

## Dominant monoallelic variant in the PAK2 gene causes Knobloch syndrome type 2

Stylianos E. Antonarakis<sup>1,8</sup>, Ales Holoubek<sup>2</sup>, Melivoia Rapti<sup>3</sup>, Jesse Rademaker<sup>3</sup>, Jenny Meylan<sup>3</sup>, Justyna Iwaszkiewicz<sup>4</sup>, Vincent Zoete<sup>4,5</sup>, Muhammad Ansar<sup>6</sup>, Christelle Borel<sup>1</sup>, Olivier Menzel<sup>7</sup>, Kateřina Kuželová<sup>2</sup>, Federico A. Santoni<sup>3,9</sup>

<sup>1</sup>Department of Genetic Medicine and Development, University of Geneva Medical Faculty, Geneva 1211, Switzerland.

<sup>2</sup>Department of Proteomics, Institute of Hematology and Blood Transfusion, Prague, Czech Republic.

<sup>3</sup>Department of Endocrinology Diabetes and Metabolism, Lausanne University Hospital, Lausanne 1011, Switzerland.

<sup>4</sup>Molecular Modeling Group, Swiss Institute of Bioinformatics, Lausanne 1015, Switzerland.

<sup>5</sup>Department of Fundamental Oncology, Ludwig Institute for Cancer Research, Lausanne University, Epalinges 1066, Switzerland.

<sup>6</sup>Institute of Molecular and Clinical Ophthalmology, Basel 4031, Switzerland

<sup>7</sup>Health 2030 Genome Center, Foundation Campus Biotech Geneva Foundation, Geneva 1202, Switzerland.

<sup>8</sup>iGE3 Institute of Genetics and Genomics of Geneva, Geneva 1211, Switzerland.

<sup>9</sup>Faculty of Biology and Medicine, University of Lausanne, Lausanne 1011, Switzerland

Correspondence to:

[stylianos.antonarakis@unige.ch](mailto:stylianos.antonarakis@unige.ch)

[federico.santoni@chuv.ch](mailto:federico.santoni@chuv.ch)

Antonarakis ORCID: <https://orcid.org/0000-0001-8907-5823>

Santoni ORCID: <https://orcid.org/0000-0002-3258-4747>

Borel ORCID: <https://orcid.org/0000-0002-7525-9418>

Ansar ORCID: <https://orcid.org/0000-0001-7299-3185>

Iwaszkiewicz ORCID: <https://orcid.org/0000-0003-3494-589X>

Zoete ORCID: <https://orcid.org/0000-0002-2336-6537>

## Abstract

Knobloch syndrome is an autosomal recessive phenotype mainly characterized by retinal detachment and encephalocele caused by biallelic pathogenic variants in the *COL18A1* gene. However, there are patients clinically diagnosed as Knobloch syndrome with unknown molecular etiology not linked to *COL18A1*. We studied an historical pedigree (published in 1998) designated as KNO2 (Knobloch type 2 syndrome with intellectual disability, autistic behavior, retinal degeneration, encephalocele). Whole exome sequencing of the two affected siblings and the normal parents resulted in the identification of a *PAK2* non-synonymous substitution p.(Glu435Lys) as a causative variant. The variant was monoallelic and apparently *de novo* in both siblings indicating a likely germline mosaicism in one of the parents; the mosaicism however could not be observed after deep sequencing of blood parental DNA. *PAK2* encodes a member of a small group of serine/threonine kinases; these P21-activating kinases (PAKs) are essential in signal transduction and cellular regulation (cytoskeletal dynamics, cell motility, death and survival signaling, and cell cycle progression). Structural analysis of the *PAK2* p.(Glu435Lys) variant which is located in the kinase domain of the protein predicts a possible compromise in the kinase activity. Functional analysis of the p.(Glu435Lys) *PAK2* variant in transfected HEK293T cells results in a partial loss of the kinase activity. *PAK2* has been previously suggested as an autism related gene. Our results show that *PAK2* induced phenotypic spectrum is broad and not fully understood. We conclude that the KNO2 syndrome in the studied family is dominant and caused by a deleterious variant in the *PAK2* gene.

## Introduction

The sequence analysis and exploration of the human genome has significantly contributed to the identification of the molecular cause of Mendelian disorders. To date, there are 4,587 protein-coding genes linked to Mendelian disorders (OMIM <https://www.omim.org/> search of 17-sep-20); there are however approximately 15,000 additional protein-coding genes that are not yet linked to a disorder of a specific trait. On the other hand, the study of Mendelian phenotypes has often revealed different genes responsible for similar phenotypes<sup>1</sup>. An extreme example is the non-specific intellectual disability with mendelian inheritance, where pathogenic variants in more than 700 different genes are responsible for these phenotypes<sup>2</sup>. Knobloch syndrome (KNO1 OMIM 267750) is an autosomal recessive phenotype mainly characterized by retinal detachment and encephalocele caused by biallelic pathogenic variants in the *COL18A1* gene on chromosome 21<sup>3</sup>. The syndrome may present with different extraocular abnormalities: a single umbilical artery, pyloric stenosis, a flat nasal bridge, midface hypoplasia, bilateral epicanthal folds, abnormal lymphatic vessels in the lung, patent ductus arteriosus, cardiac dextroversion, generalized hyperextensibility of the joints, unusual palmar creases, and unilateral duplication of the renal collecting system. Structural brain defects, epilepsy and cognitive abnormalities have also been frequently reported. However, the full phenotypic spectrum is yet to be defined.

