Functional analysis of *TLK2* variants and their proximal interactomes implicates impaired kinase activity and chromatin maintenance defects in their pathogenesis.

Lisa Pavinato [1], Marina Villamor-Payà [2], Maria Sanchiz-Calvo [2], Cristina Andreoli [3], Marina Gay [2], Marta Vilaseca [2], Gianluca Arauz-Garofalo [2], Andrea Ciolfi [4], Alessandro Bruselles [5], Tommaso Pippucci [6], Valentina Prota [3], Diana Carli [7], Elisa Giorgio [1], Francesca Clementina Radio [4] Vincenzo Antona [8], Mario Giuffrè [8], Kara Ranguin [9], Cindy Colson [9], Silvia De Rubeis [10, 11, 12, 13], Paola Dimartino [14], Joseph Buxbaum [10, 11, 12, 13, 15, 16], Giovanni Battista Ferrero [7], Marco Tartaglia [4], Simone Martinelli [5], Travis H. Stracker [*2], Alfredo Brusco [*1, 17, 18]

1. Department of Medical Sciences, University of Turin, 10126 Turin, Italy
2. Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, 08028 Barcelona, Spain
3. Department of Environment and Health, Istituto Superiore di Sanità, 00100 Rome, Italy
4. Genetics and Rare Diseases Research Division, Ospedale Pediatrico Bambino Gesù, Rome, Italy
5. Department of Oncology and Molecular Medicine, Istituto Superiore di Sanità, Rome 00161, Italy.
6. Medical Genetics Unit, Polyclinic Sant'Orsola-Malpighi University Hospital, Bologna, Italy
7. Department of Pediatrics and Public Health and Pediatric Sciences, University of Turin, 10126 Turin, Italy
8. Department of Sciences for Health Promotion and Mother and Child Care "G. D'Alessandro", University of Palermo, Palermo, Italy
9. Centre de référence Maladies rares et Anomalies du développement, Service de génétique, CAEN, France
10. Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.
11. Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.
12. The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.
13. Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.
14. Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy
15. Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
16. Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
17. "Città della Salute e della Scienza" University Hospital, Unit of Medical Genetics, 10126 Turin, Italy
18. Corresponding author: Prof. Alfredo Brusco, University of Torino, Department of Medical Sciences, via Santena 19, 10126, Torino, Italy. E-mail: alfredo.brusco@unito.it

*These authors equally contributed to the work

Running title: TLK2-associated intellectual disability

Keywords: intellectual disability, TLK2, comet assay, chromatin condensation, BioID, proximal interactome, autism spectrum disorder, tousled like kinases, MRD57

Additional information: None of the authors has any competing financial or non-financial interest to disclose.
Word count: 3902

ABSTRACT

Introduction The Tousled-Like Kinases 1 and 2 (TLK1 and TLK2) are involved in many fundamental processes, including DNA replication, cell cycle checkpoint recovery and chromatin remodelling. Mutations in *TLK2* were recently associated with "Mental Retardation Autosomal Dominant 57" (MRD57, OMIM 618050), a neurodevelopmental disorder characterized by a highly variable phenotype, including mild-to-moderate intellectual disability, behavioural abnormalities, facial dysmorphisms, microcephaly, epilepsy, and skeletal anomalies.

Methods By whole exome sequencing and array-CGH analysis, we identified three unrelated MRD57 cases. Two were sporadic and caused by a missense change (c.1652A>G; p.(Asp551Gly)) or a 39-kb deletion encompassing *TLK2*, and one was familial with three affected siblings who inherited a nonsense change from an affected mother (c.1423G>T; p.(Glu475Ter)). Using spatial proteomics (BioID) and single-cell gel electrophoresis, we investigated the proximity interaction landscape of *TLK2* and analysed the effects of p.(Asp551Gly) and a previously reported missense variant (c.1850C>T; p.(Ser617Leu)) on TLK2 interactions, localization and activity.

Results The clinical phenotypes were consistent with those of previously reported cases. The two tested mutations strongly impaired *TLK2* kinase activity. Proximal interactions between TLK2 and other factors implicated in neurological disorders, including CHD7, CHD8, BRD4, NACC1, were identified. Finally, we demonstrated a more relaxed chromatin state in lymphoblastoid cells harbouring the p.(Asp551Gly) variant compared to control cells, conferring susceptibility to DNA damage.

Conclusion Our study identified novel *TLK2* pathogenic variants, confirming, and further expanding the MRD57 related phenotype. The molecular characterization of missense variants increases our knowledge about *TLK2* function and provides new insights into its role in neurodevelopmental disorders.

INTRODUCTION

The Tousled-Like Kinases 1 and 2 (TLK1 and TLK2) are serine-threonine kinases involved in DNA replication and repair, transcription, cell cycle checkpoint recovery, chromatin maintenance and genomic stability[1–3]. Both kinases target the ASF1A and ASF1B histone H3/H4 chaperones and are regulated by DNA damage responsive checkpoint signalling[4]. Depletion of *Tlk1* and *Tlk2* in mice indicated that they are largely redundant, with the exception of an essential role for *Tlk2* in placental development[5]. Null *Tlk2* mice, generated with a conditional allele to bypass the placental defect, showed no gross developmental defects except for a slight growth delay compared to controls, suggesting that TLK1 can compensate for loss of TLK2 function after embryonic development. Further work in human cancer cells supports extensive redundancy at the cellular level. Co-depletion of TLK1 and TLK2, however, causes replication stress, DNA damage and altered chromatin maintenance, particularly affecting telomeres and other repetitive genome elements, while mild effects were observed with depletion of either proteins[6].

Mutations in *TLK2* were associated with Mental Retardation Autosomal Dominant (MRD57, MIM: 618050) by Reijnders *et al.*[7], who described 40 cases from 38 unrelated families. MRD57 is clinically characterized by autism spectrum disorder (ASD), intellectual disability (ID), behavioural problems, growth delay and facial dysmorphism, including blepharophimosis, telecanthus, prominent nasal bridge, broad nasal tip, thin vermilion of the upper lip and upslanting palpebral fissures. Other features shared by a subset of cases are gastrointestinal problems, seizures, skeletal malformations and ocular problems.

In the majority of reported cases, the disease is likely due to *TLK2* haploinsufficiency, as most of cases are heterozygous for loss-of-function (LoF) alleles, which is in line with the strong constraint against LoF variants in the gene (pLI = 1, GnomAD database). Reported missense variants cluster mainly in the C-terminal Protein Kinase Domain (PKD)[7–9], the core of TLK2 function. Four *TLK2* variants localized in this region have previously been analysed, showing a strong reduction of kinase activity on ASF1A *in vitro*[10]. TLK2 activation requires dimerization through an N-terminal coiled coil motif, suggesting that inactive mutants could have a dominant negative effect[10]. Thus, based on the available data, the predominant pathogenic mechanism of *TLK2* mutations appears to be a reduction in its overall activity. Of note, recent work suggested a possible

autosomal recessive phenotype in a proband affected by severe neurodevelopmental disease with a homozygous missense variant (c.163A>G; p.(Lys55Glu))[8]. This variant is localized outside the TLK2 PKD and coiled-coil motifs, and may be hypomorphic, since the carrier parents are not affected.

During the screening of a large survey of patients with ID by array-CGH and whole exome sequencing (WES), we identified three new cases with *TLK2*-mutations. Two subjects were heterozygous for a *de novo* 39-kb deletion encompassing *TLK2,* and a *de novo* c.1652A>G; p.(Asp551Gly) missense change. The third case was familial, with a nonsense variant (c.1423G>T; p.(Glu475Ter)) occurring in three affected siblings and their affected mother. Here, we report their clinical description, confirming and expanding the disease phenotype. We also provide data documenting an altered chromatin state in patient-derived fibroblasts and lymphoblastoid cells (LCLs), consistent with a defect in the regulation of histone chaperones. Finally, we characterized the activity and proximal interactomes of the p.(Asp551Gly) variant, as well as another missense variant (p.(Ser617Leu)), identified in a recently published study[11]. Both variants exhibited impaired kinase activity and TLK2 proximal interactomes were enriched with proteins previously implicated in ID and ASD, suggesting connections to a larger chromatin maintenance network.

MATERIALS AND METHODS

Whole exome sequencing, prioritization, and variant calling

DNA was extracted from total blood using the ReliaPrep Blood gDNA Miniprep kit (Promega, Madison, WY, USA) following manufacturer's protocol and quantified with a NanoDrop spectrophotometer (Thermo Fisher Scientifics, Waltham, MA, USA). Informed consent was obtained from participating families and the study protocol was approved by our internal Ethics Committee, according to the Declaration of Helsinki.

Array-CGH was performed using a 60K whole-genome oligonucleotide microarray (Agilent Technologies, Santa Clara, California, USA). Family 1 and 2 were enrolled in the Autism Sequencing Consortium (ASC) project and their gDNA samples were sequenced at the Broad Institute on Illumina HiSeq sequencers as previously described[11, 12]; variant calling was performed using targeted bioinformatic pipelines adapted for

different pattern of inheritance. Identified variants were confirmed by Sanger sequencing using standard conditions and the primers in table S1. Additional information is provided in Supplementary materials and methods.

All variants are referred to GRCh37 annotation and to NM_001284333.2, in line with the previously published TLK2 structure[10]. For homogeneity with the clinical work from Reijnders *et al.*[7], we specified variants also in NM_006852.6 in Supplementary table S2.

In silico prediction of missense variants impact and splicing analysis

Variants were analysed with the VarSome tool[13] as a starting point for further analysis. This allowed evaluation of at least 15 *in silico* predictors simultaneously. Variants frequencies were evaluated using Genome Aggregation Database (GnomAD) Browser version 2.1.1. Impaired splicing was predicted using Human Splicing Finder (HSF) version 3.1[14] and experimentally verified as described in Supplementary materials and methods.

Cell cultures

Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood using Histopaque®-1077 (Sigma-Aldrich) and subsequently immortalized with Epstein-Barr Virus (EBV) and cultured in RPMI medium (GIBCO, Thermo Fisher Scientific) supplemented with 10% Fetal Bovine Serum (FBS) (GIBCO), 1% Pen-Strep and 1% L-Glutamine. Primary fibroblasts were isolated from human skin biopsies after overnight incubation in Dulbecco's modified Eagle's medium (DMEM) (Sigma-Aldrich) supplemented with 10% FBS and 160 µg/ml collagenase. Fibroblasts were maintained in DMEM supplemented with 10% FBS, 1% Pen-Strep and 1 mM sodium pyruvate (Thermo Fisher Scientific) at 37°C, 5% CO_2. AD-293 cells (Stratagene) were grown in DMEM supplemented with 10% FBS (Sigma-Aldrich), 50 U/mL penicillin and 50 µg/mL streptomycin (Thermo Fisher Scientific) at 37ºC in a 5% CO_2 incubator.

RNA isolation and quantitative real time PCR

Total RNA was extracted from fibroblasts and LCLs using the Direct-Zol RNA MiniPrep system (Zymo Research, Irvine, CA, USA) and complementary DNA (cDNA) was generated using the M-MLV Reverse Transcriptase kit (Invitrogen, Thermo Fisher Scientific). The expression level of *TLK2* was measured using the Universal Probe Library system (Roche Diagnostics, Risch-Rotkreuz, Switzerland) with primers and probe in table S1 and *HMBS* as reference gene (assay number Hs00609297_m1; Applied Biosystems, Thermo Fisher Scientific). Assays were carried out in triplicate on an ABI 7500 real-time PCR instrument using the ABI 2X TaqMan™ Universal PCR Master Mix II, according to the manufacturer's protocol (Thermo Fisher Scientific). For each experiment, biological triplicates with at least two technical replicates were performed. Data were analysed with Prism-GraphPad Software performing unpaired t-test with Welch's correction. p-values are indicated as follows: ns= P value > 0.05 ; *= P value ≤ 0.05; ** P value ≤ 0.01; ***= P value ≤ 0.001; ****= P value ≤ 0.0001.

Single-cell gel electrophoresis

Samples were prepared according to the alkaline single-cell gel electrophoresis (SCGE) assay method, as previously described[15] and are briefly summarized in Supplementary material and methods, together with experimental details.

Site directed mutagenesis and *in vitro* kinase assays from cell lysates

In vitro kinase assays were performed as previously described[10] with minor modifications. Full methods provided in the Supplementary materials and methods and primers in table S3.

Western blotting

For affinity purification (AP), 40 µg of input protein and 20 µL of Strep-AP elution, with 6X SDS (0.2% Bromophenol blue and β-mercaptoethanol), were separated by SDS-PAGE and transferred to nitrocellulose membranes (0.2 or 0.45 µm pore, Amersham Protran; Sigma-Aldrich). For detection of Streptavidin, PVDF membranes (0.45 µm pore, Immobilon-P, Merck) were used. Membranes were blocked and antibodies prepared in 5% non-fat milk in PBS-T, with the exception of CHD7 and CHD8, where 5% BSA in PBS-T was used. Primary antibodies were detected with the appropriate secondary antibodies conjugated to Horseradish peroxidase (HRP) (table S4) and visualized by ECL-Plus (GE Healthcare).

Proximity-dependent biotin identification mass spectrometry (BioID-MS)

BioID-MS and BioID-Westerns were performed essentially as described in Silva *et al.* [16] with some modifications. Full methods are provided in the Supplementary materials and methods and data is available in the PRIDE database with accession number PXD019450 **(for reviewer access:** Username: **reviewer30722@ebi.ac.uk** Password: **dGfoWR6c).**

RESULTS

Identification of novel *MRD57* patients

The identification of *TLK2* as a risk gene for intellectual disability[7], prompted us to re-evaluate the genomic information available for a large cohort of patients affected by ASD and/or ID who had previously been analysed by WES. A search in the DECIPHER database[17] for novel *TLK2* deletions was also performed. We found six cases in three independent families with detrimental variants in the *TLK2* gene (figure 1A-F).

In family 1, the proband carried a *de novo* missense change c.1652A>G; p.(Asp551Gly) not reported in gnomAD and predicted to be damaging by multiple *in silico* predictors (table S5). This variant localizes in a highly Constrained Coding Region (CCR) (>93th percentile) within the PKD [18]. In family 2, three affected siblings carried a premature stop variant (c.1423G>T; p.(Glu475Ter)) inherited from their affected mother (case 2). This variant was not reported in gnomAD and was classified as pathogenic using the American College of Medical Genetics (ACMG) criteria[7, 8]. A sixth case (family 3) was identified in the DECIPHER database: a female carried a *de novo* deletion in 17q23.2, whose minimum size was 39 kb (chr17-60683462-60722398). The deletion encompassed the *TLK2*, *MRC2* and *MARCH10* genes, but the latter is predicted to be a haplosufficient gene (GnomAD pLI 0) and, to our knowledge, *MRC2* has not been linked to neurodevelopmental disorders.

Our patients (3 males and 3 females) ranged from 3 to 47 years of age, all of them were of Caucasian ethnicity. A broad range of behavioural disorders was present, including ASD (1/6), attention deficit hyperactivity disorder (ADHD) (4/6), anxiety (3/6), short attention span (2/6) and obsessive-compulsive behaviour (2/6). ID was reported for 4 patients in the borderline (IQ 70-85) or low (IQ≤70) range. Patient 5 was too young for

formal assessment of a neurodevelopmental phenotype, but a global developmental delay was reported. All affected members from family 2 had microcephaly. Dysmorphic facial features were observed in all patients and included upslanting palpebral fissures, wide nose, low hanging columella, smooth philtrum, prognathism, pointed chin and hypertelorism (figure 1G). Minor skeletal anomalies of the hands and feet were reported for four patients (figure 1H). Interestingly, some of the features we observed in a portion of our patients had not been described in previously reported cases, expanding the clinical phenotype. Among them, there were neurodevelopmental (difficulties in reading, writing, memory and transcription and pavor nocturnus), dysmorphic (prognathism, bifid tip of the nose, low hanging columella, ears without lobes, low-implant auricle, synophrys and facing down mouth corners) and skeletal anomalies (foot hexadactyly, feet big toes overlap, tapering fingers, short hands with short distal phalanx). A summary of the clinical phenotypes is reported in table S6 and a description is provided in the Supplemental Data, while a comparison with the available literature is provided in table S7.

TLK2 haploinsufficiency disrupts proper chromatin compaction

Using quantitative real time polymerase chain reaction (qRT-PCR), we analysed *TLK2* mRNA expression in LCLs from case 1 and fibroblasts from case 6. We found an approximately 50% reduction in *TLK2* mRNA expression in both cases (figure 2A). This was unexpected in case 1, who carried the c.1652A>G; p.(Asp551Gly) missense change. To verify if the reduced mRNA level observed for the missense change was due to degradation of the transcript by nonsense mediated decay (NMD)[19], we amplified the relevant portion of the cDNA from WT and mutated LCLs using primers encompassing exons 18 to 23. Sanger sequencing revealed the expression of the mutated allele (figure S1A), indicating that the missense mRNA was not completely degraded. Moreover, no differences were observed in band sizes from WT and p.(Asp551Gly) cDNAs (figure S1B), excluding that p.(Asp551Gly) led to the production of aberrantly spliced transcripts[19, 20]. Western Blot analysis on LCLs from case 1 compared to controls, uncovered a significant increase in TLK2 protein expression (figure S1C), pushing us to the necessity of a deeper characterization of the impact of p.(Asp551Gly) on TLK2 activity.

