Published in final edited form as:

Nat Med. 2016 December ; 22(12): 1421–1427. doi:10.1038/nm.4220.

Capturing the biology of mild versus severe disease in a pluripotent stem cell-based model of Familial Dysautonomia

Nadja Zeltner^{1,2,*}, Faranak Fattahi^{1,2,3}, Nicole C. Dubois⁴, Nathalie Saurat^{1,2}, Fabien Lafaille⁵, Lei Shang⁵, Bastian Zimmer^{1,2}, Jason Tchieu^{1,2}, Mohamed A. Soliman^{1,2,6}, Gabsang Lee⁷, Jean-Laurent Casanova⁵, and Lorenz Studer^{1,2,*}

¹Developmental Biology Program, Sloan Kettering Institute, New York, USA

²Center for Stem Cell Biology, Sloan Kettering Institute, New York, USA

³Cornell University Graduate School, New York, USA

⁴Developmental and Regenerative Biology, Icahn School of Medicine at Mount Sinai, New York, USA

⁵St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller University, New York, USA

⁶Weill Cornell Medical College, New York, USA

⁷Institute for Cell Engineering, Johns Hopkins University, Baltimore, USA

Abstract

Familial Dysautonomia (FD) is a debilitating disorder that affects derivatives of the neural crest (NC). For unknown reasons, FD patients show marked differences in disease severity despite carrying the identical, homozygous *IKBKAP* point mutation. Here, we present disease related phenotypes in pluripotent stem cells (PSCs) that capture FD severity. Cells from severe but not mild patients exhibit impaired specification of NC derivatives including autonomic and sensory neurons. In contrast, both severe and mild FD cells show defects in peripheral neuron survival, indicating neurodegeneration as culprit for mild FD. While genetic repair of the FD mutation reversed early developmental NC defects, sensory neuron specification was not restored indicating

Corresponding author: Dr. Lorenz Studer, The Center for Stem Cell Biology, Developmental Biology Program, Memorial Sloan Kettering Cancer Center, 1275 York Ave, Box 256, New York, NY 10065, Phone 212-639-6126, studerl@mskcc.org. Dr. Nadja Zeltner, The Center for Stem Cell Biology, Developmental Biology Program, Memorial Sloan Kettering Cancer Center, 1275 York Ave, Box 256, New York, NY 10065, Phone 212-639-8510, zeltner@mskcc.org.

Accession codes

FASTQ files for the first round of RNA sequencing are deposited at the NCBI Gene Expression Omnibus (with the accession number GSE67073 for first round and GSE77513 for second round of RNA sequencing). Whole exome sequencing files are deposited under BioProject accession number: SRP070100.

Author Contributions

N.Z.: Design and conception of the study, writing of manuscript, cell maintenance, reprogramming, GO term analysis, directed differentiation and survival assays, protocol optimization, gene targeting of PSCs, cellular/molecular assays. F.F.: AN differentiation protocol establishment, execution and survival assays in ANs. J.T.: RNAseq data analysis. N.D.: Design and execution of cardiac mesoderm differentiations. B.Z.: Data quantification of scratch assay. N.S/M.A.S.: Western blot. G.L.: Reprogramming of S3 FD PSCs, mentoring. J-L.C., L.S. and F.L.: Advice and analysis of whole exome sequencing. L.S.: Design and conception of the study, data interpretation, writing of manuscript, mentoring.

Competing Financial Interest Statement

The authors have no conflicting interests to disclose.

that other factors may contribute to disease severity. Whole exome sequencing identified candidate modifier genes for severe FD. Our study demonstrates that PSC-based modeling is sensitive in recapitulating disease severity, which presents an important step on the road towards personalized medicine.

Introduction

In vitro disease modeling using embryonic or induced pluripotent stem cells (PSCs) has shown considerable potential for applications in drug discovery as well as for the elucidation of disease mechanisms^{1,2}. Human PSC-based models can be particularly powerful where suitable animal models are not available or where access to affected primary tissues is not feasible. Furthermore, the technology promises a personalized approach to the biology and treatment of each individual patient by using cells that carry an exact copy of the patient's own DNA. While there has been considerable success in capturing disease phenotypes *in vitro* using the human PSC technology³, the ability to recapitulate disease severity in a defined patient population has not been demonstrated yet, in particular for individuals carrying an identical genetic defect such as in the case of Familial Dysautonomia (FD).

FD was one of the first genetic disorders to be faithfully modeled in PSC-derived lineages *in vitro*⁴. FD, also known as Riley-Day Syndrome or hereditary sensory and autonomic neuropathy III (HSAN-III), is a genetic disease affecting the development, function and survival of various neural crest (NC)-derived lineages such as autonomic (AN) and peripheral sensory neurons (SN)⁵⁻⁷. FD is an autosomal recessive disease and occurs almost exclusively in the Ashkenazi-Jewish population with a carrier frequency of 1 in 32⁸. From birth on, FD patients display a broad range of symptoms that include difficulties in swallowing/suckling, absence of fungiform papillae on the tongue, kyphoscoliosis, lack of tears, difficulties in regulating body temperature and blood pressure, reduced sensitivity to pain⁹ and recurring vomiting crisis (dysautonomic crisis). Autopsy data show evidence of severely reduced cell numbers in the dorsal root ganglion (DRG)⁹ and superior cervical ganglion (SCG)¹⁰. The clinical progression of FD patients indicates a degenerative component to the disease¹¹ in addition to any developmental defect. Nearly all (99.5%) FD patients carry the identical homozygous T>C point mutation in the splice donor (in intron 20) of the I- κ -B kinase complex-associated protein (*IKBKAP*) gene, leading to partial skipping of exon 20 in the mutant IKAP protein and a reduction of wild type protein. IKAP (ELP1) is a component of the transcriptional elongator complex and is ubiquitously expressed. It is not known why its impairment leads to a NC tissue-specific defect. The plant hormone kinetin¹², a food supplement, was shown to increase splicing of wild type *IKBKAP* transcript in primary and PSC-derived NC cells⁴ and is currently being tested in human FD patients¹³. FD was also the target of the first large-scale chemical screen in a PSC-based disease model that led to the identification of additional FD drug candidates including the α_2 -adrenergic receptor antagonist SKF-86466¹⁴. SKF-86466 enhances *IKBKAP* mRNA levels and IKAP protein expression in PSC-derived NC cells independent of splicing¹⁴.

While nearly all FD patients carry the same homozygous *IKBKAP* mutation, the severity of disease symptoms can vary. Here, we use FD as a platform to address whether PSC-

technology is sufficiently sensitive to predict disease severity among individual patients carrying the identical mutation in *IKBKAP*.

Results

Modeling mild and severe FD in neural crest progenitor cells *in vitro*

To model disease severity, we obtained fibroblasts from three patients each with severe (S) or mild (M) FD symptoms, from one heterozygous carrier with no obvious symptoms (A) and from one healthy control (C) subject. Supplementary Fig. 1a and Supplementary table 1 summarize available information used to classify severe versus mild patients, including decreased pain perception and frequency of dysautonomic crisis. We did not match sex and age as there are no known gender-specific defects and age-related cellular features are thought to be largely reset after reprogramming¹⁵.

We first characterized genotype (Supplementary Fig. 1b) and growth characteristics (Supplementary Fig. 1c) in the fibroblasts, but did not find any differences that discriminate between disease severity at the primary fibroblast stage. We next reprogrammed fibroblasts from two patients per severity group (S2, S3, M1, M2, A1, C1) into PSCs using non-integrating, cytoplasmic RNA Sendai-virus vectors expressing *OCT4*, *SOX2*, *KLF4* and *cMYC*¹⁶ (Supplementary Fig. 2a). At least two PSC clones per patient were reprogrammed and characterized. However, due to low inter-patient variability (**data not shown**), the majority of subsequent phenotypes were carried out with one clone per patient. PSC lines were characterized for pluripotency (Supplementary Fig. 2b, c) and splicing of *IKBKAP* (Supplementary Fig. 2d), as well as *in vitro* differentiation potential towards cell types representing the three germ layers that are unaffected in FD (Supplementary Fig. 3a, b).

To address whether the FD PSC-model can capture differences between mild and severe FD, we first revisited three previously described phenotypes (*IKBKAP* missplicing, deficit in *ASCL1* expression levels and migration defects)⁴ at the NC stage. We initially followed the protocol used in our past FD studies based on delamination and fluorescence activated cell sorting (FACS)-based enrichment of NC precursors from early neural rosette structures^{17,18} (termed here rosette-NC (rNC)). Supplementary Fig. 4a–c shows that all PSC lines generate neural rosettes and rNC cells at comparable efficiencies, with no differences in division rates as previously reported^{4,14}.

