1   **Polyalanine disease mutations impair UBA6-dependent ubiquitination**

2

3   Fatima Amer-Sarsour [1,6], Daniel Falik [2,5,6], Yevgeny Berdichevsky [1], Alina Kordonsky [3], Gali Prag

4   [3,4], Gad D Vatine [2,5,*], Avraham Ashkenazi [1,4,*]

5

6   [1] The Department of Cell and Developmental Biology, Sackler Faculty of Medicine, Tel Aviv

7   University, 6997801 Tel Aviv, Israel

8   [2] The Department of Physiology and Cell Biology, Faculty of Health Sciences and The

9   Regenerative Medicine and Stem Cell (RMSC) Research Center, Ben-Gurion University of the

10   Negev, 8410501 Beer Sheva, Israel

11   [3] School of Neurobiology, Biochemistry and Biophysics, the George S. Wise Faculty of Life

12   Sciences, Tel Aviv University, 6997801 Tel Aviv, Israel

13   [4] Sagol School of Neuroscience, Tel Aviv University, 6997801 Tel Aviv, Israel

14   [5] The Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, 8410501 Beer

15   Sheva, Israel

16   [6] Co-first authors

17   [*] Equally contributing corresponding authors: G.D Vatine (vatineg@bgu.ac.il); A. Ashkenazi

18   (ashkenaziavi@tauex.tau.ac.il)

19

20

21

22

23

1 **Abstract**

2 Expansion mutations in polyalanine stretches are now associated with a growing number of human

3 diseases with common genotypes and similar phenotypes [1-6]. These similarities prompted us to

4 query the normal function of physiological polyalanine stretches, and investigate whether a

5 common molecular mechanism is involved in these diseases. Here, we show that UBA6, an E1

6 ubiquitin-activating enzyme [7, 8], recognizes a polyalanine stretch within its cognate E2 ubiquitin-

7 conjugating enzyme, USE1. Aberrations in this polyalanine stretch reduced ubiquitin transfer to

8 USE1 and downstream target, the E3 ubiquitin ligase, E6AP. Intriguingly, we identified

9 competition for the UBA6-USE1 interaction by various proteins with polyalanine expansion

10 mutations in the disease state. In mouse primary neurons, the deleterious interactions of expanded

11 polyalanine proteins with UBA6, alter the levels and ubiquitination-dependent degradation of

12 E6AP, which in turn affected the levels of the synaptic protein, Arc. These effects could be

13 observed in induced pluripotent stem cell-derived autonomic neurons from patients with

14 polyalanine expansion mutations. Our results suggest a shared mechanism for such mutations,

15 which may contribute to the congenital malformations seen in polyalanine diseases.

16

17

Introduction

Trinucleotide repeats present within coding regions of the genome encode stretches of the same amino acid [9], and expansion mutations of such amino acid sequences have now been linked to a growing number of human diseases. The nine diseases caused by expansions of polyalanine stretches are due to the expression of aberrant nuclear proteins (eight of the nine are transcription factors). Examples include mutant paired like homeobox 2B (mutant PHOX2B) in congenital central hypoventilation syndrome (CCHS) [1], mutant RUNX family transcription factor 2 (mutant RUNX2) in cleidocranial dysplasia [3], mutant homeobox D13 (mutant HOXD13) in synpolydactyly[2], and mutant poly(A) binding protein nuclear 1 (mutant PABPN1) in oculopharyngeal muscular dystrophy [5]. The vast majority of cases are caused by *de novo* mutations and often involve congenital neurological abnormalities. The disease-causing expansions vary in length, according to the gene in question, with the severity of the associated clinical phenotype generally increasing with the length of the expanded tract [9-11]. Physiological polyalanine stretches are commonly found in proteins, although they usually do not exceed 20 residues. In contrast, disease-causing expansion mutations may contain up to 13 additional alanine residues and can result in protein misfolding [9-11]. The exact functional or structural role of the polyalanine stretches in normal proteins is unclear. One hypothesis is that they are simply protein spacer elements between functional domains. Another suggests that polyalanine stretches play roles in protein-protein recognition, but this remains unknown.

Perturbation of ubiquitin activation, conjugation, and transfer to target proteins via the E1-E2-E3 enzymatic cascade has been linked to dysregulation of neuronal homeostasis and to a growing number of neurodevelopmental disorders [12-14]. Here, we describe a polyalanine motif in the ubiquitin-conjugating E2 enzyme UBE2Z/USE1, which contributes to USE1 ubiquitin loading

3

1   by the E1 ubiquitin activating enzyme, UBA6. In addition, our results identify a domain in UBA6

2   that recognizes polyalanine-containing proteins and demonstrate that, under disease conditions,

3   UBA6 preferentially interacts with different polyalanine-expanded proteins, thereby competing

4   with USE1 binding. Since similar effects could be confirmed in neurons derived from patients with

5   disease-causing mutations, we suggest that this represents a previously undescribed vulnerability

6   caused by polyalanine expansion mutations.

7

Results

9   A search for proteins containing polyalanine stretches in the ubiquitin cascades identified a small

10 subset of E3 ubiquitin ligases and the E2 ubiquitin-conjugating enzyme, USE1 (**Extended Data**

11 **Fig. 1a**). USE1 has both N- and C-terminal extensions on top of the ubiquitin-conjugating (UBC)

12 core domain (**Fig. 1a**), classifying it as a class IV E2, which can be specifically loaded with

13 ubiquitin by the E1 ubiquitin activating enzyme, UBA6, via a transthiolation reaction [7, 8, 15]. The

14 polyalanine stretch of USE1 is located in the N-terminal extension and is well conserved in

15 primates, several other mammals, and reptiles (**Fig. 1a**, **Extended Data Fig. 1b**). In order to

16 examine the role of the alanine stretch in USE1 ubiquitin loading, we replaced two alanine residues

17 in the stretch with arginine residues. When transfected into HEK293T cells, the ubiquitin loading

18 of the USE1 2A>2R mutant was significantly reduced in comparison to wild type USE1 (**Fig. 1b**).

19 Exposure of cell lysates to the reducing agent β-mercaptoethanol, abolished USE1 loading under

20 all examined conditions, suggesting that the additional band observed is ubiquitin conjugated via

21 a thiolation reaction, similarly to a catalytic dead enzyme (USE1 C188A) that cannot be loaded

22 with ubiquitin (**Fig. 1b**). Under non-reducing conditions, the ubiquitin loading of wild type USE1

23 and USE1 2A>2R was abolished in UBA6-siRNA depleted cells (**Fig. 1b**) suggesting that the

1 difference in ubiquitin loading between the wild type and 2A>2R USE1 is dependent on UBA6

2 activity.

3   We constructed additional USE1 mutants including one with a deletion of the polyalanine

4 stretch (USE1 ΔPolyAla), and a deletion of Loop B (USE1 ΔLB), which is hyperactive [16] (**Fig.**

5 **1c**). The USE1 ΔPolyAla mutant showed reduced ubiquitin loading compared to wild type USE1

6 or to USE1 ΔLB both in cells and *in vitro* with purified recombinant UBA6 (**Fig. 1c, Fig. 1d**).

7 These findings provide further support that the polyalanine stretch contributes to the recognition

8 of USE1 by UBA6.

9

Isolated polyalanine stretches compete with USE1

11 In order to examine whether isolated polyalanine stretches (19 Ala residues) interact with UBA6

12 in mammalian cells, we transfected HEK293T cells with GFP-polyAla, and monitored the binding

13 to UBA6, or its possible competition with USE1. We found that the polyAla stretch binds UBA6

14 and that this interaction decreases the USE1-UBA6 binding (**Fig. 1e**). The ubiquitin-activating E1

15 enzymes, UBA1 and UBA6 share 40% identity of protein sequence and a strong specificity for

16 their cognate ubiquitin-like proteins [17-19]. The ubiquitin fold domain of E1 interacts with the α1

17 helix of E2 and functions to determine selectivity between E2s of different ubiquitin-like proteins

18 [20]. To identify regions in UBA6 that are likely to determine the specificity to USE1, we compared

19 the physico-chemical properties of UBA6 to those of UBA1 by structural modeling (**Fig. 2a**).

20 Interestingly, comparison of the structures of UBA1 and UBA6 at the interface with the E2-α1 as

21 well as the E2-α1 helices themselves seemed to be very similar. However, projection of calculated

22 electro-potential properties on the surfaces of the two enzymes revealed significant differences

23 within the second catalytic cysteine half (SCCH) domains. In place of the negative groove seen in

1   UBA1, UBA6 contains a large positively charged groove that is formed by the key residues

2   Lys628, Arg691, Lys709, and Lys704 (**Fig. 2a**). The structural comparison of the canonical E1

3   enzymes UBA1, 2, 3, 6 and 7 [19-22] shows significant differences in the SCCH domains (**Extended**

4   **Data Fig. 2**). Moreover, a recent study demonstrated that replacing the SCCH domain of UBA1

5   with the one of UBA2 allows the engineered UBA1 to load ubiquitin onto the E2 of SUMO

6   (UBC9) [23]. Consequently, we assumed that the SCCH domain of UBA6 may hold specificity to

7   USE1 and thus mutating the above-mentioned positive charged residues in UBA6 may affect the

8   binding to USE1. Based on this hypothesis, we constructed UBA6 mutants with the positive

9   residues replaced by Ala (UBA6 mut 4Ala) or Asp (UBA6 mut 4Asp). The USE1 binding of both

10   these mutants was indeed significantly reduced in HEK293T cells (**Fig. 2b**) as was the interaction

11   of both mutants with isolated polyalanine stretches compared to wild-type UBA6 (**Fig. 2c**). *In*

12   *vitro*, both recombinant UBA6 mut 4Ala and UBA6 mut 4Asp showed reduced ability to load

13   USE1 with ubiquitin when compared to wild type UBA6 (**Fig. 2d**). Interestingly, neither mutations

14   in the positive groove of UBA6 nor aberrations in the polyAla sequence in USE1 are sufficient to

15   abrogate ubiquitin transfer from UBA6 to USE1. This suggests that while these two regions are

16   important for specificity they are not essential for the general mechanism of ubiquitin activation

17   (*i.e*. adenylation) or transthiolation.

18

19   **Disease-causing proteins with polyalanine expansions interact with UBA6**

20   Since UBA6 is mainly localized in the cytoplasm in HEK293T cells (**Extended Data Fig. 3a**), we

21   next expressed isolated polyalanine stretches with or without a nuclear localization sequence

22   (NLS) and monitored their binding to UBA6. As expected, polyalanine stretches bound UBA6

23   effectively when expressed without NLS (**Fig. 3a**), and reduced the ubiquitin loading of USE1

6

1   (**Extended Data Fig. 4a**). However, in contrast, isolated polyalanine stretches with NLS did not

2   bind UBA6 and had no apparent effect on USE1 loading (**Fig. 3a**, **Extended Data Fig. 4a**). To

3   explore pathological UBA6 interactions, we expressed various disease-causing proteins with

4   polyalanine expansion mutations in cells (**Extended Data Fig. 3 b-e**) and compared their

5   interactions with UBA6 to those of the wild type protein (**Fig. 3 b-e**). This includes PHOX2B

6   (WT, +13 Ala, **Fig. 3b**), RUNX2 (WT, +6 Ala, +12 Ala, **Fig. 3c**), HOXD13 (WT, +10 Ala, **Fig.**

7   **3d**), and PABPN1 (WT, +7 Ala, **Fig. 3e**).

8   These pathological interactions are measurable because although UBA6 is mainly localized

9   to the cytoplasm and, as expected, wild type PHOX2B, RUNX2, HOXD13, and PABPN1 are

10   primarily localized in the nucleus, all the mutant proteins with polyalanine expansions partially

11   mislocalized to the cytoplasm and could then interact with endogenous UBA6 (**Extended Data**

12   **Fig. 3 b-e, Fig. 3 b-e**).

13

14   **Inhibition of UBA6-depndent E6AP degradation by polyalanine expansions**

15   E6AP/UBE3A is a highly potent E3 ubiquitin ligase whose regulation is critical for normal

16   development of the nervous system. Indeed, decreased activity of the ligase results in Angelman

17   syndrome while increased activity causes autism spectrum disorders [14, 24, 25]. Harper and co-

18   workers identified a unique regulatory cascade in which UBA6-USE1 ubiquitinates E6AP for

19   degradation in the proteasome [13]. Accordingly, our results indicate that depletion of UBA6, as well

20   as expression of mutant PHOX2B (+13 Ala), increase the levels of E6AP (**Extended Data Fig.**

21   **4b**). This correlates with a decrease in Lys48-linked polyubiquitination of E6AP (**Fig. 3f**),

22   suggesting that proteasome-mediated degradation of E6AP is inhibited by mutant PHOX2B.

23   Moreover, E6AP levels were increased in the polyalanine-expressing cells, which was reversed by

7

1   UBA6 overexpression (**Extended Data Fig. 4c**), compatible with a model in which the polyalanine

2   stretch impairs the regulation of E6AP levels by competing with USE1 on UBA6.

3   In order to test our model in neuronal cells, we used mouse primary cortical neurons.

4   Similarly, UBA6 was predominantly detected in the cell body and in neurites, with only a small

5   fraction localized to the nucleus (**Extended Data Fig. 5a**). Transduction of the neurons with

6   lentiviruses expressing GFP-tagged wild type or mutant PHOX2B (+13 Ala) resulted in a

7   significant difference in subcellular localization and expression pattern. While most of the wild

8   type PHOX2B could be detected in the nucleus, mutant PHOX2B (+13 Ala) was also found in the

9   cell body and in neurites, where the GFP fluorescence imaging revealed both non-aggregated and

10   aggregated patterns (**Extended Data Fig. 5b**). Neurons expressing mutant PHOX2B showed

11   increased colocalization with UBA6, which correlated with increasing lengths of the polyalanine

12   stretch (**Fig. 3g, Extended Data Fig. 5c**). Moreover, expression of mutant PHOX2B with +7 Ala

13   and +13 Ala increased the levels of E6AP in the primary neurons (**Fig. 3h**). The synaptic activity-

14   regulated cytoskeleton-associated protein (Arc) is negatively regulated by E6AP [26, 27], and

15   accordingly, the increased levels of E6AP due to ectopic expression of the PHOX2B mutants with

16   +7 Ala and +13 Ala was accompanied with a decrease in Arc levels in primary neurons (**Fig. 3h,**

17   **Extended Data Fig. 5d**).