Based on the role of TLK2 in controlling chromatin remodelling[2], we investigated possible changes in chromatin compaction, performing the single-cell gel electrophoresis (SCGE) assay with LCLs derived from case

1. While a relatively short electrophoresis time was not sufficient to unmask any difference between control and mutant cells, longer run times allowed DNA loops to stretch under the electric field and revealed a significantly more relaxed state of nucleoids in LCLs from the affected subject compared to control cells (figure 2B-C). Differences were quantified as "tail moment" values, which are defined as the product of the tail length and the percentage of DNA in the tail. LCLs harbouring the p.(Asp551Gly) variant also exhibited a higher sensitivity to γ-ray irradiation, documenting increased susceptibility to DNA damage (i.e., single and double strand breaks), which is in line with a more relaxed state of chromatin (figures 2D, E). Of note, a defective heterochromatin state was also observed in fibroblasts derived from subject 6 carrying the 17q23.2 deletion encompassing *TLK2* (figure S1D), suggesting a similar effect between the 17q23.2 deletion and the p.(Asp551Gly) variant. Overall, these findings demonstrated that TLK2 haploinsufficiency disrupts proper chromatin organization and confers susceptibility to DNA damage.

TLK2 missense mutations alter the activity and subcellular localization of the protein

All of the *TLK2* missense mutations analysed to date exhibited decreased kinase activity[10]. Given that the high conservation of the mutated residues in the PKD suggested reduced kinase activity, we examined the potential structural impact of the missense p.(Asp551Gly) (hereafter indicated as D551G) variant identified in our survey, as well as the p.(Ser617Leu) (hereafter indicated as S617L) variant, identified in a patient with ASD in the ASC exome analysis browser and recently reported in work from Satterstrom *et al.*[11]. We previously identified S617 as a TLK2 auto-phosphorylation site and demonstrated that mutations in this site could enhance or impair kinase activity[10]. We modelled each of the mutations on the crystal structure of the TLK2 PKD. TLK2-D551G was predicted to weaken hydrogen bonds with the subsequent helix and S617L introduces a hydrophobic residue in place the auto-phosphorylation site in the activation loop (figure 3A-B).

To determine if these mutations affected kinase activity, we analysed TLK2 activity using *in vitro* kinase assays. Ectopically expressed, Strep-FLAG tagged TLK2, TLK2-KD (kinase dead), TLK2-D551G and TLK2-S617L were affinity purified from AD-293 cells using an N-terminal Strep tag and incubated with purified substrate (ASF1A) for kinase assays. Both mutations led to a notable reduction in substrate modification (figures 3C-D). Quantification of multiple experiments demonstrated that TLK2-S617L severely impaired kinase

activity, comparable to the TLK2-KD (D592V) protein, while TLK2-D551G was more mildly impaired and showed slightly higher autophosphorylation levels than TLK2-WT in some experiments (figure 3D).

In previous work, we noted that loss of the coiled-coil domains of TLK2 led to perinuclear accumulation[10]. To determine if the TLK2-D551G and TLK2-S617L missense mutations altered TLK2 localization, we transfected FLAG-tagged mutants in AD-293 cells and performed immunofluorescence (IF) microscopy. Lamin A was used as inner nuclear membrane marker and nuclear DNA was stained with DAPI. TLK2-WT showed diffuse nuclear localization in transfected cells, as observed previously (figure 3E). In contrast, TLK2-D551G and TLK2-S617L showed perinuclear localization to different extents. This was particularly prominent for the TLK2-S617L mutant, where 75% of transfected cells showed a perinuclear localization of TLK2 (figures 3E-F).

The proximal interactome of TLK2 is altered by missense mutations

TLK2 is involved in many biological processes and few clear substrates aside from ASF1A and ASF1B have been well characterized[2]. As kinases often bind to substrates with low affinity, we previously used an unbiased proximity biotinylation assay coupled to mass spectrometry (BioID-MS), as it does not require high-affinity interactions that can withstand purification procedures[5, 21]. We used this approach to further characterize the cellular environment of TLK2 and determine if the missense mutations influenced its interactome.

BirA-tagged TLK2-WT, TLK2-D551G and TLK2-S617L were expressed in AD-293 cells by transient transfection (figure 4A). Network clustering of results from wild type TLK2 (TLK2-WT) identified the known TLK2 substrates ASF1A, ASF1B and TLK1, as well as the DYNLL1/2(LC8) proteins that we previously validated[5, 10]. The proximal interactome grouped into five functional clusters consistent with the known functions of TLK2, including RNA processing and splicing, transcriptional regulation, chromatin binding or remodelling, DNA repair and histone chaperones (figure 4B and table S8). We cross referenced the proximal interactome with a recent large-scale analysis of iPOND-MS data that identified TLK1 and TLK2 as high confidence interactors with nascent DNA at active replication forks[22]. Several of the TLK2-WT hits, including RAD50, BRD4, CHD8, ASF1B, SCML2 and NACC1 were also found at active forks with high confidence in iPOND studies. We next compared our TLK2-WT proximal interactome to the SFARI and DECIPHER

databases and identified 8 proteins: CHD7, CHD8, NACC1, CCNK, JMJD1C, RAD50, MSANTD2, and YEATS2. These data suggest potential functional links between TLK2 and a number of proteins involved in neurodevelopmental disorders with overlapping pathologies. Details about SFARI and OMIM classifications are provided in Supplementary table S9.

In parallel to TLK2-WT, we performed BioID analysis with TLK2-D551G and TLK2-S617L. Both mutants expressed at higher levels than TLK2-WT, consistent with what we previously observed with other inactive TLK2 mutants. Both mutants caused numerous alterations in the proximal interactions compared to TLK2-WT, despite higher expression levels (figures 4C, D and 5).

In some cases, significant differences were present between the two mutants and TLK2-WT. Of interest were CHD7 and CHD8 that are frequently mutated in CHARGE syndrome and ASD, respectively. In addition, CHD8 has been localized to active sites of DNA replication, similarly to both TLK1 and TLK2[22]. CHD8 spectra were detected at similar levels between TLK2-WT and TLK2-D551G, that retains some activity, but were reduced with the kinase dead TLK2-S617L (figures 4C, D and 5). In contrast, CHD7 spectra were highest in cells expressing TLK2-D551G. We performed BioID-Westerns to validate these proximal interactions and their relative differences. Expression of the control N-FLAG-BirA and biotin supplementation led to no detectable CHD7 or CHD8 detected in Western blots of Strep-affinity purified protein lysates (figure 6A). In contrast, both CHD7 and CHD8 were clearly validated with all TLK2 alleles, although CHD8 levels were lowest with S617L and CHD7 highest in TLK2-D551G, consistent with the MS data. As expected from the BioID-MS, TLK1 co-purified with all TLK2 alleles to a similar degree. It was also notable that the substrates, ASF1A and ASF1B, were highest with D551G and similar between TLK2-WT and TLK2-S617L, despite the much higher level of TLK2-S617L expression. These results indicate the missense mutations have differential effects on both proximal interactions and TLK-ASF1 interactions.

DISCUSSION

To our knowledge, at least 13 *TLK2* missense variants have been reported in MRD57 cases [7, 9, 11], with mutations clustering in the PKD. Functional characterization has been reported only for four variants in the PKD, (p.(His493Arg), p.(His518Arg), p.(Asp629Asn), p.(Arg720Ala)) and has shown at least a 50% reduction in enzymatic activity compared to wild type protein[10]. Our data further expands the characterization of MRD57 missense mutations and reinforces the prevailing hypothesis that the majority of these impair TLK2 kinase activity. Both TLK2-D551G and TLK2-S617L showed profoundly impaired kinase activity, as well as altered subcellular localization, the significance of which remains unclear (figure 3E-F).

Both mutants overexpressed in AD-293 cells showed higher expression compared to WT. Accordingly, LCLs carrying p.(Asp551Gly) variant showed significantly higher TLK2 protein levels compared to control. Surprisingly, the higher protein expression was counterbalanced by drastically reduced mRNA levels, suggesting a potential negative feedback control, further causing reduced *TLK2* total mRNA levels.

Despite the severe effects on kinase activity caused by the heterozygous missense mutations identified in MRD57 patients to date, monoallelic *Tlk2* loss did not cause overt phenotypes in mice, but neurodevelopment and behaviour were not assessed in these animals[5]. Given that TLK2 dimerizes with both TLK1 and TLK2 and this is important for its activity, it is also likely that kinase impaired mutants exert some dominant negative effects that contribute to an overall phenotype that is more severe than haploinsufficiency[23]. The recent identification of multiple MRD57 cases with *TLK2* haploinsufficiency suggests that an in-depth evaluation of neurodevelopment is needed in mice with reduced *Tlk2* levels to determine if they represent a model of MRD57. The placental issues observed in mice with homozygous deletion of *Tlk2* were not identified in *Tlk2* heterozygous mice, but human gestation is considerably longer, and more subtle placental issues could be present. Recent work showed that total TLK depletion leads to an innate immune secretory response in cancer cells and mice[2, 6, 24]. Maternal immune activation (MIA) has been implicated in ASD and associated with placental defects in mice, suggesting that impaired chromatin maintenance and epigenetic dysregulation could potentially underlie the pathological effects of TLK2 haploinsufficiency[2, 25, 26]. This is consistent with the increased chromatin accessibility reported here in TLK2-D551G patient cells (figure 2B-E), as well as the strong enrichment of chromatin proteins in the TLK2 interactome in the SFARI and DECIPHER databases.

Many genes encoding proteins involved in chromatin remodelling are associated with neurodevelopmental disorders. TLK2, as well as the missense mutants we tested, showed proximal interactions with many of them, including CHD8 and CHD7, that are mutated in ASD and CHARGE syndrome (figure 6A). In addition to these proteins, both TLK2 missense mutants showed altered interactions with additional proteins implicated in neurodevelopment. This included RAD50, a part of the MRE11-RAD50-NBS1 DNA repair complex that localizes to replication forks and plays a key role in DNA-double-strand break repair[27]. *RAD50* mutations, present in the DECIPHER database, underlie Nijmegen breakage syndrome-like disorder (OMIM 613078). This condition is characterized by microcephaly, which is also commonly observed in many patients with *TLK2* variants. Further, RAD50 proximal interactions were reduced with both missense mutants, potentially suggesting reduced localization to replication forks (figure 5)[28]. Similarly, YEATS2, a chromatin reader component, is suggested as an ASD-associated gene by de novo genetic risk analysis and GWAS (SFARI database criteria 3.1, suggestive evidence)[29, 30], was linked to epilepsy and was enriched with TLK2-WT compared to either missense mutant[31–33]. In contrast, ZNF148 and PAPOLG, that are also associated with neurodevelopmental disorders, were strongly enriched with both missense mutants and detected at very low levels with TLK2-WT, while other SFARI genes, including *BRD4, JMJD1C, MSANTD2, CCNK* and *NACC1* were reduced specifically with the less active TLK2-S617L variant. In future work, it will be of interest to examine the potential functional relevance of these interactions to determine if their alterations underlie the altered chromatin state we observed in LCLs from case 1 or other phenotypes associated with TLK2 loss of function. This approach may detect new candidate genes involved in neurodevelopmental disorders or help us understand the involvement of this network of SFARI genes in isolated or syndromic ASD and ID.

The knowledge of altered protein interactomes is important to understand the molecular impact of disease mutations and could be helpful in identifying pharmacological treatments to mitigate more severe phenotypes, such as epileptic seizures. It is attractive to imagine the possibility of repurposing drugs able to modulate the functions of some genes to influence their impact on disease pathology[34].

In conclusion, we provided the clinical description of six new cases carrying likely pathogenic and pathogenic *TLK2* variants and we presented new insights into the impact of *TLK2* missense variants, observing impairment

in kinase activity, localization, and interaction. Our work offers a deep characterization of two missense variants localized in a key domain of the TLK2 protein, where most mutations related to MRD57-disorder occur, providing new insights into the potential role TLK2 in neurodevelopmental disorders.

AUTHOR CONTRIBUTION

M.V.P., M.S.C., M.G., M.V. and G.A.G. performed activity, localization, and BioID-MS analysis. L.P. wrote and edited the manuscript, interpreted exome data, collected the cases, and performed variant confirmation, mRNA and splicing analysis. C.A. and V.P. performed the SCGE assay. E.G. interpreted exome data. D.C., V.A.M. G., K.R., C.C. and G.B.F. followed patients and collected their clinical information. S.D.R. and J.B. performed exome sequencing. A.B., T.P., P.D., A.C., F.C.R. and M.T. processed and analysed the WES data. T.H.S. and A.B. designed and supervised the project and manuscript writing.

ACKNOWLEDGEMENTS

We are grateful to the patients and their families for their precious collaboration. This research received funding specifically appointed to Department of Medical Sciences from the Italian Ministry for Education, University and Research (Ministero dell'Istruzione, dell'Università e della Ricerca - MIUR) under the programme "Dipartimenti di Eccellenza 2018 – 2022" Project code D15D18000410001. M.V.P. was funded by an FPI fellowship from the Ministry of Science, Innovation and Universities (MCIU) and M.S.C. by a Masters fellowship from the BIST and support from the IRB Barcelona. T.H.S. was funded by the MCIU (PGC2018-095616-B-I00/GINDATA and FEDER). M.T. was funded by Fondazione Bambino Gesù (Vite Coraggiose). The whole exome sequencing was performed as part of the Autism Sequencing Consortium and was supported by the NIMH (MH111661). Thanks to the MSPCF of IRB Barcelona, a principal unit in Proteored, PRB3, supported by PT17/0019 of the PE I+D+i 2013-2016, funded by SCIII and ERDF. This study makes use of data generated by the DECIPHER community. A full list of centres that contributed to the generation of the data is available from http://decipher.sanger.ac.uk and via email from decipher@sanger.ac.uk. Funding for the project was provided by the Wellcome Trust.

WEB RESOURCES

Autism Sequencing Consortium exome analysis browser, https://asc.broadinstitute.org/

BioGRID Database, https://thebiogrid.org/

Constrained Coding Regions Browser, https://s3.us-east-2.amazonaws.com/ccrs/ccr.html

DECIPHER, https://decipher.sanger.ac.uk/

gnomAD Browser, v2.1.1, https://gnomad.broadinstitute.org/

Human Splicing Finder, v3.1, http://www.umd.be/HSF/

OMIM, https://omim.org/

SFARI gene, https://gene.sfari.org/

STRING Database, https://string-db.org

Varsome, https://varsome.com/

LEGENDS TO FIGURES

Figure 1. Facial features and skeletal anomalies of individuals with *TLK2* variants

(A-C) Pedigrees of family 1, 2 and 3. Cases from family 1 and 3 carried respectively a heterozygous *de novo* missense variant (c.1652A>G; p.(Asp551Gly)) and a heterozygous *de novo* deletion encompassing *TLK2* gene. Cases from family 2 shared a heterozygous premature stop variant (c.1423G>T; (p.Glu475Ter)) inherited from an affected mother. Analysis of maternal grandparent genotype were not possible, but familial clinical history did not suggest a possible MRD57-like phenotype for them. wt= wild type at variant position. (D-E) Sanger validation of variants identified in family 1 and 2. Validation was performed both on gDNA from affected cases and from their unaffected relatives. (F) 17q23.2 deletion (minimum size 39 kb, chr17-60683462-60722398) identified in case 6 from family 3. The 39 kb deletion encompassed *TLK2* and *MRC2* genes. (G) Frontal and lateral face photographs of our cases, showing overlapping facial dysmorphisms. Most frequently reported features were upward slanted palpebral fissures, broad nasal tip, thin lips, low hanging columella, prognathism, wide spaced eyes and facing down mouth corners. (H) Details of reported skeletal anomalies at the level of hands and feet. Upper left panel: tapering hands fingers from case 1; upper right panels: short hands with short distal phalanx from case 6; bottom left panel: right foot hexadactyly observed in case 3;bottom right panel: big toes overlap in case 4.

Figure 2. TLK2-Asp551Gly affects chromatin density and confers susceptibility to DNA damage.

(A) *TLK2* mRNA levels in fibroblasts from case 6 and in lymphoblastoid cell line derived from case 1. *TLK2* expression was significantly reduced both in LCL carrying p.(Asp551Gly) variant and in fibroblasts carrying the 17q23.2 deletion. All experiments were performed at least in triplicate. *HMBS* mRNA expression was used as reference. Statistical analysis was performed using t-test with Welch's correction; ****= P value ≤ 0.0001. (B) SCGE assays highlighted significant differences in chromatin condensation between LCLs carrying the p.(Asp551Gly) amino acid change and control cells after 20 minutes of electrophoresis run time (**p < 0.05; two-tailed unpaired Student's t test with Welch's correction), that became more evident after 60 minutes of electrophoresis run time (***p 0.0001; two-tailed unpaired Student's t test with Welch's correction). DNA migration was quantified as Tail moment values, which is defines as the product between the tail length and the percentage of DNA in the tail. For each point, at least 100 cells were analysed. Values are represented as mean ± SEM of three independent experiments. (C) Representative images of nucleoids derived from control LCLs and LCLs from affected subject 1 referred to experiment shown in figure 2B. (D) Single and double strand breaks were induced by γ-ray irradiation (2-Gy or 4-Gy). Tail moment values specify the amount of γ-ray-induced DNA damage measured immediately after the treatment. The mutant LCLs showed a higher vulnerability to 2-Gy γ-ray irradiation (**p < 0.001). Following 4-Gy treatment, no differences were observed between control and mutant cells, which is likely explained by the observation that the overall damage, especially double strand breaks, prevails on the condensation state of chromatin at high doses of γ-ray irradiation. For each point, at least 100 cells were analysed. Values are mean ± SEM of three independent experiments. (E) Representative images of nucleoids derived from control LCLs and LCLs from affected subject 1 referred to experiment shown in figure 2D.

Figure 3. TLK2 autism mutations alter the activity and subcellular localization of TLK2.