There are patients clinically diagnosed as Knobloch syndrome with unknown molecular etiology not linked to *COL18A1* (OMIM 120328). We have studied an historical pedigree from New Zealand first published in 1998<sup>4</sup>, designated as KNO2 (Knobloch type 2 syndrome). Exome sequence of the two affected siblings and the normal parents resulted in the identification of a heterozygous *PAK2* non-synonymous substitution as a causative variant. The variant was monoallelic and apparently de novo in both siblings due to presumed parental mosaicism. *PAK2* encodes a member of a small group of serine/threonine kinases; these P21-activating kinases (PAKs) are important in signal transduction and cellular regulation (cytoskeletal dynamics, cell motility, death and survival signaling, and cell cycle progression)<sup>5</sup>. Thus, the KNO2 phenotype due to monoallelic pathogenic variants in *PAK2* has a dominant mode of inheritance.

## Methods

### *Family ascertainment and DNA sequencing*

DNA samples from the four members of the family (parents and 2 affected siblings) were stored in the DNA bank of the Geneva laboratory since 2003. These DNAs were processed by whole exome sequencing at the Health 2030 Genome Center using Twist Human Core Exome Kit (TWIST Biosciences, San Francisco, CA, USA); sequencing was performed on an Illumina HiSeq4000 platform.

### *Computational analysis*

The exome-sequencing data were analyzed using a customized pipeline already used in other studies<sup>6; 7</sup> equipped with the Burrows-Wheeler aligner tool (BWA)<sup>8</sup>, SAMtools<sup>8</sup>, PICARD, the Genome Analysis Toolkit (GATK)<sup>9</sup> and Annovar<sup>10</sup>. Briefly, sequenced reads were aligned to the GRCh38 reference human genome with an average coverage of 100x and coverage percentage of RefSeq coding region of 97.6% at 30x. Variant filtering was performed with VariantMaster<sup>11</sup> for de novo, recessive, X-linked, and dominant mode of inheritance.

### *Sanger validation and search for mosaicism*

The primer pair to amplify the sequence containing *PAK2* c.1303 G>A, p.(Glu435Lys) was designed using program Primer3 (<http://primer3.ut.ee>) and checked for SNP absence with the tool SNPCheck V3. The forward oligonucleotide primer sequence is 5'TGGTTACACGGAAAGCTTATG3' while the reverse sequence primer is: 5'TTGAACAAGGCTGGTGTTTAACT3'. An amplicon of 141 bp is amplified.

PCR conditions were set up by using a DNA control (NA12877, Coriell Institute). The polymerase GoTaq G2 Hot Start M743A of Promega was used in the recommended conditions. The PCR amplification program was: 94°C, 3min / 14x (94°C, 45sec / 63°C dT-0.5°C per cycle / 72°C, 45sec) / 35x (94°C, 45sec / 57°C, 30 sec / 72°C, 45 sec) / 72°C, 3min / 10°C, inf. The specificity of the primer pair was checked by agarose gel electrophoresis; a single band at the

expected size was observed followed by a sequencing performed by Microsynth AG. PCR products were analyzed for ultra-deep high throughput sequencing in a NovaSeq (Illumina). The coverage was >10000 reads per sample.

### ***Structural modeling***

The PAK4 kinase domain structure, stored under 5UPL code in the Protein Data Bank, was used as a template for the PAK2 modeling<sup>12</sup>. PAK4 shares 52% of sequence identity with PAK2 kinase domain. The 50 homology models of PAK2 were calculated using Modeller 9.18 software<sup>13</sup>. The best model was chosen based on its best overall DOPE score<sup>14</sup>. In the best PAK2 model the mutation of p.(Glu435Lys) was performed 5 times with BuildModel command of FoldX 5<sup>15</sup>. The average difference in structural stability between the wild type and mutants was calculated with FoldX 5. The UCSF Chimera software was used for structural visualization and figures creation<sup>16</sup>.

### ***Cell transfection***

Glu435Lys mutation was introduced by PCR-based techniques of molecular cloning (extended primers, Forward:

```
AAAAAAAAGCTTATGGCCCTAAAGTCGACATATGGTCTCTGGGTATCATGGCTATTA  
AGATGGTAGAAGGAGAGC; Reverse:  
AAAAAACGGCCGGTTTACTTGTACAGCTCGTCCATGCCGAGAGTGAT)
```

into *PAK2* sequence of a plasmid designed for exogenous expression of eGFP-tagged PAK2 using HindIII and NotI restriction sites present in the original pEGFP-N2-based construct<sup>17</sup>. The plasmids coding for the wild-type and mutated PAK2 were transfected into HEK293T cells using jetPRIME transfection reagent (Polyplus Transfection) following the manufacturer's instructions, the cells were cultivated for 24h, and harvested for western-blot analysis. An aliquot of cell preparation was used to check the transfection efficiency using BD Fortessa flow cytometer.

### ***Western-blotting***

Transfected cells were washed once with ice-cold HBS (HEPES – buffered saline; 20 mM HEPES, 150 mM NaCl, pH 7.1) and scrapped into Pierce IP Lysis Buffer (#87787) with freshly added protease and phosphatase inhibitors. The suspension was then transferred to a centrifugation tube and incubated for 10 min at 4 °C. Cellular debris was removed by centrifugation (15000 g/4°C/15 min), the lysate was mixed 1:1 (v/v) with 2x Laemmli sample buffer and incubated for 5 min at 95°C.

An equivalent of 20 µg of total protein was resolved on 7.5% polyacrylamide gel (18×18 cm) and transferred to a nitrocellulose membrane. The membrane was blocked for 1h in 3% bovine serum albumin and incubated for 1h with the primary antibody in PBS with 0.1% Tween-20 (PBST), at the room temperature. It was then washed in PBST six times and incubated with the corresponding HRP-conjugated secondary antibody for 1h. The chemiluminiscence signal from Clarity Western ECL Substrate (BioRad, #170 5060) was detected and analyzed using G:BOX iChemi XT-4 (Syngene).

