Identification of the genomic mutation in *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ mice

Siti W. Mohd-Zin[1]*, Nor-Linda Abdullah[1]*, Aminah Abdullah[2], Nicholas D.E. Greene[3], Pike-See Cheah[4], King-Hwa Ling[4,5], Hadri Yusof[4], Ahmed I. Marwan[6], Sarah M. Williams[6], Kerri T. York[6], Azlina Ahmad-Annuar[7], Noraishah M. Abdul-Aziz[1§]

[1]Department of Parasitology, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia

[2]School of Chemical Science and Food Technology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, 43600 Selangor, MALAYSIA

[3]Birth Defects Research Centre, Developmental Biology and Cancer Programme, Institute of Child Health, University College London, 30 Guilford Street, London WC1N 1EH, UK

[4]Genetics and Regenerative Medicine Research Centre, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor DE, Malaysia

[5]Department of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor DE, Malaysia

[6]Department of Surgery, University of Colorado Denver, Anschutz Medical Campus, 12700 E 17th Ave, Aurora, CO 80045, USA

[7]Department of Biomedical Science, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia

*These authors contributed equally to this work

[§]Corresponding author

Corresponding author's email address: noisha@ummc.edu.my

Fax number: +60379674754; Phone number: +60172604310

Abstract

The EphA4 receptor tyrosine kinase is involved in numerous cell-signalling activities during embryonic development. EphA4 has the ability to bind to both types of ephrin ligands, the ephrinAs and ephrinBs. The C57BL/6J-Epha4rb-2J/GrsrJ; genetically known as Epha4$^{rb\text{-}2J/rb\text{-}2J}$ is a spontaneous mouse mutant which arose at The Jackson Laboratory. These mutants exhibited a synchronous hind limb locomotion defect or 'hopping gait' phenotype, which is also characteristic of EphA4 null mice. Genetic complementation experiments suggested that Epha4$^{rb\text{-}2J}$ corresponds to an allele of EphA4 but details of the genomic defect in this mouse mutant are currently unavailable. We found a single base-pair deletion in exon 9 resulting in a frame shift mutation that subsequently resulted in a premature stop codon. Analysis of the predicted structure of the truncated protein suggests that both the kinase and sterile α motif (SAM) domains are absent. We have also developed a method to ease detection of the mutation through RFLP that will aid in studies where the true genotypes need to be ascertained. The importance of this study is underlined by the numerous isoforms attributed to the Eph-ephrin family and in this case, the identification of the type of mutation enables further functional studies such as protein-protein interactions, immunostaining and gene compensatory studies of the Eph family of receptor tyrosine kinases.

Keywords

EphA4; hopping gait; spontaneous mutation; rb-2J strain; knockout mouse

Introduction

Erythropoietin hepatocellular carcinoma receptor tyrosine kinases, more commonly known as the Ephs; are members of the largest family of receptor tyrosine kinases (RTKs) which characteristically bind to their ligand ephrins. The mammalian and chick Ephs are divided into two groups based on sequence homologies and affinity for ephrins, the A-type (EphA1-EphA10) and the B-type (EphB1-EphB6) (Pasquale, 2008). EphrinAs are known to bind to EphA receptors and ephrinBs bind to EphBs. EphA4 interacts with both ephrinAs and ephrinBs (ephrinB2 and ephrinB3) and ephrinA5 binds to EphB2 (Himanen et al., 2004; Pasquale, 2004). Eph and ephrin signalling play many important roles including remodelling of blood vessels and formation of tissue boundary (Pasquale, 2005, 2008; Wilkinson, 2015). EphA4 has been shown to be involved in cell signalling activities in numerous contexts including axon guidance and development of central nervous system vasculature (Dottori et al., 1998; Kullander et al., 2001).

During embryonic development, *EphA4* is expressed at the tips of the closing spinal neural folds, and in the developing forebrain, hindbrain and mesoderm (Abdul-Aziz et al., 2009; Nieto et al., 1992). In adult mice, *EphA4* is highly expressed in the jejunum (Islam et al., 2010) and the brain, mainly in the hippocampus (Greferath et al., 2002; Grunwald et al., 2004; Kullander et al., 2001). Targeted mouse knockouts of the *EphA4* gene display locomotor abnormalities of the hind limb resulting in a rabbit-like hopping movements (Coonan et al., 2001; Herrmann et al., 2010; Kullander et al., 2003; Kullander et al., 2001; Nieto et al., 1992) or clubfoot of the hind limb (Helmbacher et al., 2000). Interestingly, inactivation of the *EphA4* function, which causes the lack of axonal guidance, was reported to promote axonal

regeneration and improve functional recovery of a central nervous system injury (Goldshmit et al., 2011).

Two spontaneous mutations of the *EphA4* gene; *Epha4*$^{rb/rb}$ (Eph receptor A4; rabbit) and *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ (Eph receptor A4; rabbit 2 Jackson) were reported as spontaneous recessive mutants within the C57BL/6J background at The Jackson Laboratory. However, the exact genomic location of the mutated alleles has not been ascertained. Since the phenotype observed in the *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ mice was very similar to the targeted knockouts of the *EphA4* gene, complementarity testing was performed by mating a Epha4Gt(pGT1TM)38Wcs/+ female (provided by Tessier-Lavigne Laboratory of Stanford University) to an *Epha4*$^{rb\text{-}2J/+}$ male. Among the progeny of this cross, 3 pups (1 female and 2 males) out of 7 exhibited hopping gait phenotype without leaning. The information on the *Epha4*$^{rb/rb}$ and *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ are from the Mouse Mutant Resource Web Site, The Jackson Laboratory, Bar Harbour, Maine. (http://mousemutant.jax.org/) [October 9, 2010]. This finding of genetic complementation strongly suggested that the *rb-2J* mutation lies within the *EphA4* gene.