(A-B) Modelling of the D551G and S617L missense mutations on the crystal structure of the TLK2 PKD. Hydrogen bonds are shown in red dashed lines. (C) Representative *in vitro* kinase assays with Strep-purified TLK2-WT, TLK2-KD (kinase dead; D592V) and indicated missense variants. Autophosphorylation and substrate (ASF1A) phosphorylation is shown. Coomassie is shown as loading control for ASF1A. (D) Quantification of n=3 independent kinase assays. Individual results (circles) are shown for each assay on purified ASF1A substrate or affinity purified TLK2 autophosphorylation relative to corresponding TLK2-WT and bars depict mean with SEM. (E) Representative immunofluorescence microscopy of overexpressed TLK2 in AD-

293 cells is shown, indicating the 2 main localization patterns observed. Scale bar = 10 μM. (F) Quantification of TLK2 localization patterns for WT and indicated missense variants. Ten random fields were scored in 2 (D551G) or 3 (WT and S617L) biological replicates. Bars depict mean with SEM.

Figure 4. BioID based analysis of the proximal interactome of TLK2.

(A) Western blot of AD-293 lysates expressing BioID constructs: N-FLAG-BirA alone or fused to the indicated TLK2 allele. Detection with anti-TLK2, anti-FLAG or Streptavidin-HRP are shown. Ponceau stained nitrocellulose membrane is shown as a loading control. (B) Network clustering of all prey hits with a SAINT score of > or = to 0.7 in TLK2-WT samples. Physical interactions reported in Biogrid (solid lines) and functional interactions (dashed lines) reported in STRING are indicated[35, 36]. Functionally enriched clusters are indicated by color coding, Bait, TLK2 substrates, proteins found in the SFARI/DECIPHER gene database (yellow fill) or proteins enriched on nascent DNA/replication forks are indicated (red font). (C-D) Scatterplots of average spectral counts (Log_2 transformed) of bait and prey proteins identified with the TLK2-D551G and TLK2-S617L alleles compared to TLK2-WT. Previously identified TLK2 interactors, as well as proteins enriched on replication forks or found in the SFARI/DECIPHER databases are indicated (see legend).

Figure 5. Missense variants alter the proximal interactome of TLK2.

Dotplot of prey proteins with a SAINT score of > or = to 0.7 with any of the 3 baits generated using ProHits-viz[37]. Average spectral counts (SC), relative abundance and SAINT score ranges are indicated, as well as proteins enriched on replication forks [22] or found in the SFARI/DECIPHER databases (see legend). Additional details are provided in table S8.

Figure 6: Validation of proximal interactions with CHD7 and CHD8.

(A) Western blots of the indicated proteins from Strep-AP lysates from AD-293 cells transfected with the indicated BioID construct and supplemented with biotin. Input levels are shown and ponceau stained blots provided as a loading control. Representative data from 2 biological replicates is shown.

REFERENCES

1 Lee SB, Segura-Bayona S, Villamor-Payà M, Saredi G, Todd MAM, Attolini CSO, Chang TY,

Stracker TH, Groth A. Tousled-like kinases stabilize replication forks and show synthetic lethality with checkpoint and PARP inhibitors. *Sci Adv* Published Online First: 2018. doi:10.1126/sciadv.aat4985

2 Segura-Bayona S, Stracker TH. The Tousled-like kinases regulate genome and epigenome stability: implications in development and disease. Cell. Mol. Life Sci. 2019. doi:10.1007/s00018-019-03208-z

3 Bruinsma W, Berg J, Aprelia M, Medema RH. Tousled-like kinase 2 regulates recovery from a DNA damage-induced G2 arrest . *EMBO Rep* Published Online First: 2016. doi:10.15252/embr.201540767

4 Klimovskaia IM, Young C, Strømme CB, Menard P, Jasencakova Z, Mejlvang J, Ask K, Ploug M, Nielsen ML, Jensen ON, Groth A. Tousled-like kinases phosphorylate Asf1 to promote histone supply during DNA replication. *Nat Commun* Published Online First: 2014. doi:10.1038/ncomms4394

5 Segura-Bayona S, Knobel PA, Gonzalez-Buron H, Youssef SA, Peña-Blanco A, Coyaud E, Lopez-Rovira T, Rein K, Palenzuela L, Colombelli J, Forrow S, Raught B, Groth A, De Bruin A, Stracker TH. Differential requirements for Tousled-like kinases 1 and 2 in mammalian development. *Cell Death Differ* Published Online First: 2017. doi:10.1038/cdd.2017.108

6 Sandra Segura-Bayona, Marina Villamor-Payà, Camille Stephan-Otto Attolini, Travis H. Stracker. Tousled-like kinase activity is required for transcriptional silencing and suppression of innate immune signaling. 2019.

7 Reijnders MRF, Miller KA, Alvi M, Goos JAC, Lees MM, de Burca A, Henderson A, Kraus A, Mikat B, de Vries BBA, Isidor B, Kerr B, Marcelis C, Schluth-Bolard C, Deshpande C, Ruivenkamp CAL, Wieczorek D, Baralle D, Blair EM, Engels H, Lüdecke HJ, Eason J, Santen GWE, Clayton-Smith J, Chandler K, Tatton-Brown K, Payne K, Helbig K, Radtke K, Nugent KM, Cremer K, Strom TM, Bird LM, Sinnema M, Bitner-Glindzicz M, van Dooren MF, Alders M, Koopmans M, Brick L, Kozenko M, Harline ML, Klaassens M, Steinraths M, Cooper NS, Edery P, Yap P, Terhal PA, van der Spek PJ, Lakeman P, Taylor RL, Littlejohn RO, Pfundt R, Mercimek-Andrews S, Stegmann APA, Kant SG, McLean S, Joss S, Swagemakers SMA, Douzgou S, Wall SA, Küry S, Calpena E, Koelling N, McGowan SJ, Twigg SRF, Mathijssen IMJ, Nellaker C, Brunner HG, Wilkie AOM. De Novo and Inherited Loss-of-Function Variants in TLK2: Clinical and Genotype-Phenotype Evaluation of a Distinct Neurodevelopmental Disorder. *Am J Hum Genet* 2018;**102**:1195–203.

8 Töpf A, Oktay Y, Balaraju S, Yilmaz E, Sonmezler E, Yis U, Laurie S, Thompson R, Roos A, MacArthur DG, Yaramis A, Güngör S, Lochmüller H, Hiz S, Horvath R. Severe neurodevelopmental disease caused by a homozygous TLK2 variant. *Eur J Hum Genet* Published Online First: 2019. doi:10.1038/s41431-019-0519-x

9 Lelieveld SH, Reijnders MRF, Pfundt R, Yntema HG, Kamsteeg EJ, De Vries P, De Vries BBA, Willemsen MH, Kleefstra T, Löhner K, Vreeburg M, Stevens SJC, Van Der Burgt I, Bongers EMHF, Stegmann APA, Rump P, Rinne T, Nelen MR, Veltman JA, Vissers LELM, Brunner HG, Gilissen C. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat Neurosci* 2016;**19**:1194–6.

10 Mortuza GB, Hermida D, Pedersen AK, Segura-Bayona S, López-Méndez B, Redondo P, Rüther P, Pozdnyakova I, Garrote AM, Muñoz IG, Villamor-Payà M, Jauset C, Olsen J V., Stracker TH, Montoya G. Molecular basis of Tousled-Like Kinase 2 activation. *Nat Commun* Published Online First: 2018. doi:10.1038/s41467-018-04941-y

11 Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, Peng M, Collins R, Grove J, Klei L, Stevens C, Reichert J, Mulhern MS, Artomov M, Gerges S, Sheppard B, Xu X, Bhaduri A, Norman U, Brand H, Schwartz G, Nguyen R, Guerrero EE, Dias C, Aleksic B, Anney R, Barbosa M, Bishop S, Brusco A, Bybjerg-Grauholm J, Carracedo A, Chan MCY, Chiocchetti AG, Chung BHY, Coon H, Cuccaro ML, Curró A, Dalla Bernardina B, Doan R, Domenici E, Dong S, Fallerini C, Fernández-Prieto M, Ferrero GB, Freitag CM, Fromer M, Gargus JJ, Geschwind D, Giorgio E, González-Peñas J, Guter S, Halpern D, Hansen-Kiss E, He X, Herman GE, Hertz-Picciotto I, Hougaard DM, Hultman CM, Ionita-Laza I, Jacob S, Jamison J, Jugessur A, Kaartinen M, Knudsen GP, Kolevzon A, Kushima I, Lee SL, Lehtimäki T, Lim ET, Lintas C, Lipkin WI, Lopergolo D, Lopes F, Ludena Y, Maciel P, Magnus P, Mahjani B, Maltman N, Manoach DS, Meiri G, Menashe I, Miller J, Minshew N, Montenegro EMS, Moreira D, Morrow EM, Mors O, Mortensen PB, Mosconi M, Muglia P, Neale BM, Nordentoft M, Ozaki N, Palotie A, Parellada M, Passos-Bueno MR, Pericak-Vance M, Persico AM, Pessah I, Puura K, Reichenberg A, Renieri A, Riberi E, Robinson EB, Samocha KE, Sandin S, Santangelo SL, Schellenberg G, Scherer SW, Schlitt S, Schmidt R, Schmitt L, Silva IMW, Singh T, Siper PM, Smith M, Soares G, Stoltenberg C, Suren P, Susser E, Sweeney J, Szatmari P, Tang L, Tassone F, Teufel K, Trabetti E, Trelles M del P, Walsh CA, Weiss LA, Werge

T, Werling DM, Wigdor EM, Wilkinson E, Willsey AJ, Yu TW, Yu MHC, Yuen R, Zachi E, Agerbo E, Als TD, Appadurai V, Bækvad-Hansen M, Belliveau R, Buil A, Carey CE, Cerrato F, Chambert K, Churchhouse C, Dalsgaard S, Demontis D, Dumont A, Goldstein J, Hansen CS, Hauberg ME, Hollegaard M V., Howrigan DP, Huang H, Maller J, Martin AR, Martin J, Mattheisen M, Moran J, Pallesen J, Palmer DS, Pedersen CB, Pedersen MG, Poterba T, Poulsen JB, Ripke S, Schork AJ, Thompson WK, Turley P, Walters RK, Betancur C, Cook EH, Gallagher L, Gill M, Sutcliffe JS, Thurm A, Zwick ME, Børglum AD, State MW, Cicek AE, Talkowski ME, Cutler DJ, Devlin B, Sanders SJ, Roeder K, Daly MJ, Buxbaum JD. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* Published Online First: 2020. doi:10.1016/j.cell.2019.12.036

12 De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, Kou Y, Liu L, Fromer M, Walker S, Singh T, Klei L, Kosmicki J, Fu SC, Aleksic B, Biscaldi M, Bolton PF, Brownfeld JM, Cai J, Campbell NG, Carracedo A, Chahrour MH, Chiocchetti AG, Coon H, Crawford EL, Crooks L, Curran SR, Dawson G, Duketis E, Fernandez BA, Gallagher L, Geller E, Guter SJ, Hill RS, Ionita-Laza I, Gonzalez PJ, Kilpinen H, Klauck SM, Kolevzon A, Lee I, Lei J, Lehtimäki T, Lin CF, Ma'ayan A, Marshall CR, McInnes AL, Neale B, Owen MJ, Ozaki N, Parellada M, Parr JR, Purcell S, Puura K, Rajagopalan D, Rehnström K, Reichenberg A, Sabo A, Sachse M, Sanders SJ, Schafer C, Schulte-Rüther M, Skuse D, Stevens C, Szatmari P, Tammimies K, Valladares O, Voran A, Wang LS, Weiss LA, Willsey AJ, Yu TW, Yuen RKC, Cook EH, Freitag CM, Gill M, Hultman CM, Lehner T, Palotie A, Schellenberg GD, Sklar P, State MW, Sutcliffe JS, Walsh CA, Scherer SW, Zwick ME, Barrett JC, Cutler DJ, Roeder K, Devlin B, Daly MJ, Buxbaum JD. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* Published Online First: 2014. doi:10.1038/nature13772

13 Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Albarca Aguilera M, Meyer R, Massouras A. VarSome: the human genomic variant search engine. *Bioinformatics* Published Online First: 2019. doi:10.1093/bioinformatics/bty897

14 Desmet FO, Hamroun D, Lalande M, Collod-Bëroud G, Claustres M, Béroud C. Human Splicing Finder: An online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* Published Online First: 2009. doi:10.1093/nar/gkp215

15 Flex E, Martinelli S, Van Dijck A, Ciolfi A, Cecchetti S, Coluzzi E, Pannone L, Andreoli C, Radio

FC, Pizzi S, Carpentieri G, Bruselles A, Catanzaro G, Pedace L, Miele E, Carcarino E, Ge X, Chijiwa C, Lewis MES, Meuwissen M, Kenis S, Van der Aa N, Larson A, Brown K, Wasserstein MP, Skotko BG, Begtrup A, Person R, Karayiorgou M, Roos JL, Van Gassen KL, Koopmans M, Bijlsma EK, Santen GWE, Barge-Schaapveld DQCM, Ruivenkamp CAL, Hoffer MJV, Lalani SR, Streff H, Craigen WJ, Graham BH, van den Elzen APM, Kamphuis DJ, Õunap K, Reinson K, Pajusalu S, Wojcik MH, Viberti C, Di Gaetano C, Bertini E, Petrucci S, De Luca A, Rota R, Ferretti E, Matullo G, Dallapiccola B, Sgura A, Walkiewicz M, Kooy RF, Tartaglia M. Aberrant Function of the C-Terminal Tail of HIST1H1E Accelerates Cellular Senescence and Causes Premature Aging. *Am J Hum Genet* Published Online First: 2019. doi:10.1016/j.ajhg.2019.07.007

16 Silva J, Aivio S, Knobel PA, Bailey LJ, Casali A, Vinaixa M, Garcia-Cao I, Coyaud É, Jourdain AA, Pérez-Ferreros P, Rojas AM, Antolin-Fontes A, Samino-Gené S, Raught B, González-Reyes A, Ribas De Pouplana L, Doherty AJ, Yanes O, Stracker TH. EXD2 governs germ stem cell homeostasis and lifespan by promoting mitoribosome integrity and translation. *Nat Cell Biol* Published Online First: 2018. doi:10.1038/s41556-017-0016-9

17 Firth H V., Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, Vooren S Van, Moreau Y, Pettett RM, Carter NP. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet* Published Online First: 2009. doi:10.1016/j.ajhg.2009.03.010

18 Havrilla JM, Pedersen BS, Layer RM, Quinlan AR. A map of constrained coding regions in the human genome. *Nat Genet* 2019;**51**. doi:10.1038/s41588-018-0294-6

19 Manning KS, Cooper TA. The roles of RNA processing in translating genotype to phenotype. Nat. Rev. Mol. Cell Biol. 2017. doi:10.1038/nrm.2016.139

20 Siwaszek A, Ukleja M, Dziembowski A. Proteins involved in the degradation of cytoplasmic mRNA in the major eukaryotic model systems. RNA Biol. 2014. doi:10.4161/rna.34406

21 Roux KJ, Kim DI, Burke B, May DG. BioID: A Screen for Protein-Protein Interactions. *Curr Protoc Protein Sci* 2018;**91**:19.23.1-19.23.15.