### ***Material***

FRAX597 (#6029) was purchased from Tocris Biosciences, dissolved in sterile DMSO to make 10 mM stock solution, and used at 10 µM final concentration to inhibit PAK kinase activity. The antibodies against total PAK2 (ab76293) and phosphorylated Ser141/144 of PAK2/PAK1 (ab40795) were purchased from Abcam, β-actin antibody was from Sigma-Aldrich (A5441).

## **Results**

### **Family description**

We have re-studied samples from 4 members of the New Zealand family (NZKS2<sup>18</sup>, first published in <sup>4</sup>) that has been designated as a Knobloch 2 syndrome family (KNO2). This non-consanguineous family consisted of normal parents and two affected offspring. The family was studied in the Geneva laboratory in 2007 and since then no additional tests have been done, and the elusive *KNO2* gene has not been identified. In the phenotypic assessment in 2007 the older

affected brother (individual NZKS2-II-1 in <sup>18</sup>) had retinal detachment of the left eye, inferior retinal detachment of the right eye, and considerable vitreous condensation in both eyes. There was no encephalocele, but the patient had a distinct tuft of dark hair in the superior occipital midline. Chest X-ray showed a mild prominence of the main pulmonary arteries and extensive interstitial changes in the mid and lower zones with bronchial wall thickening and reticular nodular opacities. His cognition at 8 years of age was normal. The younger brother (NZKS2-II-2 in <sup>18</sup>) had a total retinal detachment on the left eye with no light perception and anterior cortical cataract. The majority of the right eye retina was flat, with a traction band running downwards from the disc to a peripheral vitreoretinal mass. There was an encephalocele at the midline in the parietal-occipital region that was surgically removed. Like his brother, he also had interstitial parenchymal pulmonary changes on chest X-rays, suggestive of dilated lymphatics. He also had severe developmental delay at age 6 years associated with an autistic spectrum disorder (ASD), with hardly any comprehensible speech. The previous molecular analysis of the family has excluded pathogenic variants in the *COL18A1* gene; moreover, linkage to chromosome 21 markers for a recessive disease was excluded<sup>18</sup>.

### **PAK2 « de novo » variant p.(Glu435Lys)**

The computational analysis of the exome sequencing from the samples of all four available members of the family (father, mother, and the two affected siblings) has identified a *de novo* variant (not present in parental DNA) in each sibling in the gene PAK2.

The *de novo* variant was a non-synonymous substitution p.(Glu435Lys) (NM\_002577.4 :c.1303G>A, chr3: 196820520, hg38). The variant was categorized as pathogenic based on the following criteria of the ACMG<sup>19</sup>: PS2 (de novo), PS3 (functional studies), PM2 (not found in databases), PP3 (computational prediction of pathogenicity). The variant was absent from the gnomAD database (<https://gnomad.broadinstitute.org/>) V2.1.1. of 282,000 alleles and V3 of 142,000 alleles from genome and exome sequences and the BRAVO Topmed2 database (<https://bravo.sph.umich.edu/freeze8/hg38/>) of 264,000 alleles from genome sequences.

The heterozygous variant *PAK2* c.1303 G>A, p.(Glu435Lys) was confirmed by Sanger sequencing in the two affected siblings and it was absent in the parents DNA (Figure 1A). The variant was classified as damaging by SIFT, Provean, mutation taster, primateAI, polyphen2; the CADD score was 32. Amino acid residue Glu435 is well conserved (Figure 1B) with a GERP score of 4.7.

We excluded an eventual mis-paternity by calculating the percentage of high quality variants (coverage >20, GATK Quality score >200, MAF>1%) shared between I-1(father) and I-2(mother) with sons (II-1 and II-2). The sharing father-sons was 52% and the sharing mother-sons was 51.5% (Figure 2). We thus hypothesized that one of the parents has germline mosaicism potentially detectable in blood. Therefore, we PCR-amplified the genomic fragment that includes the variant and subjected the PCR product to high-throughput sequencing. More than 10000 reads were analyzed in each of the four samples (father, mother, and two affected siblings). While the variant was observed in approximately half of the reads in each affected sibling, we did not find any reads with the variant in the amplicon of the blood DNA from the parents. Therefore, we concluded that one of the parents is germline mosaic for the variant, and the mosaicism is absent from blood cells.

### **3D structure prediction of the *PAK2* p.(Glu435Lys) variant**

The modeling of the *PAK2* de novo variant was performed on the homology model based on the *PAK4* kinase structure<sup>12</sup>. The p.(Glu435Lys) mutation in *PAK2* is located in the center of the kinase domain, below the ATP binding pocket (Figure 3A). The Glu435 residue seems to be structurally important as it is involved in three hydrogen bond interactions: two of them with Ser371 side-chain and backbone and one to the side-chain of Trp409. The three residues are involved in the stability and conformation of helical structures (Figure 3B).

According to FoldX 5.0 calculations, the loss of stability upon Glu435 to Lys435 mutation is not substantial (0.26 kcal/mol with SD of 0.007), as, after the simulated mutation, p.(Glu435Lys) could still make a hydrogen bond to the side chain of Ser371. Nevertheless, the mutation most likely destabilizes the local structure, as the formation of hydrogen bond with Trp409 residue is



no longer possible. The region with the mutated residue is close to the ATP binding site suggesting that the slightest changes in structure stability might affect PAK2 kinase functions (Figure 3C).