The aim of this study is to understand the nature of the *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ mutation and the extent of the mutation in the protein. Therefore it was necessary to define the mutation and determine the exact genotypes of the mice. We also reviewed the impact of the mutation with EphA4 isoforms and other reported EphA4 mutants to better understand the diversity of the EphA4 protein. In addition, we describe a rapid, simple assay for genotyping the C57BL/6J-Epha4rb-2J/GrsrJ mice, which was not previously available and will be invaluable to other groups intending to work with these mice. Therefore, we have addressed the knowledge gap for the C57BL/6J-Epha4rb-2J/GrsrJ mouse in this study.

Materials and methods

The *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ strain (003129) was obtained from The Jackson Laboratory and the colony maintained at the Universiti Kebangsaan Malaysia's Animal Biosafety Level 2 laboratory. All experimental procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of University of Malaya (approval number PAR/20/09/2011/NMAA).

Epha4$^{rb\text{-}2J/rb\text{-}2J}$ knockout mouse and dissection of the hippocampus

As the hopping gait phenotype is only evident in mice at 3 weeks old, mutational analysis was first determined using samples from adult mice in order to correlate the genotype with an affected phenotype. In adult mice, *EphA4* expression is abundant in the hippocampus (Liebl et al., 2003; Murai et al., 2003), therefore hippocampi of two aged matched control and *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ mice were isolated (modified protocol) (Fuller & Dailey, 2007). The mice were euthanized by cervical dislocation. After the the midline incision from the foramen magnum was made up to the level of eye sockets, the skull flaps were tilted to break it off using forceps, and the brain was lifted out of the skull vault gently using curved narrow patterned forceps. In the clean petri dish, the brain was cut into half at the midline, and the hemisphere of the brain was held in place by piercing through the rostral end of the brain with Dumont forceps. Then, the brain stem and cerebellum were removed before gently removing the midbrain. The hippocampus is delineated along its length in the midbrain. The hippocampus was flipped over and cut to separate it from the hemisphere and stored in -80°C.

RNA Isolation

RNA isolation was carried out by homogenisation in 1 ml of TRIzol (Invitrogen, Carlsbad, CA, USA) to 50 mg of hippocampi tissue, followed by the addition of 0.2 ml chloroform according to the manufacturer's instruction. After separation by cold centrifugation, the aqueous phase was placed in a new tube before addition of 0.5 ml of isopropanol and subsequently centrifuged. The pellet was washed with 75% ethanol and left to air-dry for 5-10 minutes. The pellet was then re-suspended in distilled, deionised water.

Mutational screening of *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$

Seven sets of primers, which amplify overlapping regions of EphA4 cDNA, were designed to allow sequencing of the entire coding region (accession number: NM_007936). After the discovery of location of the putative mutation in *EphA4*, primers for genomic DNA were also designed. All primers were designed using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/primer3/). Primer sequences with overlapping region of *EphA4* cDNA were as following: set 1 flanking exon 1 to 3 (668 bp): 5'-CACCCTCTTGGCAATGTCTT-3' and 5'-CTTTTCAGGATGTGGGTGCT-3'; set 2 flanking exon 3 to 5 (700 bp): 5'-GACATTGGTGACCGAATCAT-3' and 5'-TCCACTACACACCACAGCAGA-3'; set 3 flanking exon 5 to 8 (588 bp): 5'-GGCCGTCAGGACATTTCTTA-3' and 5'-ACTCCACTGTCCTGCTGGTC-3'; set 4 flanking exon 7 to 11 (499 bp): 5'-TCTGACTTCCTATGTTTTTCACG-3' and 5'-GAGACTTCCTGAGTGAGGCC-3'; set 5 flanking exon 10 to 14 (578 bp): 5'-CGATGCATCCTGCATTAAAA-3' and 5'-CTCAGCCAGTGATGTCTGGA-3'; set 6 flanking exon 13 to 16 (515 bp): 5'-CCGAAGCAGCCTACACTACC-3' and 5'-TAGAAGCCGTGGTTCACATG-3';

and set 7 flanking exon 15 to coding region exon 17 (471 bp): 5'-CTCCCCTGAATTCTCTGCTG-3' and 5'-ATCAGAATTAAACCTGGAGCCA-3'.

RT-PCR was performed using Transcriptor One-Step RT-PCR kit (Roche Diagnostics, Mannheim, Germany). The transcriptor enzymes mix contains transcriptor reverse transcriptase, expand system and protector RNase inhibitor. The 1x reaction buffer including Tris, $MgCl_2$, 1.5 mM dNTPs and additives for hot start PCR. We generated cDNA and amplified 50 ng of total RNA in a 50 ul of reaction mixture consisting of 0.4 uM of each primer. Glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) and Phosphoglycerate kinase 1, (*Pgk1*) mRNA were used as internal controls. The reverse transcription and amplification condition was set as follows: cDNA generation at 50^{o}C for 30 minutes, and 94^{o}C for 7 minutes, followed by 35 cycles of denaturation at 94^{o}C for 10 seconds, annealing at optimised temperature for 30 seconds, extension at 68^{o}C for 1 minute, with final extension at 68^{o}C for 7 minutes. Resulting PCR products were analysed by electrophoresis on 1% agarose gels containing ethidium bromide.

RT-PCR products from *Epha4*$^{+/+}$, *Epha4*$^{rb\text{-}2J/+}$ and *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ samples were purified using QIAquick PCR purification kit or QIAquick gel extraction kit (Qiagen, Valencia, CA, USA). All the purification steps were according to the manufacturer's directions. Subsequently, 10 ul of 30 ng/ul single band PCR products were sent for Sanger sequencing through a commercial company.