To address whether cells at the NC progenitor stage can capture aspects of disease severity we first measured *IKBKAP* splicing using qRT-PCR. We found that both severe and mild rNC progenitors express a low ratio of wild type to mutant *IKBKAP* transcripts in contrast to control cells (Fig. 1a). Thus, neither the nature of the *IKBKAP* mutation nor the extent of alternative splicing can explain mild versus severe forms of the disease. We next investigated the functionality of rNC cells. We observed that rNC cells derived from severe FD patients did not migrate efficiently in a scratch assay⁴. In contrast, cells from mild patients migrated as efficiently as control cells (Fig. 1b). To control for potentially differential proliferation, cells that migrated into the scratch were normalized to the number of total cells in the picture. Furthermore, while rNC cells from severe FD patients expressed low levels of *ASCL1*, a marker for autonomic neuron progenitors, rNC cells from mild FD showed

ASCL1 levels comparable to those in unaffected controls both at the mRNA (Fig. 1c) and the protein (Fig. 1d) levels.

To further assess global changes in gene expression between mild versus severe FD we performed two RNA sequencing studies. First, we compared PSCs and FACS-purified rNC from one biological replicate of each severe, mild and control. Unsupervised global clustering as well as GO term analysis showed that at the rNC stage mild FD samples were more similar to control and distinct from severe FD samples in contrast to cells at the PSC stage where mild and severe FD were distinct from control cells (Supplementary Fig. 5a,b). These results were confirmed in the second RNA sequencing study conducted with 2 or 3 biological replicates (Supplementary Fig. 6a). GO term analysis revealed changes in transcripts associated with extracellular region and cell adhesion terms as enriched in severe versus mild FD rNC and neurogenesis terms enriched in mild samples (Supplementary Fig. 6b, c). However, none of the top genes differentially regulated between mild and severe rNC pointed towards an obvious molecular explanation of disease severity (Supplementary Fig. 6d).

Modeling mild and severe FD in neural crest progeny *in vitro*

Next, we asked if the differentiation disparity between mild and severe FD persist in NC derivatives, i.e. in cells relevant for patients post embryonic development. FD autopsy data revealed that sensory neurons (SNs) in the DRG and autonomic neurons (ANs) in sympathetic ganglia are reduced to 20% and 34% of normal subjects, respectively^{9,10} and thus are key targets for therapeutic interventions. In fact, severe FD patients studied here have a pronounced decrease in pain sensation and report frequent vomiting episodes, likely a result of a deregulation of the autonomic nervous system. These symptoms are less commonly observed in mild FD patients (Supplementary table 1). Thus, we investigated the differentiation capacity of severe and mild FD lines into those NC derivative lineages.

We derived FD SNs employing a differentiation protocol¹⁹ that yielded 33% BRN3A⁺ SNs (Supplementary Fig. 7a). By day 12 of differentiation it was obvious that lines from severe FD patients were deficient in making SNs. In contrast, both mild FD and control cultures yielded a dense network of neuronal processes (Supplementary Fig. 7b). Immuno-staining and quantification of SNs, marked by BRN3A⁺/TUJ1⁺ at day 14, confirmed these results (Fig. 2a and b). A temporal expression analysis of *SOX10* established the NC origin of SNs (Supplementary Fig. 7c). One potential reason for the inability of severe lines to generate SNs would be a lower splicing ratio of *IKBKAP*²⁰. However, *IKBKAP* ratios were low in both severe and mild SNs (Fig. 2c). Thus, *IKBKAP* expression alone did not explain the SN phenotype or the differences between mild and severe FD.

To address FD phenotypes in autonomic-like neurons (ANs) we first established a differentiation protocol (Supplementary Fig. 7d), as no reliable published reports were available. NC cells are induced by activating WNT signaling at day 2 of differentiation using a modified dual-SMAD inhibition protocol^{19,21}. At day 11 the cells were FACS isolated based on the expression of CD49D ($\alpha 4$ -integrin receptor), which marks *SOX10*⁺ NC cells²² (Supplementary Fig. 7e). Gene expression analysis by qRT-PCR at day 30, 40, 50 and 60 and immuno-staining showed strong up-regulation of sympathetic AN markers, such as

ASCL1, *PHOX2A*, *SCG10*, *TH* and *DBH*,²³ while markers of related cell types, such as parasympathetic ANs (*CHAT*²³), enteric neurons (*EDNRB*²⁴) and sensory placode lineages (*Six1*²⁵) were only moderately induced (Supplementary Fig. 8a,b). However, analysis of mature markers such as autonomic ganglia-specific acetylcholine receptor *CHRNA3*, *CHRNA4* and vesicular monoamine transporter *VMAT1*, *VMAT2* was largely negative (Supplementary Fig. 8d), and we did not detect noradrenaline release by HPLC (**data not shown**), suggesting an immature AN-like phenotype. A recent study showed that only upon co-culture with mouse cardiomyocytes, HES-derived sympathetic neurons express such late markers, further supporting our immaturity hypothesis²⁶. When we applied the AN-like protocol to FD-PSC lines, we found that cells from severe FD patients could not effectively generate the CD49D⁺ NC, in contrast to control and mild FD cells (Fig. 2d). In fact, severe FD *SOX10*⁺ NC cells survive poorly while mild FD and control *SOX10*⁺ NC cells generated AN-like cells (*PHOX2A*⁺ and *TUJ1*⁺) by day 25 (Supplementary Fig. 9a) at efficiencies of 10–25% of total cells (*TH*⁺ and *ASCL1*⁺, Supplementary Fig. 9b). The non-AN-like cell fraction was identified as mostly non-proliferating (*Ki67*[−]), myofibroblast-like cells (α SMA⁺) with a fraction of SN contaminants (*BRN3A*⁺, *ISL1*⁺, Supplementary Fig. 9c, d).

Neurodegeneration in mild and severe FD cells accounts for disease symptoms and offers a therapeutic target

We next aimed to define cellular phenotypes that can discriminate mild from controls and explain why mild FD patients are sick at all. FD is a progressive disease as illustrated by the deterioration in pain perception with increasing age^{6,11}. Therefore, we tested whether mild FD patients exhibit primarily neurodegenerative rather than developmental phenotypes. We replated differentiated FD SNs on day 12 at equal densities (Fig. 3a, first panel) and measured survival as the number of *BRN3A*⁺/*TUJ1*⁺ SNs every 3 days. Control SNs remained healthy. In contrast SN numbers from two mild FD patients significantly diminished. Similarly, the few severe SNs progressively died (Fig. 3a and b). We further found that cleaved *Caspase3* staining increased, paralleled by loss of *BRN3A*⁺ cells in FD but not control cells (Supplementary Fig. 10a). To investigate potentially contaminating cells in the cultures, we stained for progenitor (*SOX10*, *SOX2*) and proliferation (*Ki67*) markers, which were largely negative. Most non-SNs in the cultures were α SMA⁺ non-proliferating, myofibroblast-like cells (Supplementary Fig. 10b). Finally, we assessed the survival of mild FD AN-like cells over time. Surprisingly, mild FD AN-like cells did not show any defect in survival for up to 50 days based on *TH*⁺ and *ASCL1*⁺ cells. Even when the cells were stressed by growth factor removal or adding toxin (carbonyl cyanide 3-chlorophenylhydrazone CCCP), neither mild FD nor control AN-like cells died (Supplementary Fig. 11a and b). Future studies may determine whether the failure to detect degenerative phenotypes is due to the immaturity of AN-like cells (Supplementary Fig. 8d and 11a). Together, this data suggests degeneration as the primary culprit for mild FD while both developmental and degenerative defects characterize severe FD.

We have previously identified positive effects of kinetin¹² and SKF-86466 (α_2 -adrenergic receptors antagonist) on *IKBKAP* splicing in FD rNC cells¹⁴. However, rNC phenotypes represent a neurodevelopmental paradigm that may not translate to FD patients who no longer harbor significant numbers of NC-equivalent progenitors. Thus, we asked if

neurodegeneration can be halted in mild FD SNs. After verification of expression of α_2 -adrenergic receptors (Supplementary Fig. 12a), SNs were differentiated and treated with kinetin, SKF-86466 or vehicle (DMSO), starting from day 12 and survival was assessed. Fig. 3c shows that SNs from mild FD, treated with either drug survived, while DMSO-treated cells progressively died. This rescue was correlated with an increase in *IKBKAP* splicing transcripts in kinetin but not SKF-86466 treated cells (Supplementary Fig. 12b). These results indicate potential treatment options for patients after birth and support efforts to bring kinetin into the clinic²⁷.

Genetic repair of the FD mutation yields insight into the molecular mechanism underlying disease severity

Recent evidence in FD animal models point to the importance of *ikbkap* expression levels. Conditional knockout mouse models suggest that phenotype severity is a direct result of *ikbkap* levels^{28,29}. To determine the role of *IKBKAP* expression on disease severity in human FD patient samples, we first looked at the effect on *IKBKAP* splicing. We observed that the ratio of wild type to mutant *IKBKAP* expression is equally low in severe and mild PSCs (Supplementary Fig. 2d), rNC cells (Fig. 1a), SNs (Fig. 2c), primary fibroblasts and PSC-derived endoderm, cardiomyocytes and cortical neurons (Supplementary Fig. 13a).