22 Wessel SR, Mohni KN, Luzwick JW, Dungrawala H, Cortez D. Functional Analysis of the Replication Fork Proteome Identifies BET Proteins as PCNA Regulators. *Cell Rep* Published Online First: 2019. doi:10.1016/j.celrep.2019.08.051

610 23 Sunavala-Dossabhoy G, Li Y, Williams B, De Benedetti A. A dominant negative mutant of TLK1
611 causes chromosome missegregation and aneuploidy in normal breast epithelial cells. *BMC Cell Biol*
612 Published Online First: 2003. doi:10.1186/1471-2121-4-16
613 24 Liu H, Dowdle JA, Khurshid S, Sullivan NJ, Bertos N, Rambani K, Mair M, Daniel P, Wheeler E,
614 Tang X, Toth K, Lause M, Harrigan ME, Eiring K, Sullivan C, Sullivan MJ, Chang SW, Srivastava S,
615 Conway JS, Kladney R, McElroy J, Bae S, Lu Y, Tofigh A, Saleh SMI, Fernandez SA, Parvin JD,
616 Coppola V, Macrae ER, Majumder S, Shapiro CL, Yee LD, Ramaswamy B, Hallett M, Ostrowski
617 MC, Park M, Chamberlin HM, Leone G. Discovery of Stromal Regulatory Networks that Suppress
618 Ras-Sensitized Epithelial Cell Proliferation. *Dev Cell* Published Online First: 2017.
619 doi:10.1016/j.devcel.2017.04.024
620 25 Lombardo M V., Moon HM, Su J, Palmer TD, Courchesne E, Pramparo T. Maternal immune
621 activation dysregulation of the fetal brain transcriptome and relevance to the pathophysiology of
622 autism spectrum disorder. *Mol Psychiatry* Published Online First: 2018. doi:10.1038/mp.2017.15
623 26 Carpentier PA, Dingman AL, Palmer TD. Placental TNF-α signaling in illness-induced complications
624 of pregnancy. *Am J Pathol* Published Online First: 2011. doi:10.1016/j.ajpath.2011.02.042
625 27 Taylor AMR. Chromosome instability syndromes. *Best Pract Res Clin Haematol* Published Online
626 First: 2001. doi:10.1053/beha.2001.0158
627 28 Waltes R, Kalb R, Gatei M, Kijas AW, Stumm M, Sobeck A, Wieland B, Varon R, Lerenthal Y,
628 Lavin MF, Schindler D, Dörk T. Human RAD50 Deficiency in a Nijmegen Breakage Syndrome-like
629 Disorder. *Am J Hum Genet* Published Online First: 2009. doi:10.1016/j.ajhg.2009.04.010
630 29 Ruzzo EK, Pérez-Cano L, Jung JY, Wang L kai, Kashef-Haghighi D, Hartl C, Singh C, Xu J,
631 Hoekstra JN, Leventhal O, Leppä VM, Gandal MJ, Paskov K, Stockham N, Polioudakis D, Lowe JK,
632 Prober DA, Geschwind DH, Wall DP. Inherited and De Novo Genetic Risk for Autism Impacts
633 Shared Networks. *Cell* 2019;**178**:850-866.e26.
634 30 Anney R, Klei L, Pinto D, Almeida J, Bacchelli E, Baird G, Bolshakova N, Bölte S, Bolton PF,
635 Bourgeron T, Brennan S, Brian J, Casey J, Conroy J, Correia C, Corsello C, Crawford EL, De jonge
636 M, Delorme R, Duketis E, Duque F, Estes A, Farrar P, Fernandez BA, Folstein SE, Fombonne E,
637 Gilbert J, Gillberg C, Glessner JT, Green A, Green J, Guter SJ, Heron EA, Holt R, Howe JL, Hughes
638 G, Hus V, Igliozzi R, Jacob S, Kenny GP, Kim C, Kolevzon A, Kustanovich V, Lajonchere CM,

Lamb JA, Law-Smith M, Leboyer M, Le couteur A, Leventhal BL, Liu XQ, Lombard F, Lord C, Lotspeich L, Lund SC, Magalhaes TR, Mantoulan C, McDougle CJ, Melhem NM, Merikangas A, Minshew NJ, Mirza GK, Munson J, Noakes C, Nygren G, Papanikolaou K, Pagnamenta AT, Parrini B, Paton T, Pickles A, Posey DJ, Poustka F, Ragoussis J, Regan R, Roberts W, Roeder K, Roge B, Rutter ML, Schlitt S, Shah N, Sheffield VC, Soorya L, Sousa I, Stoppioni V, Sykes N, Tancredi R, Thompson AP, Thomson S, Tryfon A, Tsiantis J, Van Engeland H, Vincent JB, Volkmar F, Vorstman JAS, Wallace S, Wing K, Wittemeyer K, Wood S, Zurawiecki D, Zwaigenbaum L, Bailey AJ, Battaglia A, Cantor RM, Coon H, Cuccaro ML, Dawson G, Ennis S, Freitag CM, Geschwind DH, Haines JL, Klauck SM, Mcmahon WM, Maestrini E, Miller J, Monaco AP, Nelson SF, Nurnberger JI, Oliveira G, Parr JR, Pericak-Vance MA, Piven J, Schellenberg GD, Scherer SW, Vicente AM, Wassink TH, Wijsman EM, Betancur C, Buxbaum JD, Cook EH, Gallagher L, Gill M, Hallmayer J, Paterson AD, Sutcliffe JS, Szatmari P, Vieland VJ, Hakonarson H, Devlin B. Individual common variants exert weak effects on the risk for autism spectrum disorders. *Hum Mol Genet* Published Online First: 2012. doi:10.1093/hmg/dds301

31 Mi W, Guan H, Lyu J, Zhao D, Xi Y, Jiang S, Andrews FH, Wang X, Gagea M, Wen H, Tora L, Dent SYR, Kutateladze TG, Li W, Li H, Shi X. YEATS2 links histone acetylation to tumorigenesis of non-small cell lung cancer. *Nat Commun* Published Online First: 2017. doi:10.1038/s41467-017-01173-4

32 Zhao D, Guan H, Zhao S, Mi W, Wen H, Li Y, Zhao Y, Allis CD, Shi X, Li H. YEATS2 is a selective histone crotonylation reader. Cell Res. 2016. doi:10.1038/cr.2016.49

33 Yeetong P, Pongpanich M, Srichomthong C, Assawapitaksakul A, Shotelersuk V, Tantirukdham N, Chunharas C, Suphapeetiporn K, Shotelersuk V. TTTCA repeat insertions in an intron of YEATS2 in benign adult familial myoclonic epilepsy type 4. *Brain* Published Online First: 2019. doi:10.1093/brain/awz267

34 Tranfaglia MR, Thibodeaux C, Mason DJ, Brown D, Roberts I, Smith R, Guilliams T, Cogram P. Repurposing available drugs for neurodevelopmental disorders: The fragile X experience. Neuropharmacology. 2019. doi:10.1016/j.neuropharm.2018.05.004

35 Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Von Mering C. STRING v11: Protein-protein association networks

with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* Published Online First: 2019. doi:10.1093/nar/gky1131

36 Oughtred R, Stark C, Breitkreutz BJ, Rust J, Boucher L, Chang C, Kolas N, O'Donnell L, Leung G, McAdam R, Zhang F, Dolma S, Willems A, Coulombe-Huntington J, Chatr-Aryamontri A, Dolinski K, Tyers M. The BioGRID interaction database: 2019 update. *Nucleic Acids Res* Published Online First: 2019. doi:10.1093/nar/gky1079

37 Knight JDR, Choi H, Gupta GD, Pelletier L, Raught B, Nesvizhskii AI, Gingras AC. ProHits-viz: A suite of web tools for visualizing interaction proteomics data. Nat. Methods. 2017. doi:10.1038/nmeth.4330

Figure S1

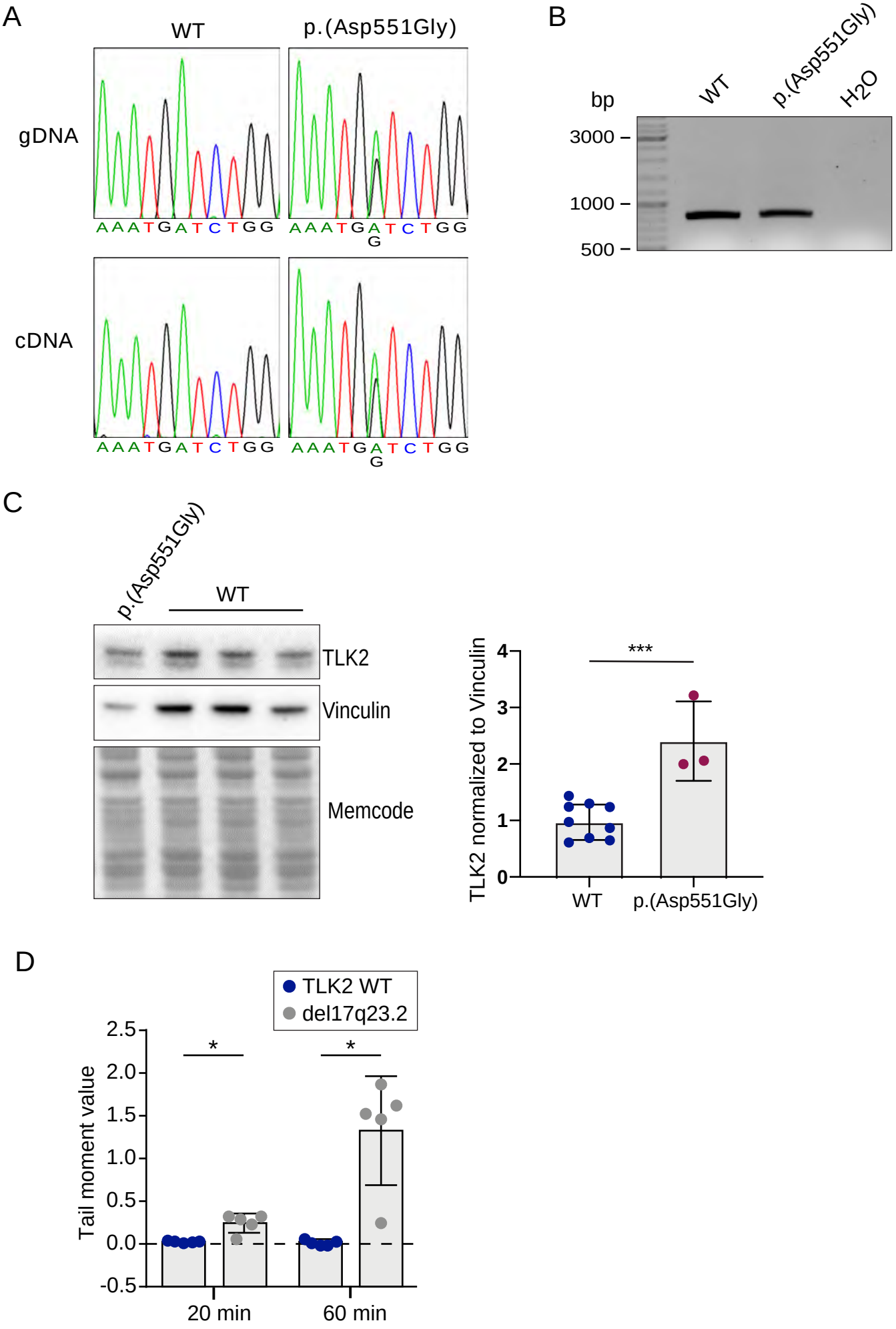
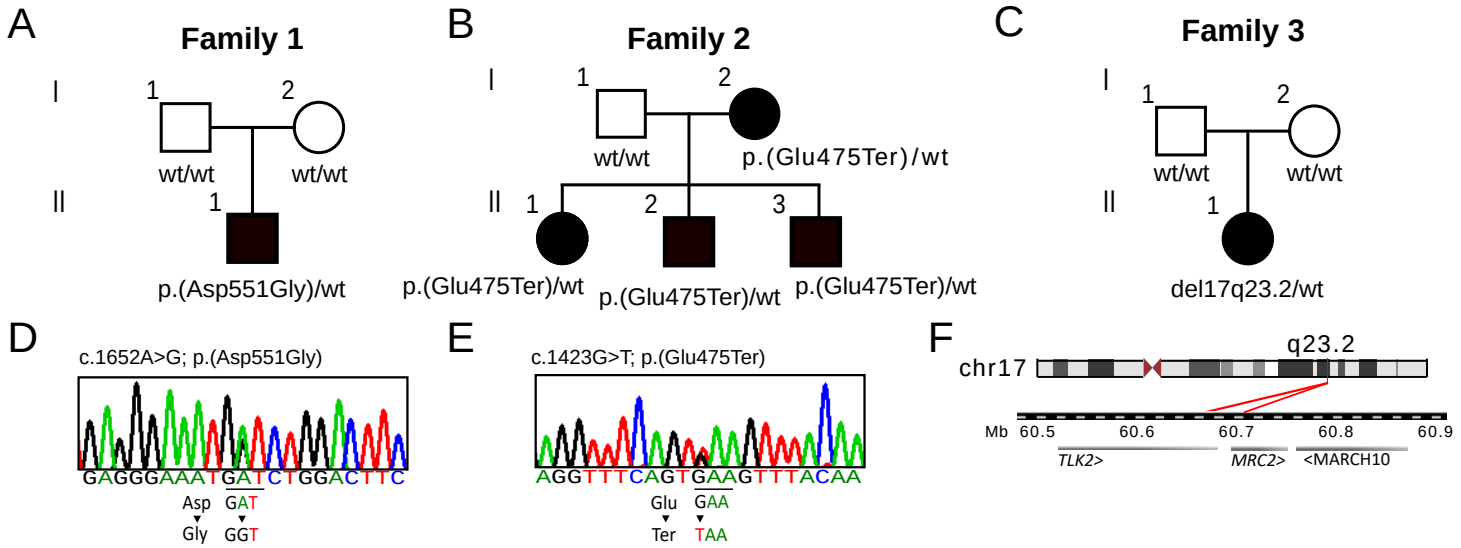
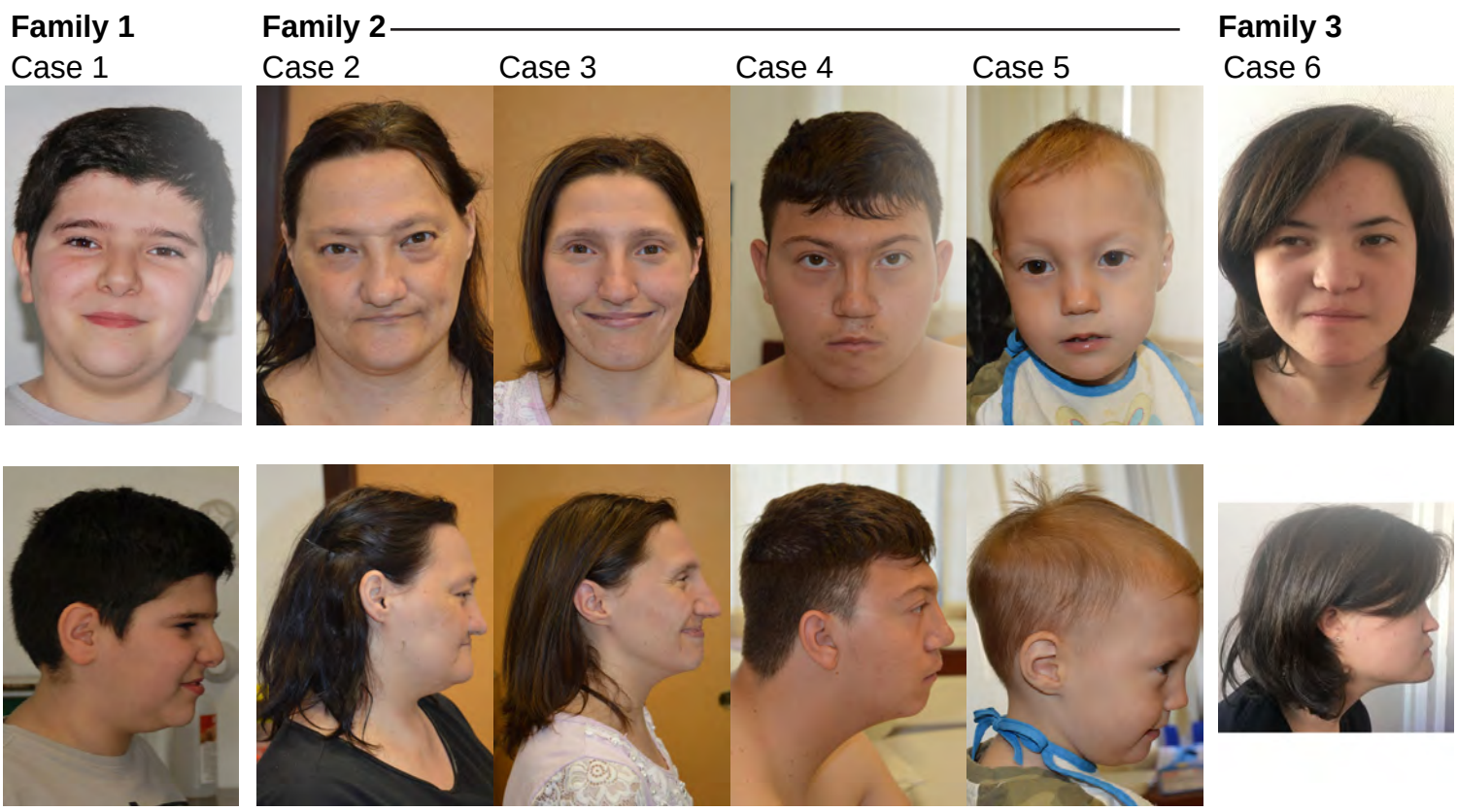

A
WT
p.(Asp551Gly)
gDNA
cDNA
AAATGATCTGG
AAATGATCTGG
G
B
bp
WT
p.(Asp551Gly)
H2O
3000
1000
500
C
p.(Asp551Gly)
WT
TLK2
Vinculin
Memcode
TLK2 normalized to Vinculin
***
WT
p.(Asp551Gly)
D
TLK2 WT
del17q23.2
Tail moment value
*
*
20 min
60 min

Figure 1


A
Family 1
I
1
2
wt/wt
wt/wt
II
1
p.(Asp551Gly)/wt
B
Family 2
I
1
2
wt/wt
p.(Glu475Ter)/wt
II
1
2
3
p.(Glu475Ter)/wt
p.(Glu475Ter)/wt
p.(Glu475Ter)/wt
C
Family 3
I
1
2
wt/wt
wt/wt
II
1
del17q23.2/wt

D
c.1652A>G; p.(Asp551Gly)
GAGGGAAATGATCTGGACTTC
Asp GAT
Gly GGT
E
c.1423G>T; p.(Glu475Ter)
AGGTTTCAGTGAAGTTTACAA
Glu GAA
Ter TAA
F
chr17
q23.2
Mb 60.5 60.6 60.7 60.8 60.9
TLK2>
MRC2>
<MARCH10