18

19   **UBA6 dysregulation in neurons of patients with polyalanine expansions**

20   In order to investigate the mechanisms underlying human autonomic neuron vulnerabilities, we

21   sought to assess the associations between endogenous polyalanine expansions and UBA6.

22   Endogenous PHOX2B is expressed in specific neuronal populations of the autonomic nervous

23   system [28, 29]. In humans, polyalanine expansion mutations within the *PHOX2B* gene cause CCHS,

1 a rare and life-threatening condition with autonomic nervous system dysfunction [30, 31]. We have

2 previously described the generation of two induced pluripotent stem cells (iPSCs) from identical

3 twins carrying a heterozygous PHOX2B +5 Ala expansion (101iCCHS 20/25 and 102iCCHS

4 20/25) [32]. Here, we obtained skin punch biopsies from an additional CCHS patient bearing a

5 heterozygous +7 Ala expansion (104iCCHS 20/27), as well as from their sex-matched healthy

6 family relatives, which serve as controls (103iCTR 20/20 and 105iCTR 20/20) (**Extended Data**

7 **Fig. 6**). These patients suffer from central sleep apnea with a more severe peripheral autonomic

8 presentation in 104iCCHS 20/27, including Hirschsprung disease. Patient-derived fibroblasts were

9 reprogrammed using non-integrating episomal plasmids as previously described [32]. All generated

10 iPSC lines expressed the pluripotency markers NANOG, SOX2, OCT3/4 TRA-1-60, and SSEA4

11 (**Extended Data Fig. 6 a-j**), and spontaneously differentiated into the three germ layers (**Extended**

12 **Data Fig. 6 l-n**). All lines showed a normal karyotype, and genetic analysis confirmed the presence

13 of a heterozygous expansion of seven alanine residues resulting with a 27 polyalanine stretch in

14 PHOX2B (**Extended Data Fig. 6 k, o**). In order to obtain endogenous PHOX2B-expressing cells,

15 we differentiated the iPSCs to neural crest progenitor cells that were further differentiated into

16 sympathetic neuroblast crest cells, and finally to peripheral autonomic neurons (**Fig. 4a**).

17 Characterization of the human autonomic neurons revealed expression of PHOX2B, the pan-

18 neuronal marker βIII-tubulin, the catecholaminergic marker tyrosine hydroxylase (TH), the

19 peripheral neuronal marker peripherin, and atonal BHLH transcription factor 1 (ATOH1), which

20 is associated with breathing and digestion (**Fig. 4a**).

21   Analysis of the PHOX2B localization showed a decrease in the nuclear fraction of

22 PHOX2B in the mutant 20/25 and 20/27 neurons, as compared to healthy control neurons

23 (**Extended Data Fig. 7 a,b**). The cytoplasmic PHOX2B was mainly perinuclear without visible

1 aggregates (**Extended Data Fig. 7c**). In addition, there was an increased association between

2 UBA6 and PHOX2B in the CCHS patient neurons (**Fig. 4b**). Some CCHS neurons exhibited

3 severe PHOX2B cytoplasmic mislocalization, with apoptotic nuclear morphology (1.9% and 3.9%

4 of the 20/25 and the 20/27 mutations, respectively) (**Extended Data Fig. 7c**). In accordance with

5 the mouse data, E6AP levels were higher in the patient neurons (**Fig. 4c**). Moreover, Arc levels

6 were lower in the CCHS neurons with the 20/25 mutation than in control cells and the levels were

7 further reduced in CCHS neurons with the 20/27 mutation (**Fig. 4d**).

8

9 **Discussion**

10 Expansion mutations in stretches of repetitive DNA sequences that encode poly amino acids such

11 as polyglutamine or polyalanine can cause various diseases. Whereas much progress has been

12 made on the molecular mechanisms of polyglutamine diseases [9, 33, 34], the consequences of

13 polyalanine expansions remain more of an enigma. Here, we demonstrate how expanded

14 polyalanine stretches can compete with the normal function of a shorter stretch to disrupt the

15 specific interaction between the E2 ubiquitin-conjugating enzyme USE1 and the E1 ubiquitin-

16 activating enzyme, UBA6. Our results indicate that UBA6 interacts with polyalanine stretches,

17 and that this interaction contributes to UBA6 recognition of USE1, but can be altered by different

18 polyalanine disease-causing proteins.

19 The correlation between polyalanine expansion mutations and a possible harmful gain-of-

20 function is supported by the observation that the respiratory abnormalities of *phox2b*[+/-] mouse

21 mutants are transient and milder than those exhibited by *phox2b*[+/27Ala] mouse mutants or those of

22 CCHS patients [35, 36]. These genetic studies suggest the possibility of additional disease contributing

23 mechanisms besides partial loss of PHOX2B function in the nucleus and alterations in

1   transcriptional programs by polyalanine expansion mutations [37]. UBA6 plays an important role in

2   mouse embryonic development, neuronal function, and survival [13, 38]. Thus, the inhibition of UBA6

3   by sequestering the enzyme into cytoplasmic mislocalized polyalanine expansions may represent

4   a deleterious mechanism for such mutations. The correlation between disease severity and the

5   length of the polyalanine expansion may be related to the greater tendency of mutants with longer

6   polyalanine stretches to mislocalize to the cytoplasm and interact there with UBA6. This tendency

7   of polyalanine-expanded proteins to mislocalize to the cytoplasm has been previously detected in

8   overexpression systems [10], but the relevance to human disease was questioned. Our results provide

9   evidence that cytoplasmic mislocalization of endogenous PHOX2B can occur in the neurons

10   affected in CCHS, which may contribute to the neuronal dysfunction seen in this disease.

11   In summary, our study provides mechanistic insights into the long-standing question of the

12   functional role of polyalanine stretches and demonstrates that they are relevant to the ubiquitin

13   system and its dysregulation in disease states.

14

Methods

Ethics statement

All mouse experiments were reviewed and approved by the Institutional Animal Care and Use

Committee of Tel Aviv University. All cell lines and protocols related to human stem cell research

in the present study were used in accordance with guidelines approved by the Institutional Review

Board (Sheba Medical Center and Tel Aviv University). Informed consent was obtained from all

donors.

22

23

1 **Antibodies**

2 The following antibodies were used in this study: mouse anti-HA (Biolegend 901501); rabbit anti-

3 HA (Cell signaling 3724); rabbit anti-Actin (SIGMA A2066); mouse anti- E6-AP (Santa Cruz

4 166689); mouse anti-FLAG (SIGMA F1804); rabbit anti-UBA6 (Cell signaling 13386); rabbit

5 anti-K48 polyUb (Cell signaling 8081S); rabbit anti-UBE2Z (USE1, (Abcam 254700); rabbit anti-

6 GFP (Abcam 6556), mouse anti-PHOX2B (Santa Cruz 376997); mouse anti-mCherry (Abcam

7 125096); mouse anti-Tau (Santa Cruz 58860); rabbit anti-MAP2a (synaptic systems 188003);

8 mouse anti-Arc (BD Transduction Laboratories 612602); rabbit anti-Arc (Abcam 183183);

9 Alexa Fluor 555 (Abcam 150114) conjugated goat anti-mouse secondary antibody,

10 Alexa Fluor® 555 (Abcam 150078) conjugated goat anti-rabbit secondary antibody, goat anti-

11 mouse (Abcam 6789) and goat anti-rabbit (Abcam 6721); horseradish peroxidase (HRP)-

12 conjugated secondary antibodies, mouse anti-MATH-1 (Santa Cruz 136173), mouse anti-PRPH

13 (Santa Cruz 377093); mouse anti-TH (Santa Cruz 25269); mouse anti-SOX10 (Abcam 155279);

14 rabbit anti-β-Tubulin III (SIGMA T2200); mouse anti-OCT3/4 (Santa Cruz 5279); mouse anti-

15 TRA 1-60 (R & D MAB4770); mouse anti-SSEA-4 (Santa Cruz 21740); mouse anti-IgG3 Isotype

16 (Bio-legend 330405); rabbit anti-SOX2 (Abcam 97959); rabbit anti-NANOG (Abcam 21624);

17 rabbit anti- 68kDa Neurofilament (Abcam 52989); rabbit anti-α-Fetoprotein (ScyTek A00058);

18 rabbit anti- α-smooth muscle actin (Abcam 32575); Alexa Fluor 488-conjugated anti-mouse

19 secondary antibody (Jackson Immuonoresearch 715545150); Alexa Fluor 488-conjugated anti-

20 mouse secondary antibody (Thermo Fisher Scientific, A-11034); Alexa Fluor 488-conjugated anti-

21 rabbit secondary antibody (Molecular Probes, A-11029); Alexa Fluor 594-conjugated anti-rabbit

22 secondary antibody (Jackson Immuonoresearch 711585152); Alexa Fluor 568-conjugated anti-

1   rabbit secondary antibody (Abcam 175471); and Alexa Fluor 647-conjugated anti-rabbit secondary

2   antibody (Abcam 150079).

3   **Cloning and constructs**

4   E1 constructs for mammalian cell expression, namely pcDNA3.1-HA-tagged UBA1 wild type and

5   pcDNA3.1-HA-tagged UBA6 (UBE1L2) wild type, were kindly provided by Dr. Marcus

6   Groettrup. pcDNA3.1-HA-UBA6 mut (4Ala) and pcDNA3.1-HA-UBA6 mut (4Asp) where

7   Lysine 628, 709,714 and Arginine 691 were changed to amino acids Alanine and Aspartate

8   respectively, were constructed by Gibson assembly using the DNA gene blocks. For *E. coli*

9   expression as a His_6-tagged proteins, the wild type and mutant UBA1 and UBA6 genes where

10   subcloned into modified and improved vector pET28a [39].

11   E2 USE1 (UBE2Z) expressing plasmids pcDNA3.1-His_6-3xFlag-USE1 was kindly

12   provided by Dr. Annette Aichem. For expression in mammalian cells, the putative Kozak sequence

13   was first added to the USE1 gene, which was subcloned into pEGFP-N1 based plasmid resulting

14   in plasmid pCMV-His_6-3xFlag-USE1 new Kozak used in this study. The ΔPolyAla region mutant,

15   where amino acids 47-56 of the USE1 gene were deleted, the 2A>2R mutant where Alanine 49

16   and Alanine 52 were changed to Arginine respectively, the ΔLB mutant where amino acids 194-

17   197 comprising a Loop B [16] were deleted, and the C188A mutant were constructed by site-directed

18   mutagenesis and Gibson assembly applying Q5® Site-Directed Mutagenesis Kit and NEBuilder®

19   HiFi DNA Assembly respectively. For *E.coli* expression as a His_6-tagged proteins, the USE1 wild

20   type and mutant genes were subcloned into modified plasmid pET30-His_6-N-AviTag.

21   His_6-Ub plasmid for expression in *E. coli* of the His-tagged ubiquitin gene was described

22   previously [40]. The HA-ubiquitin plasmid for mammalian expression was a gift from Edward Yeh

23   (Addgene plasmid # 18712 [41]).

1   pEGFP-19 Alanines and pEGFP-19 Alanines with nuclear localization sequence (NLS)

2   constructs for expression in mammalian cells as well as HA-bovine PABPN1 wild type and HA-

3   bovine PABPN1 mut +7 Ala mutant constructs were a gift of Dr. David Rubinsztein. First, a delta

4   polyalanine region mutant was constructed by Inverse PCR and then the PABPN1 gene was

5   transferred into pEGFP-N1 based vector. In order to construct human PABPN1 wild type and

6   mutant genes, a site directed mutagenesis was applied to change amino acids Asparagine 95 to

7   Serine and Serine 102 to Proline on the delta Polyalanine region construct. Humanized PABPN1

8   wild type and +7 Ala mutant plasmids were constructed by adding polyalanine regions applying

9   the DNA gene blocks and Gibson assembly.

10   Human RUNX2 carrying plasmids pcDNA3.2/GW/D-TOPO RUNX2 wild type, pcDNA

11   6.2/C-EmGFP-DEST-RUNX2 mut (+6 Ala) and pcDNA 6.2/C-EmGFP-DEST- RUNX2 mut (+12

12   Ala) were kindly provided by Dr. Yoshihito Tokita. In this study, in all aforementioned plasmids,

13   the EmGFP gene was deleted and HA-tag was introduced onto the C-terminus of the RUNX2

14   genes.

15   Mouse HOXD13 bearing plasmids pcDNA3.1-Hoxd13 wild type and pcDNA3.1-Hoxd13

16   mut +10Ala were a gift of Dr. Denes Hnisz. First, the Valine259 to Glutamate mutation was

17   corrected in both constructs and then the wild type and the mutant Hoxd13 genes were subcloned

18   into pEGFP-N1 derived vector whereas adding a putative Kozak sequence and the C-terminal HA-

19   tag.

20   PHOX2B carrying plasmids pcDNA3.0-HA-PHOX2B wild type and pcDNA3.0-HA-

21   PHOX2B mut (+13 Ala) were kindly provided Dr. Diego Fornasari. To express the PHOX2B

22   genes alone and as fusion with EGFP in neuronal cells under control of the human Synapsin I

23   promoter, the PHOX2B genes were subcloned into lentiviral pLL3.7 vector bearing the Synapsin

1 I promoter. Plasmids bearing the HA-PHOX2B mut (+7Ala) gene were constructed applying

2 Gibson assembly and resulted in plasmids pcDNA3.0-HA-PHOX2B mut (+7 Ala), pLL3.7-hSyn-

3 HA-PHOX2B mut (+7 Ala) and pLL3.7-hSyn-HA-PHOX2B mut (+7 Ala)-EGFP.

4 The integrity of every construct used in this study was verified by the sequencing analysis and the

5 detailed explanations of cloning procedures will be provided upon request.