At the protein level in PSCs both severe and mild FD showed low IKAP levels compared to controls (Supplementary Fig. 13b, **left**). However, at the rNC stage, mild and control lines expressed comparable IKAP protein levels, while severe patients showed lower levels (Supplementary Fig. 13b **right**; note, wild type IKAP is expressed in two non-disease associated splice variants (150 kDa and 111 kDa)). These results match previous *IKBKAP* splicing (Supplementary Fig. 2d, Fig. 1a) and RNA sequencing data (Supplementary Fig. 5 and 6), where at the PSC stage mild and severe cells show comparable expression levels while at the rNC stage mild cells show levels comparable to control cells distinct from severe cells. Thus, it is conceivable that increased IKAP protein levels in rNC could contribute to protecting mild FD cells from the early developmental phenotypes characteristic for severe FD cells.

To confirm that the inability of severe FD cells to generate NC precursors and derivatives is indeed specific to the *IKBKAP* mutation, we used the CRISPR/Cas9 system^{30,31} to repair of the FD-mutation. We identified 3 clones that showed heterozygous repair (targeting frequency of 10% (Supplementary Fig. 14a–c)). No Cas9 off-target cutting was detected at the top 10 predicted sites (Supplementary Fig. 15a, b). To compare those gene-corrected clones to an FD heterozygous line, we established a PSC line from the mother of S2, a heterozygous FD carrier (A1). Analysis of *IKBKAP* splicing in gene-targeted clones revealed elevated splicing ratios compared to the parental severe (S2) line and were comparable to splicing in the heterozygous FD carrier (A1, Fig. 4a,b). Analysis of IKAP protein levels yielded similar results (Supplementary Fig. 14d, e). Heterozygous FD carriers do not show FD disease symptoms⁵, so we tested whether our *in vitro* FD phenotypes reverted in gene corrected cells. Importantly, both the *ASCL1* phenotype in rNC cells (Fig. 4c) and the CD49D phenotype during early NC induction (Fig. 4d) were rescued in gene-corrected clones either to wild type or FD carrier (A1) levels. These results confirm that the

IKBKAP point mutation is responsible for the early developmental NC disease phenotypes in severe FD. Next, we investigated if genetic rescue in *IKBKAP* reverts differentiation defects of NC derivatives, i.e. SNs. Interestingly however, SN differentiation was not restored in any of the three genetic rescue lines as compared to control or the A1 heterozygous carrier lines (Fig. 4e). This data suggests that FD phenotypes downstream of the NC are not solely caused by the *IKBKAP* mutation.

Candidate modifier mutations present in severe but not mild FD may contribute to disease severity

To identify candidate modifier genes in severe FD patients, we performed whole exome sequencing in three severe (S1, S2, S3) and three mild (M1, M2, M3) FD patients. We searched for genes that display non-synonymous coding variants (not necessarily the same variants) in all 3 severe patients following a model of complete penetrance given the small sample size. We focused on rare variants, with a minor allele frequency (MAF) of less than 1% in public databases (<1% KG in www.1000genomes.org and www.exac.broadinstitute.org). We further ranked the candidate genes based on the predicted impact of each variant by variant-level metric CADD (Combined Annotation Dependent Depletion: CADD^{32,33}) and on the number of damaging mutations found in healthy individuals for each gene by gene-level metric GDI (Gene Damaging Index: GDI³⁴). We also took into account the mutation significance cutoff (MSC), a gene-level specific cutoff for CADD scores³⁵. First, testing an autosomal dominant mode of inheritance, we identified 7 candidate modifier genes harboring homozygous variations common in all severe FD patients. However, the same variations can also be found in the mild patients (Fig. 5a, Supplementary Fig. 16, Supplementary table 2, sheet 1). Second, we tested an autosomal recessive mode of inheritance and identified 23 candidate genes. 20 of these genes were also mutated in mild patients and thus filtered out. 3 candidate genes remained: *LAMB4*, *KIAA1211*, and *FAT2*. Within the severe patients, these 3 genes carried variations predicted to be damaging (CADD score > MSC score³⁵). With the lowest GDI score *LAMB4* was our top candidate (Fig. 5a, Supplementary Fig. 16 and Supplementary table 2, sheet 2). In a separate analysis, based on incomplete penetrance but taking into account close functional relatedness to *IKBKAP*, as revealed by the gene's connectome³⁶, few additional candidates were identified from the initial list of affected genes including *URH1*, *LCA10*, *MSH3*, *YEATS2*.

Laminin β 4 (*LAMB4*) harbors one missense variant in patient S1 and another shared by S2 and S3 (Supplementary table 2, sheet 2, Supplementary Fig. 17). Laminins are secreted heterotrimeric (α , β , γ -chains) proteins of the extracellular matrix. They greatly influence the differentiation, migration and adhesion of cells within tissues and have important roles in neuronal development, such as axonal growth. To date no laminin trimer, containing *LAMB4* has been identified and its function is largely unknown. This gene is expressed highest in the SCG and DRG (<http://biogps.org>), thus it may play a role in the development of tissues particularly affected in FD patients and that show clear phenotypic differences between mild and severe FD in our model. We used our RNA sequencing data conducted in rNC cells to investigate whether the three mutations lead to differential gene expression comparing severe, mild and control rNC. Only *FAT2* was significantly higher expressed in

severe compared to mild cells, particularly in exon 9, where severe patient 2 has a mutation (Supplementary Fig. 17), indicating that the candidate mechanisms of action will involve in-depth molecular and functional studies. Experiments from this study are summarized in Supplementary table 3.

Discussion

FD has been at the forefront of PSC-based disease modeling⁴ and drug discovery¹⁴ efforts. In the current study we use FD as a model to address a next major challenge for PSC research, the ability to capture disease severity, as an important step towards implementing it in personalized medicine. FD is particularly suitable to probe this question since it is caused by the identical point mutation in nearly all FD patients; yet it can present with mild or severe symptoms and a variable slope of progression.

Our study demonstrates that it is possible to recapitulate disease severity *in vitro* and that a strong neurodevelopmental phenotype is restricted to patients with severe FD. In contrast, both severe and mild FD patient cells display neurodegenerative phenotypes as illustrated in the progressive SN loss. We postulate that severe FD constitutes a genetic sub-disease within FD, as some of the secondary mutations we discovered, unique to severe patients, could be the cause of additional and/or more severe symptoms (Fig. 5b). Our work raises the question whether in human FD genetic background differences represent distinct disease subtypes or the contribution of more minor disease modifiers. Future studies will be required to address in more detail the molecular mechanism explaining the difference between mild and severe forms of FD. Such studies may also address whether it is possible to identify phenotypes representing intermediate disease severity at the cellular level, though there are currently no obvious criteria to clinically define FD patients with intermediate disease severity.

Independent of defining a detailed molecular mechanism, it is particularly exciting to report that kinetin and SKF-86466 are capable of inhibiting degeneration in neurons already born. Those results provide an additional impetus to pursue the clinical translation of those compounds in FD.

Finally, our results may catalyze similar efforts of modeling severity in other genetic disorders. One important strategy for future studies is the analysis of much larger cohorts of patient-derived PSC lines for a given disease, without “a priori” clinical information on disease severity and carried out under fully blinded conditions. Such a study may confirm stratification along existing clinical subgroups or yield additional subgroups not captured in clinical datasets. The ability to model severity in a genetically defined disease is an important step in moving PSC research forward towards a new era of personalized medicine that captures the biology of each patient individually.

Online Methods

FD fibroblasts

Fibroblasts were purchased from Coriell, under the catalog numbers S1=GM04959, S2=GM04899, S3=GM04589, M1=GM02341, M2=GM02342, M3=GM02343,

C1=AG02602, C2=GM00316, A1=GM04895, A2=GM04665, A3=GM04897 and cultured in fibroblast medium (DMEM (Gibco, 11965-092)/10% FBS (Atlanta Biologicals, S11150)/0.5% L-Glutamine (Gibco, 25030-081)/0.5% Pen Strep (Gibco, 15140-122)). For genotyping the FD mutation, gDNA was extracted using the DNeasy Blood&Tissue kit (Qiagen, 69506) and a ~500 bp long PCR product (PCR primers: 5' CCATAAGGCTCAAAGCGAAA and 5' TGAGTGTACGATTCTTTATGC) was sequenced (primers: 5' GAAAGTCAAGATAAATATAGAGAAC and 5' TCTGCTAGGAATCTCCACTA). To establish fibroblast growth, each line was seeded at 100,000 cells/6-well and cell numbers were counted every day. For *IKBKAP* splicing, fibroblast's RNA was harvested in TRIzol Reagent (Ambion, 15596018). 1 µg RNA was used to generate cDNA (QuantiTect Reverse Transcription kit, Qiagen, 205313) and SYBR green (Applied Biosystems, 4309155) was used for qRT PCR on an Eppendorf Mastercycler according to the manufacturer's instructions with primers: for wild type *IKBKAP* 5' GCAGCAATCATGTGTCCCA and 5' ACCAGGGCTCGATGATGAA for mutant *IKBKAP* 5' CACAAAGCTTGTATTACAGACT and 5' GAAGGTTTCCACATTTCCTCAAG. All results were normalized to 18S control, primers: 5' GGCCCGTAATTGGAATGAG and 5' GCTATTGGAGCTGGAATTAC. Cycling parameters correspond to those previously used^{6,14}.