Genomic DNA Isolation and PCR

While the mutational screening was performed on the coding transcripts, to enable ease of genotyping from the genomic DNA, we developed a method to genotype the mutation from genomic DNA. Tail or ear clips of adult mice were obtained and DNA

extraction for genotyping. The DNA extraction described previously (Sambrook et al., 1989). The tissues were lysed in SNET buffer containing 20 mM Tris-Cl pH 8.0 (Sigma), 5 mM EDTA pH 8.0 (Sigma), 400 mM NaCl (Sigma), 1% (w/v) SDS (Sigma) and sterilised by filtration through 0.45 µm nitrocellulose filter before the addition of 20 mg/ml Proteinase K (Sigma). The tissues were incubated in 55^{o}C until the tissue was completely lysed. The DNA was isolated by using 1:1 of phenol solution (Sigma), followed by 2:1 of absolute ethanol precipitation and 70% ethanol washing step. After left to air-dry, the pellet was re-suspended in distilled, deionised water.

The genomic DNA amplification of samples were performed with 0.6mM of EphA4 primers spanning the mutation (accession number: NC_000067.6; forward P1: 5'-GTAACATGTGCACTGCCTATCC-3' and reverse P2: 5'-CACAGGCATATTAACCAACACTTC-3') in a 50 ul total reaction of DreamTaq Green DNA polymerase (ThermoScientific), 1x buffer DreamTaq Green buffer including 2 mM $MgCl_2$, and 0.2 mM dNTP. The amplification condition was set as follows: 95°C for 4 minutes, followed by 35 cycles of denaturation at 95°C for 1 minute, annealing at 59°C for 1 minute, extension at 72°C for 1 minute, with final extension at 72°C for 10 minutes. The expected size of amplicon is 258 bp. Subsequently, 40 ul of 30 ng/ul single band PCR products were sent for DNA purification and Sanger sequencing through a commercial company.

PCR- RFLP (Restriction Fragment Length Polymorphism)

We employed PCR-RFLP to rapidly genotype the novel mutation in more samples. New primers were designed based on the *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ exon 9 nucleotide sequence (http://bioinfo.ut.ee/primer3-0.4.0/primer3/) to introduce mutations into the amplified

DNA at nucleotide 1799 and 1800 (AA→TC) of the gene to generate an *Xho*1 recognition site CTCGAG in mutant samples. The forward primer sequence (P3) is 5'- TACAGCAAAGCGAAACTCGA-3' (the altered sequence was underlined) and reverse primer sequence (P2) is 5'-CACAGGCATATTAACCAACACTTC-3'. We amplified 50ng of genomic DNA in a 20 µl of reaction mixture consisting of 0.5µM of each primer and 1X LightCycler® 480 Probe Master containing FastStart Taq DNA Polymerase, dNTP mix and 6.4mM $MgCl_2$ (Roche Diagnostics, Mannheim, Germany). The amplification condition was set as follows: 95°C for 10 minutes, followed by 35 cycles of denaturation at 95°C for 10 seconds, annealing at 53°C for 30 seconds, extension at 72°C for 10 seconds, with final extension at 72°C for 7 minutes. Approximately 15µl of amplicon was digested with 20U of *Xho*1 (New England Biolabs, Ipswich, MA, USA) according to manufacturer's protocol followed by separation on 4% agarose gel. The fragments generated were a single 149bp for the homozygote mutant, a single 166bp band for wildtype and two bands of 149bp and 166bp bands for heterozygous profiles.

Western Blot Analysis

To determine the effect of the mutation, immunoblot analysis was performed using anti-EphA4 (EphA4 Antibody S-20; sc-921) and anti-GAPDH (Santa Cruz Biotechnology) on wildtype, heterozygous and mutant protein samples isolated from the hippocampi of 31-day old mice. Protein amounting to 25ug isolated from the hippocampi was mixed with 2µl of 1M DTT in a final volume of 20µl. Protein samples were heated to 100°C for 10 minutes then immediately placed on ice prior to gel loading. Life Technologies Xcel Surelock Mini Cell system was used with a 4-12% denaturing precast protein gel according to the manufacturers guidelines.

Samples were run at 200V for 1 hour. Protein was then transferred using Life Technologies Xcel Mini Cell transfer system per manufacturers guidelines. Protein was transferred at 35V for 3 hours with transfer apparatus submerged in ice bucket. Blot was then blocked with 5% milk in TBS-T buffer (50mM Tris Base, 150mM NaCl, 01% Tween 20, pH 7.5) for 1 hour. After blocking, anti-EphA4 primary antibody was diluted to 1:100 in 5% milk in TBS-T buffer and incubated with blot overnight at 4°C. Anti-GAPDH primary antibody was added after 24 hours incubation with anti-EphA4 at a dilution of 1:50,000 and was incubated for an additional 1 hour at 4°C. Blot was washed 5 times with TBS-T buffer for 5 minutes per wash with agitation. Secondary antibody (Life Technologies; A24531) was diluted 1:10,000 in 5% milk in TBS-T buffer and incubated with blot for 1 hour at room temperature. Blot was washed 5 times with TBS-T buffer for 5 minutes per wash with agitation. Secondary antibody was detected by mixing equal volumes of ECL reagents (Pierce; 34077) and exposing the blot to the mixed ECL reagents for 5 minutes. The blot was imaged on a BioRad chemidoc station.

Results

The *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ mouse is commercially available from The Jackson Laboratory. However, the mutation was not known and there is no clear protocol on how to genotype these mice. Knowing the exact genotype is crucial when investigating the effect of the loss of the gene in the affected mice. Therefore, in this study we determined the precise mutation in the *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ strain, confirming that *EphA4* loss of function is the cause of the phenotype in this strain and facilitating further experimental studies.