6 **Cell lines and transfection**

7 Cell lines used in this study include human embryonic kidney cells, HEK293T (ATCC CRL-1573)

8 and HEK293FT (Invitrogen, R70007). The cells were authenticated by STR profiling and were

9 routinely tested for mycoplasma contamination. The cells were grown in Dulbecco's modified

10 Eagle's medium (01-052-1A, Biological Industries) supplemented with 10% heat-inactivated fetal

11 bovine serum (04-007-1A, Biological Industries), 10000 units/mL penicillin, and 10 mg/mL

12 streptomycin (03-031-1B Biological Industries) and 2 mM L-glutamine (G7513, SIGMA) at 37

13 °C with 5% CO_2. Cells were seeded and cultured for approximately 24 h until they grew to 50–

14 60% confluence before transfection. Transient transfection of indicated plasmids, was

15 accomplished using TransIT-LT1(Mirus, MIR 2300) according to the manufacturer's protocol.

16 Vector and Mirus were mixed in a reduced serum medium (Opti-MEM® 10001865, Gibco) and

17 incubated at room temperature for up to 30 min before being dripped gently onto the cell culture

18 and incubated for 24 to 48 h. Transfection efficiency was confirmed by western blot analysis. In

19 RNA interference experiments, cells were transfected 24 h after seeding with 50-100 nM

20 *SMART*pool siRNAs (Dharmacon) for gene silencing and Lipofectamine 2000 (1000186,

21 Invitrogen), with two rounds of knockdown for 5 days, according to the manufacturer's instructions

22 (Invitrogen). For this purpose, the siRNAs and Lipofectamine were diluted separately in reduced

23 serum medium Opti-MEM® (10001865, Gibco), then mixed for 15 min at room temperature and

15

1   dripped gently onto the cell culture, which was then incubated at 37 °C for 4-6 h, before restoration

2   of full medium. The following oligonucleotides (ON-TARGETplus SMARTpool, Dharmacon L-

3   006403-01-0005) were used for UBA6 depletion:

4   siRNA J-006403-09, UBA6: Target Sequence: GUGUAGAAUUAGCAAGAUU

5   siRNA J-006403-10, UBA6: Target Sequence: GCAUAGCUGUCCAAGUUAA

6   siRNA J-006403-11, UBA6: Target Sequence: CAGUGUUGUAGGAGCAAUA

7   siRNA J-006403-12, UBA6: Target Sequence: GGAAUUUGGUCAGGUUAU

8   **Isolation and culture of mouse primary cortical neurons**

9   Primary cortical neurons were isolated from wild-type C57BL/6J mouse embryos at E17. Brains

10   were harvested and placed in ice-cold HBSS under a dissection microscope. Cerebral cortices were

11   dissected and incubated in HBSS. After mechanical dissociation using sterile micropipette tips,

12   dissociated neurons were resuspended and cultured at 37°C in a humidified incubator with 5%

13   CO_2 and 95% O_2 in poly-D-lysine coated 6-well plates in neurobasal media (12349015, Gibco)

14   supplemented with 1% GlutaMAX™ Supplement (35050-061, Thermofisher), 1% Sodium

15   pyruvate (11360039, Gibco), 2% B27 supplement (17504044, Gibco) and 1% Penicillin-

16   Streptomycin (03-031-1B, Biological Industries). One-half of the culture media was changed every

17   three days until treatment. Differentiated cortical neurons were infected with indicated lentiviral

18   vectors after 5 days in culture.

19   **Lentivirus production and infection**

20   Third generation lentiviral vectors pLL3.7 (Addgene, #11795) that express shRNA under the

21   mouse U6 promoter and CMV-EGFP or hSyn-EGFP reporter cassettes were obtained from the

22   TAU Viral Core facility. Helper plasmids pMDLg/pRRE, pRSV-Rev, and pCMV-VSVG that

16

1   carry HIV regulatory protein genes as well as the pseudotyped envelope protein gene from

2   vesicular stomatitis virus envelope G (VSV G), were obtained from the same facility.

3   Briefly, HEK-293FT packaging cells growing in 15 cm dishes were transfected with a mix

4   of 7.8 µg helper vector pMDLg/pRRE, 3 µg helper vector pRSV-Rev, 4.2 µg envelope vector

5   pCMV-VSVG, and 12 µg target vector pLL3.7-hSyn-PHOX2B-EGFP carrying the wild type or

6   mutant PHOX2B gene. Polyethylenimine (PEI), (SIGMA 408727) was used as a transfection

7   reagent. At 16 hours' post-transfection, the culture media was removed and replaced with fresh

8   high-serum medium, which was harvested 48 h later and filtered through Amicon Ultra -15

9   (UFC910024) vials at 1500g for 30 minutes to obtain concentrated and purified lentiviruses for

10   transduction. On the day of transduction, half the culture media was removed from primary

11   cultured cortical neurons and was stored for later use (conditioned media), and then viral particles

12   were added and incubated for 15 hours. At the end of this time, the culture medium was replaced

13   by a 1:1 mix of fresh and the conditioned medium that was collected previously. The neurons were

14   then incubated for five more days before being harvested for either western blot or

15   immunostaining.

16   **Generation of patient-specific fibroblasts**

17   Skin biopsies were collected from a 2-year-old female CCHS patient who carries a 20/27 Ala

18   PHOX2B genotype as well as from her healthy sister and from the healthy father of two previously

19   reported identical 4-year-old twin males with CCHS who carry 20/25 Ala stretches [32]. The biopsies

20   were dissected and cultured for two weeks in 6-well plates under a coverslip in DMEM (Biological

21   industries 011701A) with 20% FBS and with half the media replaced every other day.

22

23

Reprogramming of iPSCs

One million fibroblast cells were harvested using TrypLE Express (Gibco, 12604021) and electroporated with non-integrating episomal vectors using a Neon transfection system (Invitrogen, kit MPK10096). The cells were then plated on mouse embryonic fibroblast (MEF)-coated plates and cultured in DMEM with 15% FBS, 5 ng/ml basic fibroblast growth factor (bFGF, Peprotech 10018B) and 5 μM ROCK inhibitor (Enzo, ALX270333). After two days, the medium was replaced with NutriStem (BI) supplemented with 5 ng/ml bFGF, with fresh medium added every other day. On day 22, six colonies were transferred to new MEF-coated plates and cultured in NutriStem with 5 ng/ml bFGF. Three colonies were selected and manually transferred to matrigel (Corning)-coated plates and cultured in NutriStem, which was replaced daily and with weekly passage.

iPSC characterization

iPSCs were assessed for the expression of the pluripotency markers NANOG, SOX2, OCT3/4, TRA 1-60, and SSEA by immunocytochemistry and FACS analysis. The differentiation potential was assessed by harvesting the iPSCs at confluency using TrypLE and resuspending the cells in NutriStem supplemented with 10 ng/ml bFGF and 7 μM ROCK inhibitor (Enzo, ALX270333). Embryoid bodies (EBs) spontaneously formed after 2 days, at which time, the medium was replaced with EB medium (DMEM with 15% FBS, 1% Non-Essential Amino Acids and 0.1 mM β-mercaptoethanol, Gibco 31350010). After 4–7 days, the EBs were plated on 0.1% gelatin-coated plates and cultured for 21 days with EB medium replacement twice weekly. On day 21, the cells were fixed and stained for heavy chain neurofilament, α-SMA, or α-fetoprotein as ectodermal, mesodermal and endodermal markers, respectively. G-banding karyotype analysis was used to exclude any chromosomal abnormalities that may have occurred during the reprogramming

18

1   process. Briefly, iPSCs were supplemented with 100 ng/ml colcemid (Biological Industries

2   120041D), incubated for 60 min, and harvested in Versene solution (Gibco 15040033). Cells were

3   fixed in 1:3 glacial acetic acid:methanol (Biolabs-chemicals) solution and the G-banding

4   karyotype was determined. Finally, all lines were tested for mycoplasma contamination using the

5   Hy-mycoplasma PCR kit (Hylabs, KI5034I).

6   **Differentiation of iPSCs into autonomic neurons**

7   The protocol was as described previously [42]. The iPSCs were cultured in Nutristem to a confluence

8   of 1 million cells/well. On day 0, a single cell suspension was prepared using Versene solution

9   (Gibco, 15040033), and then the cells were quickly re-aggregated in T25 flasks coated with Poly-

10   Hema (2-hydroxyethyl methacrylate) at a concentration of 350,000 cells/ml and were cultured in

11   neuromesodermal progenitor cell induction medium (NMP media) containing Essential 6 media

12   (Gibco, A1516401), 1.5 mM CHIR (Tocris biotech, 4423), 10 μM SB (Tocris biotech, 431542),

13   Penicillin-Streptomycin-Amphotericin B solution (PSA, Biological Industries, 03-033-1B), and

14   10 μM ROCK inhibitor (ROCKi Enzo, ALX270333).

15   On day 1, half the media were replaced with medium without ROCKi, and on day 3, the

16   media were replaced with sympathetic neural crest induction medium (NCi medium) containing

17   Essential 6 media, 1.5 mM CHIR, 20 ng/ml bFGF (Peprotech 10018B), 50 ng/ml BMP4 (Prospec,

18   Cyt-1093), 100 nM all trans retinoic acid (SIGMA, R2625), and PSA. On day 10, the culture was

19   dissociated into single cells using Accutase (Gibco, A1110501) for 4 min at 37°C, and cultured in

20   sympathetic neuroblast induction and propagation (NCC) medium, containing neurobasal media

21   (Gibco 21103049), 20 ng/ml bFGF, 50 ng/ml BMP4, 20 ng/ml EGF, 2 μg/ml heparin, B27 (Gibco

22   12587010), N2 (Gibco 17502048), GlutaMAX (Gibco 35050038) and PSA. On day 11, half the

23   NCC media was replaced by medium without ROCKi and half the media were subsequently

1 replaced every other day. On day 17, the medium was replaced with sympathetic neuronal

2 maturation medium (NMM medium) containing neurobasal media, B27, N2, 10 ng/ml GDNF

3 (Peprotech 45010), 10 ng/ml BDNF (Peprotech 45002), 10 ng/ml NGF (R & D 256-GF),

4 GlutaMAX, and PSA. One third of the media were subsequently replaced every other day until

5 day 31.

6 **Immunocytochemistry for iPSCs- derived autonomic neurons**

7 In order to prepare for the culture, coverslips were incubated in poly-L-ornithine solution (SIGMA,

8 P3655) overnight at 37°C followed by 3 washes with cell culture grade water and drying for

9 15 min. On seeding day, the coverslips were incubated in laminin (Merk, L2020) d for 1 h at 37°C

10 before being washed with PBS. The neurospheres were harvested manually and seeded on the

11 coverslips. After 5-4 days, the cells on the coverslips were washed with Dulbecco's Phosphate-

12 Buffered Saline (DPBS, Biological Industries, 020231A) and fixed in 4% paraformaldehyde at

13 room temperature for 15 min. The cells then were washed twice in DPBS, and blocked with DPBS

14 containing 0.1% Triton X-100 (SIGMA, T8532) and 1% bovine serum albumin (blocking solution,

15 SIGMA, A7906100G), for 1 h at RT. Primary antibodies were added to the blocking solution and

16 incubated overnight at 4°C. Following three washes with blocking solution, the cells were

17 incubated with fluorescent secondary antibodies for 2 h at room temperature. DRAQ5 was used to

18 stain the cell nuclei.

19 **Flow cytometry**

20 iPSCs were harvested and dissociated into single cells by incubation with TrypLE for 2 min at

21 37°C. For the detection of intracellular markers, samples were incubated in fixation solution

22 (Invitrogen, 00522356, 00512343) for 40 min at room temperature followed by washings with

23 permeabilization solution (Invitrogen, 00833356). For surface markers, the cells were washed once

1   with 3% FBS in DPBS. In both cases, the samples were incubated with the appropriate primary

2   antibodies for 2 h at room temperature, then were washed twice and incubated for 1 h with the

3   relevant secondary antibody followed by three more washes. Analysis was performed using a

4   NovoCyte flow cytometer (ACEA). The first gating was SSC-H/FSC-H, and the entire cell

5   population was selected (without cell debris) followed by FSC-H/FSC-A gating and single cell

6   selection. The final analysis is count (%)/FITC-H.

7   **Western blotting assay**

8   Cells were washed with PBS and harvested in Laemmli buffer containing 5% beta-

9   mercaptoethanol. For the ubiquitin loading assays, the cells were lysed on ice in lysis buffer (20

10   mM Tris-HCl, pH 6.8, 137 mM NaCl, 1 mM EGTA, 1% Triton x100, 10% glycerol, and a protease

11   inhibitors cocktail), centrifuged to discard the cell pellet and then the supernatant was added to

12   Laemmli buffer at a ratio of 1:1 without using beta-mercaptoethanol. Protein samples were boiled

13   for 5 min at 95°C, separated by SDS–PAGE, transferred onto PVDF membranes, subjected to

14   western blot analysis, and visualized using the ECL enhanced chemiluminescence reagent

15   (CYANAGEN). Protein levels in each sample were evaluated by normalization to the

16   housekeeping β-actin. The bands were quantified using ImageJ software.

17   **Immunoprecipitation and ubiquitination assays**

18   Cells were lysed on ice in IP buffer (20 mM Tris-HCl, pH 7.2, 150 mM NaCl, 2 mM $MgCl_2$,

19   0.5% NP-40), supplemented with a protease inhibitors cocktail before use. For ubiquitination

20   experiments, cells were treated with a proteasome inhibitor MG132 (10 μM) during the last 6 hours

21   before lysis with the IP buffer supplemented with 1 mM PMSF and 10 mM iodoacetamide. Whole-

22   cell lysates obtained by centrifugation were incubated with 2-5 μg of antibody overnight at 4°C,

23   followed by 2 h incubation with Protein A-Sepharose CL-4B (GE-Healthcare, 17-0780-01). The

1   immunocomplexes were then washed three times with IP buffer, and boiled at 95°C for 5 minutes

2   in Laemmli sample buffer, before being separated by SDS–PAGE for western blotting assays. For

3   experiments to examine the binding of isolated polyalanine stretches, a pre-clearing step was

4   performed by incubating the whole cell lysates with 25 µl beads for 2 h at 4°C. The beads were

5   then discarded, and the cell lysates were incubated with antibody overnight as already described.