G


Family 1
Family 2
Family 3
Case 1
Case 2
Case 3
Case 4
Case 5
Case 6

H


Case1
Case 6
Case 3
Case 4

Figure 2

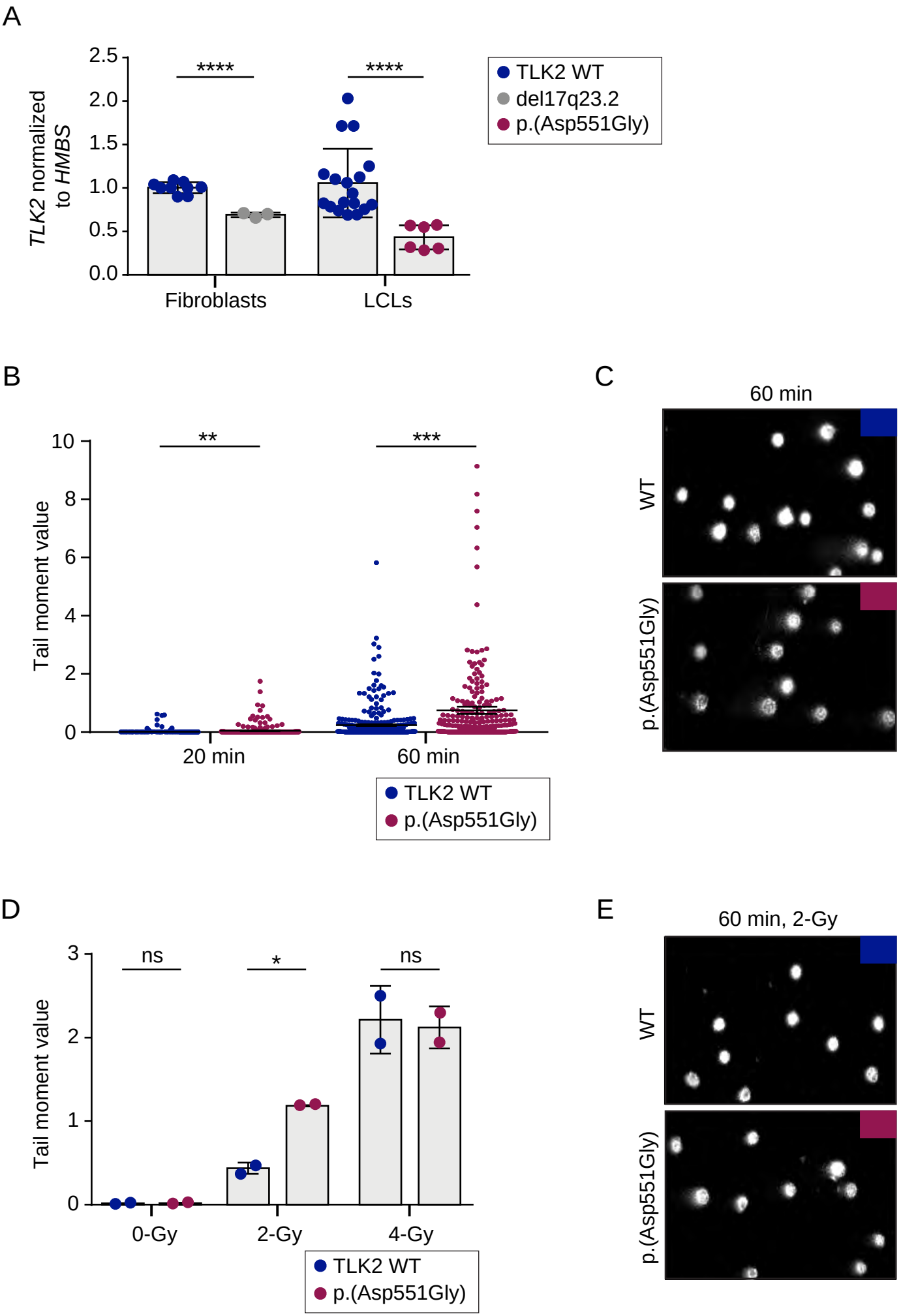

A
TLK2 normalized to HMBS
2.5
2.0
1.5
1.0
0.5
0.0
****
****
Fibroblasts
LCLs
TLK2 WT
del17q23.2
p.(Asp551Gly)
B
Tail moment value
10
8
6
4
2
0
**
***
20 min
60 min
TLK2 WT
p.(Asp551Gly)
C
60 min
WT
p.(Asp551Gly)
D
Tail moment value
3
2
1
0
ns
*
ns
0-Gy
2-Gy
4-Gy
TLK2 WT
p.(Asp551Gly)
E
60 min, 2-Gy
WT
p.(Asp551Gly)

Figure 3

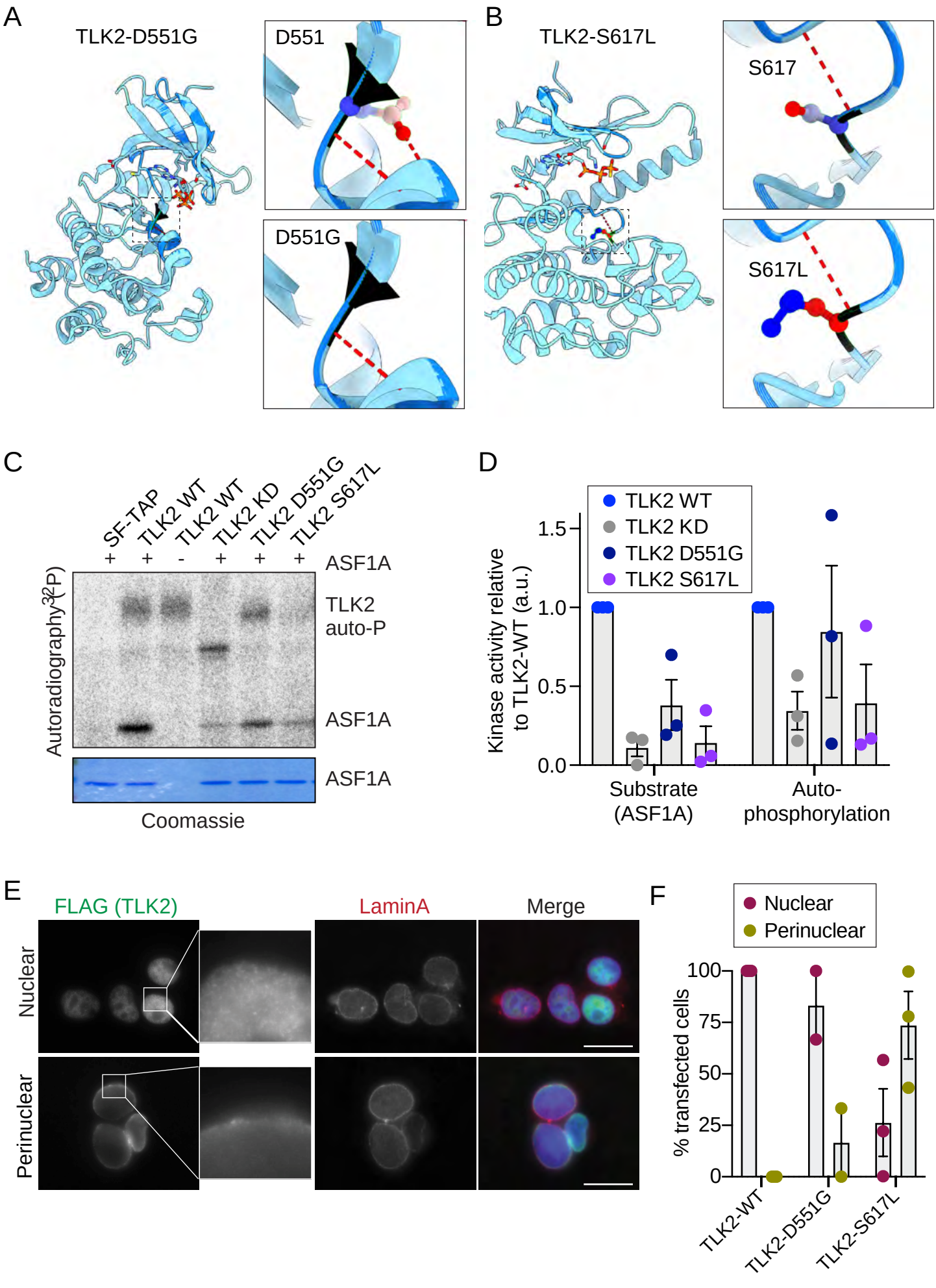

A
TLK2-D551G
D551
D551G
B
TLK2-S617L
S617
S617L
C
SF-TAP
TLK2 WT
TLK2 WT
TLK2 KD
TLK2 D551G
TLK2 S617L
+ + - + + +
ASF1A
Autoradiography (32P)
TLK2 auto-P
ASF1A
ASF1A
Coomassie
D
TLK2 WT
TLK2 KD
TLK2 D551G
TLK2 S617L
Kinase activity relative to TLK2-WT (a.u.)
1.5
1.0
0.5
0.0
Substrate (ASF1A)
Auto-phosphorylation
E
FLAG (TLK2)
LaminA
Merge
Nuclear
Perinuclear
F
Nuclear
Perinuclear
% transfected cells
100
75
50
25
0
TLK2-WT
TLK2-D551G
TLK2-S617L

Figure 4

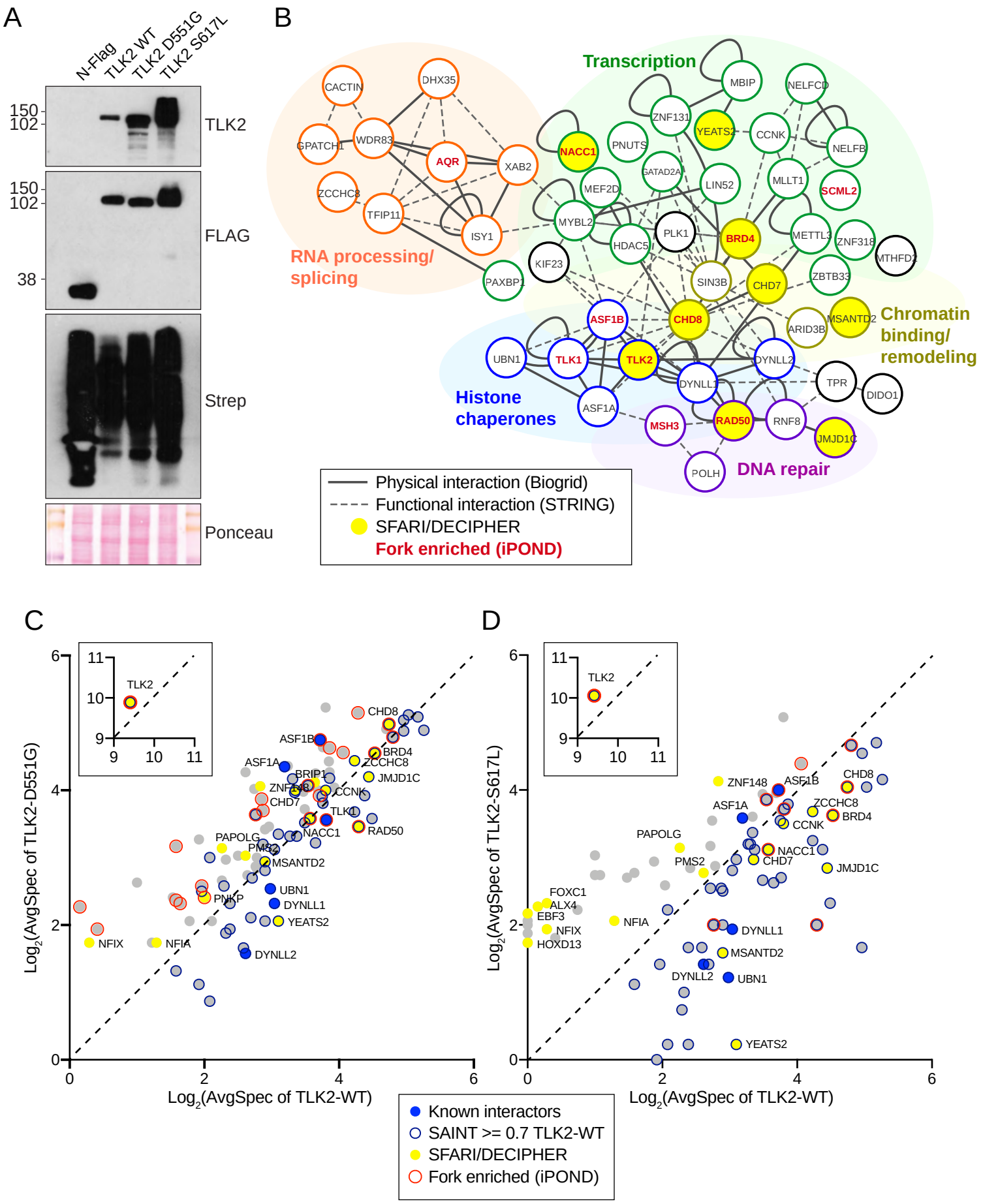

A
N-Flag
TLK2 WT
TLK2 D551G
TLK2 S617L
150
102
TLK2
150
102
FLAG
38
Strep
Ponceau
B
Transcription
RNA processing/ splicing
Chromatin binding/ remodeling
Histone chaperones
DNA repair
CACTIN
DHX35
GPATCH1
WDR83
AQR
XAB2
ZCCHC8
TFIP11
ISY1
PAXBP1
KIF23
NACC1
PNUTS
MBIP
NELFCD
ZNF131
YEATS2
CCNK
NELFB
GATAD2A
MEF2D
LIN52
MLLT1
SCML2
MYBL2
HDAC5
PLK1
BRD4
METTL3
ZNF318
MTHFD2
SIN3B
CHD7
ZBTB33
ASF1B
CHD8
ARID3B
MSANTD2
UBN1
TLK1
TLK2
DYNLL2
DYNLL1
TPR
DIDO1
ASF1A
MSH3
RAD50
RNF8
JMJD1C
POLH
Physical interaction (Biogrid)
Functional interaction (STRING)
SFARI/DECIPHER
Fork enriched (iPOND)
C
Log2(AvgSpec of TLK2-D551G)
Log2(AvgSpec of TLK2-WT)
TLK2
CHD8
ASF1B
BRD4
ASF1A
ZCCHC8
BRIP1
JMJD1C
ZNF148
CCNK
CHD7
TLK1
NACC1
RAD50
PAPOLG
PMS2
MSANTD2
UBN1
PNKP
DYNLL1
YEATS2
NFIX
NFIA
DYNLL2
D
Log2(AvgSpec of TLK2-S617L)
Log2(AvgSpec of TLK2-WT)
TLK2
ZNF148
ASF1B
CHD8
ASF1A
ZCCHC8
BRD4
CCNK
PAPOLG
NACC1
PMS2
CHD7
JMJD1C
FOXC1
ALX4
EBF3
NFIA
NFIX
DYNLL1
HOXD13
MSANTD2
DYNLL2
UBN1
YEATS2
Known interactors
SAINT >= 0.7 TLK2-WT
SFARI/DECIPHER
Fork enriched (iPOND)

