in human brain biopsies: given the obvious regulatory complexity of this novel promoter motif for *SCN1A*, additional transcription factors, transcription factor complexes, trans-regulatory promoter regulation as well as expression quantitative trait loci (eQTLs) may contribute to the observed differential expression of *SCN1A* in brain biopsy tissue of patients with epilepsy stratified according to the rs7587026 genotype, as compared to the distilled in vitro expression test system we used. In addition, one or more transcription factors that binds to the sequence may only be up-regulated in the diseased brain. The present in vitro experiments clearly indicate a strong regulatory role for *SCN1A* expression of the rs7587026-containing genomic motif, even if we were unable to define the precise mechanism (it is typically not possible to entirely replicate the complexity of in vivo promoter regulation in in vitro systems): in this context, the rs7587026 variant has a definite relevance, as our corresponding data from human epileptogenic brain tissue and MRI from healthy individuals clearly demonstrate. Whilst we note that rs7587026 and the P1b *SCN1A* promoter fall

in the same linkage disequilibrium block, the experimental results presented here suggest that a short genomic fragment around rs7587026 alone has promoter activity independent of P1b, located further away in the same linkage disequilibrium block. Finally, we note that the JASPAR and ReMap resources produced contradictory in silico results: these resources were created using different input data and pipelines (JASPAR used internally produced data, HT-SELEX, PBMs, and ChIP-seq and DAP-seq experiments from CistromeDB, ReMap 2020 [ReMap 2020 itself used data from GEO, ENCODE, ENA], ChIP-atlas and ModERN, whilst ReMap (2022) used GEO, ArrayExpress, ENCODE). The discrepancies are, therefore, not necessarily surprising and demonstrate the importance of further empirical data, which we provide, and an overarching perspective.

It remains unknown how increased *SCN1A* expression leads to reduced hippocampal volume. The generalisability to children of our human histopathology and imaging data, which were obtained from adults, is unproven. However,

**Fig. 6** Normalized tags per million (TPM) expression values of the TSSs located upstream of *SCN1A* in the hippocampus (a) and cortex (b). In red are shown the TSSs located in P1a, in blue the TSSs in P1b and in black the TSS close to rs7587026. The black bars indicate the standard error (SE) of the mean

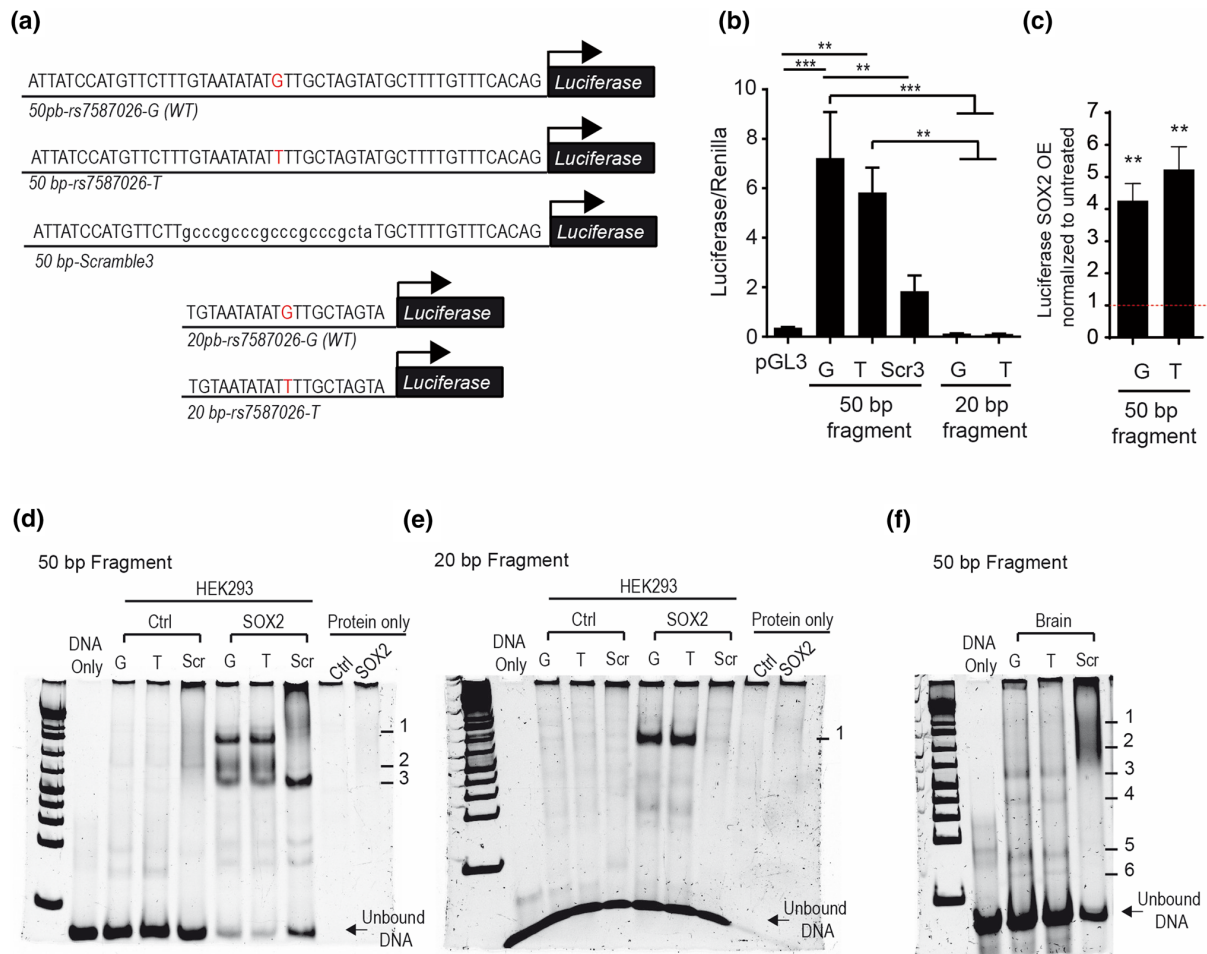


showing that genotypically driven increased *SCN1A* expression reduces hippocampal volume before the occurrence of FS in people (children) who go on to develop MTL-EHS would face insurmountable ethical and logistic implications (for example, MRI at this age would have to be conducted under general anaesthesia). Our approach provides new insights into the association between FS and MTL-EHS that would otherwise be difficult to obtain.

Our studies of *SCN1A* expression in hippocampi of individuals of different rs7587026 type, and of patterns of neuronal loss, is limited by the small numbers of minor allele homozygotes, influenced by the minor allele frequency of 0.302 [29], and the rarity of hippocampal specimens available for study. For the study of neuronal loss, we addressed this limitation by performing a replication study in a separate set of samples (Supplementary Fig. 3, online resource). Unfortunately, this was not possible for the *SCN1A* expression data.

In conclusion, having previously shown that rs7587026 is associated with increased susceptibility to MTL-EHS + FS, we demonstrate a potential mechanistic link through increased hippocampal *SCN1A* expression, smaller hippocampal volumes and an increased propensity to temperature-induced seizures in a model overexpressing *scn1a*.

There is a need for biomarkers of epileptogenesis [14]. Our results suggest that rs7587026 could become a biomarker in children for those at risk of developing epilepsy after FS, opening up the possibility of its use in studies aiming to identify anti-epileptogenic treatments after such insults. Moreover, the consequences we have shown of increased *Scn1a* expression point to the need for quantitative precision in genetic therapies intended to increase *SCN1A* transcription [10] for treatment of seizures in Dravet syndrome associated with *SCN1A* haploinsufficiency.



**Fig. 7** Genomic *SCN1A* fragment harbouring SNP rs7587026 has promoter activity. **a** DNA sequence and schematic representation of genomic fragments of *SCN1A* cloned in luciferase constructs. The 50 bp fragments are denoted as 50 bp-rs7587026-G (WT) (for the major allele/wild type sequence), 50 bp-rs7587026-T for the minor allele sequence and as 50 bp-Scramble3 (for the scramble sequence); the 20 bp fragments are denoted as 20 bp-rs7587026-G (WT) (for the major allele/wild type sequence) and 20 bp-rs7587026-T for the minor allele sequence. **b** Luciferase activity of NS20Y cells transfected with the 50 bp luciferase fragments with both genotypes and a scrambled sequence (constructs: *SCN1A*-50 bp-rs7587026-G(WT)-Luciferase, *SCN1A*-50 bp-rs7587026-T-Luciferase or *SCN1A*-50 bp-Scramble3-Luciferase) and 20 bp luciferase fragments of both genotypes (constructs: *SCN1A*-20 bp-rs7587026-G(WT)-Luciferase or *SCN1A*-20 bp-rs7587026-T-Luciferase). Luciferase values were

normalized to Renilla values. Data is represented as mean  $\pm$  SEM (one-way ANOVA, Tukey's multiple comparison test,  $n=4$ ). **c** Luciferase activity of NS20Y cells co-transfected with the 50 bp-*SCN1A* fragments (*SCN1A*-50 bp-rs7587026-G(WT)-Luciferase or *SCN1A*-50 bp-rs7587026-T-Luciferase) and pCMV-Sox2-T2A-GFP (50 ng). Luciferase values were normalized to the untreated condition (data is represented as mean  $\pm$  SEM,  $n=4$ , One sample  $t$  test). **d, e** EMSA performed with HEK293T cells overexpressing pCMV-Sox2-T2A-GFP or hypB-CAG-2A-eGFP (control) with the 50 bp *SCN1A* fragments (d) and the 20 bp *SCN1A* fragment (e). Scr denotes scrambled sequences (sequences in Supplementary Table 2, online resource). **f** EMSA reaction with nuclear protein extracts from mouse brain extract and the 50 bp *SCN1A* fragments. Numbers denote DNA-protein complexes

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00401-022-02429-0>.

**Author contributions** KS: Study coordination, interpretation of results, drafting of manuscript. KG: Zebrafish experiments (EEG, drug screening, qRT-PCR, hyperthermia experiments, manuscript writing,

interpretation of results and statistical analysis, figure generation). DT, JP, KMJvL, SS and AJB: Human genotyping and expression analysis. SA: conducted MRI-based analyses (statistical analyses, interpretations of results, figure generation and manuscript writing). RP, ZM, MT: immunohistochemistry for neuronal loss in human MTLHS. SP,

HMC and JM conducted bioinformatic analyses. WvE, KJK-S, ET: BAC construct design, cloning and preparation. CDW, GldZ, KLMM, MJW, PMT: collection, analysis and interpretation of MRI data. JMM, AF: bioinformatic analyses. CVE: Supervision of all zebrafish experiments. SMS: Conception of study, interpretation of results, revision of manuscript. Content of the manuscript contributes in part to the PhD thesis of DT.

**Funding** KS was supported by a Wellcome Trust Strategic Award (WT104033AIA). Part of this work was undertaken at University College London Hospitals, which received a proportion of funding from the NIHR Biomedical Research Centres funding scheme. Additional funding was provided by the Epilepsy Society (KS, SP, HMC, SMS). KG was supported by Deutsche Forschungsgemeinschaft (INTER/DFG/17/11583046 MechEPI; Mechanisms of Epileptogenesis) and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Individual Fellowship grant agreement No. 798703-GEMZ-H2020-MSCA-IF-2017. JP is supported by a Junior Research Group of the Medical Faculty of the University of Bonn. WvdE was supported by an ERA-NET cofund scheme (Project No.: 284365), and the MSCA-COFUND-FP scheme (EU 801133—Scientia Fellows II). JMM and AF are supported by the National Human Genome Research Institute of the National Institutes of Health [U41HG007234]; the content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health; Wellcome Trust (WT108749/Z/15/Z, WT200990/Z/16/Z); European Molecular Biology Laboratory. AJB and SS are supported by Deutsche Forschungsgemeinschaft SFB 1089; AJB is also supported by FOR2715.

**Data availability** Data that support the findings of this study are available from the corresponding author upon reasonable request and subject to requirements of the sources for individual components of the study.

## Declarations

**Conflict of interest** CDW is currently an employee of Biogen. His work was conducted while affiliated at the Department of Molecular and Cellular Therapeutics, The Royal College of Surgeons in Ireland, and the Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California. PMT also received research grant funding from Biogen, Inc., for research unrelated to this manuscript. The remaining authors report no competing interests.

**Ethical approval** The study of *SCN1A* expression and neuronal loss in hippocampal specimens from individuals with MTLAHS was approved by the Ethical Commission of University Hospital Bonn (222/16). The replication study of neuronal loss in hippocampal specimens from individuals with MTLAHS was approved by UK National Research Ethics (17/SC/0573). The QTIM study was approved by the Human Research Ethics Committees of the University of Queensland and the QIMR Berghofer Medical Research Institute. The *in vitro* promoter analysis was approved by the Ethical Commission of University Hospital Bonn (196/17).

**Research involves human or animals participants** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. All experiments involving zebrafish were performed in compliance with the European Community

Council Directive of November 2010 for Care and Use of Laboratory Animals (Directive 2010/63/EU), and the ARRIVE guidelines. The Norwegian Food Safety Authority via its experimental animal administration's supervisory and application system approved all animal experimentation (FOTS ID 15469 and 23935).