A single nucleotide deletion in *Epha4*$^{rb\text{-}2J/rb\text{-}2}$ results in a truncated EphA4 protein

The *EphA4* gene is 6,328bp in size located on mouse chromosome 1 (Ensembl; release 80, May 2015). The sequences of the translated regions, contained within exon 1-17, of the *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ strain were compared against the annotated *EphA4* gene deposited in GenBank and with control C57BL/6J mice (accession number: NM_007936). A single nucleotide deletion (del1802) was located in exon 9 at 77,390,062 on mouse chromosome 1 and confirmed by sequencing of genomic DNA. This deletion is predicted to result in a frame shift and creation of a premature stop codon (Figure 1A). The resultant protein is therefore predicted to consist of the wildtype sequence up to amino acid 582 (E582), followed by a series of six altered amino acids and a stop codon (Figure 1B). The mutation lies within a conserved protein region in several organisms including human, orang-utan, frog, pig, rat and chick (Table 1).

This truncated protein lacks a further 390 amino acids compared to the full-length wildtype protein, which has 986 amino acids (Figure 1B). The protein structure encoded by Epha4$^{rb\text{-}2J/rb\text{-}2J}$ was predicted using Simple Modular Architecture Research Tool, SMART (http://smart.embl-heidelberg.de). The predicted structure of the truncated proteins lacks the kinase and sterile alpha motif (SAM) domains (Figure 2E).

Western blot analysis (Figure 3) revealed an intact 110kDa band in samples from both wildtype and heterozygous mice. This band was absent in the homozygous mutant mice. A second band at 104kDa was detected in samples from all genotypes. We hypothesise that the lower molecular band is an alternate uncharacterised isoform.

There are no reported isoforms of EphA4 with only exon 9 (containing the truncating mutation) spliced out.

An RFLP-PCR assay was developed to rapidly genotype the mice

In addition to provide a means to identify mice carrying the deletion on exon 9, we analysed C57BL/6J-*Epha4rb-2J*/GrsrJ mouse colony with and without hopping gait characteristics using an inexpensive PCR-RFLP analysis (see Materials and Methods). The analysis showed *Epha4*$^{+/+}$ mice with wildtype profile (a single 166 bp band), *Epha4*$^{rb\text{-}2J/+}$ mice with heterozygous profile (two bands at 149bp and 166bp) and *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ mice with mutant profile (a single 149bp band) (Figure 4). All the RLFP analyses were subsequently confirmed by DNA sequencing. All mutant mice were characterised with hopping gait features whereas both wildtype and heterozygous mice were apparently normal.

Genotypes obtained correlate with the EphA4 phenotype

Among 63 samples that were sequenced, 18 were wildtype (*Epha4*$^{+/+}$), 35 were heterozygous (*Epha4*$^{rb\text{-}2J/+}$) and 10 were mutants (*Epha4*$^{rb\text{-}2J/rb\text{-}2J}$). The analysis showed that the deletion found on exon 9 of *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ mice was 100% in concordance with the features observed in the mutant mice.

Discussion

The central pattern generators (CPGs) are the neuronal networks that generate and coordinate rhythmic limb movement. The hopping gait phenotype displayed by the *EphA4* mutant was discovered due to CPG neurons aberrantly crossing the midline of

294 the spinal cord (Kullander et al., 2003). The crossing generates the synchronous
295 locomotion of the hind limbs through reciprocal over-excitation of the CPG neurons.
296 Previous studies on the understanding of the functional domains of EphA4
297 receptor in the mouse revealed the requirement of kinase function in axon guidance
298 and its formation (Dufour et al., 2006; Egea et al., 2005; Kullander et al., 2001). The
299 cytoplasmic domains of Eph receptor are the juxtamembrane (JM) domain, kinase
300 domain, sterile-α-motif (SAM) domain and PDZ domain. In functional studies, the
301 hopping gait phenotype only appeared in *EphA4* mutants (*EphA4*^{KD}*)* with a defective
302 kinase domain (Kullander et al., 2001) and in mutants (*EphA4*^{GFP}*)* with the absence of
303 the entire functional cytoplasmic domains (Egea et al., 2005). Another *EphA4* mutant
304 (*EphA4*^{EE}*)* displayed normal alternating gait when two mutations of tyrosine residues
305 (Y596E and Y602E) were introduced in the JM domain (Egea et al., 2005). Despite
306 the decrease of auto-phosphorylation in the mutant, there was an increase of basal
307 kinase activity almost similar or higher than ephrin-activated EphA4. This would
308 explain the requirement of kinase activity for normal functioning CPG neurons and
309 normal alternating gait. The findings support the hypothesis that tyrosine residues in
310 the JM domain regulate EphA4 kinase activity. The mutation in *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ is
311 located at E582 prior to the major auto-phosphorylation sites. The deletion resulted in
312 a frameshift, leading upon translation to altered six encoded codons followed by a
313 stop codon. The Epha4$^{rb\text{-}2J/rb\text{-}2J}$ protein is truncated, lacking the tyrosine auto-
314 phosphorylations sites in the JM domain, the kinase domain, SAM domain and PDZ
315 domain. The phenotype of the Epha4$^{rb\text{-}2J/rb\text{-}2J}$ mouse differs from the other EphA4
316 mutants that have been published in the past in that it not only hops but also leans.
317 Therefore, variability in phenotypic representations exists from the same gene
318 knockouts (Table 2).

The kinase activity of the Epha4$^{rb\text{-}2J/rb\text{-}2J}$ protein is expected to be low to almost non-existent similar to previously reported mutants. Truncation of EphA4 was indicated by the western blot analysis, which showed the absence of the full length EphA4 protein (110kDa) in homozygote mutant mice. A second band similar to smaller isoform reported in UniProt (http://www.uniprot.org/uniprot/?query=EphA4&sort=score) at 104kDa was detected in wildtype, heterozygous and mutant EphA4 mice. However, this band is most likely an unknown isoform, which has exon 9, spliced out but it still possesses the epitope located at the SAM domain, which is detected by the Santa Cruz antibody.