Reprogramming, PSC culture and characterization

Fibroblasts were seeded at 150,000 cells per 12-well in 5 wells per line (resulting in 5 PSC sub-clones) and transduced with 4 CytoTune Sendai viruses (Life Technologies, A1378002) encoding Oct-4, Sox2, Klf4 and c-Myc each at an MOI of 3¹⁶. Cells were fed daily with fibroblast medium. At day 7, the cells were dissociated using trypsin-EDTA (Gibco-Life Technologies, 25300-054) and replated onto 10-cm MEF (GlobalStem, GSC-6105M)-coated plates in HES medium supplemented with 0.5 mM VPA (Tocris Bioscience, 2815), as described in Zeltner et al.¹⁸. Cells were fed daily with HES medium/VPA until day 14, then with HES medium only. Between day 20 and 40 PSC colonies were manually picked, expanded and frozen. PSCs were fed daily with HES medium and passaged manually or with trypsin-EDTA¹⁸ approximately once per week onto MEFs. PSC cultures were routinely tested for mycoplasma contamination every other week. For the characterization of PSC lines, the cells were immuno-stained with Oct-4 (mIgG2b, Santa Cruz, sc-5279), Nanog (goat, R&D, AF1997) and Sendai virus (rabbit, MBL International, PD029). *IKBKAP* splicing in PSCs was conducted as described for fibroblasts. The karyotypes of the PSC clones were normal, except at high passages in S3 and M2.

***In vitro* differentiation of PSCs**—HES media, KSR differentiation media, N2 differentiation media, matrigel dish coating and PO/LM/FN dish coating, plating of PSC for differentiation and day 0 to day 11 dual-Smad inhibition differentiation were all done as described in Zeltner et al.¹⁸.

CNS neuron differentiation: For the *in vitro* differentiation to CNS neurons, PSCs were detached using Accutase (Innovative Cell Technologies, AT104) for 20 min. at 37°C, washed twice with 1X PBS and plated onto matrigel (BD, 354234)-coated 6-wells in HES medium supplemented with 10 µM Y-27632 dihydrochloride (Tocris, 1254). When 100%

confluency was reached the cells were fed daily with KSR- or N2-medium supplemented with LDN193189 (Stemgent, 04-0074) and SB431542 (Tocris, 1614)¹⁸. On day 11 they were accutased, washed twice and replated onto Poly-L Ornithine hydrobromide (Sigma, P3655)/Laminin-1 (R&D Systems, 3400-010-01)/Fibronectin (BD Biosciences, 356008)-coated plates in droplets of approximately 100–150,000 cells per 10 μ l as described in Zeltner et al.¹⁸, in N2 medium supplemented with 0.02 μ g/ml BDNF (R&D Systems, 248-BD), 0.2 mM Ascorbic Acid (Sigma, A4034) and 0.1 μ M Purmorphamine (Stemgent, 04-0009). On day 13 the cells were fed with N2 medium/AA/BDNF/Purmorphamine, on day 15, 17 and 19 they were fed with N2 medium/AA/BDNF/ 0.01 mM DAPT (R&D, 2634/50) and on day 20 they were fixed with 4% PFA (Affymetrix, 19943) and immuno-stained with Tbr1 (rabbit, Millipore, AB10554) and Tuj1 (mIgG2a, Covance, MMS-435P-250).

Endoderm differentiation: PSCs were plated as described previously¹⁸. The cells were fed daily with RPMI medium containing Glutamax (Invitrogen, 61870-036) supplemented with 0.5% FBS and 100 μ g/ml ActivinA (R&D, 338-AC), fixed on day 5 and immuno-stained for Sox17 (goat, R&D, af1924) and DAPI.

Mesoderm/cardiac differentiation: Cardiac differentiations were done as described previously^{37,38} with the following modifications: RPMI Medium 1640 (Life Technologies, 11875-093)/B27 (Life Technologies, 17504044) media was used instead of StemPro34. Embryoid bodies (EBs) were treated with 10 ng/ml ActivinA (R&D Systems, 338-AC) and 10 ng/ml of BMP4 (R&D Systems, 314-BP) from day 1–3, followed by 10 μ M of XAV939 (Stemgent, 04-0046) from day 3–5. Differentiation cultures were assessed for mesoderm formation by flow cytometry staining for KDR/VEGFR2 (Alexa Fluor 647 anti-human CD309, BioLegend, 7D4-6) and PDGFRA (PE anti-human CD140a, BioLegend, 16A-1). The differentiations were carried on until day 20 in RPMI Medium 1640 (Life Technologies, 11875-093)/B27 (Life Technologies, 17504044) media and stained for the anti-cardiac isoform of TroponinT (cTnT) (Thermo Scientific Lab Vision, clone 13–11) by flow cytometry as well as plated onto plastic and stained for cTnT via IF on day 30 of differentiation.

For the quantification of neurons, endoderm and cardiac cells, images from two wells of each of 3–5 differentiations were counted for Tbr1⁺/Tuj1⁺ cells or Sox17⁺ or TroponinT⁺, respectively and the percentage to total DAPI⁺ cells was reported.

rNC differentiation: rNC differentiations, FACS isolation, HNK1, Pax6, α AP2a, ASCL1 and Tuj1 staining were carried out as described in Zeltner et al.¹⁸. To measure the proliferation rate, FACS isolated rNC cells were plated at 30,000 cells/96-well¹⁸. 24 hrs. later they were fed with N2/FGF2/EGF media containing EdU using the Click-iT Plus EdU Alexa Fluor 555 Imaging kit (Life technologies, C10638). 48 hrs. later the cells were fixed and stained according to the manufacturer's instructions and EdU⁺/DAPI⁺ cells were counted for the quantification. The *IKBKAP* splicing PCR was conducted as described for fibroblasts. The migration (scratch assay) was conducted as described in Zeltner et al.¹⁸. In short, FACS isolated rNC cells were plated in PO/LM/FN-coated 96-wells at 20,000–60,000 cells/well and scratched 24 hrs. later using a 200 μ l pipet tip. 6–8 wells were fed with media

containing 10 ng/ml Hoechst33342 trihydrochloride, trihydrate (Invitrogen, H3570) and reference pictures (0 hrs. = before migration) were taken immediately. The media in the other wells was changed to remove dead cells. The cells were allowed to migrate for 48 hrs. and stained with Hoechst33342 before pictures were taken. The number of migrated cells was quantified automatically by the Konstanz information miner (KNIME) workflow³⁹. Results are expressed as percentage of cells migrated into a ROI (Region of interest) of the total number of cells per picture. Ascl1 qRT PCR was conducted on RNA isolated from FACS sorted rNC cells on day 18 of differentiation. Taqman gene expression assay (ABI, 4304437) for ASCL1 (primers: ABI, Hs00269932_m1 (ASCL1) and 4326321E (HPRT)) was conducted on an Eppendorf Mastercycler machine. Fold-changes were calculated using the comparative Ct method (Ct method) and then normalized to the expression in undifferentiated HES cells. Immuno-staining of rNC cells was conducted one week post FACS isolation as described in Zeltner et al.¹⁸.