Figure 5

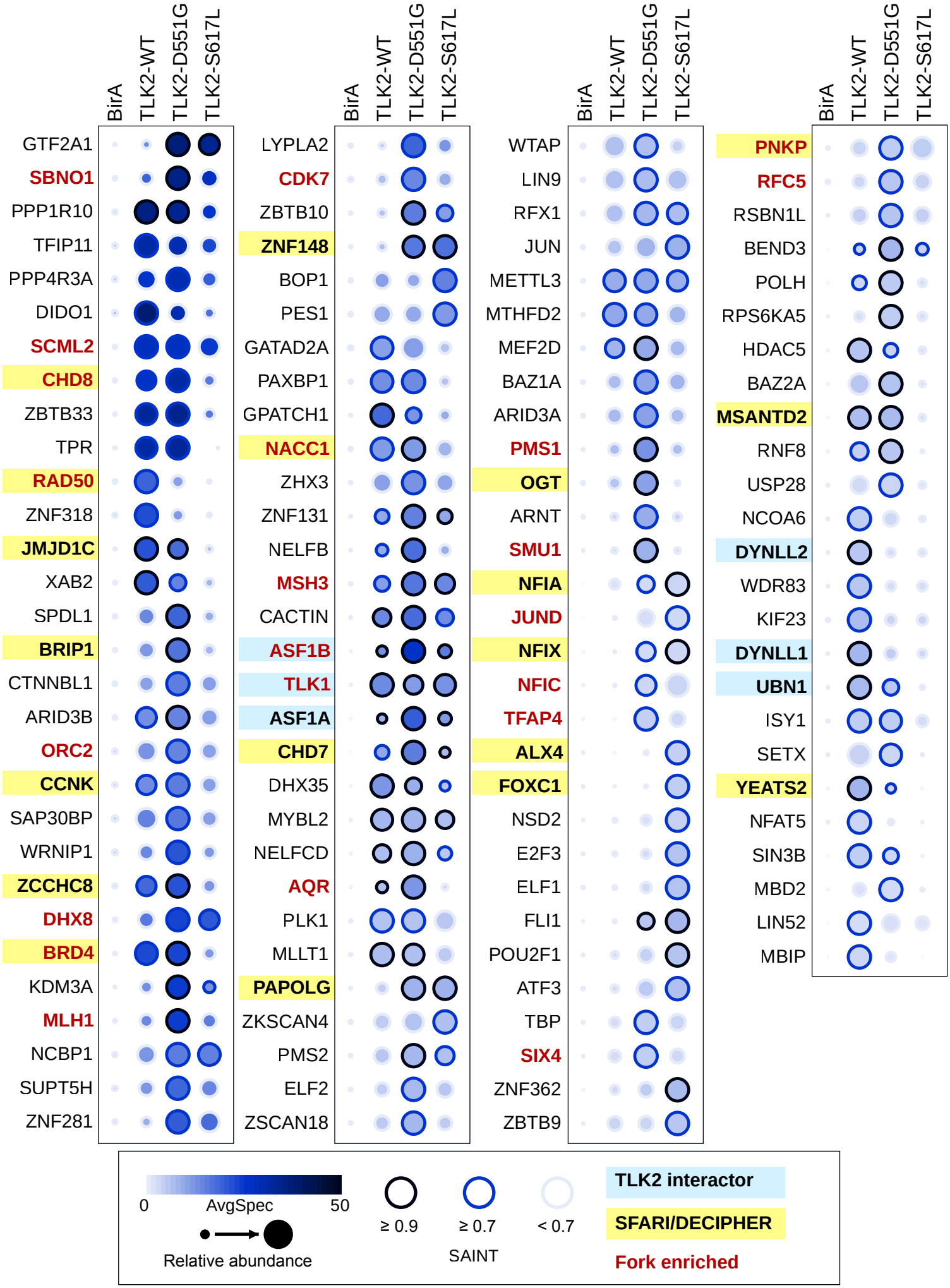
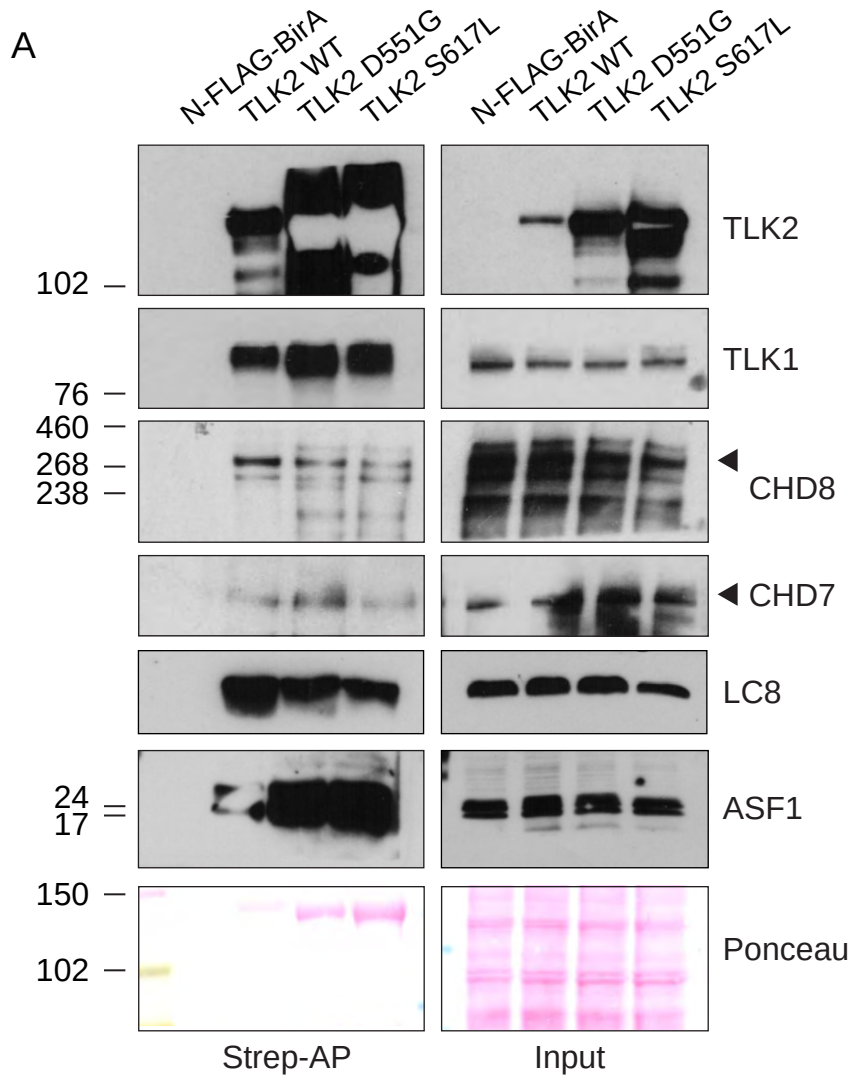

BirA
TLK2-WT
TLK2-D551G
TLK2-S617L
GTF2A1
SBNO1
PPP1R10
TFIP11
PPP4R3A
DIDO1
SCML2
CHD8
ZBTB33
TPR
RAD50
ZNF318
JMJD1C
XAB2
SPDL1
BRIP1
CTNNBL1
ARID3B
ORC2
CCNK
SAP30BP
WRNIP1
ZCCHC8
DHX8
BRD4
KDM3A
MLH1
NCBP1
SUPT5H
ZNF281
LYPLA2
CDK7
ZBTB10
ZNF148
BOP1
PES1
GATAD2A
PAXBP1
GPATCH1
NACC1
ZHX3
ZNF131
NELFB
MSH3
CACTIN
ASF1B
TLK1
ASF1A
CHD7
DHX35
MYBL2
NELFCD
AQR
PLK1
MLLT1
PAPOLG
ZKSCAN4
PMS2
ELF2
ZSCAN18
WTAP
LIN9
RFX1
JUN
METTL3
MTHFD2
MEF2D
BAZ1A
ARID3A
PMS1
OGT
ARNT
SMU1
NFIA
JUND
NFIX
NFIC
TFAP4
ALX4
FOXC1
NSD2
E2F3
ELF1
FLI1
POU2F1
ATF3
TBP
SIX4
ZNF362
ZBTB9
PNKP
RFC5
RSBN1L
BEND3
POLH
RPS6KA5
HDAC5
BAZ2A
MSANTD2
RNF8
USP28
NCOA6
DYNLL2
WDR83
KIF23
DYNLL1
UBN1
ISY1
SETX
YEATS2
NFAT5
SIN3B
MBD2
LIN52
MBIP
0
AvgSpec
50
Relative abundance
≥ 0.9
≥ 0.7
< 0.7
SAINT
TLK2 interactor
SFARI/DECIPHER
Fork enriched

Figure 6


A
N-FLAG-BirA
TLK2 WT
TLK2 D551G
TLK2 S617L
N-FLAG-BirA
TLK2 WT
TLK2 D551G
TLK2 S617L
TLK2
TLK1
CHD8
CHD7
LC8
ASF1
Ponceau
102
76
460
268
238
24
17
150
102
Strep-AP
Input

Functional analysis of *TLK2* variants and their proximal interactomes implicates impaired kinase activity and chromatin maintenance defects in their pathogenesis.

Lisa Pavinato [1], Marina Villamor-Payà [2], Maria Sanchiz-Calvo [2], Cristina Andreoli [3], Marina Gay [2], Marta Vilaseca [2], Gianluca Arauz-Garofalo [2], Andrea Ciolfi [4], Alessandro Bruselles [5], Tommaso Pippucci [6], Valentina Prota [3], Diana Carli [7], Elisa Giorgio [1], Francesca Clementina Radio [4] Vincenzo Antona [8], Mario Giuffrè [8], Kara Ranguin [9], Cindy Colson [9], Silvia De Rubeis [10, 11, 12, 13], Paola Dimartino [14], Joseph Buxbaum [10, 11, 12, 13, 15, 16], Giovanni Battista Ferrero [7], Marco Tartaglia [4], Simone Martinelli [5], Travis H. Stracker [*2], Alfredo Brusco [*1, 17, 18]

SUPPLEMENTARY MATERIAL AND METHODS

Whole exome sequencing, prioritization, and variant calling

Whole Exome Sequencing (WES) data processing, variant filtering, and prioritization by allele frequency, predicted functional impact, and inheritance models were performed using an in-house implemented pipeline, which mainly takes advantage of the Genome Analysis Toolkit (GATK v.3.7). High-quality variants with an effect on the coding sequence or affecting splice site regions were filtered against public databases (dbSNP150 and GnomAD V.2.1) to retain: i) private and clinically associated variants; ii) annotated variants with an unknown frequency or having MAF <0.1%. The functional impact of variants was analyzed by Combined Annotation Dependent Depletion V.1.3 and using ACMG/AMP 2015 guidelines.

Identified variants were confirmed by Sanger sequencing using standard conditions; primer sequences are listed in table S1.

Briefly, gDNA extracted from whole blood was amplified with touchdown PCR (annealing temperature 65-58°C) using KAPA2G Fast HotStart Taq (Merck). PCR products were further purified with FastAP Thermosensitive Alkaline Phosphatase following manufacturer instructions and subjected to Sanger sequencing, using BigDye™ Terminator v3.1 Cycle Sequencing Kit (both from ThermoFisher Scientific).

Table S1. Primers and probes used in Sanger sequencing/ RT-PCR.

Variant/assay	Forward primer	Reverse primer	Note
c.1652A>G	5'-aaaggagtgagaagctgatgacc	5'-caaagaaaacccactaatactgtctcc	
c.1423G>T	5'-agaccagcacccaagtcc	5'-tgtattctgcttggtcacttagg	
TLK2-expression assay	5'- ggcactcctaggggacataa	5'-caactggtggaacacttctgc	UPL probe #62

Table S2. Comparison between transcript NM_001284333.2 and NM_006852.6

NM_001284333.2 (Q86UE8-1; isoform 1)		**NM_006852.6 (Q86UE8-2; isoform 2)**	
c.163A>G	(p.Lys55Glu)	c.163A>G	(p.Lys55Glu)
c.1423G>T	(p.Glu475Ter)	c.1357G>T	(p.Glu453Ter)
c.1652A>G	(p.Asp551Gly)	c.1586A>G	(p.Asp529Gly)

c.1850C>T;	p.(Ser617Leu)	c.1784C>T	(p.Ser595Leu)

p.(Asp551Gly) cDNA and splicing analysis

To analyse the effect of p.(Asp551Gly) variant on complementary DNA (cDNA), we designed a primer pair encompassing exons from 18 to 23 (forward primer: 5'- GCATGCATGTAGGGAATACCG; reverse primer: 5'-TGAACAGTTTGTGGCCTTGG). cDNA derived from both WT and mutated lymphoblastoid cell lines (LCLs) was amplified as described above with touchdown PCR (annealing temperature 65-58°C) and subsequently subjected to Sanger sequencing as described in previous paragraph.

To analyse the effect of the variant on splicing, 2 µl of PCR products were loaded in 1% Agarose gel (1X Tris-Borate-EDTA Buffer, 2% agarose, 0,1% EuroSafe Nucleic Acid Stain). A homemade loading dye (sucrose 44% and 2,5% bromophenol blue, ratio 16:1) was used; base pair were detected comparing to O'GeneRuler 1 kb (ThermoFisher Scientific). Gel was run for 30 minutes and band size was observed, acquiring images under UV light using ChemiDoc Imaging System (BioRad).

TLK2 protein expression

Total proteins were extracted from cellular pellet with RIPA buffer (50mM Tris-HCl pH 7.5; 150 mM NaCl; 1% NP-40; 0.5% Sodium Deoxycholate) supplemented with DTT 0.1 M, EDTA 0.5 M and 100x Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific). Fifteen micrograms of proteins were diluted into 4X LDS sample buffer and 10X Sample Reducing Agent and were electrophoresed on 4–12% Bis-Tris Protein Gels (Thermo Fisher Scientific).

Nitrocellulose membranes (0.45 µm pore, Biorad) were reversible stained with MemCode™ Reversible Protein Stain Kit (Thermo Fisher Scientific) and immunoblotted with the antibodies as indicated in Table S4. Images were acquired with ChemiDoc Imaging System and analysed with ImageLab software (BioRad), using volume tools quantification with global subtraction method. Statistical analysis was performed using two-tailed unpaired Student's t test.

Single-cell gel electrophoresis

A LCLs harbouring the p.(Asp551Gly) variant, fibroblasts carrying the 39-kb deletion encompassing *TLK2*, and their wild-type counterparts were suspended in 0.7% low-melting agarose. Slides were prepared in duplicates with control cells and cells from affected subjects placed at the opposite sides of the same slide, and kept overnight in lysis solution at 4°C. Following lysis, slides were moved to alkaline buffer (20 min) to unwind DNA. Electrophoresis was performed for 20 or 60 minutes at 20 V, 300 mA (0.8 V/cm), at 4°C. Slides were then neutralized in 0.4 M Tris pH 7.5 (3x5 minutes), treated with absolute ethanol, stained with GelRed (Biotium), and analysed at a fluorescence microscope (Leica). Tail moments values were calculated by using a dedicated image analysis system (IAS2000 Delta Sistemi, Italy). To measure the extent of DNA damage induced by γ-ray irradiation and repair capability, LCLs were irradiated with 2-Gy or 4-Gy γ-rays from a ^{137}Cs source (0.8 Gy/min). During the treatment, cells were kept on ice to prevent DNA repair. Repair kinetics were assessed by SCGE as described above. Residual DNA damage was evaluated after 15 and 30 minutes at 37°C. For each experimental point, at least 100 nucleoids were analysed. Statistical analysis was performed using two-tailed unpaired Student's t test.

Site-directed mutagenesis

TLK2 mutations were generated using the QuickChange Lightning site-directed mutagenesis kit (Agilent Technologies) on the plasmids pcDNA3.1 N-SF-TAP-TLK2-WT[1] and BirA*-N-term-TLK2-WT[2] following manufacturers instructions. Primers used are indicated in table S3. All constructs were sequenced (Macrogen) with the primer 5'-CTTTTCACTGGATACTGAC). Constructs were subsequently transfected in AD-293 cells.

Table S3. Site-directed mutagenesis primers.

Mutant	Primer	Sequence 5'-3'
TLK2-D551G (c.1652A>G; p.(Asp551Gly))	Fw	5'-TTTCAGGTAGAAGTCCAGACCATTTCCCTCACAGTATTC
	Rv	5'-GAATACTGTGAGGGAAATGGTCTGGACTTCTACCTGAAA
TLK2- S617L	Fw	5'-CTATCATCATCCATGATCTTCAAAAGACCAAAATCTGTAATTTTTATCTC

(c.1850C>T; p.(Ser617Leu))	Rv	5'-GAGATAAAAATTACAGATTTTGGTCTTTTGAAGATCATGGATGATGATAG

Transfection and affinity purification in mammalian cells

AD-293 cells were seeded in 15 cm plates and transiently transfected the next day with 20 ug of plasmid DNA using polyethylenimine (PEI) (Polysciences Inc., Warrington, PA) and 150mM NaCl. Medium was changed 6-8 hours post-transfection. Cells were harvested 48 hours post-transfection and collected by scraping in PBS. Pelleted cells were lysed in 1 mL of ice-cold lysis buffer (50 mM Tris-HCl pH 7.5, 150mM NaCl, 1% Tween-20, 0.5% NP-40, 1X protease inhibitor cocktail (Roche) and 1X phosphatase inhibitor cocktails 2&3 (Sigma-Aldrich)) on ice for 20 min. Cells were sonicated at medium intensity for 15 mins (Bioruptor XL; Diagenode), and lysates were cleared by centrifugation at 16000g for 20 mins at 4ºC. 100 µL of the lysate were retained for inputs. 4 mg of the total protein extracts were incubated with 100 µL of pre-washed Strep-Tactin superflow resin (IBA GmbH, Gottingen, Germany) overnight at 4ºC using an overhead tumbler. The resin was washed 3 times with 500 µL wash buffer (30 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1% NP-40, 1X protease inhibitor cocktail (Roche) and 1X phosphatase inhibitor cocktails 2&3 (Sigma-Aldrich)). The proteins were eluted from the Strep-Tactin matrix in 50 µL of elution buffer (5x desthiobiotin elution buffer (IBA GmbH) in TBS buffer (30 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1% NP-40, 1X protease inhibitor cocktail (Roche) and 1X phosphatase inhibitor cocktails 2&3 (Sigma-Aldrich)) for 10 mins on ice.

Table S4. Antibodies used in this study.

Antigen	Species	Source & reference	Dilution
TLK1	Rabbit	Cell Signaling #4125	1:1000 (WB)
TLK2	Rabbit	Bethyl Laboratories A301-257A	1:1000 (WB)
ASF1A	Rabbit	Groth Laboratory[3]	1:2000 (WB)
LC8	Rabbit	Abcam ab51603, clone EP1660Y	1:1000 (WB)
CHD7	Rabbit	Bethyl Laboratories A301-223A	1:2000 (WB)
CHD8	Rabbit	Bethyl Laboratories A301-224A	1:2000 (WB)

FLAG	Mouse	Sigma-Aldrich F3165, clone M2	1:5000 (WB)
VINCULIN	Rabbit	Millipore, #AB6039	1:5000 (WB)
Strep-tag	Mouse	IBA GmbH 2-1509-001	1:1000 (WB)
M2 Flag	Mouse	Sigma-Aldrich F1804	1:500 (IF)
LaminA	Rabbit	SC-20680	1:500 (IF)
Protein A/G	anti-Rabbit HRP	Thermo Fisher Scientific 32490	1:15000 (WB Nitro) 1:30000 (WB PVDF)
Mouse IgG	goat anti-mouse HRP	Thermo Fisher Scientific 31430	1:15000 (WB Nitro) 1:30000 (WB PVDF)
Alexa Fluor 488	Goat anti-mouse IgG	Thermo Fisher Scientific A28175	1:500 (IF)
Alexa Fluor 594	Goat anti-rabbit IgG	Thermo Fisher Scientific A11012	1:500 (IF)

In vitro kinase assays from cell lysates

In vitro kinase assays were performed as previously described with minor modifications. after Strep-AP of pcDNA3.1 N-SF-TAP TLK2 from AD-293 cells. 200 µg of Strep-AP were incubated with 2 µCi ^{32}P-γ-ATP, 100 µM cold ATP, 1 µg of purified GST-ASF1A protein (kind gift from Anja Groth[4]) in 12 µL of kinase buffer (50mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 2mM DTT, 1X protein inhibitor cocktail (Roche) and 1X phosphatase inhibitor cocktails 2&3 (Sigma-Aldrich)). The reaction was incubated at 30ºC for 30 min. After that, the reaction was stopped by adding 4 µL of Sample Buffer (6x SDS, (0.2% bromophenol blue and *β*-mercaptoethanol), and boiled for 5-10 mins at 95ºC. Samples were analyzed on SDS-PAGE, stained with Coomassie Blue for 1 hour, washed 4 times with destaining buffer (10% acetic acid, 40% methanol, and 50% H_2O) and vacuum dried with an SGD2000 (Savant) for 2 hours at 60ºC . TLK2 and ASF1A phosphorylation were measured using a Typhoon 8600 Variable Mode Imager (Molecular Dynamics) and band intensity quantified using ImageJ[5].