**Consent to participate** Informed consent was obtained from all individual participants included in the study.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Abugessaisa I, Noguchi S, Hasegawa A, Harshbarger J (2017) Data Descriptor : FANTOM 5 CAGE profiles of human and mouse reprocessed for GRCh 38 and GRCh 38 genome assemblies. *Sci Data* 4:170107. <https://doi.org/10.1038/sdata.2017.107>
2. Afrikanova T, Serruys A-SK, Buenafe OEM, Clinckers R, Smolders I (2013) Validation of the zebrafish pentylentetrazol seizure model: locomotor versus electrographic responses to antiepileptic drugs. *PLoS One* 8:54166. <https://doi.org/10.1371/journal.pone.0054166>
3. Annegers JF, Hauser WA, Shirts SB, Kurland LT (1987) Factors prognostic of unprovoked seizures after febrile convulsions. *N Engl J Med* 316:493–498. <https://doi.org/10.1056/NEJM198702263160901>
4. Blümcke I, Pauli E, Clusmann H, Schramm J, Becker A, Elger C et al (2007) A new clinico-pathological classification system for mesial temporal sclerosis. *Acta Neuropathol* 113:235–244. <https://doi.org/10.1007/s00401-006-0187-0>
5. Blümcke I, Thom M, Aronica E, Armstrong DD, Bartolomei F, Bernardini A et al (2013) International consensus classification of hippocampal sclerosis in temporal lobe epilepsy: a task force report from the ILAE commission on diagnostic methods. *Epilepsia* 54:1315–1329. <https://doi.org/10.1111/epi.12220>
6. Casadei R, Pelleri MC, Vitale L, Facchin F, Lenzi L, Canaider S et al (2011) Identification of housekeeping genes suitable for gene expression analysis in the zebrafish. *Gene Expr Patterns* 11:271–276. <https://doi.org/10.1016/j.gep.2011.01.003>
7. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:7. <https://doi.org/10.1186/s13742-015-0047-8>
8. Chèneby J, Ménétrier Z, Mestdagh M, Rosnet T, Douida A, Rhalloussi W et al (2020) ReMap 2020: a database of regulatory regions from an integrative analysis of human and Arabidopsis DNA-binding sequencing experiments. *Nucleic Acids Res* 48:D180–D188. <https://doi.org/10.1093/nar/gkz945>
9. Claes L, Del-Favero J, Ceulemans B, Lagae L, Van Broeckhoven C, De Jonghe P (2001) De novo mutations in the


- sodium-channel gene *SCN1A* cause severe myoclonic epilepsy of infancy. *Am J Hum Genet* 68:1327–1332. <https://doi.org/10.1086/320609>
10. Colasante G, Lignani G, Brusco S, Di Berardino C, Carpenter J, Giannelli S et al (2019) dCas9-based *Scn1a* gene activation restores inhibitory interneuron excitability and attenuates seizures in Dravet syndrome mice. *Mol Ther* 28:235–253. <https://doi.org/10.1016/j.ymthe.2019.08.018>
  11. Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. *Neuroimage* 9:179–194. <https://doi.org/10.1006/nimg.1998.0395>
  12. Dupret B, Völkel P, Follet P, Le Bourhis X, Angrand PO (2018) Combining genotypic and phenotypic analyses on single mutant zebrafish larvae. *MethodsX* 5:244–256. <https://doi.org/10.1016/j.mex.2018.03.002>
  13. Dutton SBB, Dutt K, Papale LA, Helmers S, Goldin AL, Escayg A (2017) Early-life febrile seizures worsen adult phenotypes in *Scn1a* mutants. *Exp Neurol* 293:159–171. <https://doi.org/10.1016/j.expneurol.2017.03.026>
  14. Engel J, Pitkänen A (2020) Biomarkers for epileptogenesis and its treatment. *Neuropharmacology* 167:107735. <https://doi.org/10.1016/j.neuropharm.2019.107735>
  15. Feenstra B, Pasternak B, Geller F, Carstensen L, Wang T, Huang F et al (2014) Common variants associated with general and MMR vaccine-related febrile seizures. *Nat Genet* 46:1274–1282. <https://doi.org/10.1038/ng.3129>
  16. Fischl B, Liu A, Dale AM (2001) Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Trans Med Imaging* 20:70–80. <https://doi.org/10.1109/42.906426>
  17. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C et al (2002) Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33:341–355. [https://doi.org/10.1016/S0896-6273\(02\)00569-X](https://doi.org/10.1016/S0896-6273(02)00569-X)
  18. Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9:195–207. <https://doi.org/10.1006/nimg.1998.0396>
  19. The FANTOM Consortium and the RIKEN PMI and CLST (DGT) (2014) A promoter-level mammalian expression atlas. *Nature* 507:462–470. <https://doi.org/10.1038/NATURE13182>
  20. Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J et al (2019) GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* 47:D766–D773. <https://doi.org/10.1093/nar/gky955>
  21. French JA, Williamson PD, Thadani YVM, Darcey TTM, Mattson RH, Spencer SS et al (1993) Characteristics of medial temporal lobe epilepsy: I. Results of history and physical examination. *Ann Neurol* 34:774–780. <https://doi.org/10.1002/ana.410340604>
  22. Grasby KL, Jahanshad N, Painter JN, Colodro-Conde L, Bralten J, Hibar DP et al (2020) The genetic architecture of the human cerebral cortex. *Science*. <https://doi.org/10.1126/science.aay6690>
  23. Hezroni H, Koppstein D, Schwartz MG, Avrutin A, Bartel DP, Ulitsky I (2015) Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Rep* 11:1110–1122. <https://doi.org/10.1016/j.celrep.2015.04.023>
  24. Hon CC, Ramilowski JA, Harshbarger J, Bertin N, Rackham OJL, Gough J et al (2017) An atlas of human long non-coding RNAs with accurate 5' ends. *Nature* 543:199–204. <https://doi.org/10.1038/nature21374>
  25. International League Against Epilepsy Consortium on Complex Epilepsies (2018) Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nat Commun* 9:5269. <https://doi.org/10.1038/s41467-018-07524-z>
  26. Ishii A, Watkins JC, Chen D, Hirose S, Hammer MF (2017) Clinical implications of SCN1A missense and truncation variants in a large Japanese cohort with Dravet syndrome. *Epilepsia* 58:282–290. <https://doi.org/10.1111/epi.13639>
  27. Di Iulio J, Bartha I, Wong EHM, Yu HC, Lavrenko V, Yang D et al (2018) The human noncoding genome defined by genetic diversity. *Nat Genet* 50(5):333–337. <https://doi.org/10.1038/S41588-018-0062-7>
  28. Jaffe AE, Shin J, Collado-Torres L, Leek JT, Tao R, Li C et al (2015) Developmental regulation of human cortex transcription and its clinical relevance at single base resolution. *Nat Neurosci* 18:154–161. <https://doi.org/10.1038/nn.3898>
  29. Kasperavičiūtė D, Catarino CB, Matarin M, Leu C, Novy J, Tostevin A et al (2013) Epilepsy, hippocampal sclerosis and febrile seizures linked by common genetic variation around *SCN1A*. *Brain* 136:3140–3150. <https://doi.org/10.1093/brain/awt233>
  30. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM et al (2002) The human genome browser at UCSC. *Genome Res* 12:996–1006. <https://doi.org/10.1101/GR.229102>
  31. Kim SA, Cho CS, Kim SR, Bull SB, Yoo YJ (2018) A new haplotype block detection method for dense genome sequencing data based on interval graph modeling of clusters of highly correlated SNPs. *Bioinformatics* 34:388–397. <https://doi.org/10.1093/BIOINFORMATICS/BTX609>
  32. Kjeldsen MJ, Kyvik KO, Friis ML, Christensen K (2002) Genetic and environmental factors in febrile seizures: a Danish population-based twin study. *Epilepsy Res* 51:167–177. [https://doi.org/10.1016/S0920-1211\(02\)00121-3](https://doi.org/10.1016/S0920-1211(02)00121-3)
  33. Kral T, Clusmann H, Urbach J, Schramm J, Elger CE, Kurthen M et al (2002) Preoperative evaluation for epilepsy surgery (Bonn algorithm). *Zentralbl Neurochir* 63:106–110. <https://doi.org/10.1055/s-2002-35826>
  34. Lagarde J, Uszczyńska-Ratajczak B, Carbonell S, Pérez-Lluch S, Abad A, Davis C et al (2017) High-throughput annotation of full-length long noncoding RNAs with capture long-read sequencing. *Nat Genet* 49:1731–1740. <https://doi.org/10.1038/ng.3988>
  35. Leu C, Stevelink R, Smith AW et al (2019) Polygenic burden in focal and generalized epilepsies. *Brain* 142:3473–3481. <https://doi.org/10.1093/brain/awz292>
  36. Lévesque M, Avoli M, Bernard C (2016) Animal models of temporal lobe epilepsy following systemic chemoconvulsant administration. *J Neurosci Methods* 260:45–52. <https://doi.org/10.1016/j.jneumeth.2015.03.009>
  37. Lewis DV, Shinnar S, Hesdorffer DC, Bagiella E, Bello JA, Chan S et al (2014) Hippocampal sclerosis after febrile status epilepticus: the FEBSTAT study. *Ann Neurol* 75:178–185. <https://doi.org/10.1002/ana.24081>
  38. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* 25:402–408. <https://doi.org/10.1006/meth.2001.1262>
  39. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S et al (2013) The genotype-tissue expression (GTEx) PROJECT. *Nat Genet* 45:580–585. <https://doi.org/10.1038/ng.2653>
  40. Van Loo KJM, Schaub C, Pernhorst K, Yaari Y, Beck H, Schoch S et al (2012) Transcriptional regulation of T-type calcium channel Cav3.2: bi-directionality by early growth response 1 (Egr1) and repressor element 1 (RE-1) protein-silencing transcription factor (REST). *J Biol Chem* 287:15489–15501. <https://doi.org/10.1074/JBC.M111.310763>
  41. Löscher W (2020) The holy grail of epilepsy prevention: preclinical approaches to antiepileptogenic treatments. *Neuropharmacology* 167:107605. <https://doi.org/10.1016/j.neuropharm.2019.04.011>
  42. McCurley AT, Callard GV (2008) Characterization of housekeeping genes in zebrafish: male-female differences and effects



- of tissue type, developmental stage and chemical treatment. *BMC Mol Biol* 9:1–12. <https://doi.org/10.1186/1471-2199-9-102>
43. Pernhorst K, Herms S, Hoffmann P, Cichon S, Schulz H, Sander T et al (2013) TLR4, ATF-3 and IL8 inflammation mediator expression correlates with seizure frequency in human epileptic brain tissue. *Seizure* 22:675–678. <https://doi.org/10.1016/j.seizure.2013.04.023>
  44. Pernhorst K, Raabe A, Niehusmann P, Van Loo KMJ, Grote A, Hoffmann P et al (2011) Promoter variants determine  $\gamma$ -aminobutyric acid homeostasis-related gene transcription in human epileptic hippocampi. *J Neuropathol Exp Neurol* 70:1080–1088. <https://doi.org/10.1097/NEN.0b013e318238b9af>
  45. Pittau F, Bisulli F, Mai R, Fares JE, Vignatelli L, Labate A et al (2009) Prognostic factors in patients with mesial temporal lobe epilepsy. *Epilepsia* 50:41–44. <https://doi.org/10.1111/j.1528-1167.2008.01969.x>
  46. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575. <https://doi.org/10.1086/519795>
  47. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140. <https://doi.org/10.1093/BIOINFORMATICS/BTP616>
  48. Salgueiro-Pereira AR, Duprat F, Pousinha PA, Loucif A, Duchamps V, Regondi C et al (2019) A two-hit story: seizures and genetic mutation interaction sets phenotype severity in *SCN1A* epilepsies. *Neurobiol Dis* 125:31–44. <https://doi.org/10.1016/j.nbd.2019.01.006>
  49. Schönberger A, Niehusmann P, Urbach H, Majores M, Grote A, Holthausen H et al (2009) Increased frequency of distinct TSC2 allelic variants in focal cortical dysplasias with balloon cells and mineralization: original article. *Neuropathology* 29:559–565. <https://doi.org/10.1111/j.1440-1789.2009.01018.x>
  50. Scott RC, King MD, Gadian DG, Neville BGR, Connelly A (2003) Hippocampal abnormalities after prolonged febrile convulsion: a longitudinal MRI study. *Brain* 126:2551–2557. <https://doi.org/10.1093/brain/awg262>
  51. Semah F, Picot MC, Adam C, Broglin D, Arzimanoglou A, Bazin B et al (1998) Is the underlying cause of epilepsy a major prognostic factor for recurrence? *Neurology* 51:1256–1262. <https://doi.org/10.1212/wnl.51.5.1256>
  52. Stein JL, Medland SE, Vasquez AA, Hibar DP, Senstad RE, Winkler AM et al (2012) Identification of common variants associated with human hippocampal and intracranial volumes. *Nat Genet* 44:552–561. <https://doi.org/10.1038/ng.2250>
  53. Stephen LJ, Kwan P, Brodie MJ (2001) Does the cause of localisation-related epilepsy influence the response to antiepileptic drug treatment? *Epilepsia* 42:357–362. <https://doi.org/10.1046/j.1528-1157.2001.29000.x>
  54. Tai XY, Bernhardt B, Thom M, Thompson P, Baxendale S, Koepf M et al (2018) Review: neurodegenerative processes in temporal lobe epilepsy with hippocampal sclerosis: clinical, pathological and neuroimaging evidence. *Neuropathol Appl Neurobiol* 44:70–90. <https://doi.org/10.1111/nan.12458>
  55. Team RC (2015) R: a language and environment for statistical computing. Foundation for Statistical Computing, Vienna
  56. Thijs RD, Surges R, O'Brien TJ, Sander JW (2019) Epilepsy in adults. *Lancet* 393:689–701. [https://doi.org/10.1016/S0140-6736\(18\)32596-0](https://doi.org/10.1016/S0140-6736(18)32596-0)
  57. Thom M (2014) Review: hippocampal sclerosis in epilepsy: a neuropathology review. *Neuropathol Appl Neurobiol* 40:520–543. <https://doi.org/10.1111/nan.12150>
  58. Tilgner H, Jahanbani F, Blauwkamp T, Moshrefi A, Jaeger E, Chen F et al (2015) Comprehensive transcriptome analysis using synthetic long-read sequencing reveals molecular co-association of distant splicing events. *Nat Biotechnol* 33:736–742. <https://doi.org/10.1038/nbt.3242>
  59. Tsortouktzidis D, Schulz H, Hamed M, Vatter H, Surges R, Schoch S et al (2021) Gene expression analysis in epileptic hippocampi reveals a promoter haplotype conferring reduced aldehyde dehydrogenase 5a1 expression and responsiveness. *Epilepsia* 62:e29–e34. <https://doi.org/10.1111/epi.16789>
  60. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A et al (2015) Tissue-based map of the human proteome. *Science* 347:1260419. <https://doi.org/10.1126/science.1260419>
  61. Whelan CD, Altmann A, Botía JA, Jahanshad N, Hibar DP, Absil J et al (2018) Structural brain abnormalities in the common epilepsies assessed in a worldwide ENIGMA study. *Brain* 141:391–408. <https://doi.org/10.1093/brain/awx341>
  62. Wieser HG, ILAE Commission on Neurosurgery of Epilepsy (2004) Mesial temporal lobe epilepsy with hippocampal sclerosis. *Epilepsia* 45:695–714. <https://doi.org/10.1111/j.0013-9580.2004.09004.x>
  63. Woitecki AMH, Müller JA, van Loo KMJ, Sowade RF, Becker AJ, Schoch S (2016) Identification of synaptotagmin 10 as effector of NPAS4-mediated protection from excitotoxic neurodegeneration. *J Neurosci* 36:2561–2570. <https://doi.org/10.1523/JNEUROSCI.2027-15.2016>
  64. Zhang YH, Burgess R, Malone JP, Glubb GC, Helbig KL, Vadlamudi L et al (2017) Genetic epilepsy with febrile seizures plus. *Neurology* 89:1210–1219. <https://doi.org/10.1212/WNL.00000000000004384>
  65. Zuberi SM, Brunklaus A, Birch R, Reavey E, Duncan J, Forbes GH (2011) Genotype-phenotype associations in *SCN1A*-related epilepsies. *Neurology* 76:594–600. <https://doi.org/10.1212/WNL.0b013e31820c309b>
  66. de Zubicaray GI, Chiang MC, McMahon KL, Shattuck DW, Toga AW, Martin NG et al (2008) Meeting the challenges of neuroimaging genetics. *Brain Imaging Behav* 2:258–263. <https://doi.org/10.1007/s11682-008-9029-0>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