### **Functional analysis of the PAK2 p.(Glu435Lys) variant**

To uncover the possible impact of the p.(Glu435Lys) variant on the PAK2 kinase activity, we analyzed the extent of PAK2 phosphorylation at Ser141. This site is autophosphorylated when the kinase domain is active, and it can thus serve as a marker of PAK2 kinase activity<sup>12; 20</sup>. HEK293T cells were transfected with plasmids for exogenous expression of the wild-type PAK2 (PAK2-wt) or of the mutated form (PAK2-E435K). The exogenous proteins were tagged with the green fluorescent protein (eGFP) and the transfection efficiency could thus be checked by flow cytometry: both PAK2-wt-eGFP and PAK2-E435K-eGFP were expressed in about 30% of cells, and eGFP-positive cells had closely similar mean fluorescence intensity comparing the PAK2 variants. The level of PAK2 phosphorylation at Ser141 was determined by western-blotting, using previously characterized antibodies<sup>17</sup>. The PAK2 p.(Glu435Lys) variant was associated with substantial reduction of phosphorylation at Ser141 of the exogenous protein, labeled as PAK2-GFP in Figure 4, panel A (pSer141/144 antibody, lane 1 versus lane 3). On the other hand, no change was induced in the total PAK2 level (panel A and B, PAK2-GFP and PAK2 endogenous). For comparison, we included samples treated with an inhibitor of PAK kinase activity, FRAX597 (Figure 4, panel A, lanes 2 and 4). As expected, the inhibitor also decreased Ser141/144 phosphorylation of the endogenous PAK2 and PAK1. We concluded that the PAK2 p.(Glu435Lys) variant results in a substantial loss of the kinase activity.

### **Discussion**

#### *Mosaicism in de novo*

The occurrence of the same pathogenic *PAK2* variants in the two affected siblings, that was absent from blood DNA of the parents, is likely to be due to parental mosaicism, i.e. a fraction of cells in the gametes of one of the parents that harbor the pathogenic variant. This phenomenon is not rare in cases of “de novo” dominant phenotypes. In a recent study of 120 apparently “de

novo” cases of epileptic encephalopathy<sup>21</sup>, there were 10 instances of proven parental mosaicism (6 from fathers and 4 from mothers; minor allele frequency, 1.4 to 30.6%; mean number of mosaic reads after sequencing of blood or saliva DNA was 12.9%). In addition there were 2 of these 120 cases with two affected siblings in which there was an inferred mosaicism undetected in the parents. These two last cases are similar to the findings described in our study. In another study of Dravet syndrome<sup>22</sup> there were 20 of 174 “de novo” SCN1A mutation cases with parental mosaicism detected in blood DNA (13 from fathers and 7 from mothers). In the years before high throughput, sequencing mosaicism has been extensively discussed. For example parental germline mosaicism has been documented in up to 5% of mothers of patients with Duchenne muscular dystrophy<sup>23</sup>, and in 11% of mothers of patients with hemophilia A<sup>24</sup>. The detection of parental mosaicism is clinically relevant since it modifies the recurrence risk.

#### *PAK gene family related Mendelian disorders*

PAK2 is a member of a small group of serine/threonine kinases; these P21-activating kinases (PAKs) are essential in signal transduction and cellular regulation (cytoskeletal dynamics, cell motility, death and survival signaling, and cell cycle progression). They were initially discovered as binding proteins of small GTPases<sup>25</sup>, and are effectors of CDC42<sup>26</sup> and RAC1 small GTPases (see <sup>5</sup> for review). This family of protein kinases is divided in two groups: group 1 includes PAK1, PAK2 and PAK3, while group 2 consists of PAK4, PAK5 (also called PAK7), and PAK6<sup>5</sup>. Mendelian disorders have been identified to date for: i/ PAK1 OMIM 602590), an autosomal dominant intellectual developmental disorder with microcephaly, seizures and speech delay (OMIM 618158)<sup>27</sup>; ii/ PAK3 (OMIM 300142), an X-linked intellectual disability type 30 (OMIM 300558)<sup>28</sup>. PAK2 is highly intolerant to Loss of Function (LoF) mutations (pLI =0.97) and to missense ( $Z_{miss}=3.4$ ). According to BrainSpan (<https://www.brainspan.org/>), PAK2 is mainly expressed during the embryonic development in most of developing cerebral structures. PAK2 expression decreases significantly after 25 pcw (post conception weeks) and reaches a plateau at 2-3 years of age (Figure 5).

#### *Functional consequence of the PAK2 p.(Glu435Lys) variant*

Glu435 is located within the PAK2 kinase domain. We provide evidence of a large decrease of phosphorylation at the autophosphorylation site Ser141, reflecting a loss of kinase activity<sup>12; 20</sup> due to p.Glu435Lys point mutation. Thus, the acidic character of PAK2 residue 435 seems to be

important for PAK2 catalytic function. Similarly, an acidic substitution of Thr402 (T402E), which is another autophosphorylation site, rendered PAK2 catalytically active<sup>29</sup>.

PAK group I, including PAK2, are involved in a variety of cellular processes, and mice with PAK2 knock-out do not survive an early embryonic stage of development<sup>30</sup>. However, some of the PAK2 functions are not kinase-dependent<sup>31</sup> and might be preserved in the mutated PAK2. Presumably, considering the expression profile in early prenatal development (Figure 5), the mutation could affect cell adhesion, migration, or apoptosis, but the importance of PAK2 kinase activity for those processes is currently not fully understood.