There exists 7 isoforms of the EphA4 protein (available from UniProt; Table 3). It seems likely that the Epha4$^{rb\text{-}2J/rb\text{-}2J}$ protein exists in two spliced variations. The first being that the full length version is mutated and lacks the kinase domain; therefore the signal from the full length version is not picked up by the western blot in Figure 3 but possesses an unknown shorter isoform at 104 kDa (similar in size but differs from Figure 2C). The truncated full length version of Epha4$^{rb\text{-}2J/rb\text{-}2J}$ is 63 kDa in size when analysed using SMART prediction (Figure 2E) and it is similar in size to the 63 kDa short isoform of EphA4 (Figure 2D). Ephs and ephrins are known to carry spliced variations of its protein as evidenced in both man (Finne et al. 2004) and mouse (Holmberg et al. 2000).

Conclusions

We identified a single nucleotide deletion of adenine in exon 9 of EphA4 in the *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ mouse mutant. This deletion results in a frameshift, which is predicted to cause premature truncation of the protein, and lack of key cytoplasmic domains. This data is important in validating experiments that utilises the *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$

knockout mouse to look at the function of the *EphA4* gene. Furthermore, this study underlines the numerous isoforms of the *EphA4* gene and the variability in phenotype arising from which isoform becomes the major gene product.

Acknowledgements

Supported by High Impact Research Grant UM.C/625/1/HIR/062 to Noraishah M. Abdul-Aziz and UM.C/625/1/HIR/148/2 from the University of Malaya to Noraishah M. Abdul-Aziz, High Impact Research Grant UM.C/625/1/HIR/MOHE/MED/08/04 from the Ministry of Higher Education Malaysia to Noraishah M. Abdul-Aziz, Fundamental Research Grant Scheme FP040/2010, and postgraduate grant from University of Malaya PPP PV002-2011A to Nor-Linda Abdullah and Noraishah M. Abdul-Aziz. Nicholas Greene was supported by the National Institute for Health Research Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London. Nicholas Greene is also High Impact Research Icon of University of Malaya, Malaysia.

Conflict of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Abdul-Aziz, N. M., Turmaine, M., Greene, N. D., & Copp, A. J. (2009). EphrinA-EphA receptor interactions in mouse spinal neurulation: implications for neural fold fusion. *Int J Dev Biol, 53*(4), 559-568. doi:10.1387/ijdb.082777na

Cook SA, et al., Two spontaneous mutant alleles, rb and rb-2J, of the Epha4 gene on mouse Chromosome 1. Mouse Mutant Resources Web Site, The Jackson Laboratory, Bar Harbor, Maine. MGI Direct Data Submission. 2004

Coonan, J. R., Greferath, U., Messenger, J., Hartley, L., Murphy, M., Boyd, A. W., . . . Bartlett, P. F. (2001). Development and reorganization of corticospinal projections in

EphA4 deficient mice. *J Comp Neurol, 436*(2), 248-262. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11438928

Dottori, M., Hartley, L., Galea, M., Paxinos, G., Polizzotto, M., Kilpatrick, T., . . . Boyd, A. W. (1998). EphA4 (Sek1) receptor tyrosine kinase is required for the development of the corticospinal tract. *Proc Natl Acad Sci U S A, 95*(22), 13248-13253. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9789074

Dufour, A., Egea, J., Kullander, K., Klein, R., & Vanderhaeghen, P. (2006). Genetic analysis of EphA-dependent signaling mechanisms controlling topographic mapping in vivo. *Development, 133*(22), 4415-4420. doi:10.1242/dev.02623

Egea, J., Nissen, U. V., Dufour, A., Sahin, M., Greer, P., Kullander, K., . . . Klein, R. (2005). Regulation of EphA 4 kinase activity is required for a subset of axon guidance decisions suggesting a key role for receptor clustering in Eph function. *Neuron, 47*(4), 515-528. doi:10.1016/j.neuron.2005.06.029

Filosa, A., Paixao, S., Honsek, S.D., Carmona, M.A., Becker, L., Feddersen, B.,…, Klein, R. (2009). Neuron-glia communication via EphA4/ephrinA3 modulates LTP through glial glutamate transport. Nat Neurosci 12(10):1285-92. doi: 10.1038/nn.2394

Finne, E. F., Munthe, E., & Aasheim, H. C. (2004). A new ephrin-A1 isoform (ephrin-A1b) with altered receptor binding properties abrogates the cleavage of ephrin-A1a. *Biochem J, 379*(Pt 1), 39-46. doi:10.1042/BJ20031619

Fuller, L., & Dailey, M. E. (2007). Preparation of rodent hippocampal slice cultures. *CSH Protoc, 2007*, pdb prot4848. doi:10.1101/pdb.prot4848

Goldshmit, Y., Spanevello, M. D., Tajouri, S., Li, L., Rogers, F., Pearse, M., . . . Turnley, A. M. (2011). EphA4 blockers promote axonal regeneration and functional recovery following spinal cord injury in mice. *PLoS One, 6*(9), e24636. doi:10.1371/journal.pone.0024636

Greferath, U., Canty, A. J., Messenger, J., & Murphy, M. (2002). Developmental expression of EphA4-tyrosine kinase receptor in the mouse brain and spinal cord. *Gene Expr Patterns, 2*(3-4), 267-274. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12617813

Grunwald, I. C., Korte, M., Adelmann, G., Plueck, A., Kullander, K., Adams, R. H., . . . Klein, R. (2004). Hippocampal plasticity requires postsynaptic ephrinBs. *Nat Neurosci, 7*(1), 33-40. doi:10.1038/nn1164

Helmbacher, F., Schneider-Maunoury, S., Topilko, P., Tiret, L., & Charnay, P. (2000). Targeting of the EphA4 tyrosine kinase receptor affects dorsal/ventral pathfinding of limb motor axons. *Development, 127*(15), 3313-3324. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10887087

Herrmann, J. E., Pence, M. A., Shapera, E. A., Shah, R. R., Geoffroy, C. G., & Zheng, B. (2010). Generation of an EphA4 conditional allele in mice. *Genesis, 48*(2), 101-105. doi:10.1002/dvg.20587