Immunofluorescence

AD-293 cells were seeded and transiently transfected with the indicated plasmids. The next day, cells were trypsinized and seeded on poly-L-Lysine-coated coverslips. 48 hours post-transfection, cells were fixed with 4% formaldehyde (Santa Cruz Biotechnology) for 10 mins and permeabilized in 0.2% Triton X-100 (Sigma-Aldrich) in 1x PBS for 10 mins at room temperature (R/T). Coverslips were washed twice with PBS and blocked with PBS-BT (0.1% Triton X-100, 3% BSA (Sigma-Aldrich) in PBS) for 30 min at R/T. The coverslips were incubated with the corresponding primary antibodies (Table S4) for 4 hrs at 4ºC in a humid chamber. After three washes with PBS-BT, the coverslips were incubated with the secondary antibody (Table S4) for 1 hour at R/T in a dark humid chamber. The coverslips were washed 3 times in PBS-BT and 4′,6-diamidino-2-phenylindole (DAPI) was added diluted 1:3000 in the first wash. Fluorescent images were acquired with an Orca AG camera (Hamamatsu) mounted on a Leica DMI6000B microscope equipped with 1.4 numerical aperture 100X oil immersion objective. The phenotypic distribution was quantified in 10 different fields of view in each condition.

Proximity-dependent biotin identification mass spectrometry (BioID-MS)

AD-293 cells were seeded in 15 cm plates and transiently transfected the next day with 20 µg of BirA* plasmids using PEI and 150 mM NaCl as described. Medium was changed 6-8 hours post-transfection. 24 hours post-transfection, 50 µM of biotin (IBIAN Biotechnology; 2-1016-002) were added per plate. For mass spectrometry, 5x15cm plates were used per condition. 48 hours post-transfection, the cells were harvested with Trypsin-EDTA (Sigma-Aldrich) and the 5 plates per condition were pooled together. Cell pellets were washed twice in cold PBS and lysed in 5 mL of cold lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1% SDS, 2 mM Mg_2Cl, 1% Triton X-100 (Sigma-Aldrich), 1mM EDTA (Sigma-Aldrich), 1mM EGTA (Sigma-Aldrich), 1:2000 benzonase 25 U/mL (Sigma-Aldrich), 1x protease inhibitor cocktail (Roche) and 1x phosphatase inhibitor cocktails 2&3 (Sigma-Aldrich)). 100 µL of the lysate were retained for Western blotting analysis. The remaining lysate was incubated with streptavidin-sepharose beads (GE Healthcare 2-1206-010) during 3 hours in an end-over-end rotator at 4ºC in order to isolate the biotinylated proteins. The beads were washed once in lysis buffer and three times in 50 nM ammonium bicarbonate pH 8.3 buffer. Samples were snap-frozen and sent to the Mass Spectrometry & Proteomics Core Facility at IRB Barcelona for tryptic digestion and analysis.

Tryptic digestion was performed directly on beads by incubation with 2 ug of trypsin in 50 mM NH_4HCO_3 at 37ºC overnight. The next morning, an additional 1 µg of trypsin was added and incubated for 2 h at 37°C. The digestion was stopped by adding formic acid to 1% final concentration. Samples were cleaned through C18 tips (polyLC C18 tips) and peptides were eluted with 80% acetonitrile, 1% TFA. Samples were diluted to 20% acetonitrile, 0.25% TFA, loaded into strong cation exchange columns (SCX) and peptides were eluted in 5% NH_4OH, 30% methanol. Finally, samples were evaporated to dry, reconstituted in 50 µL and diluted 1:8 with 3% acetonitrile, 1% formic acid aqueous solution for nanoLC-MS/MS analysis.

The nano-LC-MS/MS was set up as follows. Digested peptides were diluted in 3% ACN/1% FA. Sample was loaded to a 300 µm × 5 mm PepMap100, 5 µm, 100 Å, C18 µ-precolumn (Thermo Scientific) at a flow rate of 15 µl/min using a Thermo Scientific Dionex Ultimate 3000 chromatographic system (Thermo Scientific). Peptides were separated using a C18 analytical column Acclaim PEPMAP 100 75 µm x50 cm nanoviper C18 3 µm 100A (Thermo Scientific) with a 90 min run, comprising three consecutive steps with linear gradients from 3 to 35% B in 60 min, from 35 to 50% B in 5 min, and from 50% to 85% B in 2 min, followed by isocratic elution at 85% B in 5 min and stabilization to initial conditions (A= 0.1% FA in water, B= 0.1% FA in CH_3CN). The column outlet was directly connected to an Advion TriVersa NanoMate (Advion) fitted on an Orbitrap Fusion Lumos™ Tribrid (Thermo Scientific). The mass spectrometer was operated in a data-dependent acquisition (DDA) mode. Survey MS scans were acquired in the Orbitrap with the resolution (defined at 200 m/z) set to 120,000. The lock mass was user-defined at 445.12 m/z in each Orbitrap scan. The top speed (most intense) ions per scan were fragmented by CID and detected in the linear ion trap. The ion count target value was 400,000 and 10,000 for the survey scan and for the MS/MS scan respectively. Target ions already selected for MS/MS were dynamically excluded for 15s. Spray voltage in the NanoMate source was set to 1.60 kV. RF Lens were tuned to 30%. Minimal signal required to trigger MS to MS/MS switch was set to 5,000. The spectrometer was working in positive polarity mode and singly charge state precursors were rejected for fragmentation.

We performed a twin database search with two different softwares, Thermo Proteome Discoverer v2.3.0.480 (PD) and MaxQuant v1.6.6.0 (MQ). The search engine nodes used were Sequest HT for PD and Andromeda for MQ. The databases used in the search was SwissProt Human (release 2019 01) including contaminants and TLK1 and TLK2 proteins. We run the search against targeted and decoy databases to determine the false discovery rate (FDR). Search parameters included trypsin enzyme specificity, allowing for two missed cleavage sites, oxidation in M and acetylation in protein N-terminus as dynamic modifications. Peptide mass tolerance was 10 ppm and the MS/MS tolerance was 0.6 Da. Peptide were filtered at a false discovery rate (FDR) of 1 % based on the number of hits against the reversed sequence database.

For the quantitative analysis, contaminant identifications were removed and unique peptides (peptides that are not shared between different protein groups) were used for the quantitative analysis with SAINTexpress-spc v3.6.1[6]. SAINTexpress compares the prey control spectral counts with the prey test spectral counts for all available replicates. For each available bait and for each available replicate, we took as prey count the maximum count result between PD and MQ. Once obtained this combined dataset, we ran the SAINTexpress algorithm with TLK2 samples and a number of controls samples from previous experiments in the same cell type (n=45 total). High confidence interactors were defined as those with a SAINT score of 0.7 or greater. Output data from SAINTexpress is provided in Supplementary table S8 and raw data is available in the PRIDE repository, accession number PXD019450.

Table S5. In silico predictors for p.(Asp551Gly) variant

Tool	Predicted impact
MutationTaster	Disease causing
Mutation assessor	Medium
FATHMM	Tolerated
FATHMM-MKL	Damaging
FATHMM-XF	Damaging
LRT	Deleterious
DEOGEN2	Tolerated
EIGEN	Pathogenic
EIGEN PC	Pathogenic
SIFT	Damaging
SIFT4G	Damaging
PROVEAN	Damaging
PolyPhen-2	Probably damaging
MVP	Pathogenic
REVEL	Pathogenic
PrimateAI	Damaging
MetaSVM	Tolerated
MetaLR	Tolerated

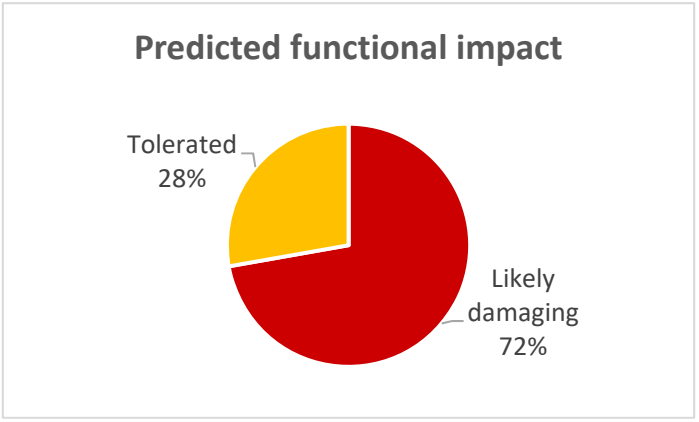


Table S6. Clinical characterization of patients carrying TLK2 variants

	Case 1	**Case 2**	**Case 3**	**Case 4**	**Case 5**	**Case 6**
	c.1652A>G; p.(Asp551Gly)	**(c.1423G>T; p.(Glu475Ter))**	**(c.1423G>T; p.(Glu475Ter))**	**(c.1423G>T; p.(Glu475Ter))**	**(c.1423G>T; p.(Glu475Ter))**	**del17q23.2**
Sex	Male	Female	Female	Male	Male	Female
Age at last examination	11	47	26	16	3	16
Ethnicity	Caucasian	Caucasian	Caucasian	Caucasian	Caucasian	Caucasian
ASD	+	-	-	-	-	-
ID	+	+	+	+	N.A.	N.A.
IQ≤70	-	-	+	+	-	-
IQ 71-85	+	+	-	-	-	-
IQ 86-100	-	-	-	-	-	-
Social-emotional problems	+	-	+	+	-	-
Tantrums	-	-	-	-	-	-
ADHD	+	-	+	+	+	-
Brain abnormality	-	-	-	-	+	-
Hypotonia	-	-	-	+	+	-
Speech delay	+	-	+	+	+	-
Bipolar disorder	-	-	-	-	-	-
Epilepsy	-	-	-	-	-	-
Obsessive compulsive behaviour	+	-	+	-	-	-
Anxiety	+	-	+	+	-	-
Aggressiveness	-	-	+	+	-	-
Difficulties in memory and transcription	-	-	+	+	-	+
Difficulties in reading and writing	+	-	+	+	-	-
Short attention span	+	-	+	-	-	-
Delayed motor development	+	-	-	-	-	+
Microcephaly	-	+	+	+	+	-
Gastrointestinal problems	-	-	-	-	-	-
Plagiocephaly	-	-	+	+	+	-
Skeletal anomalies of the hands	+	-	-	-	-	+
Skeletal anomalies of the feet	-	-	+	+	-	-
Joint hypermobility	+	-	+	+	-	-

Other minor skeletal anomalies	-	-	+	+	-	-
Dysmorphic facial features	+	+	+	+	+	+

+ = observed; -= not observed; N.A.= not available

SUPPLEMENTAL NOTE: CASE REPORTS

Family 1 – Case 1

The patient was the only child of healthy unrelated Caucasian parents. The mother presented one spontaneous miscarriage of the first trimester and one pregnancy ended with intrauterine death at 21 weeks of gestation for severe cardiomyopathy. The maternal uncle presented schizophrenia and the maternal aunt presented depression. Remaining family history is unremarkable.

Pregnancy, physiologically conceived, started bigeminal and was complicated by loss of the twin at 9 weeks of gestation, threat of abortion at the 4th and 5th month, oligoamnios and decreased fetal movements.

He was born at 41+0 weeks of gestation by spontaneous delivery with weight 2850 gr (5° centile, -0.68 SD), length 48.7 cm (10° centile, -1.3 SD), OFC 33.9 cm (20° centile, -0.84 SD). APGAR score was 8 and 9 at 1st and 5th minute, respectively.

Slight feeding difficulties with poor growth have been reported in the first months of life.

He presented mild motor delay with sitting at 9 months of age and walking at 20 months of age and language delay with first words at 12 months and first sentences at 3 years.

He presented tendency to isolation and was diagnosed with autism at 6.5 years.

He presented pavor nocturnus until 7 years of age, short attention span, good memory, anxiety, compulsive obsessive behavior and difficulties in writing.

EEG, brain MRI and abdominal ultrasound at 5 years of age were normal. IQ evaluation with WISC-III scale showed a total IQ of 83 (borderline).

At last physical examination at 11 years of age he presented height 143 cm (32° centile, -0.46 SD), weight 43 kg (67° centile, +0.43 SD), OFC 54 cm (64° centile, +0.35 SD). The calculated body mass index (BMI) was in the range of normal (BMI 21).

Dysmorphic features included upslanting palpebral fissures, prominent nasal bridge, broad nasal tip, low hanging columella, thin lips, prognatism, pointed chin. He also presented tapering fingers and mild joint hypermobility.

Family 2– Cases 2-5

The patients came from a non-consanguineous Italian family and came to attention thanks to genetic investigations required for affected patients 3 and 4, while patient 5 born meanwhile.

The mother (case 2) familial history was unremarkable. Her brother, her two sisters and her parents were never suspected of having intellectual disability, although detailed clinical information about them were not available. She had mild intellectual disability, that came to light only after a retrospective clinical evaluation after the identification of *TLK2* c.1423G>T; p.(Glu475Ter) variant by trio whole exome sequencing.

At last physical examination at 46 years of age she presented height 155 cm (10° centile, -1.27 SD), weight 68 kg (74° centile, +0.64 SD), OFC 52 cm (1° centile, -2.5 SD). She was overweight (BMI 28.3).

Dysmorphic features included lateral thin eyebrows, hypertelorism, upslanting palpebral fissures, prominent columella, broad nasal bridge and tip, short philtrum, thin upper lip, facing down mouth angles, mandibular prognatism.

Patient 3 was the first female child, characterised by intellectual disability (IQ<70) and other behavioural and neurological issues, including social-emotional problems, attention deficit hyperactivity disorder, aggressiveness, anxiety, obsessive compulsive disorder and speech delay. She showed also difficulties in memory, transcription, reading and writing and a short attention span.

At last physical examination at 25 years of age she presented height 154 cm (8° centile, -1.43 SD), weight 70 kg (78° centile, +0.77 SD), OFC 52.5 cm (4° centile, -1.7 SD). She was overweight (BMI 29.5).

She presented strabismus, joint hypermobility, scoliosis, plagiocephaly and left foot hexadactyly.

Dysmorphic features included lateral thin eyebrows, hypertelorism, upslanting palpebral fissures, prominent columella, broad nasal tip, macrostomia, thin upper lip, facing down mouth angles, mandibular prognatism.

Patient 4 has been subjected to whole exome sequencing as his older sister. He was characterized by intellectual disability (IQ<70), social-emotional problems, attention deficit hyperactivity disorder, anxiety and aggressiveness. Moreover, a speech delay and generalized hypotonia were observed. He had difficulties in memory, transcription, reading and writing.

At last physical examination at 15 years of age she presented height 185 cm (91° centile, +1.33 SD), weight 75 kg (91° centile, +1.33 SD), OFC 52 cm (3° centile, -1.95 SD). His BMI was in the range of normal (BMI 21.9).

As additional features, he was characterized by plagiocephaly, joint hypermobility, pes planus and strabismus. Feet big toes overlap was observed.

Dysmorphic features included telecanthus, prognatism, upslanting palpebral fissures, epicanthal folds, posteriorly rotated ears, low-implant auricles, synophrys, prominent columella, broad nasal tip, thick filter, long face, ptosis, thin upper lip and facing down mouth angles.

Patient 5 was the last born of the family and *TLK2* c.1423G>T; p.(Glu475Ter) variant was detected after target Sanger sequencing following diagnosis of patient 3 and 4. He was too young to assess intellectual disability but he presented a global psychomotor delay, speech delay, hypotonia and attention deficit hyperactivity disorder.

At last physical examination at 2 years of age she presented height 80 cm (2° centile, -2.06 SD), weight 8,2 kg (<1° centile, -3.67 SD), OFC 43 cm (<1° centile, -3.96 SD), presenting an underweight situation (BMI 12.8).

He showed plagiocephaly and conductive hearing loss and brain MRI showed a mild hypoplasia of the corpus callosum.

His facial features were similar to the ones observed for the other affected relatives, including long face, prominent nasal bridge and broad nasal tip, thin vermilion of the upper lip, upslanting palpebral fissures, lateral thin eyebrows, facing down mouth angles and high palate.

Family 3– Case 6

The 17-year-old Caucasian patient has a history of intellectual deficiency with thumb anomalies. She was born at 35 weeks by spontaneous vaginal birth with a weight of 2320g (2° centile, -2.10 SD), height 43.5cm (<1° centile, -2.67 SD), OFC 32.5cm (6° centile, -1.56 SD), with absence of flexure fold of the thumbs.

A persistent arterial canal was found and cured surgically. Feeding difficulties were present in the early age. She walked at 14 months. At 5 years of age, the patient weighted 13kg (1° centile, -2.54 SD) for a height of 96 cm (1° centile, -2.55 SD), OFC 48.5 cm (10° centile, -1.3 SD). Her BMI brought out an underweight (BMI 14.1); at 7 years, she was treated by growth hormone.

Dysmorphic characteristics included slightly short forehead, hypertelorism, a large nose root with epicanthi, bifid tip of the nose. There was an absence of earlobes, the palpebral fissures were narrow and oriented upward. She had short hands with an impossibility of bending the thumbs, short and brittle nails. X rays of the thumbs were normal.

Academically, in 2nd grade school she was able to read and write. She had memory and transcription difficulties. She had no trouble concentrating. She benefited from treatment in speech therapy as well as physiotherapy and psychomotricity.