**Katri Silvennoinen**<sup>1,2</sup>  · **Kinga Gawel**<sup>3,4</sup> · **Despina Tsortouktzidis**<sup>5,6</sup> · **Julika Pitsch**<sup>5,6</sup> · **Saud Alhusaini**<sup>7,8</sup> · **Karen M. J. van Loo**<sup>5,9</sup> · **Richard Picardo**<sup>10</sup> · **Zuzanna Michalak**<sup>10</sup> · **Susanna Pagni**<sup>1,2</sup> · **Helena Martins Custodio**<sup>1,2</sup> · **James Mills**<sup>1,2</sup> · **Christopher D. Whelan**<sup>7,11</sup> · **Greig I. de Zubicaray**<sup>12</sup> · **Katie L. McMahon**<sup>13</sup> · **Wietske van der Ent**<sup>3</sup> · **Karolina J. Kirstein-Smardzewska**<sup>3</sup> · **Ettore Tiraboschi**<sup>3</sup> · **Jonathan M. Mudge**<sup>14</sup> · **Adam Frankish**<sup>14</sup> · **Maria Thom**<sup>10</sup> · **Margaret J. Wright**<sup>15</sup> · **Paul M. Thompson**<sup>11</sup> · **Susanne Schoch**<sup>5,6</sup> · **Albert J. Becker**<sup>5</sup> · **Camila V. Esguerra**<sup>3</sup> · **Sanjay M. Sisodiya**<sup>1,2</sup>

<sup>1</sup> Department of Clinical and Experimental Epilepsy, UCL Queen Square Institute of Neurology, Box 29, Queen Square, London WC1N 3BG, UK

<sup>2</sup> Chalfont Centre for Epilepsy, Chalfont St Peter, Bucks SL9 0RJ, UK

<sup>3</sup> Chemical Neuroscience Group, Centre for Molecular Medicine Norway (NCMM), Faculty of Medicine, University of Oslo, 0349 Oslo, Norway

<sup>4</sup> Department of Experimental and Clinical Pharmacology, Medical University of Lublin, 20-090 Lublin, Poland

<sup>5</sup> Institute of Neuropathology, Medical Faculty, University of Bonn, Section for Translational Epilepsy Research, 53127 Bonn, Germany

<sup>6</sup> Department of Epileptology, Medical Faculty, University of Bonn, 53127 Bonn, Germany

<sup>7</sup> Department of Molecular and Cellular Therapeutics, The Royal College of Surgeons in Ireland, Dublin 2, Ireland

<sup>8</sup> Department of Neurology, Yale University School of Medicine, New Haven, CT 06520, USA

<sup>9</sup> Department of Epileptology and Neurology, RWTH Aachen University, 52074 Aachen, Germany

<sup>10</sup> Department of Neuropathology, UCL Queen Square Institute of Neurology, London WC1N 3BG, UK

<sup>11</sup> Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Marina del Rey, Los Angeles, CA 90292, USA

<sup>12</sup> School of Psychology, Faculty of Health, Queensland University of Technology (QUT), Brisbane, QLD 4059, Australia

<sup>13</sup> School of Clinical Sciences, Faculty of Health, Queensland University of Technology (QUT), Brisbane, QLD 4029, Australia

<sup>14</sup> European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Cambridge CB10 1SD, UK

<sup>15</sup> Queensland Brain Institute, University of Queensland, St Lucia 4072 QLD, Australia

3.3 Publication 3: A Versatile Clustered Regularly Interspaced Palindromic Repeats Toolbox to Study Neurological Cav3.2 Channelopathies by Promoter-Mediated Transcription Control

**Tsortouktzidis D**, Tröscher AR, Schulz H, Opitz T, Schoch S, Becker AJ, van Loo KMJ. A Versatile Clustered Regularly Interspaced Palindromic Repeats Toolbox to Study Neurological Cav3.2 Channelopathies by Promoter-Mediated Transcription Control. **Front Mol Neurosci.** 2022 Jan 6;14:667143. doi: 10.3389/fnmol.2021.667143. PMID: 35069110; PMCID: PMC8770422.



# A Versatile Clustered Regularly Interspaced Palindromic Repeats Toolbox to Study Neurological Ca<sub>v</sub>3.2 Channelopathies by Promoter-Mediated Transcription Control

Despina Tsortouktzidis<sup>1</sup>, Anna R. Tröscher<sup>1,2</sup>, Herbert Schulz<sup>3</sup>, Thoralf Opitz<sup>4</sup>, Susanne Schoch<sup>1</sup>, Albert J. Becker<sup>1†</sup> and Karen M. J. van Loo<sup>1,5\*†</sup>

## OPEN ACCESS

### Edited by:

Felix Viana,  
 Institute of Neurosciences, Spanish  
 National Research Council (CSIC),  
 Spain

### Reviewed by:

Amaud Montell,  
 Centre National de la Recherche  
 Scientifique (CNRS), France  
 Luis Quintino,  
 Lund University, Sweden

### \*Correspondence:

Karen M. J. van Loo  
 kvanloo@ukaachen.de

<sup>†</sup> These authors have contributed  
 equally to this work

### Specialty section:

This article was submitted to  
 Methods and Model Organisms,  
 a section of the journal  
 Frontiers in Molecular Neuroscience

**Received:** 12 February 2021

**Accepted:** 15 December 2021

**Published:** 06 January 2022

### Citation:

Tsortouktzidis D, Tröscher AR,  
 Schulz H, Opitz T, Schoch S,  
 Becker AJ and van Loo KMJ (2022) A  
 Versatile Clustered Regularly  
 Interspaced Palindromic Repeats  
 Toolbox to Study Neurological Ca<sub>v</sub>3.2  
 Channelopathies by  
 Promoter-Mediated Transcription  
 Control.  
 Front. Mol. Neurosci. 14:667143.  
 doi: 10.3389/fnmol.2021.667143

<sup>1</sup> Institute of Neuropathology, Medical Faculty, Section for Translational Epilepsy Research, University of Bonn, Bonn, Germany, <sup>2</sup> Department of Neurology, Kepler University Hospital, Johannes Kepler University Linz, Linz, Austria, <sup>3</sup> Department of Microgravity and Translational Regenerative Medicine, Clinic for Plastic, Aesthetic and Hand Surgery, Otto von Guericke University, Magdeburg, Germany, <sup>4</sup> Institute of Experimental Epileptology and Cognition Research, Medical Faculty, University of Bonn, Bonn, Germany, <sup>5</sup> Department of Epileptology and Neurology, RWTH Aachen University, Aachen, Germany

Precise genome editing in combination with viral delivery systems provides a valuable tool for neuroscience research. Traditionally, the role of genes in neuronal circuits has been addressed by overexpression or knock-out/knock-down systems. However, those techniques do not manipulate the endogenous loci and therefore have limitations. Those constraints include that many genes exhibit extensive alternative splicing, which can be regulated by neuronal activity. This complexity cannot be easily reproduced by overexpression of one protein variant. The CRISPR activation and interference/inhibition systems (CRISPRa/i) directed to promoter sequences can modulate the expression of selected target genes in a highly specific manner. This strategy could be particularly useful for the overexpression of large proteins and for alternatively spliced genes, e.g., for studying large ion channels known to be affected in ion channelopathies in a variety of neurological diseases. Here, we demonstrate the feasibility of a newly developed CRISPRa/i toolbox to manipulate the promoter activity of the *Cacna1h* gene. Impaired, function of the low-voltage-activated T-Type calcium channel Ca<sub>v</sub>3.2 is involved in genetic/mutational as well as acquired/transcriptional channelopathies that emerge with epileptic seizures. We show CRISPR-induced activation and inhibition of the *Cacna1h* locus in NS20Y cells and primary cortical neurons, as well as activation in mouse organotypic slice cultures. In future applications, the system offers the intriguing perspective to study functional effects of gain-of-function or loss-of-function variations in the *Cacna1h* gene in more detail. A better understanding of Ca<sub>v</sub>3.2 channelopathies might result in a major advancement in the pharmacotherapy of Ca<sub>v</sub>3.2 channelopathy diseases.

**Keywords:** ion channelopathies, *Cacna1h* promoter modulation, CRISPR-induced activation, CRISPR-induced inhibition, T-Type calcium channel Ca<sub>v</sub>3.2

## INTRODUCTION

The human genome encodes approximately 400 ion channel genes, encompassing both voltage-gated and ligand-gated ion channels (Hutchings et al., 2019). Ion channels are pore forming membrane proteins that allow ionic flows across membranes and are crucial for normal functioning of many tissues, including the central and peripheral nervous system, heart, kidney, and liver (Cannon, 2007). Ion channel dysfunctions, also coined as ion “channelopathies,” have been associated with a large variety of diseases. Besides disorders of the nervous system (e.g., epilepsy, ataxia, Alzheimer’s disease, and Autism spectrum disorders), also cardiac arrhythmia and several muscle, endocrine, and renal disorders are linked to dysfunction of ion channels (Wang et al., 1996; Heeringa et al., 2009; Ryan et al., 2010). Still, many molecular and structural mechanisms of how channelopathies convert cells from a health to disease state are not fully understood, including critical time-windows during development as well as the potential reversibility by reconstituting normal expression/function of affected molecules. Major obstacles to manipulate ion channels are given by their large size, which limits options for widespread and *in vivo* overexpression and their diversification by alternative splicing.

Genetic editing using a modified CRISPR (clustered regularly interspaced palindromic repeats) system could be a powerful approach for the manipulation of large proteins that cannot be done by conventional techniques. Recently, substantial progress has been made in genomic editing by using specific applications of the CRISPR technique, including the CRISPR *activation* (CRISPRa) and CRISPR *interference/inhibition* (CRISPRi) technology. CRISPRa uses a catalytically dead Cas9 (dCas9) enzyme fused with a highly efficient transcriptional activator complex consisting of the tripartite transcriptional activator VP64-p65-Rta (VPR), shown to increase target gene expression even up to 320-fold (Chavez et al., 2015). CRISPRi also uses the dCas9 enzyme and is fused with the transcriptional repressor KRAB (Krüppel-associated box) protein, which results in up to 60–80% reduction in the expression of endogenous eukaryotic genes (Gilbert et al., 2013). Using these enhanced CRISPR technology systems, the expression of genes can be modified in their native context with utmost precision. Therefore, this strategy will be particularly useful for (a) the overexpression of large proteins, which is difficult to accomplish by conventional techniques and (b) for alternatively spliced genes.