Himanen, J. P., Chumley, M. J., Lackmann, M., Li, C., Barton, W. A., Jeffrey, P. D., . . . Nikolov, D. B. (2004). Repelling class discrimination: ephrin-A5 binds to and activates EphB2 receptor signaling. *Nat Neurosci, 7*(5), 501-509. doi:10.1038/nn1237

Holmberg, J., Clarke, D. L., & Frisen, J. (2000). Regulation of repulsion versus adhesion by different splice forms of an Eph receptor. *Nature, 408*(6809), 203-206. doi:10.1038/35041577

Islam, S., Loizides, A. M., Fialkovich, J. J., Grand, R. J., & Montgomery, R. K. (2010). Developmental expression of Eph and ephrin family genes in mammalian small intestine. *Dig Dis Sci, 55*(9), 2478-2488. doi:10.1007/s10620-009-1102-z

Kullander, K., Butt, S. J., Lebret, J. M., Lundfald, L., Restrepo, C. E., Rydstrom, A., . . . Kiehn, O. (2003). Role of EphA4 and EphrinB3 in local neuronal circuits that control walking. *Science, 299*(5614), 1889-1892. doi:10.1126/science.1079641

Kullander, K., Mather, N. K., Diella, F., Dottori, M., Boyd, A. W., & Klein, R. (2001). Kinase-dependent and kinase-independent functions of EphA4 receptors in major axon tract formation in vivo. *Neuron, 29*(1), 73-84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11182082

Leighton PA; Mitchell KJ; Goodrich LV; Lu X; Pinson K; Scherz P; Skarnes WC; Tessier-Lavigne M (2001). Defining brain wiring patterns and mechanisms through gene trapping in mice. Nature. 410 (6825) 174-9

Liebl, D. J., Morris, C. J., Henkemeyer, M., & Parada, L. F. (2003). mRNA expression of ephrins and Eph receptor tyrosine kinases in the neonatal and adult mouse central nervous system. *J Neurosci Res, 71*(1), 7-22. doi:10.1002/jnr.10457

Milstein, O., Sun, L., Du, X., Moresco, E.M.Y., Beutler B. Record for frog, updated Dec 12, 2013. MUTAGNETIX TM, B. Beutler and collegues. Center for the Genetics of Host Defense, UT Southwestern Medical Center, Dallas, TX. URL: mutagenetix.utsouthwestern.edu (Generated on Jul 03, 2015)

Murai, K. K., Nguyen, L. N., Irie, F., Yamaguchi, Y., & Pasquale, E. B. (2003). Control of hippocampal dendritic spine morphology through ephrin-A3/EphA4 signaling. *Nat Neurosci, 6*(2), 153-160. doi:10.1038/nn994

Nieto, M. A., Gilardi-Hebenstreit, P., Charnay, P., & Wilkinson, D. G. (1992). A receptor protein tyrosine kinase implicated in the segmental patterning of the hindbrain and mesoderm. *Development, 116*(4), 1137-1150. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1295734

443 Pasquale, E. B. (2004). Eph-ephrin promiscuity is now crystal clear. *Nat Neurosci*, *7*(5), 417-
444 418. doi:10.1038/nn0504-417
445 Pasquale, E. B. (2005). Eph receptor signalling casts a wide net on cell behaviour. *Nat Rev*
446 *Mol Cell Biol*, *6*(6), 462-475. doi:10.1038/nrm1662
447 Pasquale, E. B. (2008). Eph-ephrin bidirectional signaling in physiology and disease. *Cell*,
448 *133*(1), 38-52. doi:10.1016/j.cell.2008.03.011
449 Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning. A laboratory manual.*
450 (2nd ed.). Plainview, New York.: Cold Spring Harbor Laboratory Press.
451 Wilkinson, D. G. (2015). Balancing cell behavior at boundaries. *J Cell Biol*, *208*(6), 659-660.
452 doi:10.1083/jcb.201501107

Tables

Table 1. ***Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ mutation lies within a conserved region.** Alignment of Epha4$^{rb\text{-}2J/rb\text{-}2J}$ protein sequences of different organisms including human, rat and *Xenopus*, revealed the spontaneous mutation was in a conserved region.

Organism	Amino acids
EphA4 wildtype (mouse)	AKQEADEEKHL…..
EphA4 mutant (mouse)	AKQ**E**QMKRNI-stop
Human	AKQEADEEKHL…..
Orangutan	AKQEADEEKHL…
Xenopus	AKQEADEEKHL….
Pig	AKQEADEEKHL….
Rat	AKQEADEEKHL….
Chick	AKQEADEEKHL…

457 **Table 2. Phenotypes of *EphA4* mutation mice based on position of mutation**

Position of mutation	EphA4 mutations	Type of knockout and method	Phenotype and anatomical defects	References
Exon 1 knockout and exone 3 frame shift mutation	*EphA4*$^{-}$	Replacement vector; lac-Z reporter fusion	**Hind limb phenotype ("club foot") high penetrance** Loss of dorsal hindlimb innervation (peroneal nerve); absent anterior commissure	Helmbacher et al. 2000
Ligand binding domain (exon 3)	*EphA4*^{0}	Gene replacement *pgk-neo*	**Kangaroo-like (ROO) hopping gait.** Abnormal corticolspinal tract (CST) axons; absent anterior commissure	Dottori et al. 1998; Coonan et al., 2001
Extracellular region (exon 3)	*EphA4*^{EGFP}	Insertion of EGFP (reporter)	**Hopping gait**	Grunwald et al. 2004
Deletion of exon 3	*EphA4 conditional;* *EphA4*^{Flox}	Knock-in mCFP reporter gene; targerted (floxed/Frt)	**Hind limb hopping gait**	Herrmann et al. 2010
Deletion of exon 3	*EphA4 null*	Knock-in mCFP reporter gene	**Hind limb hopping gait**	Herrmann et al. 2010
Entire intracellular is missing	*EphA4*^{GFP}	Entire intracellular part was replaced by green fluorescent protein (GFP)	**Hopping gait**	Egea et al., 2005
Fibronectin domain (abberant amino acids after position 439._ and truncation at position 442. T>C intron 6 at position 113891 mutation (NC_000067). Possibilty, splicing of exon 6 and frameshift that create stop codon early in exon 7.	*Frog*	Chemical induced (ENU)	**Hopping gait**	Milstein et al., 2010
Juxtamembrane domain (Y596E and Y602E)	*EphA4*^{EE}	Knock-in strategy (similar to Kullander 2001), replacement vector—glutamic acid residues replace juxtamembrane tyrosines	Normal alternating gait. No discernible phenotype. Abnormal thalamocortical topography; and partly defective central paatern generator (CPG) rhythmicity.	Egea et al., 2005