Peripheral sensory neuron differentiation: The SN differentiations were carried out as described in Chambers et al.,¹⁹ with few modifications. In short, PSC were plated on matrigel coated dishes and differentiations were induced once 100% confluency was reached, using a gradient of KSR and N2 differentiation media, supplemented with LDN193189 and SB431542¹⁸. From day 2 to 11, 1.5 μ M CHIR99021 (GSK3 β -inhibitor, activation of Wnt signaling; Stemgent, 04-0004), 10 μ M SU5402 (FGF inhibition; BioVision, 1645-1) and 10 μ M DAPT (Notch inhibition; Tocris, 2634) were added to the media. On day 12, the cells were detached using Accutase, washed twice and replated into PO/LM/FN coated 96-wells at 280,000 cells per well in N2 differentiation media supplemented with 50X B-27 supplement (Life technologies, 12587-010), 20 ng/ml BDNF (R&D Systems, 248-BD), 20 ng/ml GDNF (Peprotech 450-10), 25 ng/ml β -NGF (PeproTech, 450-01), 100 μ g/ml Primocin (InvivoGen, ant-pm-2), 0.6 μ g/ml Laminin-1 (R&D Systems, 3400-010-1), 0.6 μ g/ml Fibronectin (Fisher Scientific, CB40008A-BD) and 10 μ M Y-27632 (Tocris, 1254). For maintenance, the cells were fed with the same media without Y-27632 every 2–3 days. On day 14, 17, 20, 23 and 26 the cells were fixed in 4% PFA and stained with Brn3a (mIgG1, 1:100, Millipore, MAB1585), Tuj1 (rabbit, 1:1000, Covance, MRB-435P-100) and DAPI (1:1000, Sigma-Aldrich, D9542-5M), cleaved Caspase3 (rabbit, 1:100, Cell Signaling, 9661S), Sox10 (goat, 1:100, Santa Cruz, sc-17342), Ki67 (rabbit, 1:500, SP6, RM9106-S1), Sox2 (rabbit, 1:400, Cell Signaling, 3579P) and α SMA (mIgG2a, 1:1000, Sigma, A5228-200UL). For the quantification, Brn3a⁺/Tuj1⁺ cells in pictures taken from 2–3 wells from 3–9 independent differentiations were counted (Figure 2b) or the percentage of DAPI⁺ of Brn3a⁺ cells were quantified using the Meta Express software (Figure 5b). *IKBKAP* splicing was done as described for fibroblasts. For the Sox10 expression time course, cells were harvested for RNA extraction at day 4, 8, 12 and 16 without replating. The real-time RT PCR was conducted using SYBR SsoFast EvaGreen Supermix (BioRad, 172–5202) and the QuantiTect Primer Assay for Sox10 (Qiagen, QT00055405). The α_2 adrenergic receptor PCR was previously described⁴⁰. Drug treatment of SNs for the rescue of FD phenotypes were done either from day 0 throughout the entire differentiation or from replating at day 12. Kinetin (Sigma-Aldrich, K0753-5G) and SKF86466 hydrochloride (Tocris, 3866) were used at 10 μ M.

Autonomic neuron differentiation: PSCs were plated and dual-Smad inhibition differentiation until day 11 was conducted as reported in Zeltner et al.¹⁸. For Wnt-based NC induction the culture was supplemented with 3 μ M CHIR99021 from day 2 to 14 and the cells were sorted for CD49d (1:800, PE-Cy7 anti-human CD49d, BioLegend, 304314) on day 11. The cells were aggregated in Ultra Low Attachment 24-well culture plates (Fisher Scientific, 3471) as 3D spheroids (500,000 cells/well) in Neurobasal (NB, Life Technologies, 21103-049) media supplemented with N2 supplement (Stem Cell Technologies, 07156) and B27, 3 μ M CHIR99021 and 10 ng FGF2 (R&D Systems, 233-FB-001MG/CF). On day 14, the spheroids were plated onto PO/LM/FN coated 24-wells in NB+N2 supplement+B27 containing GDNF and Ascorbic Acid (Sigma, A8960-5G). On day 25, cells were fixed in 4% PFA and stained for Phox2a (rabbit, 1:800, kindly gifted from Dr. J-F Brunet, France), Tuj1 and DAPI. For the AN marker expression, RNA was extracted at day 30, 40, 50 and 60 and real-time RT-PCR was conducted either using SYBR SsoFast EvaGreen Supermix (QuantiTect primer assays for ASCL1: QT0023775, Phox2a: QT00215467, ChAT: QT00029624, EDNRB: QT00014343, Six1: QT00010584, CRNA3: QT00015386, CRNB4: QT00010570, VMAT1: QT00028420, VMAT2: QT00059857) or Taqman gene expression assays (primers for Scg10: Hs00199796_m1, TH: Hs00165941_m1, DBH: Hs01089840_m1). For the stress conditions in the AN survival assay, the cells were maintained either in the above mentioned media or in NB only or in NB only supplemented with 0.1 μ M Carbonyl cyanide 3-chlorophenylhydrazone (CCCP, Sigma, C2759-100MG). For intracellular FACS analysis, the cells were dissociated using Accutase, then fixed and permeabilized using BD Cytofix/Cytoperm (BD Bioscience, 554723) according to the manufacturer's instructions. Primary antibody staining for ASCL1 (mIgG1, 1:200, BD Pharmingen, 556604) and TH (rabbit, 1:700, Pel-Freez, P40101-0) was done at 4 °C over night and secondary antibody staining was done at RT for 30 min. using AlexaFluor 647 goat anti mouse IgG1 (Invitrogen, A21240) for ASCL1 and AlexaFluor 647 goat anti rabbit (Invitrogen, A21245) for TH. The analysis was done on a flow Cytometer and using the FlowJo software. Intracellular FACS and immunohistochemistry analysis was done as described above. ISL1 (mouse IgG2b, 1:200, DSHB, 39.4D5-c).

Genetic rescue

Strategy: The 3' homology arm spans from intron18 to intron 20 of the *IKBKAP* gene. The FD point mutation is located in intron 20, 5bps downstream of exon 20. The gRNA binding site is 38bps upstream of the FD mutation. The selection cassette is located in intron 20, thus upon correction of the splicing site (FD mutation) it will be removed from the protein. The 3' homology arm spans intron 20.

Cloning of *IKBKAP* donor plasmid: The 3' and 5' homology arms are 697 bps and 682 bps long, respectively and were amplified from C1.2 PSC genomic DNA using primers 5' CCCAGGGAGATTGTACACAC and 5' CGACTCTAGAGGATCCCAATGCCAAGCT or 5' CGGTACCCGGGGATCAGTGGGAGAGCTAAGGCACA and 5' TATGGATCCGCTAGCTTCACAGGCCACCTAAAACC, respectively. The selection cassette consists of the human PGK promoter followed by the puromycin resistance gene. The three fragments were ligated into a linearized pUC19 vector backbone using the Gibson

Assembly Master Mix (New England Biolabs, E2611S) and the plasmid insert was sequenced entirely.

gRNA design: The CRISPR/Cas9 technology³¹ was used to design the gRNA and induce the double strand break to enhance homologous recombination to repair the FD mutation in *IKBKAP*. Option A from the hCRISPR gRNA synthesis protocol (deposited www.addgene.org/41824) was used to generate the gRNA. The website http://www.genome-engineering.org/crispr/?page_id=41 was used to generate the gRNA target sequence as well as to check for potential off-target effects. The gRNA with the best score was chosen (shown in bold below), synthesized as a gBlock from IDT and cloned into a TOPO vector (Invitrogen, 450245). The target sequence is located 39 bps upstream of the FD mutation.

U6 promoter/**target sequence**/gRNA scaffold/termination signal

TGTACAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCA
AGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCTTCATATTTGCATATACGAT
ACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTA
GTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAA
ATTATGTTTTAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTTCGATTTC
TTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG**AGTTGTTTCATCATCGAGC**
CCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGA
AAAAGTGGCACCGAGTCGGTGCTTTTTTCTAGACCCAGCTTTCTTGACAAAGTT
GGCATTAA

Homologous recombination in severe FD PSCs: 10 million S2 FD PSCs were nucleofected with 10 µg *IKBKAP* donor plasmid, 5 µg gRNA-TOPO vector and 5 µg Cas9-GFP plasmid (Addgene, plasmid #44719) using solution I and program A23 from the Amaxa Human Stem Cell Nucleofector Kit 1 (Lonza, VPH-5012). The cells were plated on MEFs in half fresh HES media and half HES media conditioned on S2 FD PSCs supplemented with 10 µM Y-27632 dihydrochloride. 48 hours later the cells were detached with Accutase and GFP⁺ cells were isolated by FACS using low pressure/µm nozzle settings. The 122,000 GFP⁺ cells were plated on fresh puromycin-resistant MEFs in a 10 cm dish. 72 hours post sort puromycin selection was initiated (1 µg/ml Millipore, 540411-100MG) and the cells were fed daily until colonies appeared, which were manually picked, expanded and frozen.

Screening: Genomic DNA was extracted from 21 clones (DNeasy Blood & Tissue kit (Qiagen, 69506) and screened using a 3 primer PCR strategy. Primers 5'GCTAGCGGATCCATAACT and 5'TGCCATTTGATTTCAGTCTCC bind on either side of the homology cassette and primer 5'AGGGGGCTGGAAGAGCTA binds inside the selection cassette. The expected amplicon sizes are shown in Fig. 4a. To sequence the FD mutation, a 483 bp amplicon surrounding the FD mutation was PCR amplified and sequenced.