### *Knobloch syndrome*

The autosomal recessive Knobloch syndrome (KNO1 OMIM 267750) has been caused by biallelic, pathogenic variants in the *COL18A1* gene<sup>3</sup>. Menzel et al. concluded that pathogenic variants in the *COL18A1* gene did not cause the phenotype of the family from New Zealand; thus it was proposed to name the syndrome of this family as KNO2, Knobloch syndrome 2<sup>18</sup>). In this study we have identified an autosomal dominant variant in the *PAK2* gene as the likely cause of KNO2. A possible functional connection might exist between *COL18A1* and *PAK2* since both are associated with integrin complexes (*COL18A1* through proteolytic cleavage into endostatin<sup>32</sup>, *PAK2* localizes with integrin associated adhesion complexes<sup>17</sup> and is activated by RAC1 downstream of integrin matrix receptors<sup>33</sup>). The hypothesis that both genes are involved in the same integrin-driven pathway might explain the similarity of the phenotypic spectrum between KNO1 and KNO2.

To date, one patient with a de novo damaging mutation in *PAK2* has been reported<sup>30</sup>. This patient was part of a Chinese ASD (Autism spectrum disorder) cohort but, unfortunately, we were not able to obtain his complete phenotype profile. The study does not provide statistical evidence that the de novo *PAK2* variant could be considered as causative to the Autism phenotype. In the same study, it was also shown that mice with heterozygous *PAK2* missense mutations display autism-related behaviors.

The family described here adds to the suggestion of the implication of *PAK2* likely pathogenic variants in ASD; the phenotype of the family is mainly that of a retinal disorder and skull bone fusions defects. Therefore, at this stage the extent of the phenotypic expression of pathogenic

variants in the *PAK2* gene has not yet been completed. We propose to keep the KNO2 phenotypic designation in this family for historical reasons, but to reconsider the syndrome name upon addition of more patients due to *PAK2* high impact pathogenic variants.

## Acknowledgments

The study was supported by a grant from the BLACKSWAN Foundation, Swiss foundation for research on orphan diseases (<https://www.blackswanfoundation.ch>), and by the Ministry of Health, Czech Republic (project for conceptual development of the research organization No 00023736 to KK and AH). We thank the Protein Modeling Facility of University of Lausanne for the bioinformatic support, and E. Ranza for assistance in the phenotypic classification. The study was partially supported by the Swiss National Science Foundation (310030\_185292) and Horizon2020 (847941) to F.A.S, and the ChildCare foundation to S.E.A.

## Figure Legends

*Figure 1. A) Family pedigree, segregation and Sanger validation of the variant PAK2: c.1303G>A. B) Glu435 is a highly conserved amino acid from human to zebrafish.*

*Figure 2. Percentage of variants (MAF>1%) shared between A) father (I-1) B) mother (I-2) and sons. In black, the sharing with 100 unrelated individuals (average father-unrelated sharing  $0.33\pm 0.026$ , average mother-unrelated sharing  $0.32\pm 0.027$ , the best 7 are represented).*

*Figure 3. A) General structure of wild type PAK2 kinase domain model and interactions of Glu435 residue with Ser371 and Trp409 and detailed view of Glu435 residue interactions. B) Detailed view of predicted p.(Glu435Lys) mutant interactions. PAK2 is shown in ribbon representation (magenta), Glu435Lys (orange), Ser371, Thr409 residues are shown in ball and stick representation. The hydrogen bonds are shown as thin blue lines.*

*Figure 4. Impact of Lys435Glu mutation on Ser141 PAK2 phosphorylation. A) HEK293T cells were transfected with plasmids for exogenous expression of eGFP-tagged wild-type PAK2*

(PAK2-wt) or mutated PAK2 (PAK2-E435K), and the phosphorylation status of Ser141 was analyzed using western-blotting. Some samples were treated for 1h with PAK kinase inhibitor, FRAX597, at 10  $\mu$ M concentration, prior to cell lysis. B) Average pPAK2/PAK2 band intensity and standard error from four replicates. Relative band intensities from the phospho-specific antibody (pSer141/144) were normalized to PAK2-total antibody intensities.

Figure 5. PAK2 expression (log2 RPKM) in cortical and subcortical structures across the full course of human brain development, as provided by BrainSpan. The gene is mainly expressed during the early stages of brain development; its expression decreases during pregnancy and reaches a plateau at 2-3yrs of age. For better visualization of the expressivity trend over time, the age distances have been equispaced. The red line and shaded area represent the time loess expression average across brain structures and the 95% confidence intervals respectively. pcw – post-conception weeks. Brain structure acronyms in <https://www.brainspan.org>

## Bibliography

1. Antonarakis, S.E., and Beckmann, J.S. (2006). Mendelian disorders deserve more attention. *Nature reviews Genetics* 7, 277-282.
2. Vissers, L.E., Gilissen, C., and Veltman, J.A. (2016). Genetic studies in intellectual disability and related disorders. *Nature reviews Genetics* 17, 9-18.
3. Sertie, A.L., Sossi, V., Camargo, A.A., Zatz, M., Brahe, C., and Passos-Bueno, M.R. (2000). Collagen XVIII, containing an endogenous inhibitor of angiogenesis and tumor growth, plays a critical role in the maintenance of retinal structure and in neural tube closure (Knobloch syndrome). *Human molecular genetics* 9, 2051-2058.
4. Wilson, C., Aftimos, S., Pereira, A., and McKay, R. (1998). Report of two sibs with Knobloch syndrome (encephalocoele and vitreoretinal degeneration) and other anomalies. *Am J Med Genet* 78, 286-290.