In this study, we have developed a CRISPRa/i toolbox for manipulating the expression of a well-described ion channelopathy gene, *Cacna1h*, encoding the low-voltage-activated T-Type calcium channel Cav3.2. Ion channelopathies for *Cacna1h* have been described particularly for epilepsy variants (Khosravani et al., 2004, 2005; Powell et al., 2009; Souza et al., 2019), but were also reported for other neurological diseases including autism spectrum disorders (Splawski et al., 2006), amyotrophy lateral sclerosis (ALS; Rzhetsky et al., 2016) and pain disorders (Souza et al., 2016). By using the unique CRISPRa/i toolbox, we demonstrate to specifically modulate *Cacna1h* gene expression in different cell types in order to closely recapitulate Cav3.2 channelopathies.

## MATERIALS AND METHODS

### Design of sgRNAs

sgRNA design was performed using the computational software Benchling (Cloud-Based Informatics Platform for Life Sciences R&D | Benchling, 2021). The previously validated *Cacna1h* promoter (Van Loo et al., 2012) was used to set the sgRNAs targeting sequence. The sgRNAs were selected based on their local off-target scores, proximity to the start-ATG, a distance of at least 50 bp between each other and a 100% match in targeting both the mouse and rat genome.

### Cloning

All primer sequences used for cloning are shown in **Supplementary Table 1**. As a basis for our cloning strategies, we exchanged the hybrid cytomegalovirus-actin-globin promoter of pAAV-MCS (Agilent) by the human synapsin (hSyn) promoter (Kügler et al., 2003) using the *MluI/Bsu15I* restriction sites. In order to generate the pAAV-U6-sgRNA plasmids, we cloned a U6-*BbsI/BbsI* cassette into the pAAV-hSyn-MCS backbone. For this, the U6-*BbsI/BbsI* cassette was PCR-amplified from px458 [Addgene #48138, Ran et al. (2013)] and inserted by in-fusion cloning (IFC; Takara Bio Europe/Clontech) into *MluI/AsiSI*-digested pAAV-hSyn-MCS. Next, the polyA was removed by digestion with *AfeI/PmlI* and subsequent self-ligation. Finally, the sgRNAs were annealed and cloned into the *BbsI* sites (Cloud-Based Informatics Platform for Life Sciences R&D | Benchling, 2021). Briefly, the annealing was performed using 10  $\mu$ M of each oligo, 1X T4 ligation buffer and 5U of T4 PNK in a total volume of 10  $\mu$ L, following an incubation at 37°C for 30 min and 95°C for 5 min, the reaction was cooled at room temperature. The annealed oligos (1  $\mu$ L) were cloned using 25 ng of construct, 1X T4 ligation buffer, T4 ligase (200 U) and *BbsI* (2.5 U) in a total volume of 10  $\mu$ L, the reaction was performed by 30 cycles at 37°C for 5 min and 23°C for 5 min.

The all-in-one CRISPRa lentiviral system was generated by replacing the promoter of pLenti-Ef1a-dCas9-VPR [Addgene#114195, Savell et al. (2019)] by the hSyn promoter using IFC. For this, pAAV-hSyn-MCS was used as PCR template for the hSyn promoter and cloned into *AfeI/KpnI*-digested pLenti-Ef1a-dCas9-VPR. In a second step, the *Cacna1h* and *LacZ* sgRNA sequences from the pAAV-U6-sgRNA plasmids (see above) were cloned by IFC into the *PacI* site of pLenti-syn-dCas9-VPR, resulting in pLenti-U6-sgRNA<sub>(Cacna1h/LacZ)</sub>-Syn-dCas9-VPR.

The all-in-one CRISPRi lentiviral system was produced by replacing the hU6-sgRNA-hUbc cassette of pLenti-hU6-sgRNA-hUbc-dCas9-KRAB-T2a-eGFP [Addgene#71237, Thakore et al. (2015)] for the hSyn promoter by IFC. The backbone was digested with *XbaI/PacI* and the hSyn promoter PCR amplified from pAAV-hSyn-MCS, resulting in pLenti-hSyn-dCas9-KRAB-T2A-eGFP. Subsequently, the U6-sgRNA<sub>Cacna1h/LacZ</sub> cassettes from the pAAV-U6-sgRNA plasmids (see above) were inserted into the *PacI* site of pLenti-hSyn-dCas9-KRAB-T2A-eGFP by IFC.

pTRE-dCas9-VPR was generated by replacing the promoter of pLenti-Ef1a-dCas9-VPR [Addgene#114195, Savell et al. (2019)] for the tetracycline response element (TRE) of the pK031.TRE-Cre [Addgene#69136 (Mizuno et al., 2014)] by IFC using the *AfeI/KpnI* restriction sites. pTRE-dCas9-KRAB-T2A-eGFP was produced by replacing the promoter of the previously generated pLenti-hSyn-dCas9-KRAB-T2A-eGFP with the TRE of pK031.TRE-Cre [Addgene#69136 (Mizuno et al., 2014)] by IFC, using the *PacI/XbaI* restriction sites. pAAV-hSyn-rtTA was generated by IFC using *EcoRI/SalI*-digested pAAV-hSyn-MCS and the rtTA sequence PCR-amplified from mCreb1 in pInducer20, kindly provided by Dasgupta and co-workers (Chhipa et al., 2018).

### Cell Culture, Transfection, and Luciferase Assay

NS20Y cells (Sigma, #08062517) were maintained at 37°C and 5% CO<sub>2</sub> in DMEM (Sigma, D6546) supplemented with 10% (v/v) heat inactivated FBS, 2 mM L-Glutamine, 100 units/mL penicillin/streptomycin. Cells were transfected in 24 well plates using Lipofectamine (Invitrogen) following the manufacturer's instructions. The DNA concentration used per well: *Cacna1h*-Luciferase or *Cacna1h*-mRuby 100 ng, CMV-VPR [Addgene#63798, (Chavez et al., 2015)] or CMV-KRAB [Addgene#110821, (Yeo et al., 2018)] 200 ng, pAAV-sgRNA 200 ng and hypB-CAG-2A-eGFP 100 ng. Following 48 h after transfection, the cells were imaged or collected for luciferase assays. Luciferase assays were performed using the Dual Luciferase Reporter Assay System (Promega) according to the manufacturer's specifications. Firefly luciferase activity was determined using the Glomax Luminometer (Promega).

### Viral Production and Neuronal Transduction

AAV1/2 viruses were produced by large-scale triple CaPO<sub>4</sub> transfection of HEK293-AAV cells (Agilent, #240073) as described previously (Van Loo et al., 2012). Lentiviruses were produced in HEK293T cells using a second-generation lentiviral packaging system. The procedure was performed as described before (Van Loo et al., 2019). Dissociated primary neurons were prepared from mouse cortex (C57Bl6/N) at embryonic day 15–19 as described before (Woitecki et al., 2016). All animals were handled according to government regulations as approved by local authorities (LANUV Recklinghausen). All procedures were planned and performed in accordance with the guidelines of the University Hospital Bonn Animal-Care-Committee as well as the guidelines approved by the European Directive (2010/63/EU) on the protection of animals used for experimental purposes. Cells were kept in BME medium (Gibco) supplemented with 1% FBS, 0.5 mM L-glutamine, 0.5% glucose and 1X B27 at 37°C and 5% CO<sub>2</sub>. Neuronal transduction was performed at days *in vitro* (DIV) 4 in 24 well plates containing 70,000 cells/well with a multiplicity of infection (MOI) of approximately 14. Cells were collected on DIV15 in lysis/binding buffer (Invitrogen, A33562) and stored at –80°C until mRNA extraction or at DIV20 for Western blot.

### RNA Isolation and Real Time RT-PCR

RNA isolation and cDNA production was performed using Dynabeads mRNA Direct Micro Kit (Invitrogen, 61021) and RevertAid H Minus First strand cDNA Synthesis Kit (Thermo Fisher Scientific, K1632) following the manufacturer's instructions. Quantitative PCR was performed in a Thermal Cycler (BioRad C1000 Touch, CFX384 Real-Time system). The reaction was performed using 1X Maxima SYBR Green/Rox qPCR Master Mix (Thermo Fisher Scientific, K0223), 0.3 μM of each primer (for primer sequences see **Supplementary Table 2**) and 1/10 synthesized cDNA (for NS20Y and 1/2 for Neurons) for a total volume of 6.25 μL. The qPCR conditions were as follow: 2 min at 50°C, 10 min at 95°C, 40 cycles of 15 s at 95°C and 1 min at 59°C. mRNA quantification was performed by real-time RT-PCR using the  $\Delta\Delta C_t$ -method. Quantification was based on *synaptophysin* (Chen et al., 2001).

### Protein Extraction and Western Blotting

Transduced cortical neurons (4,20,000 cells) were lysed in RIPA buffer (150 mM sodium chloride, 1% NP40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris-HCl, pH8, protease inhibitor 1X, and phosphatase inhibitor 1X), loaded on 8% SDS polyacrylamide gels and transferred onto nitrocellulose membranes. Blots were blocked in 5% milk for 1 h and incubated overnight with a primary antibody against Cav3.2 (1:200, Sigma C1868). Following washing steps in PBS-T (0.1% Tween) the membrane was incubated with IRDye800 Goat anti-Rabbit (1:10,000, LI-COR Biosciences) for 45 min, washed and imaged. The membrane was subsequently incubated with anti-Tubulin antibody (1:5,000, ab6160) for 2 h at RT, washed and incubated with IRDye680 Goat anti-Rat (1:10,000, LI-COR Biosciences). Bands were detected with infrared Odyssey system (LI-COR Biosciences) and quantified using the software Image Studio Lite.

### Mouse Organotypic Slices

Organotypic hippocampal slices were prepared from mice (C57Bl6/N) at postnatal day 3–6 as described before (Biermann et al., 2014). In brief, the hippocampus was isolated and cut in 350 μM thickness using a McIlwain tissue chopper. The slices were cultured in 6 well plates containing cell culture inserts 0.4 μM, 30 mm (Millipore) and medium (50% Neurobasal medium, 25% Hank's balanced salt solution (without MgCl<sub>2</sub>, without CaCl<sub>2</sub>), 25% horse serum, 0.65% D(+)-glucose, 0.01 M Hepes, 2 mM L-glutamine and 0.5X B27). Cultured slices were kept at 37°C, 5% CO<sub>2</sub>.

### Cell Imaging

Images were obtained using an inverted phase contrast fluorescent microscope (Zeiss Axio Observer A1 with objectives 20X, LD A-Plan and 5X, Fluor) and processed using Fiji. The cell surface area, set by the GFP fluorescence, was determined using the Weka trainable segmentation plugging. The integrated density, determined by the mRuby fluorescence was normalized to the cell surface area to obtain the reported values of integrated density/cell surface (IntDE/cell surface).

## Statistical Analyses

Statistical analyses were performed using GaphPad software. One sample *t*-test, Student *t*-test, and two way ANOVA followed by multiple comparison tests were used to compare significance of the results. Values were considered significant a  $p < 0.05$ . For all graphs data are displayed as mean  $\pm$  SEM.

## RESULTS

To modulate the promoter activity of the *Cacna1h* gene, we first designed sgRNAs sequences targeting the mouse and rat *Cacna1h* gene. Using the Benchling computational tool, we selected two sgRNAs binding 66 (sgRNA1) and 131 (sgRNA2) base pairs (bp) upstream of the start-ATG of the *Cacna1h* gene (Figure 1A). Next, we tested the efficiency of the two sgRNAs to target the *Cacna1h* promoter and regulate its activity in neuroblastoma NS20Y cells. For this, we transfected the two sgRNAs together with (i) a minimal reporter unit expressing mRuby under control of the rat *Cacna1h* promoter, (ii) a ubiquitous promoter expressing eGFP, and (iii) either a CRISPRa construct [dCas9-VPR, Chavez et al. (2015)] or a CRISPRi construct [dCas9-KRAB, (Yeo et al., 2018)] in NS20Y cells and analyzed the fluorescence intensity 2 days after transfection (Figures 1B,C). A strong activation of the *Cacna1h* promoter was observed after co-transfection with the CRISPRa construct and an inhibition after co-transfection with the CRISPRi construct (Figure 1C). No modulation was observed for the ubiquitous CAG-eGFP construct, indicating that the sgRNAs only affected the activity of the *Cacna1h* promoter (Figure 1C).

Next, we compared the modulatory effects of the two sgRNAs in both the CRISPRa and CRISPRi systems. A significant increase in *Cacna1h*-mRuby fluorescence activity was observed for the two single sgRNAs in the CRISPRa system (mean  $\pm$  SEM = sgRNA1:  $3.41 \pm 0.26$ -fold increase,  $p = 0.0026$ ; sgRNA2:  $2.93 \pm 0.16$ -fold increase,  $p = 0.0012$ ), and reached similar levels as observed for the combination of the two sgRNAs (Figure 1D, left panel; mean  $\pm$  SEM = sgRNA1 + 2:  $3.60 \pm 0.51$ -fold increase,  $p = 0.015$ ). Also for the CRISPRi system, comparable modulatory effects were observed for sgRNA1 (mean  $\pm$  SEM =  $0.52 \pm 0.078$ -fold decrease,  $p = 0.0088$ ), sgRNA2 ( $0.51 \pm 0.046$ -fold decrease,  $p = 0.0017$ ) and the combination of the two sgRNAs ( $0.52 \pm 0.10$ -fold decrease,  $p = 0.017$ ; Figure 1D, middle panel). A control sgRNA targeting *LacZ* did not have an effect on *Cacna1h* fluorescence intensity in the CRISPRa or CRISPRi system (Figure 1D, right panel), indicating that the modulatory effects observed for the two *Cacna1h*-sgRNAs were highly specific.