CRISPR/Cas9 off target site sequencing: The top 5 predicted off-target sites in coding and non-coding regions each were Sanger sequenced. For each of the 10 sites, two PCR primer pairs were chosen. The PCR was performed on genomic DNA from S2-iPSCs, S2rescue1-

iPSCs, S2rescue5-iPSCs and S2rescue7-iPSCs. The resulting PCR product was cut from the agarose gel, purified and sequenced with two sequencing primers each. Primers: ID#1: NZ50f-5' CACAGAGGACCTGTGCTCAA, NZ50r-5' CCTGTGAAGTTGTGGGACCT, NZ51f-5' TGCCCACTTTCTCTGTTCCT and NZ51r-5' -CCTGTGAAGTTGTGGGACCT, NZ52fseq-5' CCTCCCCCTTCTATTCCGAAG, NZ53fseq-5' CCTGCAACTGACGCTGTATC. ID#2: NZ54f-5' ATCCCATGCTCTTCCTCCTT, NZ54r-5' GACAATCCAGGCAACCTTGT, NZ55f-5' CCTCCTTCTCTCTGGCCTTT, NZ55r-5' TAGCATTTTGGCATGAGTGG, NZ56fseq-5' CCCTGGGAGCTGTCTGTAAG, NZ57fseq-5' TGGTTTCTCCCTGACACTT. ID#3: N58f-5' TGAACTTTTCCGGACTGACC, NZ58r-5' GGAGACCTGGTTCGTAGTGA, NZ59f-5' CGGACTACGATCCTTTTCCTG, NZ59r-5' TGACAGTCGATGGAGACCTG, NZ60fseq-5' ACTACAGAGCCCGGAATG, NZ61fseq-5' GGGAAAGGTGGGACTCTGTG. ID#4: NZ62f-5' GTGCTGGGGGAGCTCTAAG, NZ62r-5' CTGCCCTGTTCCTTGATGT, NZ63f-5' CTAGAGAAGCTGGGGTGCTG, NZ63r-5' TGTCCCAGAGCTGAGACTCC, NZ64fseq-5' CACCATGTTCTGCACTGAC, NZ65fseq-5' GAGCTCTAAGGGGGCTGTTC. ID#5: NZ66f-5' TGTCTTTCCCTCCCTTCCTT, NZ66r-5' TGTTTATGGAGCCCGGATAC, NZ67f-5' TGTCTTTCCCTCCCTTCCTT, NZ67r-5' ATGAGCCAGACTGCCATCTC, NZ68fseq-5' AGGCAGTTTGTGGTGA, NZ69fseq-5' TGTTTTAGGGGGAAGC. ID#6: NZ70f-5' CAGTGGAAGGAGAGGAGCAG, NZ70r-5' TGGGTGGAAGGTCAAGAAAG, NZ71f-5' CAGTGGAAGGAGAGGAGCAG, NZ71r-5' CCAACACAAGGAGTGGGACT, NZ72fseq-5' CTCTGAGGAAGCCAGTCCAC, NZ73fseq-5' CCCTTACCCAGGAGGTCAC. ID#7: NZ74f-5' GGCTCTCTGCCCAATTCATA, NZ74r-5' TGGGGATTTTTCCCTTTAG, NZ75f-5' AACCTCCATCCCACCTTAC, NZ75r-5' CAGGCACACCCCATTTTAAG, NZ76fseq-5' AGAAAGCTTGCCATCCTTGA, NZ77fseq-5' TTCTGGAAGCAGTAGCACCA. ID#8: NZ78f-5' CAGCTTTGGGTCTATTTGC, NZ78r-5' AATGGCCCCTTGGTGAATAC, NZ79f-5' TCTGGCTGGAATGATGTTTG, NZ79r-5' TCTTGGTCCATCCAAAATGG, NZ80fseq-5' CTGCCTCATGGTGAAGAGGT, NZ81fseq-5' TCTCCTTATCCATCTGCCTCA. ID#17: NZ82f-5' GCAGCTGCAGGATAGGGTAG, NZ82r-5' GACAGCTCCTGCATCAGGTT, NZ83f-5' ATAGGGTAGCCCTCCTCCAC, NZ83r-5' GAGGACAGCTCCTGCATCA, NZ84fseq-5' TACTGCGCTTGCTCAAAAAC, NZ85fseq-5' ACATTCTCGCGAAGGGTCT. ID#22: NZ86f-5' GCCTCAGCCTGTTGTTCTTC, NZ86r-5' CGATCTCGTGGCTGATAACC, NZ87f-5' TGTTCCTTCGTCTGGCCTAC, NZ87r-5' GTCGATCTCGTGGCTGATAA, NZ88fseq-5' GCCTTCCTCTTCTCCATCG, NZ89fseq-5' TCAACAACCTCAACGGCTTC. No mutations were detected in any of the sequences. In ID#3, S2rescue1 and in ID#22, all three clones did not yield PCR products or sequences that would allow the determination of mutations.

Western blot analysis

For protein analysis of neural crest cells, HNK1/p75 double positive cells were isolated by cell sorting and cells were pelleted and snap frozen. Pluripotent stem cells were harvested by scrapping after removal of the MEF layer. To extract protein, the cell pellet was resuspended in 1x RIPA buffer (Sigma # R 0278) containing 1x Halt Protease and Phosphatase inhibitor cocktail (Thermo #87786). In few samples to obtain sufficient protein for analysis, cell pellets from two separate neural crest inductions were combined or FACS isolation was conducted on single positive cells. For western blotting, a total of 10 µg of protein was mixed with 4x Novex sample buffer (Thermo #NP0007) and 10x NuPAGE reducing agent (Thermo #NP0004) to a final concentration of 1x. Lysates were run on 4–12% pre-cast Bis-Tris gels (Novex, NP0322BOX), transferred onto PVDF membranes (Thermo, 88518) and blocked with 5% milk. Antibodies used for western blotting analysis were: β-Actin (Sigma, A2228) and IKAP (GeneTex, GTX110274). Band intensity was measured using NIH ImageJ and normalized to β-actin. For the quantification of WBs in Supplementary Fig. 13b and Supplementary Fig. 14d, e, the control lines (S2 and HES) represent the same data points, run on the same gels, for easier comparison.

RNA sequencing

Global gene clustering

First round: Paired-end RNA sequencing was performed on H9-ESCs, mild (M)-FD PSCs, severe (S)-FD PSCs and subsequent NC progenitors with no replicates. Reads were mapped to the human genome build hg19 using the TopHat software⁴¹. Reads that aligned with no more than two base pair mismatches were accepted for downstream analysis. Second round: The FASTQ files were mapped to the target genome (UCSC HG19) using the rnaStar aligner⁴² that maps reads genomically and resolves reads across splice junctions. We used the 2 pass mapping method⁴³, in which reads are passes twice. The first mapping pass uses a list of known annotated junctions from Ensemble. Novel junctions found in the first pass are then added to the known junctions and a second mapping pass is done. After mapping we compute the expression count matrix from the mapped reads using HTSeq (www.huber.embl.de/users/anders/HTSeq) and one of several possible gene model databases. The raw count matrix generated by HTSeq are then be processed using the R/Bioconductor package DESeq (www.huber.embl.de/users/anders/DESeq), which is used to both normalize the full dataset and analyze differential expression between sample groups. A heatmap was generated using the heatmap.2 function from the gplots R package. The data plot was the mean centered normalized log2 expression of the top 100 significant genes. For simple hierarchical clustering the correlation metric was used ($D_{ij} = 1 - \text{cor}(X_i, X_j)$) with the Pearson correlation on the normalized log2 expression values.

GO term analysis: Genes that were expressed significantly different when compared in the rNC state to the PSC state were subjected to GO term analysis using the DAVID functional annotation tool (<http://david.abcc.ncifcrf.gov/tools.jsp>). The occurrence of GO terms with a Bonferroni score 0.05 were quantified. Certain GO terms were combined in order to score 10 GO terms per line, the numbers were displayed as percentages of the total, in pie charts. The analysis was done independently for each of the FD severity groups (S, M, C).

Whole-exome sequencing

500ng of DNA was captured by hybridization using the SureSelect XT HumanAllExon V4 (Agilent). Samples were prepared according to the manufacturer instructions. PCR amplification of the libraries was carried out for 6 cycles in the pre-capture step and for 7 cycles post capture. Samples were barcoded and run on a HiSeq 2500 in a 50bp/50bp Paired end run, using the TruSeq SBS Kit v3 (Illumina). The average number of read pairs per sample was 77 million, the average duplication rate was 2.1%, and the mean target coverage was over 100x. Intergenic, UTRs, intronic and synonymous variations were removed in our initial search for variants in the severe patients. In an autosomal recessive approach, 26 genes carried homozygous variations in all 3 severe patients with a MAF of less than 1% in the 1000 genomes database. 63 variations were dismissed due to their high frequency (MAF>1%) in the Exac database, leaving 7 genes carrying 14 variations as candidates. All of these 7 genes carry the same variations in the mild patients and therefore were discarded as candidate modifier genes. In an autosomal dominant approach, 77 genes carried heterozygous variations in all 3 severe patients with a MAF of less than 1% in the 1000 genomes database. 446 variations were removed based on their high frequency (MAF>1%) in the Exac database, leaving 23 genes carrying 51 variations as candidate genes commonly mutated between all severe patients. Of these, 15 carry the same variations in the mild patients and therefore were discarded as candidate modifier genes. 5 of the remaining 8 genes harbor variations with the frequency lower than 1% in the 1000 genome database predicted to be damaging as therefore not prioritized as candidates under a complete penetrance hypothesis. *FAT2*, *KIA1211* and *LAMB4* thus are the primary hit candidates. A connectome analysis using *IKBKAP* as the core gene on the website (<http://hgc.rockefeller.edu/>) and including all of the initial candidate genes identified *UR11*, *LCA10*, *MSH3*, *YEATS2* as being closely related to *IKBKAP* and could therefore be of interest for further studies.