5. Kumar, R., Sanawar, R., Li, X., and Li, F. (2017). Structure, biochemistry, and biology of PAK kinases. *Gene* 605, 20-31.
6. Ansar, M., Chung, H., Waryah, Y.M., Makrythanasis, P., Falconnet, E., Rao, A.R., Guipponi, M., Narsani, A.K., Fingerhut, R., Santoni, F.A., et al. (2018). Visual impairment and progressive phthisis bulbi caused by recessive pathogenic variant in MARK3. *Human molecular genetics*.
7. Ansar, M., Chung, H.L., Al-Otaibi, A., Elagabani, M.N., Ravenscroft, T.A., Paracha, S.A., Scholz, R., Abdel Magid, T., Sarwar, M.T., Shah, S.F., et al. (2019). Bi-allelic Variants in IQSEC1 Cause Intellectual Disability, Developmental Delay, and Short Stature. *American journal of human genetics* 105, 907-920.
8. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754-1760.
9. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics* 43, 491-498.
10. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38, e164.
11. Santoni, F.A., Makrythanasis, P., Nikolaev, S., Guipponi, M., Robyr, D., Bottani, A., and Antonarakis, S.E. (2014). Simultaneous identification and prioritization of variants in familial, de novo, and somatic genetic disorders with VariantMaster. *Genome research* 24, 349-355.
12. Ha, B.H., and Boggon, T.J. (2018). CDC42 binds PAK4 via an extended GTPase-effector interface. *Proc Natl Acad Sci U S A* 115, 531-536.
13. Webb, B., and Sali, A. (2016). Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics* 54, 5 6 1-5 6 37.
14. Shen, M.Y., and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Sci* 15, 2507-2524.
15. Delgado, J., Radusky, L.G., Cianferoni, D., and Serrano, L. (2019). FoldX 5.0: working with RNA, small molecules and a new graphical interface. *Bioinformatics* 35, 4168-4169.
16. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25, 1605-1612.
17. Grebenova, D., Holoubek, A., Roselova, P., Obr, A., Brodska, B., and Kuzelova, K. (2019). PAK1, PAK1Delta15, and PAK2: similarities, differences and mutual interactions. *Sci Rep* 9, 17171.
18. Menzel, O., Bekkeheien, R.C., Reymond, A., Fukai, N., Boye, E., Kosztolanyi, G., Aftimos, S., Deutsch, S., Scott, H.S., Olsen, B.R., et al. (2004). Knobloch syndrome: novel mutations in

- COL18A1, evidence for genetic heterogeneity, and a functionally impaired polymorphism in endostatin. *Hum Mutat* 23, 77-84.
19. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 17, 405-424.
  20. Gatti, A., Huang, Z., Tuazon, P.T., and Traugh, J.A. (1999). Multisite autophosphorylation of p21-activated protein kinase gamma-PAK as a function of activation. *J Biol Chem* 274, 8022-8028.
  21. Myers, C.T., Hollingsworth, G., Muir, A.M., Schneider, A.L., Thuesmann, Z., Knupp, A., King, C., Lacroix, A., Mehaffey, M.G., Berkovic, S.F., et al. (2018). Parental Mosaicism in "De Novo" Epileptic Encephalopathies. *N Engl J Med* 378, 1646-1648.
  22. Xu, X., Yang, X., Wu, Q., Liu, A., Yang, X., Ye, A.Y., Huang, A.Y., Li, J., Wang, M., Yu, Z., et al. (2015). Amplicon Resequencing Identified Parental Mosaicism for Approximately 10% of "de novo" SCN1A Mutations in Children with Dravet Syndrome. *Hum Mutat* 36, 861-872.
  23. Helderma-van den Enden, A.T., de Jong, R., den Dunnen, J.T., Houwing-Duistermaat, J.J., Kneppers, A.L., Ginjaar, H.B., Breuning, M.H., and Bakker, E. (2009). Recurrence risk due to germ line mosaicism: Duchenne and Becker muscular dystrophy. *Clin Genet* 75, 465-472.
  24. Erickson, R.P. (2010). Somatic gene mutation and human disease other than cancer: an update. *Mutat Res* 705, 96-106.
  25. Manser, E., Leung, T., Salihuddin, H., Zhao, Z.S., and Lim, L. (1994). A brain serine/threonine protein kinase activated by Cdc42 and Rac1. *Nature* 367, 40-46.
  26. Jung, J.H., and Traugh, J.A. (2005). Regulation of the interaction of Pak2 with Cdc42 via autophosphorylation of serine 141. *J Biol Chem* 280, 40025-40031.
  27. Harms, F.L., Kloth, K., Bley, A., Denecke, J., Santer, R., Lessel, D., Hempel, M., and Kutsche, K. (2018). Activating Mutations in PAK1, Encoding p21-Activated Kinase 1, Cause a Neurodevelopmental Disorder. *American journal of human genetics* 103, 579-591.
  28. Allen, K.M., Gleeson, J.G., Bagrodia, S., Partington, M.W., MacMillan, J.C., Cerione, R.A., Mulley, J.C., and Walsh, C.A. (1998). PAK3 mutation in nonsyndromic X-linked mental retardation. *Nature genetics* 20, 25-30.
  29. Pulkkinen, K., Renkema, G.H., Kirchhoff, F., and Saksela, K. (2004). Nef associates with p21-activated kinase 2 in a p21-GTPase-dependent dynamic activation complex within lipid rafts. *J Virol* 78, 12773-12780.

30. Wang, Y., Zeng, C., Li, J., Zhou, Z., Ju, X., Xia, S., Li, Y., Liu, A., Teng, H., Zhang, K., et al. (2018). PAK2 Haploinsufficiency Results in Synaptic Cytoskeleton Impairment and Autism-Related Behavior. *Cell Rep* 24, 2029-2041.
31. Kosoff, R.E., Aslan, J.E., Kostyak, J.C., Dulaimi, E., Chow, H.Y., Prudnikova, T.Y., Radu, M., Kunapuli, S.P., McCarty, O.J., and Chernoff, J. (2015). Pak2 restrains endomitosis during megakaryopoiesis and alters cytoskeleton organization. *Blood* 125, 2995-3005.
32. Heljasvaara, R., Aikio, M., Ruotsalainen, H., and Pihlajaniemi, T. (2017). Collagen XVIII in tissue homeostasis and dysregulation - Lessons learned from model organisms and human patients. *Matrix Biol* 57-58, 55-75.
33. del Pozo, M.A., Alderson, N.B., Kiosses, W.B., Chiang, H.H., Anderson, R.G., and Schwartz, M.A. (2004). Integrins regulate Rac targeting by internalization of membrane domains. *Science* 303, 839-842.