Juxtamembrane domain (Y596F and Y604F)	*EphA4*^{2F}	EphA4 Knock-in strategy (targeting vectors)	**Hopping gait**	Kullander et al. 2001
Kinase domain and truncated at E582. Deletion adenine at 1802 (NM_007936)	*EphA4*$^{rb\text{-}2J}$	Spontaneous mutation	**Hopping gait and leaning**	Mohd-Zin et al., this study; Cook et al., 2004
Kinase domain (K653M)	*EphA4*^{KD}	EphA4 Knock-in strategy (targeting vectors); kinase dead	**Hopping gait** Signalling mutants: abnormal signalling mutants hopping gait; anterior commissure	Kullander et al. 2001
SAM domain (905-974 amino acids deletion); leaving 12 last amino acids residues intact	*EphA4*$^{\Delta SAM}$	EphA4 Knock-in strategy (targeting vectors)	Normal alternating gait	Kullander et al. 2001
PDZ-binding motif and 12 last amino acids of EphA4 gene	*EphA4*$^{\Delta PBM}$	Knock-in	Not described	Dufour et al. 2006
Not described	*EphA4*^{PLAP}	Gene trapped, insertion of PLAP vector	**Hopping kangaroo gait** Guidance defects in the CST (crossing defects of axons) and anterior commissure. However, low expression of *EphA4* in CST neurons.	Leighton et al., 2001
Not described	*PGK-cre;EphA4*^{lx}	Recombination of PGK-Cre and EphA4$^{lox/lox}$	**Kangaroo-like (ROO) hopping gait.** Abnormal formation of anterior commissure and significant reduction of CST axons	Filosa et al. 2009
Not desribed	*Conditional EphA4; EphA4*^{lx}	Targeted (floxed/frt)	Normal alternating gait EphA4 expression in control EphA4$^{lx/lx}$ mice is reduced to 15–20% compared to +/+ mice	Filosa et al. 2009
Not studied	*EphA4*^{rb}	Spontaneous mutation	Normal alternating gait	Cook et al, 2004

458 Different type of mutations showed most with hind limb hopping gait phenotypes (bold) and few with normal alternating gait phenotypes

459 (underlined).

Table 3

UniProt ID***	Size of transcript (bp)*	Protein (aa, amino acids)	Size (Da)	Aligned to Q03137 (position)	Domain(s)
Q03137 Long isoform (Experimental evidence at protein level)	6328	986 aa	109,814	1-986 aa	Ligand binding domain (30-209 aa), cysteine rich domain (191-325 aa), fibronectin type III domain (328-439 aa & 440-537 aa), juxtamembrane domain, protein kinase domain (621-882 aa), SAM domain (911-975 aa), PDZ-binding motif (984-986 aa)**
Q03137-2 Shorter isoform (No experimental confirmation available)	No information	936 aa	103,984	Missing 783-832 aa (missing protein kinase domain). Missing whole exon 14.	Ligand binding domain, fibronectin type III domain, juxtamembrane domain, SAM domain
Q3V1W9 Short isoform (Experimental evidence at transcript level)	No information	572 aa	63,034 Da	Missing whole exon 9-17 (missing juxtamembrane domain, kinase domain, SAM domain and PDZ-binding motif)	Ligand binding domain, fibronectin type III domain
A0A087WRH4 (Experimental evidence at protein level)	614	117 aa	13,206	833-949 aa (whole exon 15&16)	SAM domain
A0A087WQW6 (Experimental evidence at protein level)	584	177 aa	19,075	328 to 473 aa (exon 5&6)	Fibronectin type III
Q99KA8 (Experimental evidence at transcript level)	No information	927 aa	103,444	Missing 1-59 aa (ligand binding domain is not affected)	Ligand binding domain, fibronectin type III domain, juxtamembrane domain, protein kinase domain, SAM domain
A0A087WQZ6 (Protein Predicted)	3043	38 aa	4,137	1-38 aa (missing the rest of domains)	No information

*Information obtained from Ensembl (http://asia.ensembl.org/Mus_musculus/Gene/Summary?db=core;g=ENSMUSG00000026235;r=1:77367185-77515088)

**Information obtained from UniProt (http://www.uniprot.org/uniprot/Q03137#showFeatures)

***Note: (taken from UniProt last modified January 8, 2015)
The value **'Experimental evidence at protein level'** indicates that there is clear experimental evidence for the existence of the protein. The criteria include partial or complete Edman sequencing, clear identification by mass spectrometry, X-ray or NMR structure, good quality protein-protein interaction or detection of the protein by antibodies.
The value **'Experimental evidence at transcript level'** indicates that the existence of a protein has not been strictly proven but that expression data (such as existence of cDNA(s), RT-PCR or Northern blots) indicate the existence of a transcript.
The value **'Protein Predicted'** is used for entries without evidence at protein, transcript, or homology levels.

Legends

Figure 1. A single nucleotide deletion in exon 9 of *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$

A, The deletion of adenine at nucleotide 1,802 or 1,803 in exon 9 of *EphA4* gene (indicated by an arrowhead). **B,** Nucleotide and deduced amino acid sequence of the *Epha4*$^{+/+}$ and *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$. The deletion in *EphA4* (highlighted by the red box) resulted in a frame shift of downstream codons and a premature termination.