Statistical analysis—All statistical comparisons between PSC lines were done using one-way ANOVA. Multiple comparisons were done among samples in each subfigure. The test specifics described in the figure legends were compiled from pooling data from the two severe, two mild or two control lines, respectively. Normal distribution of the raw data was not assumed, since the sample sizes were too small ($N < 5$), instead individual data points were plotted in the figures. Similar variance among patient lines was calculated using the Brown-Forsythe test and was only assumed in Figure 1b, 2d, 3d, 3e, Supplementary Figure 3b, 4b, 4d, 12a, 12d, 13b (left and right), 14d.

Reproducibility definitions—Biological replicates (N) are defined as differentiations performed at least one cell split apart, which is generally one week. Number of replicates are indicated in each figure legend as well as represented by individual plotted data points in the graphs. Sample sizes were not predefined, but generally at least 3 biological replicates were performed; for more subtle phenotypes the sample size was subsequently increased. Exclusion criteria: In each differentiation experiment the HES control line was included and experiments were excluded from analysis if the HES control line did not perform as expected, probably due to technical difficulties. Researchers were never blinded. None of the

cell lines used here are currently listed in the database of commonly misidentified cell lines maintained by ICLA.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank J-F. Brunet, Ecole Normale Supérieure, Paris, France for providing us with the *PHOX2A* antibody. S. Irion and M. Tomishima for critical reviewing of the manuscript. A. Hudon for general help. H. Ralph from the Weill Cornell Cell Screening Core for image quantification. Bioinformatics, Integrated Genomics Operations (funded by the NCI Cancer Center Support Grant (CCSG, P30 CA08748), Cycle for Survival and the Marie-Josée and Henry R. Kravis Center for Molecular Oncology) and Flow Cytometry Core Facilities at Sloan Kettering Institute. This work was supported by a fellowship for advanced researchers from the Swiss National Science Foundation (N.Z.) and through contracts from NYSTEM (C026446; C026447) and the Tri-institutional stem cell initiative (Starr Foundation) (L.S.). Robertson Investigator Award from New York Stem Cell Foundation (G.L.), Maryland Stem Cell Research Funding (MSCRF/TEDCO) (G.L.) and the Adrienne Helis Malvin Medical Research Foundation (G.L.).

References

1. Inoue H, Nagata N, Kurokawa H, Yamanaka S. iPS cells: a game changer for future medicine. *The EMBO journal*. 2014; 33:409–417. [PubMed: 24500035]
2. Zeltner N, Studer L. Pluripotent stem cell-based disease modeling: current hurdles and future promise. *Current opinion in cell biology*. 2015; 37:102–110. [PubMed: 26629748]
3. Bellin M, Marchetto MC, Gage FH, Mummery CL. Induced pluripotent stem cells: the new patient? *Nat Rev Mol Cell Biol*. 2012; 13:713–726. [PubMed: 23034453]
4. Lee G, et al. Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature*. 2009; 461:402–406. [PubMed: 19693009]
5. Axelrod FB. Familial dysautonomia. *Muscle & nerve*. 2004; 29:352–363. [PubMed: 14981733]
6. Slaugenhaupt SA, et al. Tissue-specific expression of a splicing mutation in the *IKBKAP* gene causes familial dysautonomia. *Am J Hum Genet*. 2001; 68:598–605. [PubMed: 11179008]
7. Axelrod FB, Goldberg JD, Ye XY, Maayan C. Survival in familial dysautonomia: Impact of early intervention. *J Pediatr*. 2002; 141:518–523. [PubMed: 12378191]
8. Dong J, Edelmann L, Bajwa AM, Kornreich R, Desnick RJ. Familial dysautonomia: detection of the *IKBKAP* IVS20(+6T --> C) and R696P mutations and frequencies among Ashkenazi Jews. *American journal of medical genetics*. 2002; 110:253–257. [PubMed: 12116234]
9. Pearson J, Pytel BA, Grover-Johnson N, Axelrod F, Dancis J. Quantitative studies of dorsal root ganglia and neuropathologic observations on spinal cords in familial dysautonomia. *Journal of the neurological sciences*. 1978; 35:77–92. [PubMed: 624961]
10. Pearson J, Pytel BA. Quantitative studies of sympathetic ganglia and spinal cord intermedio-lateral gray columns in familial dysautonomia. *Journal of the neurological sciences*. 1978; 39:47–59. [PubMed: 731273]
11. Axelrod FB, et al. Progressive sensory loss in familial dysautonomia. *Pediatrics*. 1981; 67:517–522. [PubMed: 7254974]
12. Slaugenhaupt SA, et al. Rescue of a human mRNA splicing defect by the plant cytokinin kinetin. *Hum Mol Genet*. 2004; 13:429–436. [PubMed: 14709595]
13. Axelrod FB, et al. Kinetin improves *IKBKAP* mRNA splicing in patients with familial dysautonomia. *Pediatr Res*. 2011; 70:480–483. [PubMed: 21775922]
14. Lee G, et al. Large-scale screening using familial dysautonomia induced pluripotent stem cells identifies compounds that rescue *IKBKAP* expression. *Nat Biotechnol*. 2012; 30:1244–1248. [PubMed: 23159879]

15. Miller JD, et al. Human iPSC-Based Modeling of Late-Onset Disease via Progerin-Induced Aging. *Cell Stem Cell*. 2013; 13:691–705. [PubMed: 24315443]
16. Fusaki N, Ban H, Nishiyama A, Saeki K, Hasegawa M. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc Jpn Acad Ser B Phys Biol Sci*. 2009; 85:348–362.
17. Lee G, et al. Isolation and directed differentiation of neural crest stem cells derived from human embryonic stem cells. *Nature biotechnology*. 2007; 25:1468–1475.
18. Zeltner N, Lafaille FG, Fattahi F, Studer L. Feeder-free derivation of neural crest progenitor cells from human pluripotent stem cells. *Journal of visualized experiments: JoVE*. 2014
19. Chambers SM, et al. Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. *Nature biotechnology*. 2012; 30:715–720.
20. Wainger BJ, et al. Modeling pain in vitro using nociceptor neurons reprogrammed from fibroblasts. *Nature neuroscience*. 2015; 18:17–24. [PubMed: 25420066]
21. Mica Y, Lee G, Chambers SM, Tomishima MJ, Studer L. Modeling neural crest induction, melanocyte specification, and disease-related pigmentation defects in hESCs and patient-specific iPSCs. *Cell reports*. 2013; 3:1140–1152. [PubMed: 23583175]
22. Fattahi F, et al. Deriving human ENS lineages for cell therapy and drug discovery in Hirschsprung disease. *Nature*. 2016; 531:105–109. [PubMed: 26863197]
23. Howard MJ. Mechanisms and perspectives on differentiation of autonomic neurons. *Dev Biol*. 2005; 277:271–286. [PubMed: 15617674]
24. Grundy D, Schemann M. Enteric nervous system. *Current opinion in gastroenterology*. 2007; 23:121–126. [PubMed: 17268239]
25. Schlosser G. Induction and specification of cranial placodes. *Dev Biol*. 2006; 294:303–351. [PubMed: 16677629]
26. Oh Y, et al. Functional Coupling with Cardiac Muscle Promotes Maturation of hPSC-Derived Sympathetic Neurons. *Cell Stem Cell*. 2016
27. Axelrod FB, et al. Kinetin improves IKBKAP mRNA splicing in patients with familial dysautonomia. *Pediatr Res*. 2011; 70:480–483. [PubMed: 21775922]
28. Dietrich P, Alli S, Shanmugasundaram R, Dragatsis I. IKAP expression levels modulate disease severity in a mouse model of familial dysautonomia. *Hum Mol Genet*. 2012; 21:5078–5090. [PubMed: 22922231]
29. Morini E, et al. Sensory and Autonomic Deficits in a New Humanized Mouse Model of Familial Dysautonomia. *Hum Mol Genet*. 2016
30. Ding Q, et al. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell stem cell*. 2013; 12:393–394. [PubMed: 23561441]
31. Mali P, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013; 339:823–826. [PubMed: 23287722]
32. Adzhubei IA, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010; 7:248–249. [PubMed: 20354512]
33. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009; 4:1073–1081. [PubMed: 19561590]
34. Itan Y, et al. The human gene damage index as a gene-level approach to prioritizing exome variants. *Proc Natl Acad Sci U S A*. 2015; 112:13615–13620. [PubMed: 26483451]
35. Itan Y, et al. The mutation significance cutoff: gene-level thresholds for variant predictions. *Nat Methods*. 2016; 13:109–110. [PubMed: 26820543]
36. Itan Y, et al. HGCS: an online tool for prioritizing disease-causing gene variants by biological distance. *BMC Genomics*. 2014; 15:256. [PubMed: 24694260]

Figure 1.