6   **Protein expression, purification, and labeling**

7   All proteins used in this study were overexpressed in Rosetta (DE3, pLysS) *Escherichia coli*

8   (MERCK) cells using 0.4 mM isopropyl 1-thio-D-galactopyranoside (Inalco Pharmaceuticals)

9   induction overnight at 16 °C. Purification of proteins was performed on Ni Sepharose 6 Fast Flow

10 sepharose (Cytiva) in 50 mM HEPES (pH 7.5), 300 mM NaCl, 1 mM TCEP, 10 mM imidazole

11 and 10% (w/v) glycerol and eluted using 250 mM imidazole (pH 7.5) in the same buffer. UBA6

12 and E2 variants were further desalted using PD-10 Columns (Cytiva) into Loading buffer

13 containing 20 mM HEPES (pH 7.5), 150mM NaCl, and 10% (w/v) glycerol. Concentrated protein

14 aliquots were stored at 80 °C. All protein concentrations indicated correspond to total protein and

15 are based on UV absorbance at 280 nm.

16   Fluorescein-5-Maleimide (AnaSpec) was attached to ubiquitin following the directions as

17 previously described [40]. Briefly, proteins in 20mM Tris (pH 7.5), 150mM NaCl and 1mM TCEP

18 were incubated for 2 hours in the presence of the fluorophore at room temperature such that the

19 label : protein ratio would be 4. To quench the reaction, beta-mercaptoethanol was added at a ratio

20 of 10:1 to fluorescein. Fluorescein-labeled ubiquitin was then separated from free dye on PD10

21 desalting columns (Cytiva) and was stored at 80 °C.

22

23

E1 and E2 loading assays

All loading assays were performed at 32 °C in a buffer containing 20 mM HEPES (pH 7.5), 150 mM NaCl and 10% (w/v) glycerol. E1 and E2 loading assays were performed using 3 µM fluorescein ubiquitin, 10 nM to 1 µM E1, 1 µM E2, and 2.5mM concentrations each of ATP and MgCl$_2$. Reactions were stopped using nonreducing SDS-PAGE loading buffer. Samples were separated on 4–20% Tris-Glycine NuPAGE gels (Thermo) in Tris-Glycine buffer. Band detection in E1 and E2 loading assays was performed using an Alliance Q9 imager (Uvitec Cambridge). The different qualitative end-point assays were performed using freshly thawed protein aliquots, and the results obtained were reproducible across at least three different protein batches. All constructs and conditions were carried out in triplicate.

In silico **structural analysis**

We constructed models for UBA6 using the Protein Homology/analogY Recognition Engine V 2.0 (PHYRE2) and the AlphaFold server [43, 44]. Further idealization of the geometry was achieved by five cycles of minimization with Refmac5. The model of full length USE1 was downloaded from the AlphaFold server. Structural alignment of USE1 and UBA6 on Ubc4:UBA1:Ub complex (PDB code 4II2) [45] provided a structural model for the USE1:UBA6 complex. Structure visualization and figures preparations were performed with PyMOL (Molecular Graphics System, http://www.pymol.org). We employed the continuum solvation method APBS (Adaptive Poisson–Boltzmann Solver) with the CHARMM force field to calculate the electro-potential surface of UBA1 and UBA6 [46].

Confocal microscopy

The cells were grown on coverslips, and then washed and fixed in 4% Paraformaldehyde for 10-15 minutes before being permeabilized with 0.1% Triton X-100. A solution of, 1% BSA in PBS

1 was used to block both primary and secondary antibodies. The primary antibody was added at a

2 ratio of 1:100 and incubated for at least 1 h at room temperature while the secondary antibody

3 (1:300 Invitrogen) was allowed to incubate with the sample for 30 min at room temperature.

4 Neurons were permeabilized in 2% BSA + 0.1% Triton, and blocked with 2% BSA and the primary

5 antibodies were incubated overnight at 4°C, at a ratio of 1:150, with the secondary antibody

6 incubated for 2 h at room temperature at 1:500. A Zeiss 710 confocal microscope was utilized for

7 confocal imaging with a 63X oil-immersion lens. Nuclear staining was detected by staining with

8 DRAQ5. For quantification, the operator was blinded to the outcome of the experiment when

9 selecting suitably similar fields to image for subsequent computerized analysis. For the

10 colocalization experiments, the association of PHOX2B and UBA6 outside the nucleus was

11 measured by selecting PHOX2B positive cells manually using Fiji, and excluding the nuclei by

12 segmenting and removing the nuclear channel. The colocalization between the channels of interest

13 was then measured using the JACops plugin in Fiji with the default parameters, and Pearson's

14 correlation coefficient was calculated. The cytoplasmic intensity was measured for the channel of

15 interest and the mean gray value was recorded after excluding the nuclei as already described. For

16 the localization experiments related to UBA6 in neurons, the association of UBA6 with the cell

17 body was examined by selecting the neuronal cell bodies manually using Fiji [47]. For the

18 localization of UBA6 with neurites, the cell bodies were selected manually and removed. Arc

19 intensity in the neurons was measured by circulating the neuronal cell body together with the first

20 neurite junction manually using Fiji. The mean integrated value and the area for each neuron were

21 recorded. The value of the mean divided by the area was used for statistical analysis.

22

23

24

Bioinformatics analysis

USE1/UBE2Z human homologs were searched against the Uniprot (ref pubmed id 29425356) and NCBI databases using BLAST (ref pubmed id 2231712). Prosite (ref pubmed id 23161676) was used to scan for alanine residues motifs with between 6 and 10 continuous alanine residues in the BLAST results (a search for proteins containing polyalanine stretches in the ubiquitin cascades). Alignments of the E2 family in vertebrates and across all databases were calculated using MAFFT (ref pubmed id 28968734). The figures were generated using Jalview (ref pubmed id 19151095).

Statistics

Basic data handling was performed in Microsoft Excel. For single comparisons, the statistical significance of the difference between experimental groups was determined using two-tailed Student's t-test with the Prism GraphPad software v.9. Comparisons of multiple means were made by one-way or two-way analysis of variance (ANOVA) followed by the Tukey's, Dunnett's and Sidak's post hoc tests to determine statistical significance. Differences were considered statistically significant for $P < 0.05$. Sample sizes were chosen based on extensive experience with the assays we have performed. The experiments were appropriately randomized. For primary neurons transduced with lentiviral vectors, we used independent cultures prepared from brains of mouse embryos taken from different females. For iPSC-derived neurons, three independent cultures from different differentiation days were considered for analysis, and for cell-line based experiments, we considered replicates performed in different days. Errors bars shown in the figures are standard errors of the mean (s.e.m).

Data availability

Available protein structures were from PDB code 4II2, 6DC6, 1Y8Q, 2NVU. Structural analysis was supported by AlphaFold Protein Structure Database, protein sequences were from UniProt.

Contributions

A.A developed the study rationale. A.A, G.D.V and G.P wrote the manuscript with inputs from all authors. F.A.S performed and analyzed cell line and primary neuron experiments and was involved in all other experiments of the study. G.D.V and D.F designed iPSC experiments, D.F performed the experiments involving iPSCs, Y.B performed the molecular biology and in vitro ubiquitination experiments with assistance from A.K, G.P generated and analyzed the structural models. A.A. supervised the study.

Competing interests

The authors have no competing interests

References

1. Amiel, J. *et al.* Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome. *Nat Genet* **33**, 459-461 (2003).

2. Muragaki, Y., Mundlos, S., Upton, J. & Olsen, B.R. Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13. *Science* **272**, 548-551 (1996).

3. Mundlos, S. *et al.* Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. *Cell* **89**, 773-779 (1997).

4. Laumonnier, F. *et al.* Transcription factor SOX3 is involved in X-linked mental retardation with growth hormone deficiency. *Am J Hum Genet* **71**, 1450-1455 (2002).

5. Richard, P. *et al.* Correlation between PABPN1 genotype and disease severity in oculopharyngeal muscular dystrophy. *Neurology* **88**, 359-365 (2017).

6. Brown, L.Y. *et al.* Holoprosencephaly due to mutations in ZIC2: alanine tract expansion mutations may be caused by parental somatic recombination. *Hum Mol Genet* **10**, 791-796 (2001).

7. Jin, J., Li, X., Gygi, S.P. & Harper, J.W. Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. *Nature* **447**, 1135-1138 (2007).

8. Pelzer, C. *et al.* UBE1L2, a novel E1 enzyme specific for ubiquitin. *J Biol Chem* **282**, 23010-23014 (2007).

9. Usdin, K. The biological effects of simple tandem repeats: lessons from the repeat expansion diseases. *Genome Res* **18**, 1011-1019 (2008).

10. Albrecht, A. & Mundlos, S. The other trinucleotide repeat: polyalanine expansion disorders. *Curr Opin Genet Dev* **15**, 285-293 (2005).

11. Amiel, J., Trochet, D., Clement-Ziza, M., Munnich, A. & Lyonnet, S. Polyalanine expansions in human. *Hum Mol Genet* **13 Spec No 2**, R235-243 (2004).

12. Amer-Sarsour, F., Kordonsky, A., Berdichevsky, Y., Prag, G. & Ashkenazi, A. Deubiquitylating enzymes in neuronal health and disease. *Cell Death Dis* **12**, 120 (2021).

27

1   13.   Lee, P.C. *et al.* Altered social behavior and neuronal development in mice lacking the Uba6-Use1

2   ubiquitin transfer system. *Mol Cell* **50**, 172-184 (2013).

3   14.   Matsuura, T. *et al.* De novo truncating mutations in E6-AP ubiquitin-protein ligase gene (UBE3A)

4   in Angelman syndrome. *Nat Genet* **15**, 74-77 (1997).

5   15.   Ye, Y. & Rape, M. Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol* **10**, 755-

6   764 (2009).

7   16.   Schelpe, J., Monte, D., Dewitte, F., Sixma, T.K. & Rucktooa, P. Structure of UBE2Z Enzyme

8   Provides Functional Insight into Specificity in the FAT10 Protein Conjugation Machinery. *J Biol*

9   *Chem* **291**, 630-639 (2016).

10 17.   Bohnsack, R.N. & Haas, A.L. Conservation in the mechanism of Nedd8 activation by the human

11   AppBp1-Uba3 heterodimer. *J Biol Chem* **278**, 26823-26830 (2003).

12 18.   Whitby, F.G., Xia, G., Pickart, C.M. & Hill, C.P. Crystal structure of the human ubiquitin-like

13   protein NEDD8 and interactions with ubiquitin pathway enzymes. *J Biol Chem* **273**, 34983-34991

14   (1998).

15 19.   Schulman, B.A. & Harper, J.W. Ubiquitin-like protein activation by E1 enzymes: the apex for

16   downstream signalling pathways. *Nat Rev Mol Cell Biol* **10**, 319-331 (2009).

17 20.   Lee, I. & Schindelin, H. Structural insights into E1-catalyzed ubiquitin activation and transfer to

18   conjugating enzymes. *Cell* **134**, 268-278 (2008).

19 21.   Huang, D.T. *et al.* Basis for a ubiquitin-like protein thioester switch toggling E1-E2 affinity. *Nature*

20   **445**, 394-398 (2007).

21 22.   Lois, L.M. & Lima, C.D. Structures of the SUMO E1 provide mechanistic insights into SUMO

22   activation and E2 recruitment to E1. *EMBO J* **24**, 439-451 (2005).

23 23.   Akimoto, G., Fernandes, A.P. & Bode, J.W. Site-Specific Protein Ubiquitylation Using an

24   Engineered, Chimeric E1 Activating Enzyme and E2 SUMO Conjugating Enzyme Ubc9. *ACS Cent*

25   *Sci* **8**, 275-281 (2022).

24. Kishino, T., Lalande, M. & Wagstaff, J. UBE3A/E6-AP mutations cause Angelman syndrome. *Nat Genet* **15**, 70-73 (1997).

25. Glessner, J.T. *et al.* Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* **459**, 569-573 (2009).

26. Greer, P.L. *et al.* The Angelman Syndrome protein Ube3A regulates synapse development by ubiquitinating arc. *Cell* **140**, 704-716 (2010).

27. Kuhnle, S., Mothes, B., Matentzoglu, K. & Scheffner, M. Role of the ubiquitin ligase E6AP/UBE3A in controlling levels of the synaptic protein Arc. *Proc Natl Acad Sci U S A* **110**, 8888-8893 (2013).

28. Dauger, S. *et al.* Phox2b controls the development of peripheral chemoreceptors and afferent visceral pathways. *Development* **130**, 6635-6642 (2003).

29. Pattyn, A., Morin, X., Cremer, H., Goridis, C. & Brunet, J.F. The homeobox gene Phox2b is essential for the development of autonomic neural crest derivatives. *Nature* **399**, 366-370 (1999).

30. Vanderlaan, M., Holbrook, C.R., Wang, M., Tuell, A. & Gozal, D. Epidemiologic survey of 196 patients with congenital central hypoventilation syndrome. *Pediatr Pulmonol* **37**, 217-229 (2004).

31. Trang, H. *et al.* Proceedings of the fourth international conference on central hypoventilation. *Orphanet J Rare Dis* **9**, 194 (2014).

32. Falik, D. *et al.* Generation and characterization of iPSC lines (BGUi004-A, BGUi005-A) from two identical twins with polyalanine expansion in the paired-like homeobox 2B (PHOX2B) gene. *Stem Cell Res* **48**, 101955 (2020).