To confirm and precisely quantify the *Cacna1h*-specific CRISPRa and CRISPRi regulation, we next exchanged the mRuby reporter for a luciferase reporter gene (Figure 1E, left panel). As expected, transfection of the CRISPRa and CRISPRi components into NS20Y cells, resulted in an increase of luciferase activity for the CRISPRa system (Figure 1E, middle panel; sgRNA1:  $5.44 \pm 0.84$ -fold increase,  $p = 0.034$ ; sgRNA2:  $4.79 \pm 0.80$ -fold increase,  $p = 0.042$ ; sgRNA1 + 2:  $6.75 \pm 0.45$ -fold increase,  $p = 0.006$ ) and a decrease for the CRISPRi system (Figure 1E,

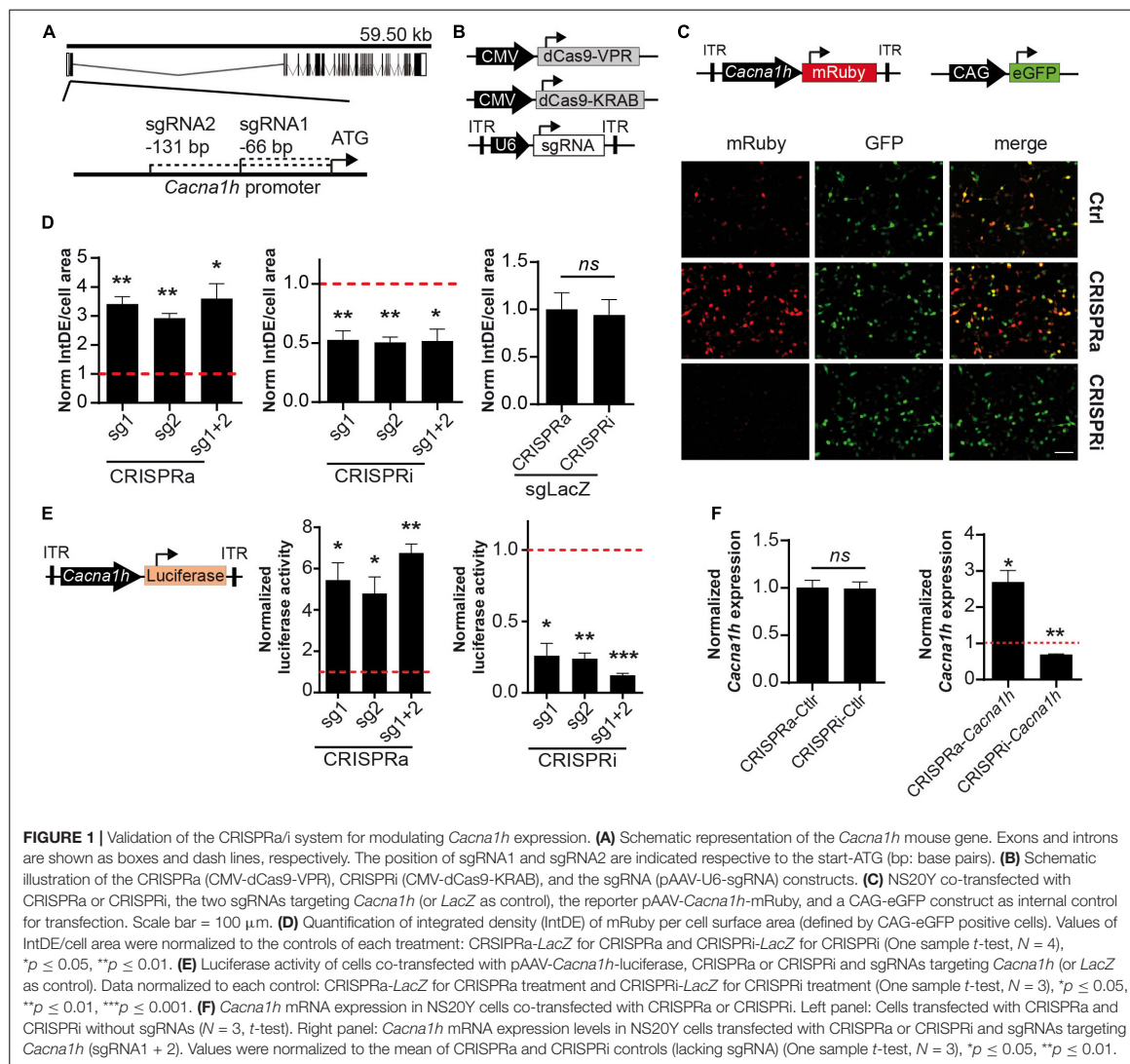
right panel; sgRNA1:  $0.26 \pm 0.087$ -fold decrease,  $p = 0.013$ ; sgRNA2:  $0.24 \pm 0.039$ -fold decrease,  $p = 0.0026$ ; sgRNA1 + 2:  $0.12 \pm 0.014$ -fold decrease,  $p = 0.0003$ ). Also here, no significant differences were observed between the two sgRNAs or the combination of the two sgRNAs.

We next examined the endogenous *Cacna1h* mRNA expression levels after manipulation with the CRISPRa and CRISPRi systems using quantitative real-time RT-PCR. Since both sgRNAs showed similar effects on the *Cacna1h* reporter constructs (Figures 1D,E) we decided to use the combination of the two sgRNAs for this experiment. Activation of the system in NS20Y cells, resulted in augmentation of endogenous *Cacna1h* mRNA expression levels, whereas inhibition significantly decreased the *Cacna1h* mRNA expression levels (Figure 1F, right panel; CRISPRa:  $2.69 \pm 0.33$ -fold increase,  $p = 0.036$ ; CRISPRi:  $0.69 \pm 0.018$ -fold decrease,  $p = 0.0034$ ). No changes in *Cacna1h* expression were observed for the CRISPRa and CRISPRi controls (no sgRNA; Figure 1F, left panel). Altogether, these results confirmed that the present *Cacna1h*-CRISPR modulatory toolbox successfully can activate or inhibit endogenous *Cacna1h* expression in NS20Y cells.

We then probed whether the modulatory effects observed in NS20Y cells could also be observed in primary cultured neurons. For this, we transduced primary mouse cortical neurons at DIV4 with all-in-one CRISPRa/i lentiviruses (Figure 2A) and measured endogenous *Cacna1h* mRNA expression levels at DIV15 and protein levels at DIV20. Since the two sgRNAs displayed similar effects in NS20Y cells (Figures 1D,E), we decided to proceed with only one sgRNA (sgRNA2) to minimize unspecific effects in the neuronal cultures. Intriguingly, a strong activation was observed after transduction with CRISPRa lentiviruses both at the mRNA (Figure 2B, left panel,  $4.37 \pm 0.37$ -fold increase,  $p = 0.0016$ ) and at the protein level (Figure 2C,  $2.18 \pm 0.42$ -fold increase,  $p = 0.0046$ ). In addition, also inhibition of the system using CRISPRi in neuronal cultures resulted in a reduced *Cacna1h* expression at the mRNA (Figure 2B, right panel,  $0.12 \pm 0.049$ -fold change,  $p = 0.032$ ) as well as at the protein level (Figure 2D,  $0.39 \pm 0.14$ -fold change,  $p = 0.042$ ), indicating that sgRNA2 can efficiently modulate *Cacna1h* promoter activity in primary neurons. Although we observed additional non-specific bands in our Western blot experiments (Supplementary Figure 1), only the band of approximately 260 kDa corresponding to Cav3.2 increases or decreases in intensity after treatment with *Cacna1h*-CRISPRa and *Cacna1h*-CRISPRi, respectively.

To prove unequivocally that the *Cacna1h*-CRISPRa/i toolbox is specific for *Cacna1h* modulation, we next examined the mRNA expression levels of other calcium channel family members (*Cacna1g*, *Cacna1i*, and *Cacna1e*) after *Cacna1h*-CRISPRa/i targeting using quantitative real-time RT-PCR. No alterations were observed for any of the other calcium channels tested (Figure 2E), indicating that our newly developed *Cacna1h*-CRISPRa/i toolbox is highly specific for *Cacna1h* modulation.

In order to test if our system has the potential to be delivered in neuronal network structures *in vivo* we infected mouse organotypic hippocampal slices with recombinant adeno-associated (rAAV) and lentiviruses encoding our *Cacna1h*-CRISPRa toolbox and a fluorescent reporter. After

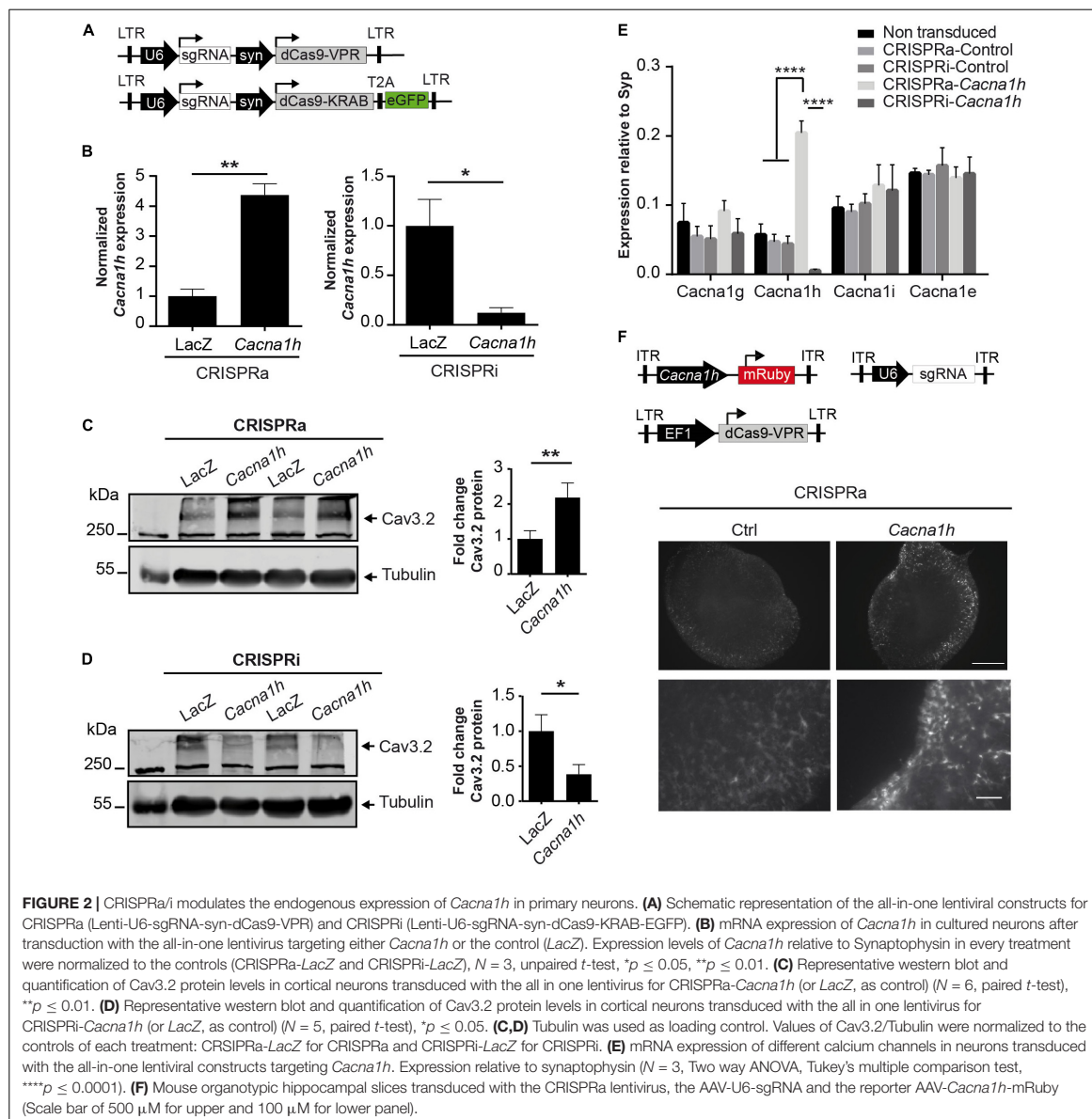


co-transduction with AAV-sgRNA, lenti-EF1a-dCas9-VPR (Savell et al., 2019) and the reporter AAV-*Cacna1h*-mRuby we observed a stronger signal of the red fluorescent protein when using the sgRNA targeting *Cacna1h* than in the control indicating an activation of the *Cacna1h* promoter *in vivo* (Figure 2F).