At 17 years of age, she weighed 43 kg (4° centile, -1.4 SD), height 155 cm (12° centile, -1.16 SD), OFC 53.8 cm (32° centile, -0.47 SD); as during the childhood, the patient maintained an underweight condition (BMI 17.9).

The tests performed for this patient included a chromosomal analysis and array-CGH. Array-CGH showed a *de novo* deletion in 17q23.2 with a size of 39-87 kb including *TLK2*.

Table S7. Comparison between features observed in our cases and previously reported ones[7, 8]

Observed Features		Reijnders *et al.*	Töpf *et al.*	Our cases
	Low normal ID (IQ 85-100)	72%	0%	0% (2/4)

Neurological and behavioural				
	Borderline ID (IQ 70-85)	14%	0%.	50% (2/4)
	ID (IQ ≤70)	6%	100%	50% (2/4)
	Social-emotional problems	18%	n.r.	50% (3/6)
	Tantrums	32%	n.r.	0% (0/6)
	ASD	32%	n.r.	17% (1/6)
	ADHD/ADD	15%	n.r.	67% (4/6)
	Short attention span	5%	n.r.	33% (2/6)
	Anxiety	12%	n.r.	50% (3/6)
	Obsessive-compulsive behaviour	6%	n.r.	33% (2/6)
	Aggressiveness	6%	100%	33% (2/6)
	Difficulties in reading and writing	n.r.	n.r.	50% (3/6)
	Difficulties in memory and transcription	n.r.	n.r.	50% (3/6)
	Motor delay	89%	100%	33% (2/6)
	Language delay	92%	100%	67% (4/6)
	Pavor nocturnus	n.r.	n.r.	17% (1/6)
	Epilepsy	14%	100%	0% (0/6)
Growth parameters	Short stature	37%	n.r.	0% (0/6)
	Underweight	14%	n.r.	33% (2/6)
	Overweight	8%	n.r.	33% (2/6)
	Microcephaly	24%	100%	67% (4/6)
Gastro-intestinal problems	Constipation	60%	100%	0% (0/6)
	Diarrhoea	9%	n.r.	0% (0/6)
Skeletal	Scoliosis	9%	n.r.	17% (1/6)
	Contractures hands	9%	n.r.	0% (0/6)

	Foot hexadactyly	n.r.	n.r.	17% (1/6)
	Feet big toes overlap	n.r.	n.r.	17% (1/6)
	Tapering fingers	n.r.	n.r.	17% (1/6)
	Short hands with short distal phalanx	n.r.	n.r.	17% (1/6)
Dysmorphic facial features	Abnormal palpebral fissures	55%	100%	100% (6/6)
	Prominent nasal bridge	68%	100%	67% (4/6)
	Broad nasal tip	66%	n.r.	83% (5/6)
	Bifid tip of the nose	n.r.	n.r.	17% (1/6)
	Low hanging columella	n.r.	n.r.	33% (2/6)
	Thin lips	62%	100%	83% (5/6)
	Prognathism	n.r.	n.r.	83% (5/6)
	Pointed and tall chin	42%	n.r.	17% (1/6)
	Blepharophimosis	82%	n.r.	0% (0/6)
	Telecanthus	n.r.	100%	33% (2/6)
	Epicanthal folds	42%	n.r.	33% (2/6)
	Narrow mouth	32%	n.r.	0% (0/6)
	High palate	30%	n.r.	50% (3/6)
	Microtia	29%	n.r.	0% (0/6)
	Posteriorly rotated ears	29%	n.r.	17% (1/6)
	Ears without lobes	n.r.	n.r.	17% (1/6)
	Low-implant auricle	n.r.	n.r.	17% (1/6)
	Long face	27%	n.r.	67% (4/6)
	Ptosis	21%	n.r.	17% (1/6)
	Wide spaced eyes/Hypertelorism	74%	n.r.	67% (4/6)
	Synophrys	n.r.	n.r.	17% (1/6)
	Macrostomia/Big mouth	n.r.	100%	17% (1/6)
	Facing down mouth angles	n.r.	n.r.	67% (4/6)

Other features	Neonatal feeding difficulties	44%	100%	17% (1/6)
	Hypotonia	40%	100%	33% (2/6)
	Refraction abnormality	31%	n.r.	0% (0/6)
	Strabismus	29%	n.r.	33% (2/6)
	Recurrent otitis media	27%	n.r.	0% (0/6)
	Conductive hearing loss	16%	100%	17% (1/6)
	Joint hypermobility	24%	n.r.	50% (3/6)
	Pes Planus	25%	n.r.	17% (1/6)
	Plagiocephaly	16%	n.r.	50% (3/6)
	Hypertrichosis	17%	n.r.	0% (0/6)
	Brain abnormality	25%	100%	17% (1/6)
	Craniosynostosis	11%	n.r.	0% (0/6)
	Hoarse voice	9%	n.r.	0% (0/6)

n.r.= not reported; ID= intellectual disability; ASD= autism spectrum disorder; ADHD= attention deficit hyperactivity disorder; ADD= attention deficit disorder.

Features not previously reported in other cases are in bold.

Table S8. BioID-MS results. The first tab contains the unfiltered results for all baits. The second tab is the raw output of the SAINTexpress software (http://saint-apms.sourceforge.net/Main.html) and the third tab contains the lists of proteins used for comparison to iPOND experiments or the SFARI (gene.sfari.org)/DECIPHER (decipher.sanger.ac.uk) databases.

Table S9. Classification of proximal interactors. The first tab is an alphabetical list of the proximal interactors identified for all baits. Their statistical significance is colour coded as in the legend and those found in the OMIM (omim.org) or SFARI (gene.sfari.org)/DECIPHER (decipher.sanger.ac.uk) databases are

indicated. The second tab indicates only the SFARI genes identified with additional details about their molecular functions.

Figure S1.

(A) Sanger sequencing of *TLK2* WT versus p.(Asp551Gly) variant. Results from sequencing of both gDNA and cDNA are showed, to allow a simply comparison. No differences in the peaks between p.(Asp551Gly) gDNA and cDNA were observed. (B) Splicing analysis of *TLK2* WT versus *TLK2* p.(Asp551Gly) variant. After 30 minutes of run on 1% agarose gel, LCLs from WT control and from patient showed the same base pair size. (D) Western Blot analysis showed a significantly increased TLK2 protein expression compared to controls (***p 0,0005; two-tailed unpaired Student's t test). TLK2 expression was normalized on Vinculin; MemCode image is reported as loading control. (C) SCGE assays documented that *TLK2* haploinsufficiency affects proper chromatin compaction. Fibroblasts carrying the 17q23.2 deletion encompassing TLK2 showed a more relaxed state of chromatin compared to control cells. Nucleoids of cells derived from subjects 6 showed a significantly higher tail moment value after 60 minutes of electrophoresis run time. For each point, at least 100 cells were analysed. Values are represented as mean ± SEM of three independent experiments.

REFERENCES

1 Mortuza GB, Hermida D, Pedersen AK, Segura-Bayona S, López-Méndez B, Redondo P, Rüther P, Pozdnyakova I, Garrote AM, Muñoz IG, Villamor-Payà M, Jauset C, Olsen J V., Stracker TH, Montoya G. Molecular basis of Tousled-Like Kinase 2 activation. *Nat Commun* Published Online First: 2018. doi:10.1038/s41467-018-04941-y

2 Segura-Bayona S, Knobel PA, Gonzalez-Buron H, Youssef SA, Peña-Blanco A, Coyaud E, Lopez-Rovira T, Rein K, Palenzuela L, Colombelli J, Forrow S, Raught B, Groth A, De Bruin A, Stracker TH. Differential requirements for Tousled-like kinases 1 and 2 in mammalian development. *Cell Death Differ* Published Online First: 2017. doi:10.1038/cdd.2017.108

3 Groth A, Ray-Gallet D, Quivy JP, Lukas J, Bartek J, Almouzni G. Human Asf1 regulates the flow of S phase histones during replicational stress. *Mol Cell* Published Online First: 2005. doi:10.1016/j.molcel.2004.12.018

4 Klimovskaia IM, Young C, Strømme CB, Menard P, Jasencakova Z, Mejlvang J, Ask K, Ploug M, Nielsen ML, Jensen ON, Groth A. Tousled-like kinases phosphorylate Asf1 to promote histone supply

during DNA replication. *Nat Commun* Published Online First: 2014. doi:10.1038/ncomms4394

5 Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. Fiji: An open-source platform for biological-image analysis. Nat. Methods. 2012. doi:10.1038/nmeth.2019

6 Teo G, Liu G, Zhang J, Nesvizhskii AI, Gingras AC, Choi H. SAINTexpress: Improvements and additional features in Significance Analysis of INTeractome software. *J Proteomics* Published Online First: 2014. doi:10.1016/j.jprot.2013.10.023

7 Reijnders MRF, Miller KA, Alvi M, Goos JAC, Lees MM, de Burca A, Henderson A, Kraus A, Mikat B, de Vries BBA, Isidor B, Kerr B, Marcelis C, Schluth-Bolard C, Deshpande C, Ruivenkamp CAL, Wieczorek D, Baralle D, Blair EM, Engels H, Lüdecke HJ, Eason J, Santen GWE, Clayton-Smith J, Chandler K, Tatton-Brown K, Payne K, Helbig K, Radtke K, Nugent KM, Cremer K, Strom TM, Bird LM, Sinnema M, Bitner-Glindzicz M, van Dooren MF, Alders M, Koopmans M, Brick L, Kozenko M, Harline ML, Klaassens M, Steinraths M, Cooper NS, Edery P, Yap P, Terhal PA, van der Spek PJ, Lakeman P, Taylor RL, Littlejohn RO, Pfundt R, Mercimek-Andrews S, Stegmann APA, Kant SG, McLean S, Joss S, Swagemakers SMA, Douzgou S, Wall SA, Küry S, Calpena E, Koelling N, McGowan SJ, Twigg SRF, Mathijssen IMJ, Nellaker C, Brunner HG, Wilkie AOM. De Novo and Inherited Loss-of-Function Variants in TLK2: Clinical and Genotype-Phenotype Evaluation of a Distinct Neurodevelopmental Disorder. *Am J Hum Genet* 2018;**102**:1195–203.

8 Töpf A, Oktay Y, Balaraju S, Yilmaz E, Sonmezler E, Yis U, Laurie S, Thompson R, Roos A, MacArthur DG, Yaramis A, Güngör S, Lochmüller H, Hiz S, Horvath R. Severe neurodevelopmental disease caused by a homozygous TLK2 variant. *Eur J Hum Genet* Published Online First: 2019. doi:10.1038/s41431-019-0519-x

List of all genes in the proximal interactome of wild-type and mutated TLK2

Gene	SFARI score	Disease	OMIM ID	Inheritance
ALX4		{Craniosynostosis 5, susceptibility to}	615529	AD
		Frontonasal dysplasia 2	613451	AR
		Parietal foramina 2	609597	AD
AQR				
ARID3A				
ARID3B				
ARNT				
ASF1A				
ASF1B				
ATF3				
BAZ1A				
BAZ2A				
BEND3				
BOP1				
BRD4	3		608749	
BRIP1		{Breast cancer, early-onset, susceptibility to}	114480	AD, SMu
		Fanconi anemia, complementation group J	609054	
CACTIN				
CCNK	S	?Intellectual developmental disorder with hypertelorism and distinctive facies	618147	
CDK7				
CHD7	1	CHARGE syndrome	214800	
CHD7	1	Hypogonadotropic hypogonadism 5 with or without anosmia	612370	
CHD8	1	{Autism, susceptibility to, 18}	615032	AD
CTNNBL1				
DHX35				
DHX8				
DIDO1				
DYNLL1				
DYNLL2				
E2F3				
EBF1				
EBF3		Hypotonia, ataxia, and delayed development syndrome	617330	AD
ELF1				
ELF2				
FLI1		Bleeding disorder, platelet-type, 21	617443	AD, AR
FOXC1				
GATAD2A				
GPATCH1				

GTF2A1				
HDAC5				
HOXD13				
ISY1				
JMJD1C	3		604503	
JUN				
JUNB				
JUND				
KDM3A				
KIF23				
LIN52				
LIN9				
LYPLA2				
MBD2				
MBIP				
MEF2D				
METTL3				
MLH1		Colorectal cancer, hereditary nonpolyposis, type 2	609310	
		Mismatch repair cancer syndrome	276300	AR
		Muir-Torre syndrome	158320	AD
MLLT1				
MSANTD2	3			
MSH3		Endometrial carcinoma, somatic	608089	
		Familial adenomatous polyposis 4	617100	AR
MTHFD2				
MYBL2				
NACC1	1	Neurodevelopmental disorder with epilepsy, cataracts, feeding difficulties, and delayed brain myelination	617393	AR
NCBP1				
NCOA6				
NELFB				
NELFCD				
NFAT5				
NFIA	3	Brain malformations with or without urinary tract defects	613735	
NFIC				
NFIX	S	Marshall-Smith syndrome	602535	AD
		Sotos syndrome 2	614753	AD
NSD2				
OGT				
ORC2				
PAPOLG	3			
PAXBP1				
PES1				

PLK1				
PMS1				
PMS2		Colorectal cancer, hereditary nonpolyposis, type 4	614337	
		Mismatch repair cancer syndrome	276300	AR
PNKP				
POLH		Xeroderma pigmentosum, variant type	278750	AR
POU2F1				
PPP1R10				
PPP4R3A				
RAD50		Nijmegen breakage syndrome-like disorder	613078	
RFC5				
RFX1				
RNF8				
RPS6KA5				
RSBN1L				
SAP30BP				
SBNO1				
SCML2				
SETX				
SIN3B				
SIX4				
SMU1				
SPDL1				
STPG4				
SUPT5H				
TBP				
TFAP4				
TFIP11				
TLK1				
TPR				
UBN1				
USP28				
WDR83				
WRNIP1				
WTAP				
XAB2				
YEATS2	3	?Epilepsy, myoclonic, familial adult, 4	615127	
ZBTB10				
ZBTB33				
ZBTB9				
ZCCHC8		?Pulmonary fibrosis and/or bone marrow failure, telomere-related, 5	618674	AD
ZHX3				
ZKSCAN4				

ZNF131				
ZNF148		Global developmental delay, absent or hypoplastic corpus callosum, and dysmorphic facies	617260	AD
ZNF281				
ZNF318				
ZNF362				
ZSCAN18				

Legend

SFARI (release 2020Q1) / DECIPHER (v.9.32)

OMIM

AD=Autosomal Dominant; AR=Autosomal Recessive; SMu= Somatic Mutation

Significance defined as SAINT > or = 0.7

Significant with all alleles

Significant only in TLK2-WT

Significant in TLK2-WT and TLK2-D551G

Significant in TLK2-D551G and TLK2-S617L

Significant only in TLK2-D551G

Significant only in TLK2-S617L

List of all SFARI genes in the proximal interactome of wild-type and mutated TLK2

Gene	SFARI score	Disease	OMIM ID	Molecular function
BRD4	3	-	608749	BRD4 encodes a chromatin reader protein that recognizes and binds acetylated histones and plays a key role in transmission of epigenetic memory across cell divisions and transcription regulation. The protein encoded by the BRD4 gene remains associated with acetylated chromatin throughout the entire
CCNK	S		603544	The protein encoded by this gene is a member of the transcription cyclin family. These cyclins may regulate transcription through their association with and activation of cyclin-dependent kinases (CDK) that phosphorylate the C-terminal domain (CTD) of the large subunit of RNA polymerase II. This gene
CHD7	1	CHARGE syndrome Hypogonadotropic hypogonadism 5 with or without anosmia	214800 612370	This gene encodes a protein that contains several helicase family domains. Mutations in this gene have been found in some patients with the CHARGE syndrome.
CHD8	1	{Autism, susceptibility to, 18}	615032	This gene encodes a DNA helicase that functions as a transcription repressor by remodeling chromatin structure. It binds beta-catenin and negatively regulates Wnt signaling pathway, which plays a pivotal role in vertebrate early development and morphogenesis. Alternatively spliced transcript variants
JMJD1C	3	-	604503	histone demethylase, tumor suppressor
MSANTD2	3	-	-	This gene encodes a protein of unknown function.
NACC1	1	Neurodevelopmental disorder with epilepsy, cataracts, feeding difficulties, and delayed brain myelination	617393	This gene encodes a member of the BTB/POZ protein family. BTB/POZ proteins are involved in several cellular processes including proliferation, apoptosis and transcription regulation. The encoded protein is a transcriptional repressor that plays a role in stem cell self-renewal and pluripotency maintenance.
NFIA	3	Brain malformations with or without urinary tract defects	613735	This gene encodes a member of the NF1 (nuclear factor 1) family of transcription factors. Multiple transcript variants encoding different isoforms have been found for this gene. This protein recognizes and binds the palindromic sequence 5'-TTGGCNNNNNGCCAA-3' present in viral and cellular promoters
NFIX	S	Marshall-Smith syndrome, Sotos 2/Malan syndrome	602535	The protein encoded by this gene is a transcription factor that binds the palindromic sequence 5'-TTGGCNNNNNGCCAA-3 in viral and cellular promoters.
PAPOLG	3	-	-	This gene encodes a member of the poly(A) polymerase family which catalyzes template-independent extension of the 3' end of a DNA/RNA strand. Responsible for the post-transcriptional adenylation of the 3'-terminal of mRNA precursors and several small RNAs including signal recognition particle
YEATS2	3	?Epilepsy, myoclonic, familial adult, 4	615127	YEATS2 is a scaffolding subunit of the ADA2A (TADA2A; MIM 602276)-containing (ATAC) histone acetyltransferase complex.