33. Ashkenazi, A. *et al.* Polyglutamine tracts regulate beclin 1-dependent autophagy. *Nature* **545**, 108-111 (2017).

34. Paulson, H.L., Shakkottai, V.G., Clark, H.B. & Orr, H.T. Polyglutamine spinocerebellar ataxias - from genes to potential treatments. *Nat Rev Neurosci* **18**, 613-626 (2017).

35. Dubreuil, V. *et al.* A human mutation in Phox2b causes lack of CO2 chemosensitivity, fatal central apnea, and specific loss of parafacial neurons. *Proc Natl Acad Sci U S A* **105**, 1067-1072 (2008).

36. Ramanantsoa, N. *et al.* Ventilatory response to hyperoxia in newborn mice heterozygous for the transcription factor Phox2b. *Am J Physiol Regul Integr Comp Physiol* **290**, R1691-1696 (2006).

37. Basu, S. *et al.* Unblending of Transcriptional Condensates in Human Repeat Expansion Disease. *Cell* **181**, 1062-1079 e1030 (2020).

38. Chen, Z. *et al.* Down-regulation of UBA6 exacerbates brain injury by inhibiting the activation of Notch signaling pathway to promote cerebral cell apoptosis in rat acute cerebral infarction model. *Mol Cell Probes* **53**, 101612 (2020).

39. Shilling, P.J. *et al.* Improved designs for pET expression plasmids increase protein production yield in Escherichia coli. *Commun Biol* **3**, 214 (2020).

40. Keren-Kaplan, T. *et al.* Synthetic biology approach to reconstituting the ubiquitylation cascade in bacteria. *EMBO J* **31**, 378-390 (2012).

41. Kamitani, T., Kito, K., Nguyen, H.P. & Yeh, E.T. Characterization of NEDD8, a developmentally down-regulated ubiquitin-like protein. *J Biol Chem* **272**, 28557-28562 (1997).

42. Kirino, K., Nakahata, T., Taguchi, T. & Saito, M.K. Efficient derivation of sympathetic neurons from human pluripotent stem cells with a defined condition. *Sci Rep* **8**, 12865 (2018).

43. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. & Sternberg, M.J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858 (2015).

44. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).

45. Olsen, S.K. & Lima, C.D. Structure of a ubiquitin E1-E2 complex: insights to E1-E2 thioester transfer. *Mol Cell* **49**, 884-896 (2013).

46. Baker, N.A., Sept, D., Joseph, S., Holst, M.J. & McCammon, J.A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **98**, 10037-10041 (2001).

47. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012).

1

Figure 1

2

3

4

1
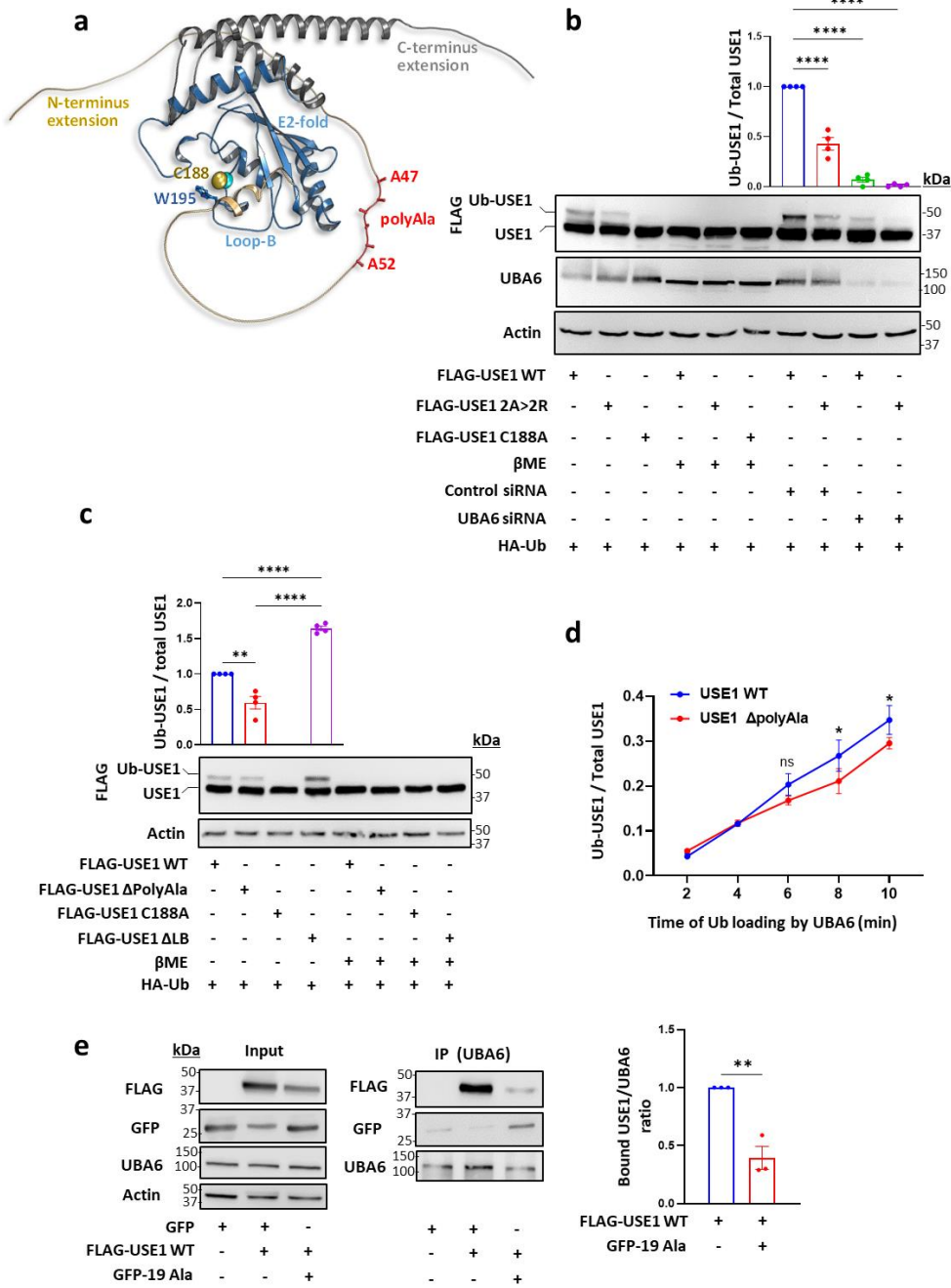
2   **Figure 1. A polyalanine stretch in USE1 regulates UBA6-USE1 ubiquitin transfer. a**,

3   Structure of the USE1 enzyme (PDB 5A4P) indicating the catalytic cysteine (Cys188), loop B

4   (LB) with Trp195 masking Cys188, and model of the C-and N-terminal extensions including the

5   alanine repeats (created by AlphaFold). **b-c**, FLAG-wild type (WT) USE1, FLAG-USE1 mutants

6   with aberrations in the polyalanine (2A> 2R, ΔPolyAla), FLAG-USE1 catalytic dead mutant

7   (C188A), and FLAG-USE1 ΔLB were co-expressed with HA-Ub in control or UBA6-depleted

8   (*Uba6* small inhibitory RNA, siRNA) HEK293T cells. Cell lysates were incubated with or without

9   β mercaptoethanol (βME) and analyzed for ubiquitin loading. Results are mean ± s.e.m normalized

10 to control WT USE1. One-way ANOVA Dunnet's test (Fig.1b), Tukey's test (Fig.1c), n=4. **d**,

11 Time-dependent *in vitro* ubiquitin loading of WT and ΔPolyAla USE1 by UBA6. Quantification

12 of USE1 ubiquitin loading is presented as mean ± s.e.m. Two-way ANOVA Sidak's test, n=3. **e**,

13 FLAG-WT USE1, empty GFP and GFP-polyAla (19Ala) constructs were transfected into

14 HEK293T cells. Cell lysates were immunoprecipitated (IP) with anti-UBA6 antibodies and the

15 immunocomplexes were analyzed with anti-FLAG, anti-UBA6 and anti-GFP antibodies. The

16 bound USE1/UBA6 ratio is shown. Unpaired 2-tailed t-test, n=3. ns non-significant, * $P < 0.05$,

17 ** $P < 0.01$, **** $P < 0.0001$.

18

19

32
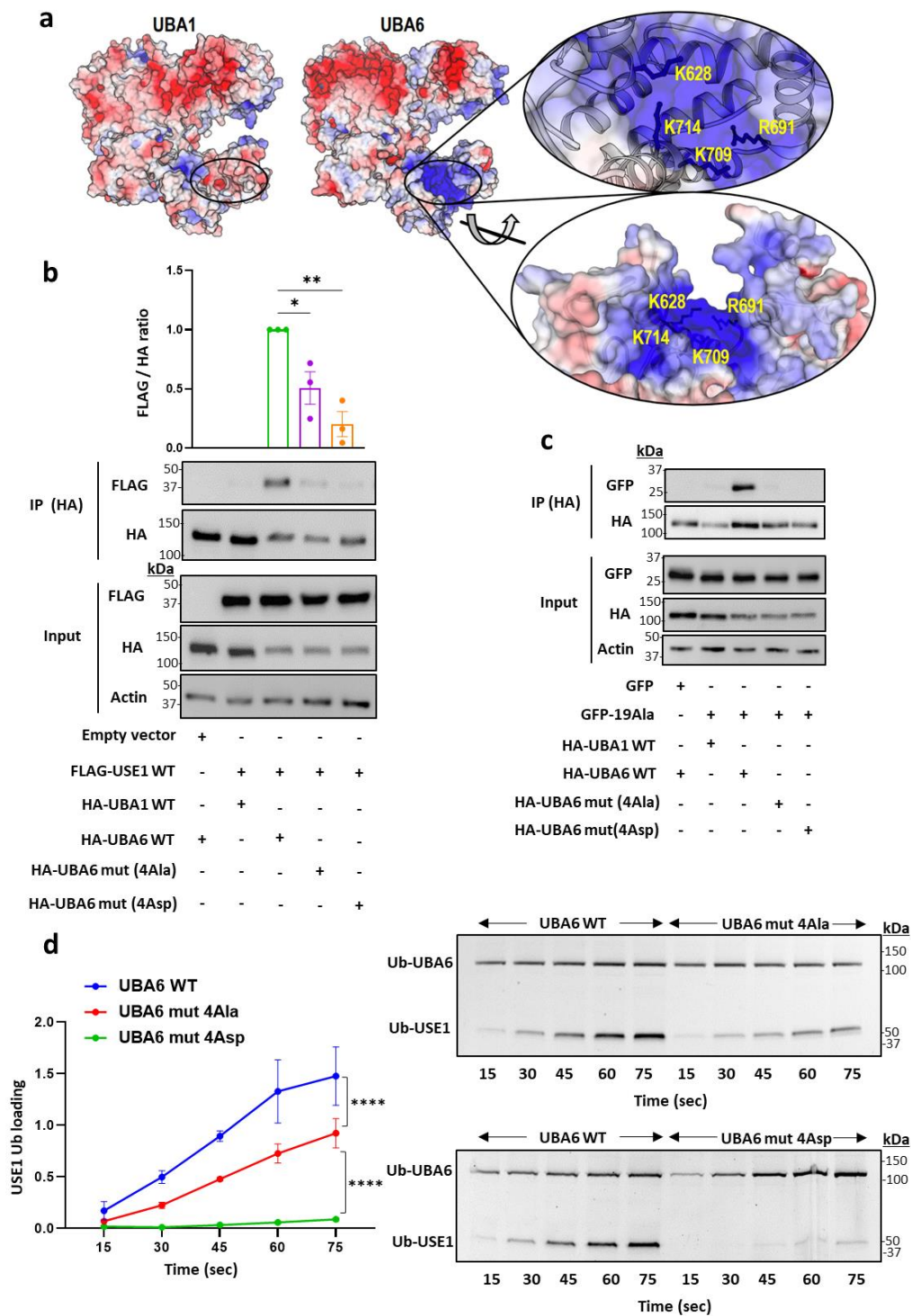
Figure 2

1   **Figure 2. UBA6 interacts with polyalanine stretches. a**, Electrostatic surface representation of

2   the UBA6 structural model in comparison to UBA1. The location of key Arg and Lys residues

3   forming the positively charged patch in UBA6 is presented. **b-c**, HA-tagged constructs of WT

4   UBA1, WT UBA6, and UBA6 mutants with Ala or Asp substitution mutations in Lys628, Arg691,

5   Lys709, and Lys714 (UBA6 mut 4Ala or UBA6 mut 4Asp) were co-transfected with FLAG-USE1

6   (**b**) or empty GFP and GFP-polyAla (19Ala) (**c**) into HEK293T cells. Cell lysates were

7   immunoprecipitated with anti-HA antibodies and the immunocomplexes were analyzed with anti-

8   FLAG, anti-HA, and anti-GFP antibodies. Results are mean ± s.e.m. One-way ANOVA Dunnett's

9   test, n=3. **d**, Representative gel of time-dependent *in vitro* ubiquitin loading of USE1 by WT