The *Cacna1h*-CRISPR toolbox we describe here was made for constitutive activation or inhibition of the *Cacna1h* transcripts. We next adapted the system for induced gene expression control by incorporating doxycycline-controlled Tet-On gene expression systems (Colasante et al., 2019; Zhang et al., 2019). In this way, *Cacna1h* expression can be activated or inhibited at a precise temporal resolution. We tested the system in primary cortical neurons with a combination of viruses that allow doxycycline-inducible CRISPR-*Cacna1h* activation (denoted

as TET-ON-CRISPRa-*Cacna1h*; Figure 3A) or doxycycline-inducible CRISPR-*Cacna1h* inhibition (denoted as TET-ON-CRISPRi-*Cacna1h*; Figure 3B). Doxycycline was administered 4 days after viral transduction and mRNA was analyzed at DIV15. Interestingly, we observed an increase in *Cacna1h* expression after transduction with the TET-ON-CRISPRa-*Cacna1h* system following doxycycline treatment (Figure 3C;  $1.85 \pm 0.12$ -fold change,  $p = 0.0015$ ), and a decrease in *Cacna1h* expression after TET-ON-CRISPRi-*Cacna1h* transduction and doxycycline treatment (Figure 3D;  $0.28 \pm 0.11$ -fold change,  $p = 0.027$ ). No effect on *Cacna1h* expression was observed when using a control sgRNA targeting *LacZ* (Supplementary Figure 2). We thus present a proof of principle that *Cacna1h* expression can be activated or inhibited at a precise temporal resolution.



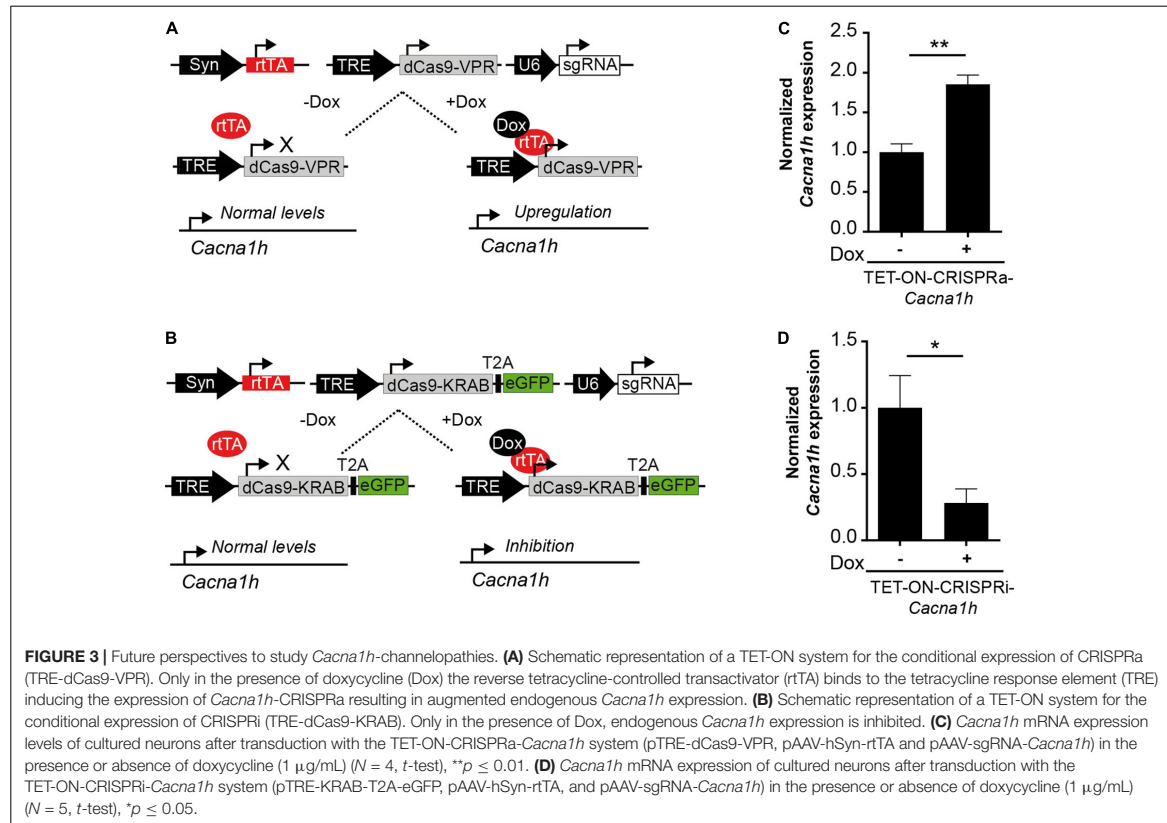


## DISCUSSION

Detailed analyses of the dynamic contribution of individual ion channels in the context of neuronal network function remains challenging, particularly in the context of mutational and transcriptional “channelopathies,” in which for example the fine-tuned regulation of mRNA expression levels of affected genes, including *Cacna1h* plays a major role. Traditionally, the functional characterization of individual ion channels in terms of gain- and loss-of-function approaches has been addressed by the

exogenous overexpression of genetically encoded proteins and by knock-down systems *via* RNA interference (RNAi)/genetic gene ablation (knock-out), respectively. Although very valuable, the applicability of these techniques has limitations (Kampmann, 2018).

Here, we present a modular method using a CRISPRa/i system that allows to manipulate the endogenous expression of *Cacna1h* *in vitro* and *in vivo*. This CRISPRa/i system applies specific sgRNAs directed to the *Cacna1h* promoter and a dCas9 fused to the transcriptional activator VPR or the transcriptional



inhibitor KRAB (Gilbert et al., 2013; Chavez et al., 2015) to induce changes in the endogenous expression of the *Cacna1h* gene. We provide evidence that both the CRISPRa and CRISPRi systems can specifically manipulate the endogenous expression of *Cacna1h* in dividing cells as well as in primary neuronal cultures. In addition, we demonstrate the possibility to activate the *Cacna1h* promoter in mouse organotypic hippocampal slices. Interestingly, the fold-change induction by CRISPRa occurred within a range also observed for Cav3.2 channelopathies (Becker et al., 2008), making our *Cacna1h*-CRISPR system highly suitable for analyzing this particular and also other channelopathies at the functional level.

Increasing the gene expression levels of *Cacna1h* in a high percentage of cultured cells, which is required for many downstream analyses, or *in vivo* requires the application of viral transduction systems. Here, our *Cacna1h*-CRISPRa toolbox has a clear advantage over conventional virally mediated overexpression approaches, where the limiting packaging capacity of the commonly used and easily applicable recombinant adeno-associated- and lenti-viruses, excludes the possibility to overexpress large proteins like *Cacna1h*. On the other hand, recombinant viruses with a higher packaging capacity, like Adeno- and Herpes Simplex virus, cannot be easily generated in the lab and are expensive to purchase. Since augmentation of proteins using the CRISPRa system is independent of their

transcript size, even the large *Cacna1h* gene (mRNA transcript length up to 8.2 Kb) can be easily augmented using rAAV and lentiviruses, which can be produced in most molecular biology labs. Another advantage of the *Cacna1h*-CRISPRa system over conventional gain-of-function approaches is the endogenous modulation of the *Cacna1h* genomic locus, providing the opportunity to study the effects of abundant alternative splicing variants of *Cacna1h*. For *Cacna1h*, at least 12–14 alternative splicing sites have been described, resulting in the possibility to generate more than 4,000 alternative *Cacna1h* transcripts (Zhong et al., 2006). Such a complexity cannot be achieved by conventional overexpression of only one variant of *Cacna1h*. Also in terms of loss-of-function approaches, our *Cacna1h*-CRISPRi has clear advantages over conventional shRNA/siRNA approaches. A siRNA approach for *Cacna1h* in mouse embryonic stem cells reduced *Cacna1h* expression levels to approximately 40% but also caused a non-specific decrease in *Cacna1g* mRNA expression levels (Rodríguez-Gómez et al., 2012). By interfering directly with gene promoters, CRISPRa/i modifies corresponding transcript abundance in a way that strongly recapitulates physiological promoter regulation – compared to the above mentioned molecular manipulations, which typically target one specific mRNA variant only. In contrast, CRISPRa/i leads to abundance changes of the respective RNA, which will then be subjected to alternative splicing. Thereby, the abundance

of the full complement of alternatively spliced variants will be increased.

The fact that we did not observe any alterations in the mRNA expression levels of the calcium channel family members *Cacna1g*, *Cacn1i*, and *Cacna1e* clearly indicates that our newly developed *Cacna1h*-CRISPRa/i toolbox is highly specific for *Cacna1h* modulation. Importantly, since the catalytic inactive Cas9 (dCas9) used in our system does not cleave the DNA, the possible off-target effects are more likely to be less deleterious than using the conventional Cas9 to induce a genetic knock-out (Colasante et al., 2020). Compared to genetic knock-out approaches, the *Cacna1h*-CRISPRa/i toolbox, which is based on viral-transduction, allows to decrease *Cacna1h* expression in selected neuron types, by putting dCas expression under the control of cell type specific promoters, in localized brain regions and under temporal control, without time consuming breeding of animals.

Our Tet-On *Cacna1h*-CRISPR toolbox system might open new roads for investigating the role of *Cacna1h* in several channelopathies in more detail. Our present study has some limitations. Firstly, we did not test the present CRISPRa/i systems parallel to cell culture in an *in vivo* approach. However, AAV-constructs acting properly in cultured neurons have been successfully translated into *in vivo* applications (Van Loo et al., 2012, 2015). Secondly, we could not scrutinize the effects of CRISPRa/i on a functional level such as by T-type  $Ca^{2+}$ -current density or neuronal discharge analyses. Although neuronal cultures constitute a well-known system for studying electrophysiological properties, measuring calcium currents is challenging, especially after modulation of the system. However, all of our previous studies indicate that cellular changes of Cav3.2 mRNA expression are reflected on the protein level and by changes of current densities and neuronal discharge behavior (Su et al., 2002; Becker et al., 2008; Van Loo et al., 2015).

Also in the context of mutational/genetic ion channelopathies, the potential of inducible *Cacna1h*-CRISPR approaches are immense. For example, in the Genetic Absence Epilepsy Rat from Strasburg (GAERS) a mutation in *Cacna1h* has been described, which augments the expression of Cav3.2 at the cell surface and increases calcium influx (Proft et al., 2017). Here, the *Cacna1h*-CRISPRa could be used as a gain-of-function approach to mimic the enhanced expression seen in GAERS. In addition, a transient inhibition of *Cacna1h* via the Tet-on inducible *cacna1h*-CRISPRi system in the GAERS rats at different stages during development could also reveal potentially interesting time windows for *Cacna1h*-associated channelopathies.

## REFERENCES

- Becker, A. J., Pitsch, J., Sochivko, D., Opitz, T., Staniek, M., Chen, C. C., et al. (2008). Transcriptional upregulation of Cav3.2 mediates epileptogenesis in the pilocarpine model of epilepsy. *J. Neurosci.* 28, 13341–13353. doi: 10.1523/JNEUROSCI.1421-08.2008
- Biermann, B., Sokoll, S., Klueva, J., Missler, M., Wiegert, J. S., Sibarita, J. B., et al. (2014). Imaging of molecular surface dynamics in brain slices using single-particle tracking. *Nat. Commun.* 5:3024. doi: 10.1038/ncomms4024

Taken together, our here described *Cacna1h*-CRISPRa/i modular approach could thus be used to model transient gain-of-function or loss-of-function effects in the *Cacna1h* gene and to study Cav3.2 channelopathies in more detail *in vivo*.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

The animal study was reviewed and approved by the LANUV Recklinghausen, Leibnizstr. 10, 45659 Recklinghausen.

## AUTHOR CONTRIBUTIONS

DT, SS, AB, and KL conceived and planned the study. DT and AT carried out the molecular and cellular experiments. HS, TO, and SS contributed to the study design and analysis. DT wrote the initial draft. AB, SS, and KL contributed to the review and editing. All authors contributed to the article and approved the submitted version.

## FUNDING

Our work was supported by the Deutsche Forschungsgemeinschaft (SFB 1089 to AB, SS, and KL, FOR 2715 to AB and HS, SCHO 820/7-2; SCHO 820/5-2; SCHO 820/6-1; SCHO 820/4-1; SCHO 820 5-2 to SS) and BONFOR. AT was supported by the Humboldt Foundation.