## Figures

Figure 1

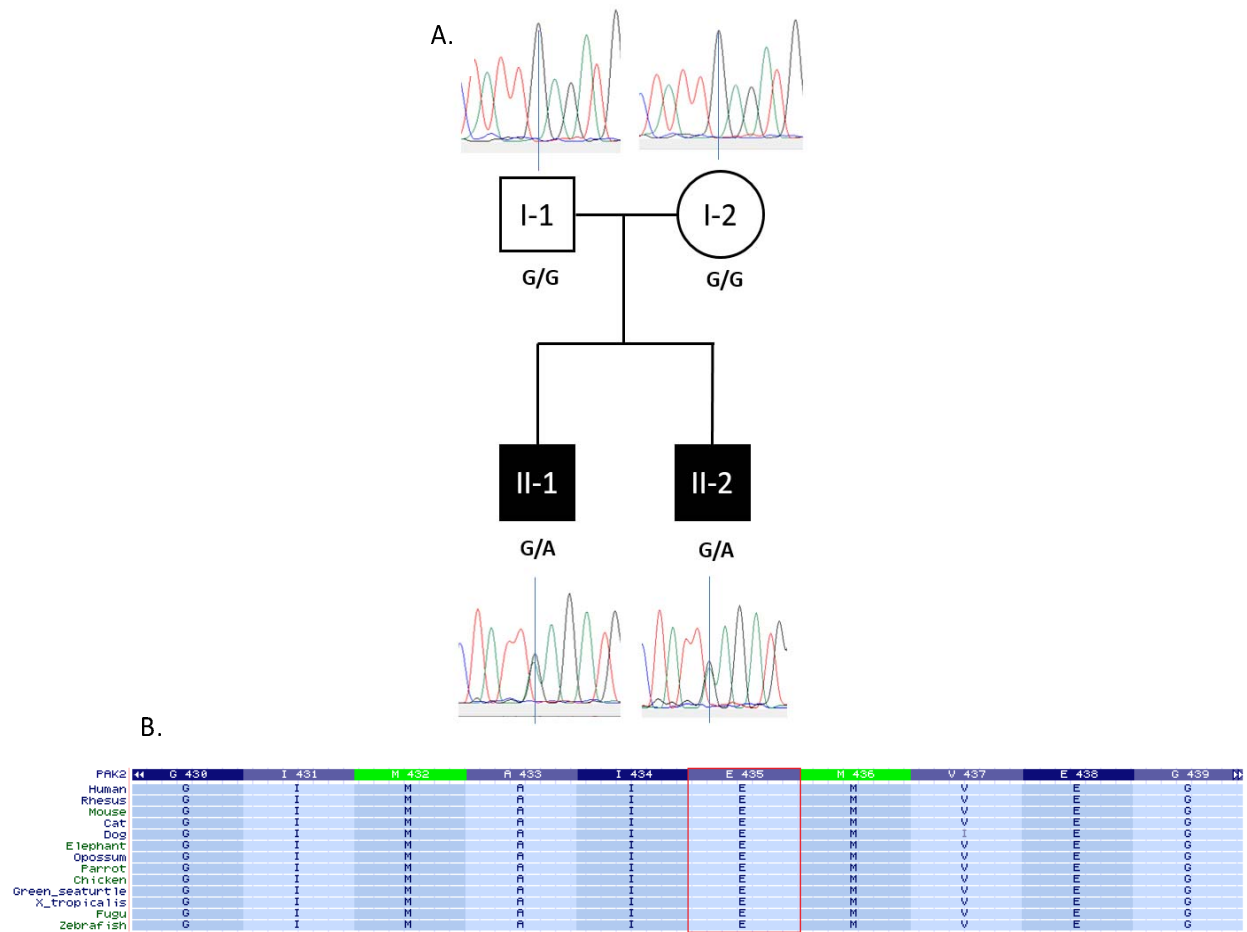


Figure 2

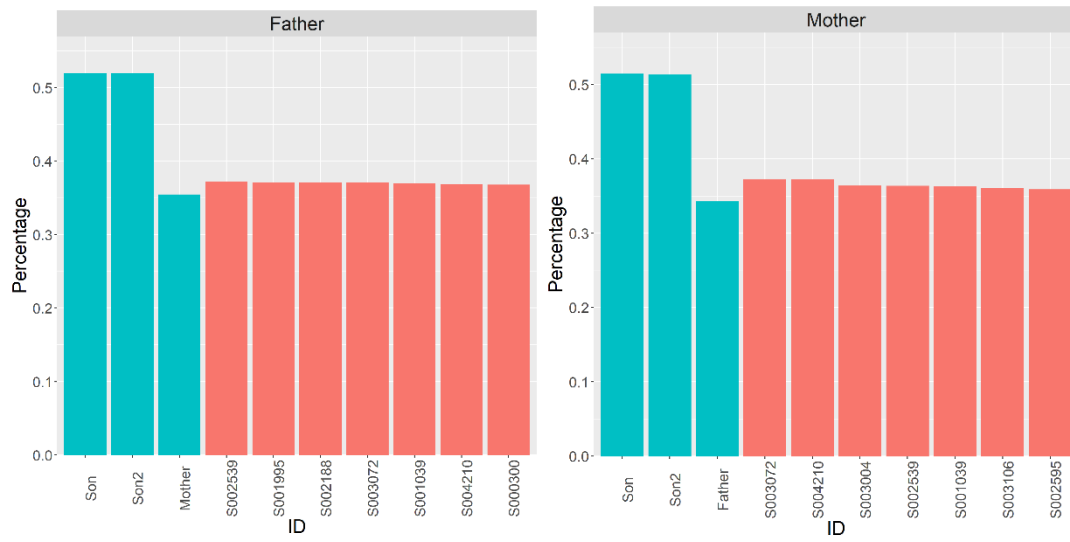
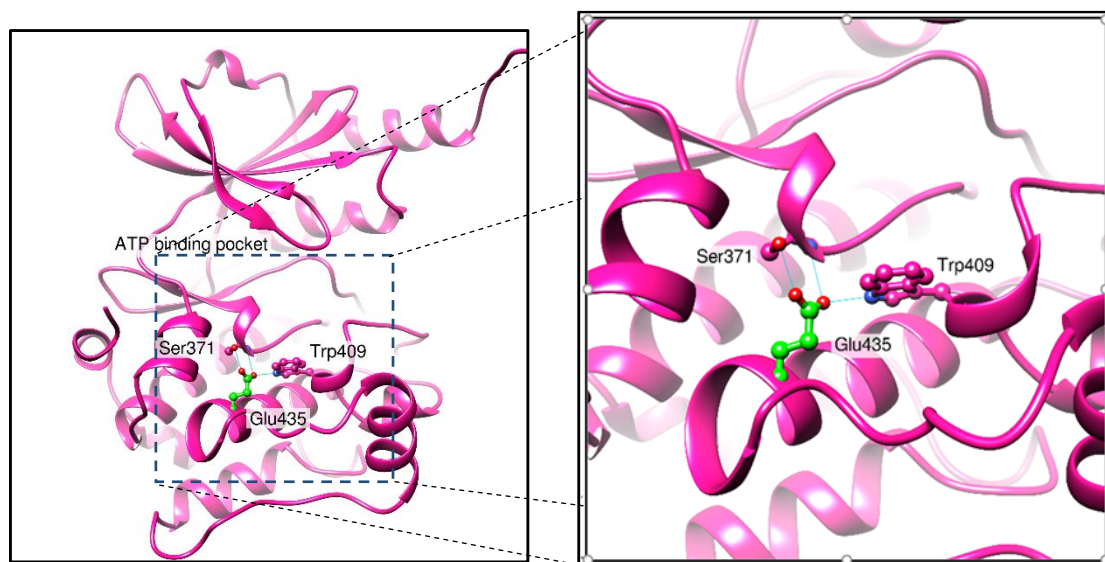