10 UBA6, UBA6 mut 4Ala, and UBA6 mut 4Asp in the presence of fluorescein-labelled ubiquitin.

11 Imaging of fluorescent ubiquitin conjugates was carried out with a laser scanner at 488 nm.

12 Quantification of USE1 ubiquitin loading is presented as mean ± s.e.m. Two-way ANOVA

13 Tukey's test, n=3. *$P < 0.05$, **$P < 0.01$, **** $P < 0.0001$.

Figure 3
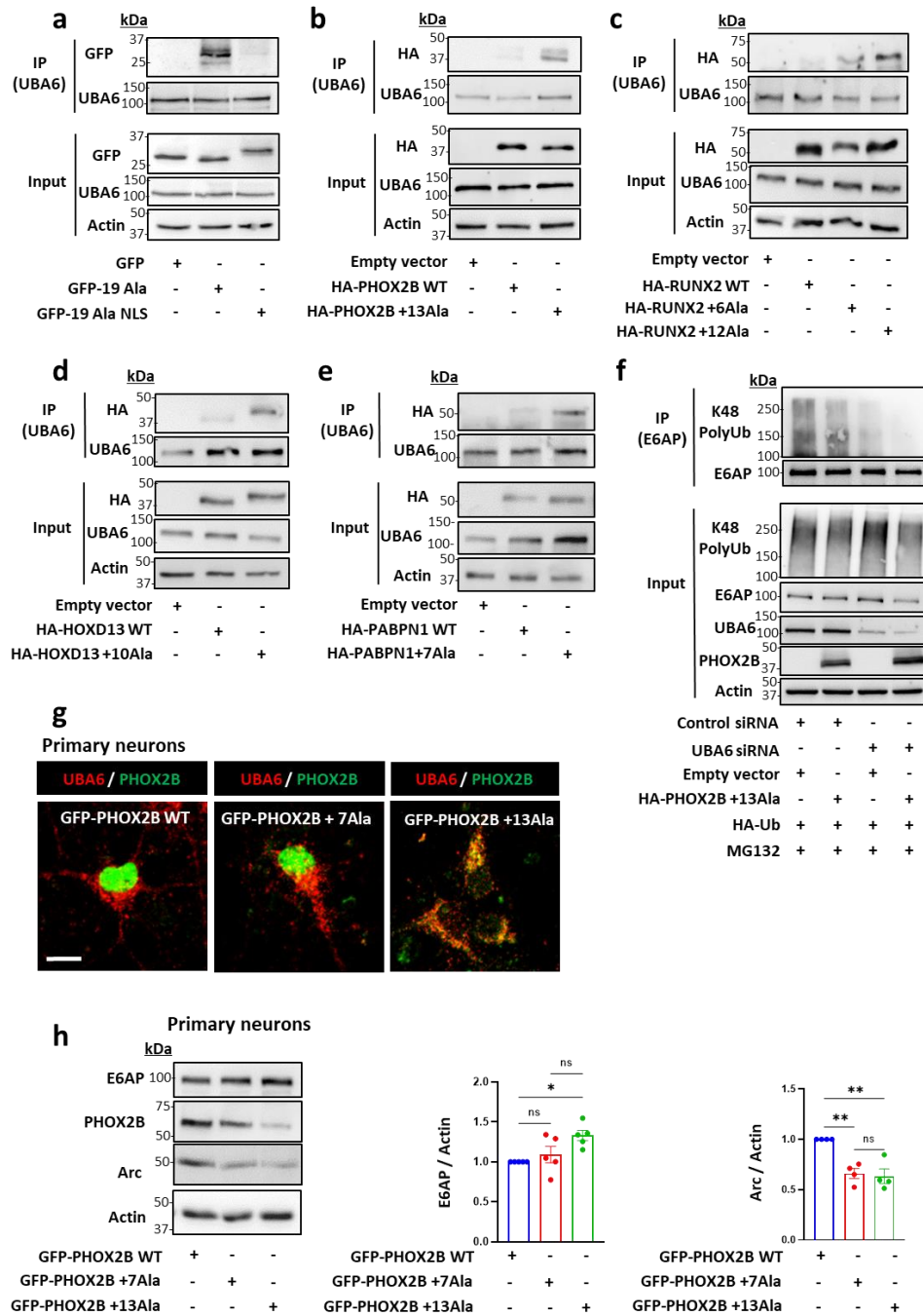
1   **Figure 3. Polyalanine-expanded disease proteins interact with UBA6 and inhibit E6AP**

2   **degradation. a-e**, HEK293T cells were transfected with the indicated constructs and

3   immunoprecipitated for endogenous UBA6. **a**, Empty GFP, and GFP-polyAla with or without a

4   nuclear localization sequence (NLS). **b**, WT and mutant PHOX2B (+13Ala). **c**, WT and mutant

5   RUNX2 (+6Ala and +12Ala). **d**, WT and mutant HOXD13 (+10Ala). **e**, WT and mutant PABPN1

6   (+7 Ala). **f**, Control and UBA6-depleted HEK293T cells were transfected with HA-Ub, mutant

7   PHOX2B or empty vector and incubated for the last 6 h with the proteasome inhibitor MG132

8   (10 μM). Endogenous E6AP was immunoprecipitated from cell lysates for ubiquitination analysis.

9   **g-h,** Mouse primary cortical neurons were transduced with lentiviral vectors expressing GFP-

10   tagged WT PHOX2B or mutant PHOX2B (+7 Ala and +13Ala). **g,** Association of endogenous

11   UBA6 with GFP-PHOX2B. Scale bar 10 μm. Quantification is shown in **Extended Data Fig. 5c**.

12   **h**, Analysis of E6AP and Arc levels in the WT and mutant PHOX2B-expressing neurons (n=5 and

13   n=4, respectively). Results are mean ± s.e.m normalized to WT PHOX2B. One-way ANOVA

14   Tukey's test, ns non-significant, $*P < 0.05$, $**P < 0.01$.
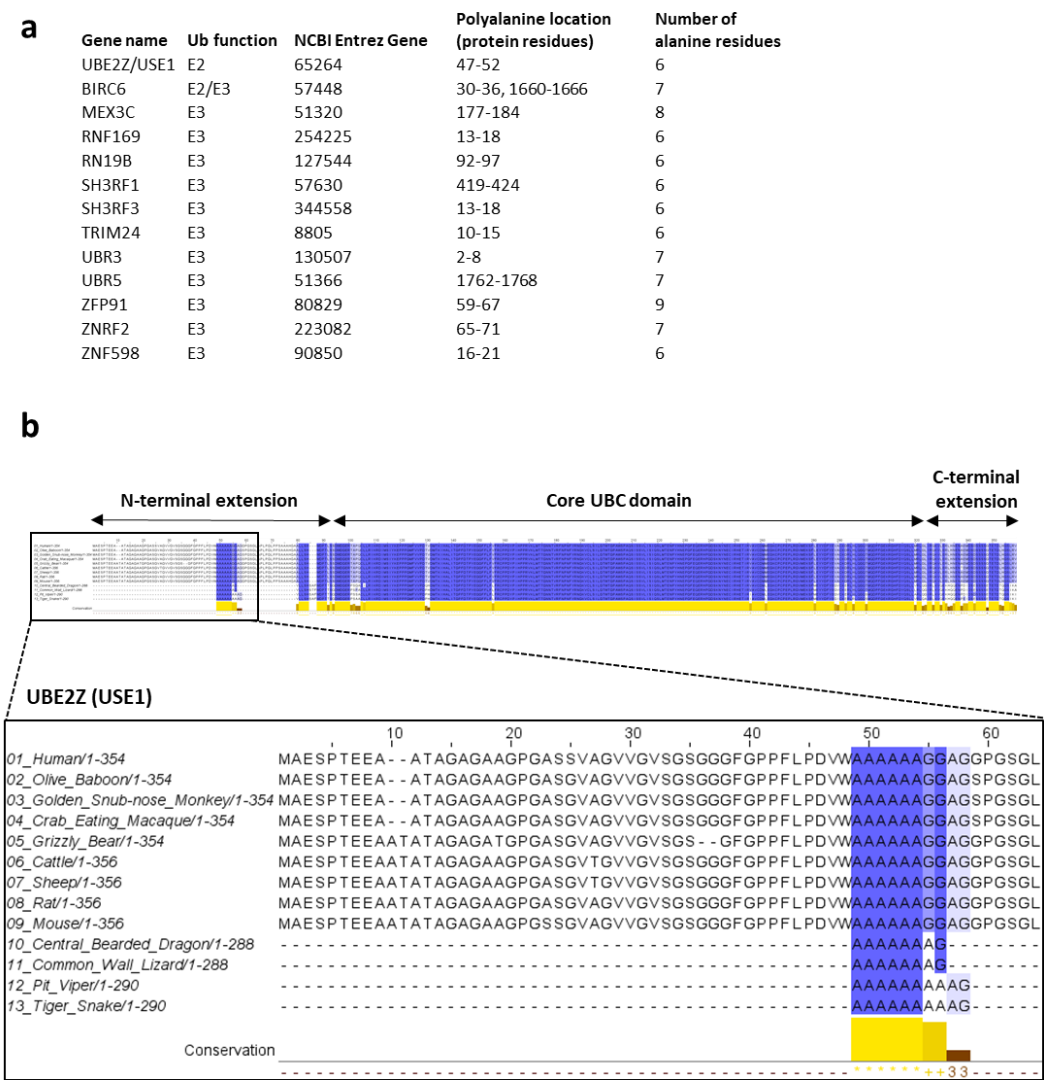
15

16
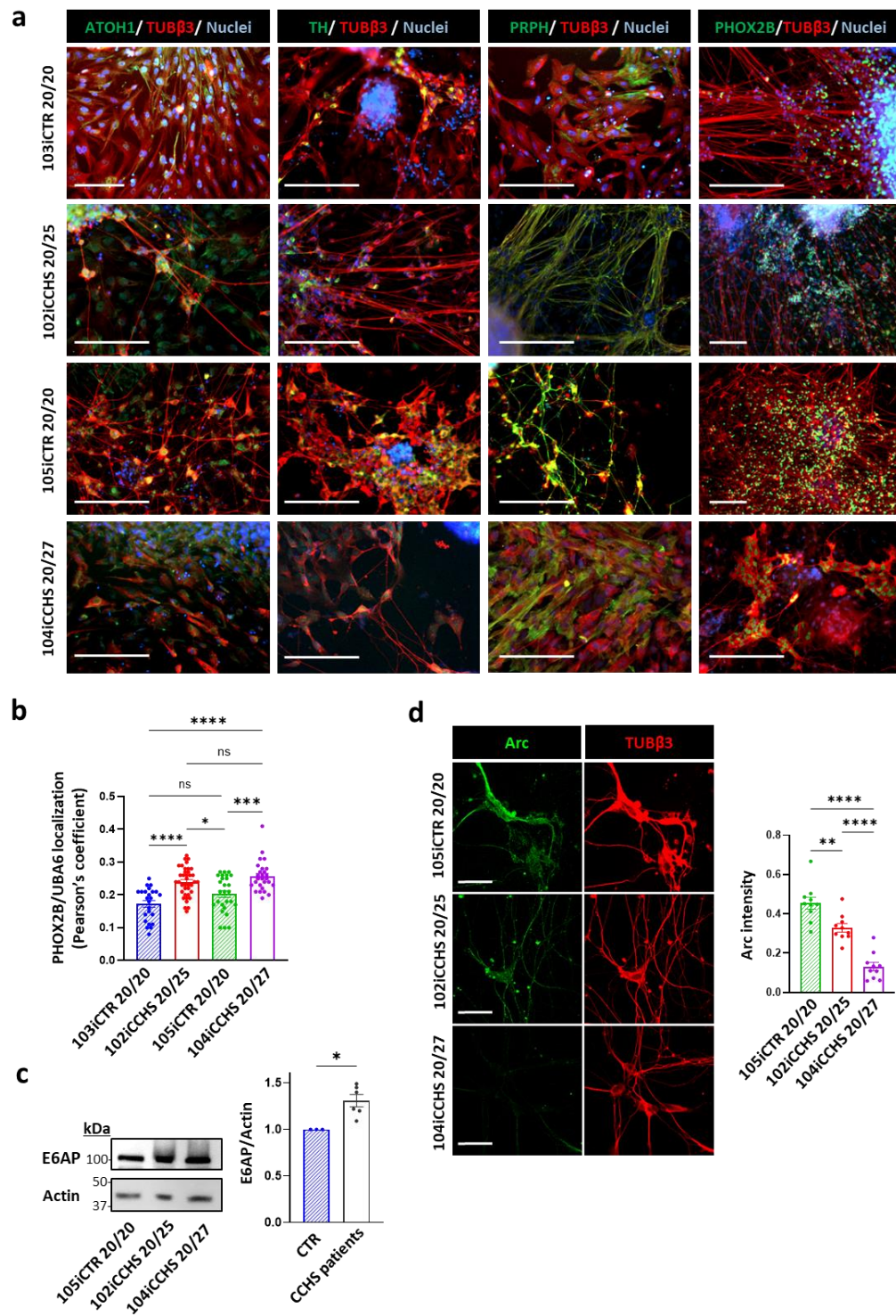
Figure 4

1   **Figure 4. iPSC-derived autonomic neurons from CCHS patients showing UBA6-dependent**

2   **alterations in E6AP and Arc levels. a,** Characterization of iPSC-derived autonomic neurons at

3   day 31 of differentiation from healthy controls and CCHS patients. Immunocytochemistry of

4   PHOX2B, βIII-tubulin (TUBβ3), tyrosine hydroxylase (TH) peripherin (PRPH), and atonal BHLH

5   Transcription Factor 1 (ATOH1). Scale bar 200 μm. **b**, Quantification of the association of

6   endogenous UBA6 with endogenous PHOX2B (Pearson's coefficient) in autonomic neurons from

7   control and CCHS patients. Results are the average values from neurons in different imaged fields.

8   Total number of neurons analyzed was 160 for 103iCTR 20/20, 350 for 102iCCHS 20/25, 380 for

9   105iCTR 20/20 and 550 for 104iCCHS 20/27. One-way ANOVA Tukey's test. Images are shown

10   in **Extended Data Fig. 7c. c,** Analysis of E6AP levels in control (105iCTR 20/20) and CCHS

11   patients (102iCCHS 20/25 and 104iCCHS 20/27). Results are mean ± s.e.m normalized to control

12   from three independent differentiation days. Unpaired 2-tailed t-test. **d,** Immunostaining of Arc

13   and TUBβ3 in autonomic neurons from control and CCHS patients. Scale bar 50μm. For

14   quantification, Arc intensity was normalized to TUBβ3. Results are mean ± s.e.m from 40

15   analyzed neurons. One-way ANOVA Tukey's test, ns non-significant, $*P < 0.05$, $**P < 0.01$,

16   $***P < 0.001$, $****P < 0.0001$.