Figure 2. A schematic representation of Eph receptor, Epha4$^{+/+}$and Epha4$^{rb\text{-}2J/rb\text{-}2J}$ structure

A, A general structure of an Eph receptor. **B,** The predicted structures using Simple Modular Architecture Research Tool, SMART for Epha4$^{+/+}$. **C,** SMART structure of EphA4 isoform at 104 kD lacking kinase domain but possesses SAM domain. **D,** SMART structure of EphA4 isoform at 63 kDa which is truncated downstream of the juxtamembrane domain. **E,** The Epha4$^{rb\text{-}2J/rb\text{-}2J}$ structure by SMART showing truncation downstream of the juxtamembrane domain, lacking the kinase and SAM domains.

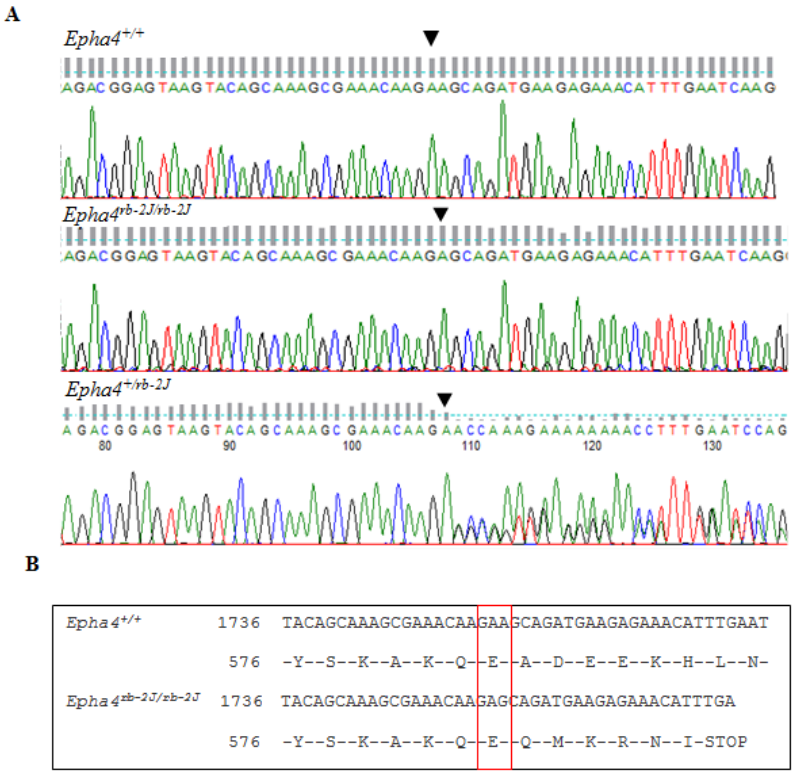
Figure 3. EphA4 protein expression

Western blot analysis of EphA4 immunostaining showed that the EphA4 mutant mice lacked the expression of the 110kDa full-length protein but detected a second band similar to EphA4 isoform at 104 kDa (UniProt). The 110kDa full-length protein was detected in the EphA4 wildtype profile and heterozygous mice.

Figure 4. A representative PCR-RFLP analysis of the *Epha4*$^{rb\text{-}2J/rb\text{-}2J}$ mice genotypes

PCR-RFLP samples were resolved in 4% agarose gel. Lanes 1-3 represent the wildtype profiles (166bp band only), Lanes 4-6 represent the heterozygote profiles (149bp and 166bp

bands), Lanes 7-9 represents the homozygote mutant profiles (149bp band only) and Lane 10 represents the 50bp ladder.



498 Figure 1
499

500

501

502

503

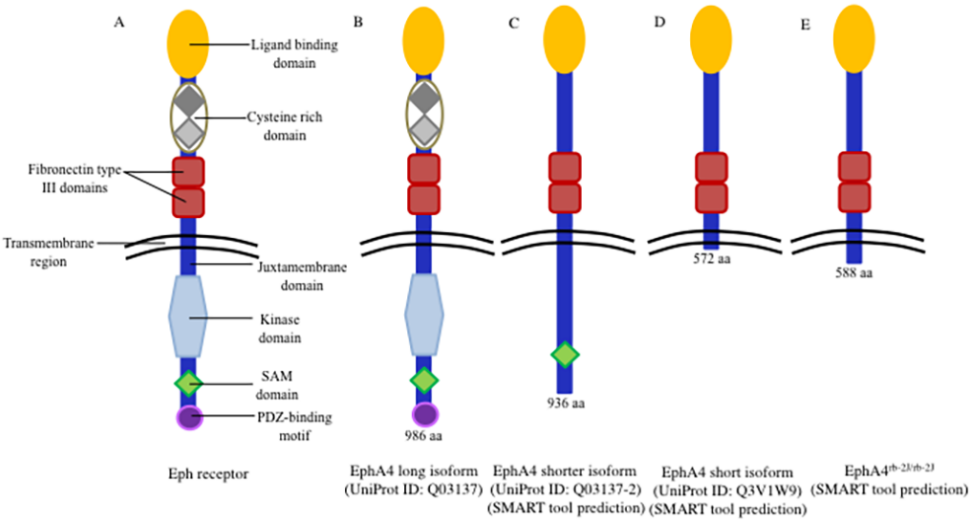
504

505

506

507

508

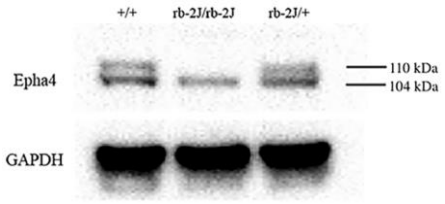


509

510 Figure 2

511

512



513

514

515 Figure 3

516

517

518 Figure 4

519