## ACKNOWLEDGMENTS

We thank Sabine Optiz for excellent technical assistance.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnmol.2021.667143/full#supplementary-material>

- Cannon, S. C. (2007). Physiologic Principles Underlying Ion Channelopathies. *Neurotherapeutics* 4, 174–183. doi: 10.1016/j.nurt.2007.01.015
- Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., Iyer, E. P. R., et al. (2015). Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* 12, 326–328. doi: 10.1038/nmeth.3312
- Chen, J., Sochivko, D., Beck, H., Marechal, D., Wiestler, O. D., and Becker, A. J. (2001). Activity-induced expression of common reference genes in individual CNS neurons. *Lab. Invest.* 81, 913–916. doi: 10.1038/labinvest.3780300

- Chhipa, R. R., Fan, Q., Anderson, J., Muraleedharan, R., Huang, Y., Ciruolo, G., et al. (2018). AMP kinase promotes glioblastoma bioenergetics and tumour growth. *Nat. Cell Biol.* 20, 823–835. doi: 10.1038/s41556-018-0126-z
- Cloud-Based Informatics Platform for Life Sciences R&D | Benchling (2021). Available online at: <https://www.benchling.com/> [Accessed February 10, 2021].
- Colasante, G., Lignani, G., Brusco, S., Di Berardino, C., Carpenter, J., Giannelli, S., et al. (2019). dCas9-Based Scn1a Gene Activation Restores Inhibitory Interneuron Excitability and Attenuates Seizures in Dravet Syndrome Mice. *Mol. Ther.* 28, 235–253. doi: 10.1016/j.ythme.2019.08.018
- Colasante, G., Qiu, Y., Massimino, L., Di Berardino, C., Cornford, J. H., Snowball, A., et al. (2020). In vivo CRISPRa decreases seizures and rescues cognitive deficits in a rodent model of epilepsy. *Brain* 143, 891–905. doi: 10.1093/brain/awaa045
- Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451. doi: 10.1016/j.cell.2013.06.044
- Heeringa, S. F., Möller, C. C., Du, J., Yue, L., Hinkes, B., Chernin, G., et al. (2009). A novel TRPC6 mutation that causes childhood FSGS. *PLoS One* 4:e7771. doi: 10.1371/journal.pone.0007771
- Hutchings, C. J., Colussi, P., Parker, D. B., and Clark, T. G. (2019). Ion channels as therapeutic antibody targets. *MABS* 11, 265–296. doi: 10.1080/19420862.2018.1548232
- Kampmann, M. (2018). CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chem. Biol.* 13, 406–416. doi: 10.1021/acscchembio.7b00657
- Khosravani, H., Altier, C., Simms, B., Hamming, K. S., Snutch, T. P., Mezeyova, J., et al. (2004). Gating Effects of Mutations in the Cav3.2 T-type Calcium Channel Associated with Childhood Absence Epilepsy. *J. Biol. Chem.* 279, 9681–9684. doi: 10.1074/jbc.C400006200
- Khosravani, H., Bladen, C., Parker, D. B., Snutch, T. P., McRory, J. E., and Zamponi, G. W. (2005). Effects of Cav3.2 channel mutations linked to idiopathic generalized epilepsy. *Ann. Neurol.* 57, 745–749. doi: 10.1002/ana.20458
- Kügler, S., Kilic, E., and Bähr, M. (2003). Human synapsin 1 gene promoter confers highly neuron-specific long-term transgene expression from an adenoviral vector in the adult rat brain depending on the transduced area. *Gene Ther.* 10, 337–347. doi: 10.1038/sj.gt.3301905
- Mizuno, H., Luo, W., Tarusawa, E., Saito, Y. M., Sato, T., Yoshimura, Y., et al. (2014). NMDAR-regulated dynamics of layer 4 neuronal dendrites during thalamocortical reorganization in neonates. *Neuron* 82, 365–379. doi: 10.1016/j.neuron.2014.02.026
- Powell, K. L., Cain, S. M., Ng, C., Sirdesai, S., David, L. S., Kyi, M., et al. (2009). A Cav3.2 T-type calcium channel point mutation has splice-variant-specific effects on function and segregates with seizure expression in a polygenic rat model of absence epilepsy. *J. Neurosci.* 29, 371–380. doi: 10.1523/JNEUROSCI.5295-08.2009
- Proft, J., Rzhpetskiy, Y., Lazniewska, J., Zhang, F. X., Cain, S. M., Snutch, T. P., et al. (2017). The Cacna1h mutation in the GAERS model of absence epilepsy enhances T-type Ca<sub>2+</sub> currents by altering calnexin-dependent trafficking of Cav3.2 channels. *Sci. Rep.* 7:11513. doi: 10.1038/s41598-017-11591-5
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308. doi: 10.1038/nprot.2013.143
- Rodríguez-Gómez, J. A., Levitsky, K. L., and López-Barneo, J. (2012). T-type Ca<sub>2+</sub> channels in mouse embryonic stem cells: modulation during cell cycle and contribution to self-renewal. *Am. J. Physiol.* 302, C494–C504. doi: 10.1152/ajpcell.00267.2011
- Ryan, D. P., Dias da Silva, M. R., Wah Soong, T., Fontaine, B., Donaldson, M. R., Kung, A. W. C., et al. (2010). Mutations in a potassium channel (Kir2.6) causes susceptibility to thyrotoxic hypokalemic periodic paralysis. *Cell* 140, 88–98. doi: 10.1016/j.cell.2009.12.024
- Rzhpetskiy, Y., Lazniewska, J., Blesneac, I., Pamphlett, R., and Weiss, N. (2016). CACNA1H missense mutations associated with amyotrophic lateral sclerosis alter Cav3.2 T-type calcium channel activity and reticular thalamic neuron firing. *Channels* 10, 466–477. doi: 10.1080/19336950.2016.1204497
- Savell, K. E., Bach, S. V., Zipperly, M. E., Revanna, J. S., Goska, N. A., Tuscher, J. J., et al. (2019). A neuron-optimized CRISPR/dCas9 activation system for robust and specific gene regulation. *Eneuro* 6, ENEURO.0495–18.2019. doi: 10.1523/ENEURO.0495-18.2019
- Souza, I. A., Gandini, M. A., Wan, M. M., and Zamponi, G. W. (2016). Two heterozygous Cav3.2 channel mutations in a pediatric chronic pain patient: recording condition-dependent biophysical effects. *Pflugers Arch. Eur. J. Physiol.* 468, 635–642. doi: 10.1007/s00424-015-1776-3
- Souza, I. A., Gandini, M. A., Zhang, F. X., Mitchell, W. G., Matsumoto, J., Lerner, J., et al. (2019). Pathogenic Cav3.2 channel mutation in a child with primary generalized epilepsy. *Mol. Brain* 12:86. doi: 10.1186/s13041-019-0509-5
- Splawski, I., Yoo, D. S., Stotz, S. C., Cherry, A., Clapham, D. E., and Keating, M. T. (2006). CACNA1H mutations in autism spectrum disorders. *J. Biol. Chem.* 281, 22085–22091. doi: 10.1074/jbc.M603316200
- Su, H., Sochivko, D., Becker, A., Chen, J., Jiang, Y., Yaari, Y., et al. (2002). Upregulation of a T-Type Ca<sub>2+</sub> Channel Causes a Long-Lasting Modification of Neuronal Firing Mode after Status Epilepticus. *J. Neurosci.* 22, 3645–3655. doi: 10.1523/jneurosci.22-09-03645.2002
- Thakore, P. I., D'Ippolito, A. M., Song, L., Safi, A., Shivakumar, N. K., Kabadi, A. M., et al. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* 12, 1143–1149. doi: 10.1038/nmeth.3630
- Van Loo, K. M. J., Rummel, C. K., Pitsch, J., Müller, J. A., Bikbaev, A. F., Martinez-Chavez, E., et al. (2019). Calcium channel subunit  $\alpha 2\delta 4$  is regulated by early growth response 1 and facilitates epileptogenesis. *J. Neurosci.* 39, 3175–3187. doi: 10.1523/JNEUROSCI.1731-18.2019
- Van Loo, K. M. J., Schaub, C., Pernhorst, K., Yaari, Y., Beck, H., Schoch, S., et al. (2012). Transcriptional regulation of T-type calcium channel Ca<sub>v</sub>3.2: bidirectionality by early growth response 1 (Egr1) and repressor element 1 (RE-1) protein-silencing transcription factor (REST). *J. Biol. Chem.* 287, 15489–15501. doi: 10.1074/jbc.M111.310763
- Van Loo, K. M. J., Schaub, C., Pitsch, J., Kulbida, R., Opitz, T., Ekstein, D., et al. (2015). Zinc regulates a key transcriptional pathway for epileptogenesis via metal-regulatory transcription factor 1. *Nat. Commun.* 6, 1–12. doi: 10.1038/ncomms9688
- Wang, Q., Curran, M. E., Splawski, I., Burn, T. C., Millholland, J. M., VanRaay, T. J., et al. (1996). Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias. *Nat. Genet.* 12, 17–23. doi: 10.1038/ng10196-17
- Woitecki, A. M. H., Müller, J. A., van Loo, K. M. J., Sowade, R. F., Becker, A. J., and Schoch, S. (2016). Identification of synaptotagmin 10 as effector of NPAS4-mediated protection from excitotoxic neurodegeneration. *J. Neurosci.* 36, 2561–2570. doi: 10.1523/JNEUROSCI.2027-15.2016
- Yeo, N. C., Chavez, A., Lance-Byrne, A., Chan, Y., Menn, D., Milanova, D., et al. (2018). An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat. Methods* 15, 611–616. doi: 10.1038/s41592-018-0048-5
- Zhang, J., Chen, L., Zhang, J., and Wang, Y. (2019). Drug Inducible CRISPR/Cas Systems. *Comput. Struct. Biotechnol. J.* 17, 1171–1177. doi: 10.1016/j.csbj.2019.07.015
- Zhong, X., Liu, J. R., Kyle, J. W., Hanck, D. A., and Agnew, W. S. (2006). A profile of alternative RNA splicing and transcript variation of CACNA1H, a human T-channel gene candidate for idiopathic generalized epilepsies. *Hum. Mol. Genet.* 15, 1497–1512. doi: 10.1093/hmg/ddl068

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Tsourtouktzidis, Tröscher, Schulz, Opitz, Schoch, Becker and van Loo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## 4. Discussion

### 4.1 Allelic variants of rSNP confer differential gene expression in epileptic hippocampi

A main finding of the present thesis was given by the fact that two genes involved in monogenic epilepsy contain rSNPs that influence gene abundance in acquired forms of epilepsies (TLE). Considering the role of those genes in the CNS, our results support the idea of a trait-associated genetic vulnerability.

Patients homozygous for the minor or less frequent promoter haplotype presented lower abundant mRNA expression levels of *ALDH5A1* than patients homozygous for the major haplotype. Considering (i) that the lack of *ALDH5A1* leads to an imbalance in GABA recycling, (ii) that mutations in *ALDH5A1* are causative of SSADH where half of the affected individuals present with epilepsy (Akaboshi et al., 2003), and (iii) that the *ALDH5A1* deletion mouse model develops seizures (Hogema et al., 2001), it appears possible that the lower abundance of *ALDH5A1* transcripts given by the minor promoter haplotype might confer an increased risk for developing epilepsy related to an allelic variant. Although no overrepresentation of the minor haplotype was present in the TLEHS cohort in comparison to a control group, potentially explained by a still limited number of patients included in the study, our complementary luciferase assay supports that the SNP affects promoter activity.

An additional rSNP included in this work locates in an intron of *SCN1A*, also a key relevant gene in monogenic epilepsies. A previous GWAS study showed an association between the minor allele and the pathology MTLEHS and a history of FS (Kasperavičiute et al., 2013). The same study reported a lack of association between genotype and expression levels of *SCN1A* in the neocortex (Kasperavičiute et al., 2013). We found that hippocampi of MTLEHS patients homozygous for the minor allele presented with higher *SCN1A* expression levels compared to homozygous patients for the major allele, suggesting a tissue-specific effect. Interestingly, our observations suggest that higher levels of *SCN1A* might associate with TLE, whereas it is known that loss-of-function variants are the causative factor in cases of Dravet syndrome (Ding et al., 2021). Importantly, we observed that overexpressing *scn1a* in zebrafish leads to spontaneous seizures, suggesting that abundance or gain-of-function of *scn1a* also associates with seizure events. TLE affects mostly adults while Dravet syndrome manifests in infancy. Understanding the role of

*SCN1A* in both pathologies in terms of loss and gain of function requires detailed knowledge of the temporal expression of *SCN1A*.

In contrast to our study on *ALDH5A1*, where the SNP haplotype lies in the core promoter of the gene (Pernhorst et al., 2011), the *SCN1A* candidate SNP locates in a genomic region with intronic architectural characteristics. Intriguingly, a 50 bp SNP-containing region depicted a strong promoter-like activity that may be compatible with the functional characteristics of an enhancer element. However, differential luciferase activity driven by allelic variants of *SCN1A* was not present in *in vitro* assays. Although reporter assays are well accepted as a complementary method in identifying differences in allelic expression they have several limitations: only a limited DNA fragment is analyzed, mechanisms that involve other genomic sequences are excluded (Buckland, 2006); also the environmental context from cell lines might differ from the *in vivo* scenario. Besides, one cannot rule out the possibility that another SNP in total LD with the one studied is the causative or functional variant (Buckland, 2006).

#### 4.2 Mechanisms that mediate transcriptional differences associated with SNPs

Key TFs have been identified, which control dynamic ion channel expression and function in experimental models of epilepsy, data which is supported in biopsy tissue of pharmacoresistant TLE patients (McClelland et al., 2011; van Loo et al., 2015). A well described example is the metal regulatory transcription factor 1 (MTF1). MTF1 is increased in expression soon after pilocarpine-induced SE and precedes the augmentation of Cav3.2 mRNA and corresponding protein levels (van Loo et al., 2015). In the context of a 'genetic predisposition' to a TF-mediated promoter activation response, it is known that one of the molecular mechanisms by which rSNP may confer distinct allelic expression is by altering a transcription factor binding site (TFBS) sequence and therefore the binding affinity of the TF (Ramírez-Bello et al., 2013). We found that Egr1 increases the activity of the *ALDH5A1* promoter and that the minor haplotype presented a less-sensitive response to changes mediated by Egr1. Considering that Egr1 is a regulator of epileptogenesis (van Loo et al., 2019) we suggest a potential mechanism through this TF for *ALDH5A1* promoter regulation in epilepsy. Thus, *ALDH5A1* regulation may be part of the concerted epileptogenic effects of Egr1.