17

Extended Data Fig. 1

a

Gene name	Ub function	NCBI Entrez Gene	Polyalanine location (protein residues)	Number of alanine residues
UBE2Z/USE1	E2	65264	47-52	6
BIRC6	E2/E3	57448	30-36, 1660-1666	7
MEX3C	E3	51320	177-184	8
RNF169	E3	254225	13-18	6
RN19B	E3	127544	92-97	6
SH3RF1	E3	57630	419-424	6
SH3RF3	E3	344558	13-18	6
TRIM24	E3	8805	10-15	6
UBR3	E3	130507	2-8	7
UBR5	E3	51366	1762-1768	7
ZFP91	E3	80829	59-67	9
ZNRF2	E3	223082	65-71	7
ZNF598	E3	90850	16-21	6

b

Extended Data Fig. 1. Analysis of alanine repeats in the ubiquitin system. **a,** Analysis of alanine repeats domains in the human ubiquitin cascades comprising E1, E2, and E3 enzymes. **b,** A multiple sequence alignment of USE1 homologs from different vertebrates. The alignment is colored according to sequence identity including the N-terminus containing the polyalanine stretch.

39

Extended Data Fig. 2

1
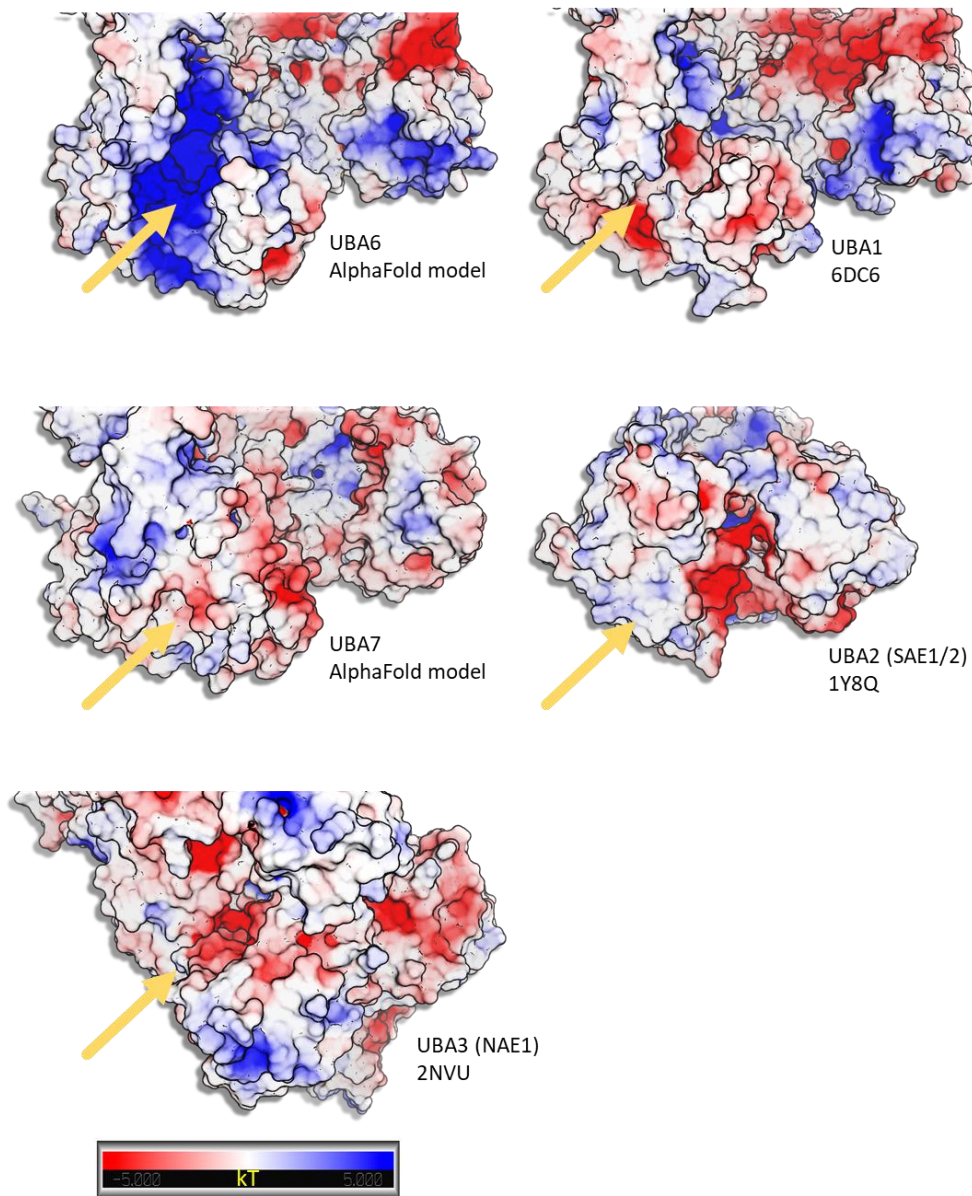
1 **Extended Data Fig. 2. Structures and models of the SCCH domains of the cannonical E1**

2 **ubiquitin-like activating enzymes.** AlphaFold models of UBA6 and UBA7 and the crystal

3 structures of UBA1, UBA2 and UBA3 are shown. The structures were aligned and electrostatic

4 potential was clulated as descibed in the methods. The yellow arrows indicate the location of the

5 groove within the SCCH domains. UBA6, UBA1 and UBA7 form an extended lobe within the

6 SCCH, which is missing in UBA2 and UBA3. The groove in UBA7 is covered and do not exist in

7 UBA2 and UBA3. The grooves in UBA1 and UBA6 are highly similar in terms of structure but

8 present significantly different electro potential surfaces. The gradiant from negative (red) to

9 positve (blue) charge is shown. The figure was prepared by PyMol.

10
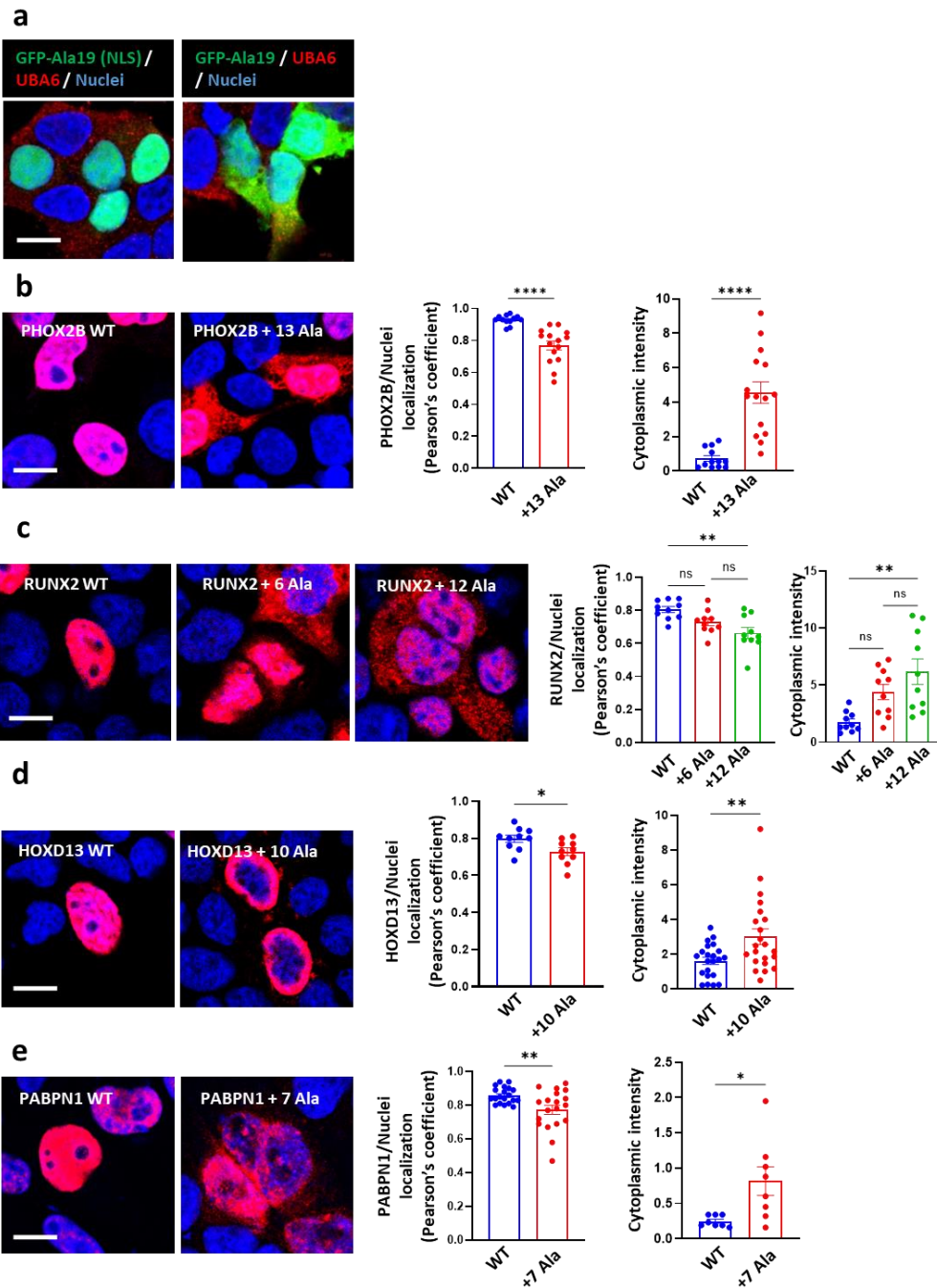
Extended Data Fig. 3

1   **Extended Data Fig. 3. Polyalanine expansion mutations cause cytoplasmic mislocalization of**

2   **different nuclear proteins.** HEK293T cells were transfected with the indicated constructs, and

3   were subjected to immunostaining. **a**, GFP-19Ala with nuclear localization sequence (NLS) or

4   GFP-19Ala, labeled for endogenous UBA6. **b**, HA-PHOX2B WT and HA-PHOX2B +13Ala. **c**,

5   HA-RUNX2 WT, HA-RUNX2 +6Ala, and HA-RUNX2+ 12Ala. **d**, HA-HOXD13 WT and HA-

6   HOXD13 +10Ala. **e**, HA-PABPN1 WT and HA-PABPN1 +7Ala. Image scale bar 20 μm.

7   Quantification of the association of HA-tagged proteins (labeled in red) with the nucleus (labeled

8   in blue, Pearson's coefficient) is indicated as well as the cytoplasmic intensity. Results are the

9   average values from cells in different imaged fields. Between 60-100 transfected cells were

10   analyzed. Results are mean ± s.e.m. Unpaired 2-tailed t-test (**b,d,e**) and one-way ANOVA Tukey's

11   test (**c**). ns non-significant, $*P < 0.05$, $**P < 0.01$, $****P < 0.0001$.