Figure 3

A.



B.

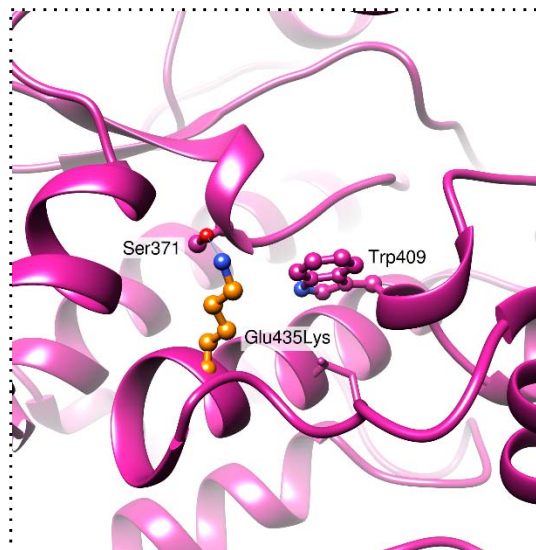


Figure 4

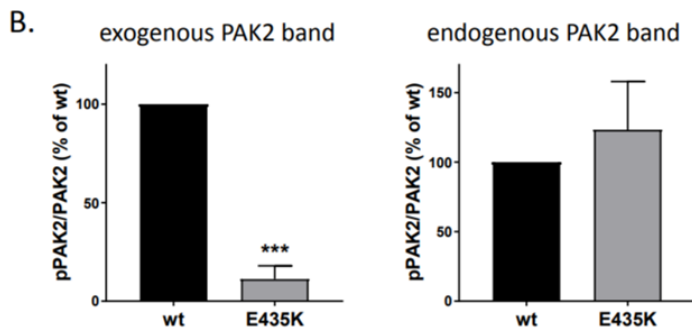
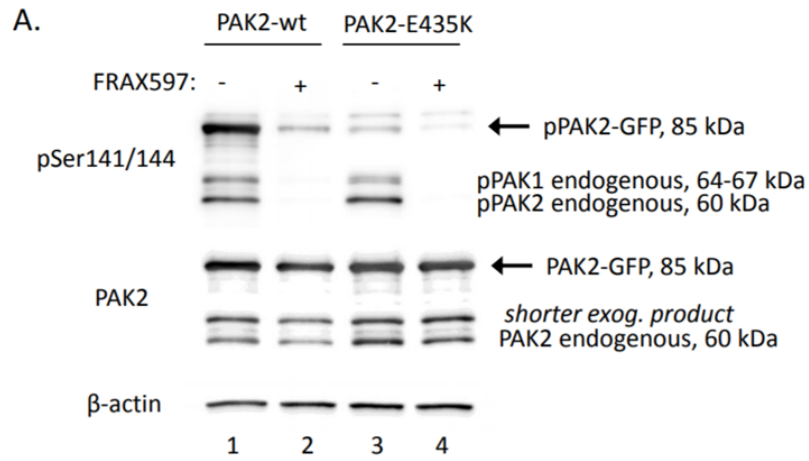


Figure 5