In contrast to *ALDH5A1*, we did not find a TF involved in the regulation of *SCN1A* SNP variants. Although we observed that the TF SOX2 physically binds to the genomic region containing the SNP and that it increases luciferase activity, we did not find differences between allelic variants. A recent study suggests that by far not all transcription factor binding motifs are described; this concept is based on the fact that only a small proportion (~30 %) of functional variants fall within consensus TF binding sequences (Buckland et al., 2005). In line with this idea, our EMSA experiment using nuclear protein extract from mouse brain, an experimental context that resembles the biological complexity of the in vivo situation, suggested that other transcription factors interact with the genomic region containing the *SCN1A* SNP under study.

The rSNPs reported here influence gene promoter activity and corresponding mRNA abundance in hippocampi under epileptic conditions. Importantly, we do not claim a causative relationship between the rSNPs and the pathology but rather a possible contribution. Of note, the obvious lack of availability of brain tissue from healthy individuals for control purposes is a limitation to clarify if the observed allelic expression differences are epilepsy specific. Furthermore, replicative studies will be of great value to identify the SNPs as potential biomarkers.

#### 4.3 CRISPRa/i for identifying, validating, and functionally assessing regulatory elements of the genome

Despite their shared role as regulatory relevant SNPs, the localization of the analyzed SNPs within the gene architecture strongly differs; they locate at the promoter for *ALDH5A1* and in an intron for *SCN1A*. In comparison to SNPs within experimentally validated promoters, elucidating the impact of SNPs lying in other non-coding regions is an extra challenge. Their study might require additional technologies to clarify if the SNP-containing region is functionally relevant. Thus, in an attempt to study if the recently developed CRISPR toolbox for transcriptional modulation can be used for the validation of non-coding genomic relevant regions, we focused on the *Cacna1h* promoter, which has been structurally and functionally analyzed. In a study van Loo, et al. investigated several genomic fragments of different lengths upstream of the starting ATG of *Cacna1h* and identified a genomic segment (105 bp) as the putative core promoter (van Loo et al., 2012). Remarkably, our findings showed that delivering sgRNAs within this sequence

results in a transcriptional modulation by both approaches CRISPRa and CRISPRi, thereby validating the significance of this genomic area and the importance of understanding promoter architecture as a precondition for a successful manipulation.

Compared to classical approaches, where functional characterization of promoters mostly relied on reporter assays, multiple cloning, and evaluation of individual genomic fragments, CRISPR has clear advantages. For instance, (i) using different sgRNAs can more effectively identify regulatory regions while avoiding multiple cloning, (ii) the effect is directly produced in the endogenous genomic context, and (iii) possibly many genomic regions can be interrogated at the same time when pooling sgRNAs. Our recent data and work from others support that CRISPRa/i-screening is well suited for deciphering functional aspects of DNA motifs located in non-coding genomic regions (Ray et al., 2020). Recent studies have identified enhancers by introducing CRISPR in their pipeline (Fulco et al., 2016; Zhou et al., 2022). Strikingly, a recent study conducted a CRISPRa/i screening and found that 1 out of 32 SNPs is located in a relevant regulatory region (Zhou et al., 2022).

The CRISPRa-mediated upregulation of Cav3.2 mRNA recapitulated transcript levels present after SE in an animal model of TLE (Becker et al., 2008). This suggests the potential of CRISPR in mimicking transcriptional modulation after a brain insult.

Based on the implications of rSNPs in epilepsy and knowing the potential of CRISPR-based strategies to functionally interfere with DNA motifs located in non-coding genome areas, this work rises a prospective question: Is it possible to experimentally mimic the functional effect of a rSNP? Two factors critical in this context are given by (i) levels of CRISPRa induction, as rSNPs are generally of small effects, and (ii) allele-specific targeting. Concerning induction levels, a recent study has shown that titration of gene promoter activation mediated by CRISPRa is possible by applying different sgRNAs (Savell et al., 2019). Concerning our present work, additional sgRNAs may result in differential levels of expression of Cav3.2. Allele-specific targeting by CRISPR is still extremely challenging. However, a strategy relying on a SNP-derived protospacer adjacent motif (PAM) approach enables unique allele recognition based on the presence of a new PAM selectively at the target allele and will enable the broader application of such experiments in the future (Rabinowitz et al., 2020). If proven possible, mimicking the



effect of rSNPs in an organism would fill-up the gap in understanding genetic vulnerability to traits.

Overall, the present study contributes to a better understanding of the involvement of non-coding genomic regions in the field of epilepsy, rises potential biomarkers, and encourages the use of newly described CRISPR approaches in the field.

#### 4.4 References

Akaboshi S, Hogema BM, Novelletto A, Malaspina P, Salomons GS, Maropoulos GD, Jakobs C, Grompe M, Gibson KM. Mutational spectrum of the succinate semialdehyde dehydrogenase (ALDH5A1) gene and functional analysis of 27 novel disease-causing mutations in patients with SSADH deficiency. *Hum Mutat* 2003; 22: 442-450

Becker AJ, Pitsch J, Sochivko D, Opitz T, Staniek M, Chen CC, Campbell KP, Schoch S, Yaari Y, Beck H. Transcriptional upregulation of Cav3.2 mediates epileptogenesis in the pilocarpine model of epilepsy. *J Neurosci* 2008; 28: 13341-13353

Buckland PR. The importance and identification of regulatory polymorphisms and their mechanisms of action. *Biochim Biophys Acta* 2006; 1762: 17-28

Buckland PR, Hoogendoorn B, Coleman SL, Guy CA, Smith SK, O'Donovan MC. Strong bias in the location of functional promoter polymorphisms. *Hum Mutat* 2005; 26: 214-223

Ding J, Li X, Tian H, Wang L, Guo B, Wang Y, Li W, Wang F, Sun T. SCN1A Mutation-Beyond Dravet Syndrome: A Systematic Review and Narrative Synthesis. *Front Neurol* 2021; 12: 743726

Fulco CP, Munschauer M, Anyoha R, Munson G, Grossman SR, Perez EM, Kane M, Cleary B, Lander ES, Engreitz JM. Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* 2016; 354: 769-773

Hogema BM, Gupta M, Senephansiri H, Burlingame TG, Taylor M, Jakobs C, Schutgens RBH, Froestl W, Snead OC, Diaz-Arrastia R, Bottiglieri T, Grompe M, Gibson KM. Pharmacologic rescue of lethal seizures in mice deficient in succinate semialdehyde dehydrogenase. *Nat Genet* 2001; 29: 212-216

Kasperavičiute D, Catarino CB, Matarin M, Leu C, Novy J, Tostevin A, Leal B, Hessel EVS, Hallmann K, Hildebrand MS, Dahl HHM, Ryten M, Trabzuni D, Ramasamy A, Alhusaini S, Doherty CP, Dorn T, Hansen J, Krämer G, Steinhoff BJ, Zumsteg D, Duncan S, Kälviäinen RK, Eriksson KJ, Kantanen AM, Pandolfo M, Gruber-Sedlmayr U, Schlachter K, Reinthaler EM, Stogmann E, Zimprich F, Théâtre E, Smith C, O'Brien TJ, Meng Tan K, Petrovski S, Robbiano A, Paravidino R, Zara F, Striano P, Sperling MR, Buono RJ, Hakonarson H, Chaves J, Costa PP, Silva BM, da Silva AM, de Graan PNE, Koeleman BPC, Becker A, Schoch S, von Lehe M, Reif PS, Rosenow F, Becker F, Weber

Y, Lerche H, Rössler K, Buchfelder M, Hamer HM, Kobow K, Coras R, Blumcke I, Scheffer IE, Berkovic SF, Weale ME, UK Brain Expression Consortium, Delanty N, Depondt C, Cavalleri GL, Kunz WS, Sisodiya SM. Epilepsy, hippocampal sclerosis and febrile seizures linked by common genetic variation around SCN1A. *Brain* 2013; 136: 3140-3150

McClelland S, Flynn C, Dubé C, Richichi C, Zha Q, Ghestem A, Esclapez M, Bernard C, Baram TZ. Neuron-restrictive silencer factor-mediated hyperpolarization-activated cyclic nucleotide gated channelopathy in experimental temporal lobe epilepsy. *Ann Neurol* 2011; 70: 454-465

Pernhorst K, Raabe A, Niehusmann P, van Loo KMJ, Grote A, Hoffmann P, Cichon S, Sander T, Schoch S, Becker AJ. Promoter variants determine  $\gamma$ -aminobutyric acid homeostasis-related gene transcription in human epileptic hippocampi. *J Neuropathol Exp Neurol* 2011; 70: 1080-1088

Rabinowitz R, Almog S, Darnell R, Offen D. CrisPam: SNP-Derived PAM Analysis Tool for Allele-Specific Targeting of Genetic Variants Using CRISPR-Cas Systems. *Front Genet* 2020; 11: 851

Ramírez-Bello J, Vargas-Alarcón G, Tovilla-Zárate C, Fragoso JM. Single nucleotide polymorphisms (SNPs): functional implications of regulatory-SNP (rSNP) and structural RNA (srSNPs) in complex diseases. *Gac Med Mex* 2013; 149: 220-228

Ray JP, de Boer CG, Fulco CP, Lareau CA, Kanai M, Ulirsch JC, Tewhey R, Ludwig LS, Reilly SK, Bergman DT, Engreitz JM, Issner R, Finucane HK, Lander ES, Regev A, Hacohen N. Prioritizing disease and trait causal variants at the TNFAIP3 locus using functional and genomic features. *Nat Commun* 2020; 11: 1237

Savell KE, Bach S v., Zipperly ME, Revanna JS, Goska NA, Tuscher JJ, Duke CG, Sultan FA, Burke JN, Williams D, Ianov L, Day JJ. A Neuron-Optimized CRISPR/dCas9 Activation System for Robust and Specific Gene Regulation. *ENeuro* 2019; 6: ENEURO.0495-18

van Loo KMJ, Schaub C, Pernhorst K, Yaari Y, Beck H, Schoch S, Becker AJ. Transcriptional regulation of T-type calcium channel CaV3.2: Bi-directionality by early growth response 1 (Egr1) and repressor element 1 (RE-1) protein -silencing transcription factor (REST). *Journal of Biological Chemistry* 2012; 287: 15489-15501

van Loo KMJ, Schaub C, Pitsch J, Kulbida R, Opitz T, Ekstein D, Dalal A, Urbach H, Beck H, Yaari Y, Schoch S, Becker AJ. Zinc regulates a key transcriptional pathway for epileptogenesis via metal-regulatory transcription factor 1. *Nat Commun* 2015; 6: 8688

van Loo KMJ, Rummel CK, Pitsch J, Müller JA, Bikbaev AF, Martinez-Chavez E, Blaess S, Dietrich D, Heine M, Becker AJ, Schoch S. Calcium Channel Subunit  $\alpha 2\delta 4$  Is Regulated by Early Growth Response 1 and Facilitates Epileptogenesis. *J Neurosci* 2019; 39: 3175-3187

Zhou T, Zhu X, Ye Z, Wang YF, Yao C, Xu N, Zhou M, Ma J, Qin Y, Shen Y, Tang Y, Yin Z, Xu H, Zhang Y, Zang X, Ding H, Yang W, Guo Y, Harley JB, Namjou B, Kaufman KM, Kottyan LC, Weirauch MT, Hou G, Shen N. Lupus enhancer risk variant causes dysregulation of IRF8 through cooperative lncRNA and DNA methylation machinery. *Nat Commun* 2022; 13: 1855

## 5. Acknowledgements

I would like to thank my supervisor Albert Becker for guiding me all these years. Thanks for being always there, and for giving me the possibility to work with what I like the most. Many thanks for providing me with the tools to grow professionally, I will always appreciate it.

To Susanne Schoch, who has always been kind and helpful to me; and to whom is always nice to share our interests in molecular biology.

To Karen van Loo who has been a second mentor to me, and to whom working has been always a pleasure.

To Julika Pitsch for all her support and unconditional help through all these years.

To all the committee members, who have taken the time and effort to participate in the PhD process.

To all members of the Forschergruppe, for the great collaborations that made this and upcoming works possible.

Many thanks to all the members of the AG Schoch/Becker, who created through all the years a great place to be in, full of joy, kindness, and friendship. Thanks for being the best colleagues one can have: Sabine, Idil, Annika, Annachiara, Silvia, Anne, Delara, Jorge, Shayne, Pia, Aya, Maksim, Aleks, Sanya, Alex Müller, Julia Betzin, Eva, all the rest that contributed daily to creating that beautiful environment. Special thanks to Anne and Silvia, that have been always there for all of us and to whom I grew up in the lab.

Many thanks to Sabine who has made every part of this project possible and is always willing to create and try all new ideas. Many thanks for all happiness and companionship you provided to us every day.

A very special thanks to the former members of my office who are now very valuable friends, and that were always the main reason for tons of laughs every morning: Idil Arioiz (ah!), Annika Mayer (It's okay, It's okay), and Annachiara Meconi (No, No, No). Also, many thanks to Jorge, for all the chocolates he kindly provided us.

Many thanks to my friends Gabriel, Sara, Nadlie, Nicky, Iki, Kerstin, Gaby. To my beautiful doñas: Eu, Katty, Lili, May, Vero, Moncho, Yessika, Mariana, Desi, and Andrea. To Luijo, Pedro, Eric, Andres, Mattias and Eberle. It has always been a joy having you around. Special thanks to those siblings that life constantly put on my way Katty, Pedro and Eugenia, thanks for your unconditional friendship and for contributing daily to my personal growth. Therefore, thanks to “la Simón”.

Finally, I would like to thank my family, that has always supported, trusted, and accompanied me through all this way. I have no words to express my infinite gratitude to you.