12

43
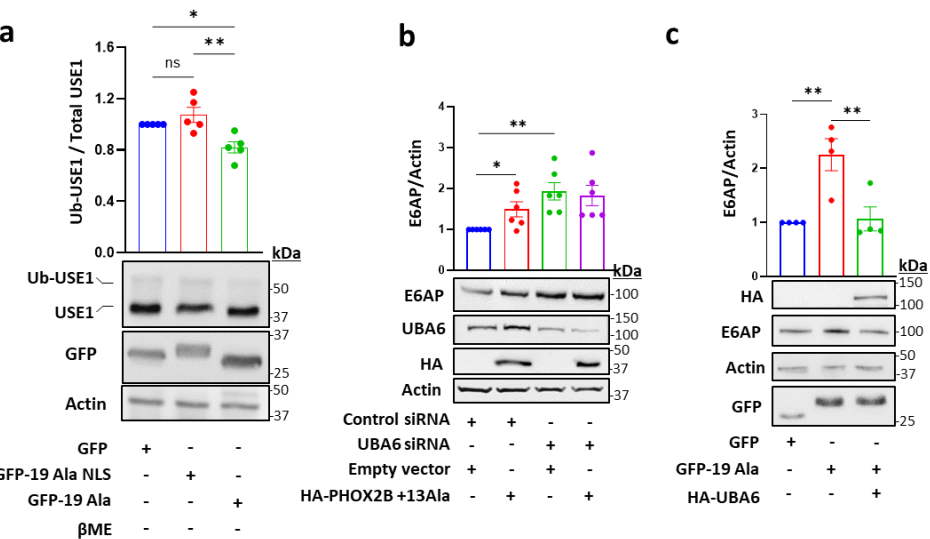
Extended Data Fig. 4

Extended Data Fig. 4. Isolated polyalanine stretches regulate USE1 ubiquitin loading and E6AP levels. a, HEK293T cells were transfected with empty GFP, GFP-polyAla, and GFP-polyAla with a nuclear localization sequence (NLS). Endogenous USE1 ubiquitin loading was analyzed in βME-untreated cell lysates. **b**, Control and UBA6-depleted HEK293T cells were transfected with empty vector or mutant PHOX2B and analyzed for E6AP levels. **c,** Cells were transfected with empty GFP, or GFP-polyAla with or without HA-UBA6, and analyzed for E6AP levels. Results are mean ± s.e.m. (**a**) One-way ANOVA Tukey's test, n=5 (**b**) Unpaired 2-tailed t-test, n=6. (**c**) One-way ANOVA, Tukey's test, n=4. n.s not significant, $*P < 0.05$, $**P < 0.01$.
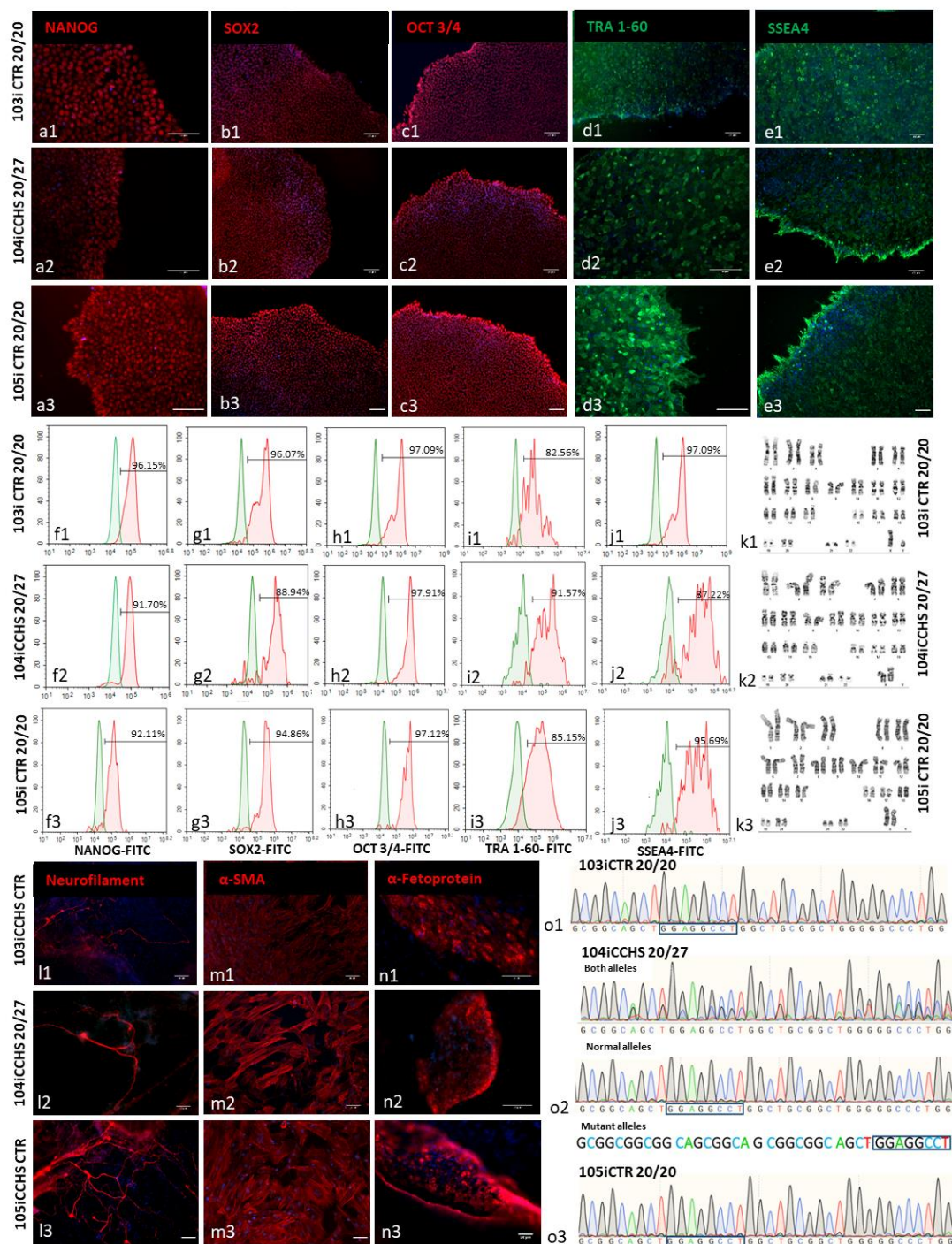
Extended Data Fig. 5
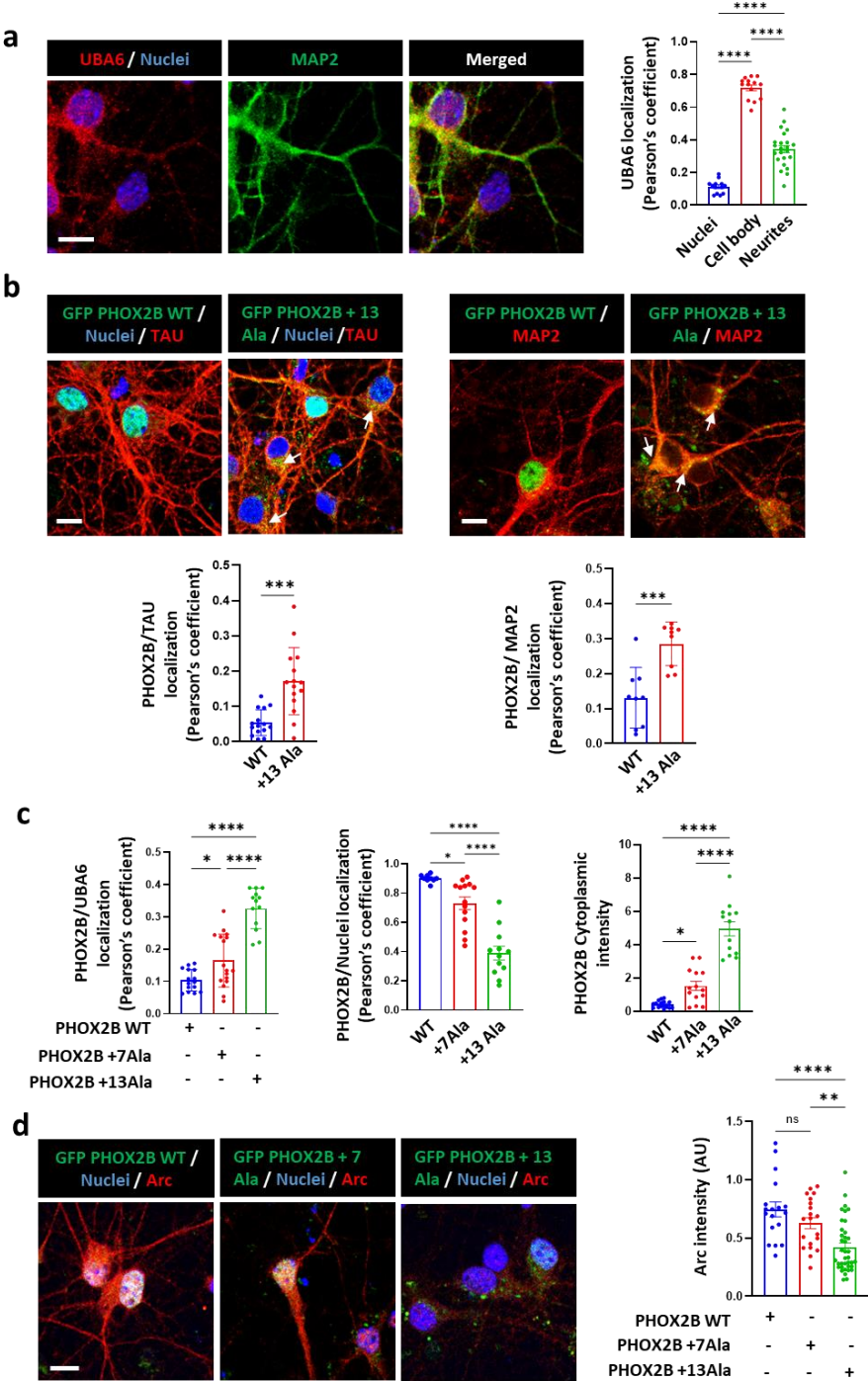
1 **Extended Data Fig. 5. Analysis of UBA6 and mutant PHOX2B association and effects on Arc**

2 **levels in primary neurons. a,** Mouse primary cortical neurons were labeled for endogenous

3 UBA6 and MAP2. Quantification of the association of UBA6 with the neuronal nucleus, cell body,

4 and neurites is shown (Pearson's coefficient). **b**, Mouse primary cortical neurons were transduced

5 with lentiviral vectors expressing GFP-tagged WT PHOX2B or mutant PHOX2B (+13Ala), and

6 were labeled for endogenous MAP2 and TAU (non-nuclear fraction of mutant PHOX2B marked

7 with arrows). The quantification of the association of GFP-PHOX2B with the nucleus, MAP2, and

8 TAU is shown (Pearson's coefficient), as well as GFP-PHOX2B cytoplasmic intensity. **c**.

9 Quantification of UBA6 association with WT and mutant (+7 Ala, +13 Ala) GFP-PHOX2B

10 (Pearson's coefficient) related to **Fig. 3g**. **d,** The primary neurons were transduced with the GFP-

11 PHOX2B lentiviral vectors and were labeled for Arc. The quantification of Arc intensity in the

12 GFP-PHOX2B expressing neurons is presented. Image scale bar 10 μm. Results are mean ± s.e.m

13 representing the average values from neurons in different imaged fields. (**a**) ****$P < 0.0001$ One–

14 way ANOVA Tukey's test. n= 50 neurons analyzed. (**b**) *** $P < 0.001$ unpaired 2-tailed t-test,

15 n=30-50 neurons analyzed for WT PHOX2B and mutant PHOX2B +13Ala. (**c**) *$P < 0.05$,

16 ****$P < 0.0001$ One–way ANOVA Tukey's test. n=30, n=60, and n=90 neurons analyzed for WT

17 PHOX2B, mutant PHOX2B +7Ala and +13Ala, respectively. (**d**) ns non-significant, **$P < 0.01$,

18 ****$P < 0.0001$, one-way ANOVA Tukey's test. n=20, n=30 and n=100 neurons analyzed for WT

19 PHOX2B, mutant PHOX2B +7Ala and +13Ala, respectively.

20

21

22

23
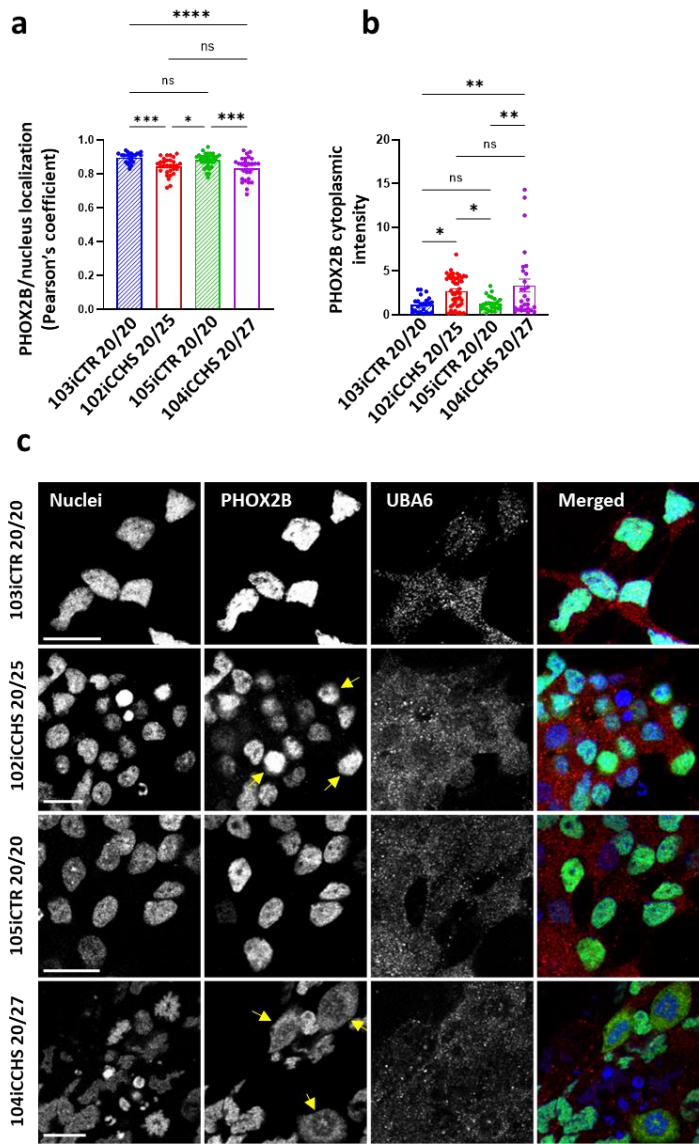
Extended Data Fig. 6

1

2

1

2 **Extended Data Fig. 6. Generation and characterization of CCHS patient and family relative-**

3 **derived iPSCs.** Skin punch biopsies were collected from a 2-year-old female patient with CCHS

4 who harbors a heterozygous 27 polyalanine expansion in PHOX2B (104iCCHS 20/27), and from

5 her healthy sister (105iCTR 20/20). A biopsy was also collected from the healthy father (103iCTR

6 20/20) of the 4-year-old male patients who harbors a heterozygous 25 polyalanine expansion in

7 PHOX2B (102iCCHS 20/25). Patient-specific fibroblasts were electroporated with non-

8 integrating reprogramming episomal plasmids. **a-e**, Immunocytochemistry for pluripotency

9 markers (NANOG, SOX2, OCT3/4 TRA-1-60, SSEA4). **f-j**, Flow cytometry analysis for

10 pluripotency markers (red-NANOG, SOX2, OCT3/4 TRA-1-60, SSEA4). **k**, Karyotype. **l-n**,

11 Embryoid bodies (EBs) were generated and allowed to spontaneously differentiate for 21 days.

12 Differentiated EBs express the ectoderm marker neurofilament, the mesoderm marker α-SMA, and

13 the endoderm marker α-fetoprotein. **o**, Sequencing of the 3rd PHOX2B exon confirms the

14 heterozygous +7 polyalanine expansion in the CCHS patient, but not in the healthy controls. Image

15 scale bar 100 µm.

16

17

48

Extended Data Fig. 7

1 **Extended Data Fig. 7. PHOX2B cytoplasmic mislocalization in patient-derived autonomic**

2 **neurons a-b**, Quantification of the association of endogenous PHOX2B with the nucleus

3 (Pearson's coefficient) in autonomic neurons from control and CCHS patients. Quantification is

4 shown also for PHOX2B cytoplasmic intensity. Results are mean ± s.e.m. *$P < 0.05$, **$P < 0.01$,

5 ***$P < 0.001$ ****$P < 0.0001$ one-way ANOVA Tukey's test. The results represent additional

6 analysis from the same neurons analyzed in **Fig. 4b**. **c**, iPSC-derived human autonomic neurons

7 from control and CCHS patients were labeled by nuclear staining (colored blue), and for

8 endogenous PHOX2B (colored green) and endogenous UBA6 (colored red). Images indicate

9 events of severe cytoplasmic mislocalization of PHOX2B (marked with arrows). Scale bar 10 μm.

10

11

12

13

14

15