Phenotypes in rNC cells derived from FD-PSCs recapitulate differences between mild and severe FD patients. **a.** Quantification of wild type to mutant *IKBKAP* transcript expression ratios in severe (S2 and S3) and mild (M1 and M2) FD PSC-derived rNC cells compared to controls (C1 and HES-rNC). $N=3-6$ biological replicates for each, one-way ANOVA, $**p=0.001$, $F=8.85$, Sidak's multiple comparison test. S versus M: n.s. $t=0.004$, S versus C: $*p<0.05$, $t=3.1$. S versus HES: $**p<0.005$, $t=3.84$. M versus C: $*p<0.05$, $t=3.31$. M versus HES: $**p<0.005$, $t=4.12$. C versus HES: n.s. $t=0.72$. **b.** Quantification of rNC cell migration in the scratch assay. Left, example (Hoechst-stained) images depicting cell migration 48h after scratch. Dashed lines indicate the scratched region of interest. Scale bar: 500 μm . Right, quantification of the percentage of cells located in the scratched ROI compared to all cells in the field. $N=3-5$ biological replicates, one-way ANOVA, $*p<0.016$, $F=5.52$, Tukey's multiple comparison test. S versus M: $*p<0.05$, $q=5.1$. S versus C: $*p<0.05$, $q=4.63$. S versus HES: $*p<0.05$, $q=4.53$. M versus C: n.s. $q=1.06$. M versus HES: n.s. $q=0.56$. C versus HES: n.s. $q=0.42$. Mean $\pm$ S.E.M.. **c.** Quantification of *ASCL1* transcript expression in FACS isolated rNC cells analyzed by qRT-PCR. $N=2-6$ biological replicates, one-way ANOVA, $****p<0.0001$, $F=14.93$, Tukey's multiple comparison test. S versus M: $***p<0.0005$, $q=6.3$. S versus C/HES: $****p<0.00005$, $q=7.24$. M versus C/HES: n.s. $q=0.14$. **(d)** Example images depicting immuno-staining of rNC cultures 4 days after isolation for ASCL1, TUJ1 and DAPI Scale bar: 100 μm . Data are presented as mean $\pm$ S.E.M. and circles indicate biological replicates from individual differentiations, throughout.

Figure 2.

FD phenotypes in peripheral sensory and autonomic neurons. **a.** SNs were differentiated from FD PSCs *in vitro* using an established protocol¹⁹. At day 12 of differentiation the cells were replated at equal densities and immuno-stained for BRN3A, TUJ1 and DAPI at day 14. Scale bar: 100 μ m. **b.** The number of BRN3A⁺/TUJ1⁺ cells was counted in day 14 cultures of SNs derived from FD PSCs and controls. N=3–13 biological replicates, one-way ANOVA, **** p <0.0001, F =17.62, Tukey's multiple comparison test. S versus M: **** p <0.00005, q =7.68. S versus C/HES: **** p <0.00005, q =7.1. M versus C/HES: ns, q =1.17. **c.** The ratio of wild type to mutant *IKBKAP* splicing was measured in day 12 SNs. N=3–8 biological replicates, one-way ANOVA, **** p <0.0001, F =30.46, Tukey's multiple comparison test. S versus M: ns q =0.02. S versus C/HES: **** p <0.00005, q =9.56. M versus C/HES: **** p <0.00005, q =9.76. **d.** The percentage of CD49D⁺ NC progenitors derived from control and FD PSCs were quantified by FACS at day 11 of the AN-like differentiation protocol. N=3–6 biological replicates, one-way ANOVA, **** p <0.0001, F =24.83, Tukey's multiple comparison test. S1 versus M: *** p <0.0005, q =6.82. S versus C/HES: **** p <0.00005, q =9.61. M versus C/HES: ns, q =2.73.

Figure 3.

Mild and severe FD SNs have impaired survival capacity *in vitro*, which can be rescued pharmacologically. **a.** PSC-derived SNs were replated at equal density (DAPI staining (first panel) and maintained for at least 14 more days in optimal survival conditions. Every three days cultures and immuno-stained for BRN3A, TUJ1 and DAPI (DAPI not shown for easier visibility, but can be seen in Supplementary Fig. 10b) to visualize the death of SNs. Scale bar: 200 μ m. **b.** At each time point the number of BRN3A⁺/TUJ1⁺ cells was counted and normalized to day 14 of the respective experiment. N=2–8 biological replicates, two-way ANOVA, interaction ****p<0.0001, F=3.98, Dunnett's multiple comparison test. Day 26: C1 versus HES: ns, q=1.72. C1 versus M1/M2/S2/S3: each ****p<0.00005, q=4.8/5.24/4.92/4.77. **c.** The survival defect in mild FD SNs can be rescued with kinetin and SKF-86466. M1-PSC-derived SNs were treated with 10 mM kinetin, 10 mM SKF-86466 or DMSO-vehicle beginning at day 12. Quantified was performed as in b. N=3 biological replicates, two-way ANOVA, interaction *p<0.05, F=2.34, Dunnett's multiple comparison test. Day 26: DMSO versus kinetin: ***p<0.0005, q=4.19. DMSO versus SKF-86466: ***p<0.0005, q=4.55.

Figure 4.

Genetic rescue of severe FD PSCs and *in vitro* FD phenotypes. **a, b.** Ratio of wild type to mutant *IKBKAP* splicing in PSCs (N=3–5 biological replicates, one-way ANOVA, *** $p < 0.0001$, $F = 9.4$, Dunnett's multiple comparison test. S versus S2-rescue1/S2-rescue5/S2-rescue7/A1: ns, $q = 0.3/0.2/0.1/0.2$. S versus HES *** $p < 0.005$, $q = 5.3$ (a)) and rNC (N= 3–6 biological replicates, one-way ANOVA, ** $p < 0.001$, $F = 4.4$, Dunnett's multiple comparison test. S versus S2-rescue1/S2-rescue5/S2-rescue7/A1: ns, $q = 0.008/0.06/0.009/0.06$. S versus HES * $p < 0.05$, $q = 3.2$ (b)). **c.** Rescue of the *ASCL1* phenotype in rNC cells. N=2–6 biological replicates, one-way ANOVA, **** $p < 0.0001$, $F = 16.12$, Dunnett's multiple comparison test. S versus S2-rescue1: **** $p < 0.00005$, $q = 7.21$. S versus S2-rescue5: ** $p < 0.005$, $q = 3.62$. S versus S2-rescue7: ns, $q = 1.14$. S versus A: ns, $q = 1.04$. S versus HES: ** $p < 0.005$, $q = 4.4$. **d.** Rescue of the CD49D phenotype in the NC precursor at day 11 of the AN-like protocol. N=3–5 biological replicates, one-way ANOVA, ** $p < 0.0098$, $F = 4.18$, Dunnett's multiple comparison test. S versus S2-rescue1: ns, $q = 1.35$. S versus S2-rescue5: ** $p < 0.005$, $q = 3.9$. S versus S2-rescue7: * $p < 0.05$, $q = 3.25$. S versus A: ns, $q = 1.87$. S versus HES: * $p < 0.05$, $q = 3.02$. **e.** Genetic reversal in *IKBKAP* does not reverse the SN generation defect. SNs were differentiated and immuno-stained for BRN3A and DAPI at day 12 (left). The percentage of BRN3A⁺ cells over DAPI⁺ cells was quantified (right). N=3 biological replicates, one-way ANOVA, **** $p < 0.0001$, $F = 32.3$, Dunnett's multiple

comparison test. S2 versus S2-rescue1/5/7: ns, $q=2.7/2.6/1.7$. S versus A1: $**p<0.0005$, $q=5.4$. S2 versus HES: $***p<0.00005$ $q=11.3$. Scale bar=100 μm .

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 5. Mechanistic insight into the molecular difference between mild and severe FD. **a.** Representation of whole exome sequencing of three mild and three severe FD patients. Red triangles represent mutations that passed all filtering criteria, i.e. exonic mutations, non-synonymous, MAF<1% in 1KG and Exac databases. The three final modifier candidate genes *LAMB4*, *FAT2*, *KIAA1211* were not mutated in mild patients and predicted to be damaging and rare based on their CADD, MSC and GDI scores. **d.** Model summarizing the findings of our study.