1 **Developmental regulation of neuronal gene expression by Elongator complex protein 1**

2 **dosage.**

3

4 Elisabetta Morini[1,2,8], Dadi Gao[1,2,3,8], Emily M. Logan[1], Monica Salani[1], Aram J. Krauson[1], Anil

5 Chekuri[1,2], Yei-Tsung Chen[4], Ashok Ragavendran[1,3], Probir Chakravarty[5], Serkan Erdin[1,3], Alexei

6 Stortchevoi[1,3], Jesper Q. Svejstrup[6,7], Michael E. Talkowski[1,2,3], Susan A. Slaugenhaupt[1,2]*

7

8 [1]Center for Genomic Medicine, Massachusetts General Hospital Research Institute, Boston, MA.

9 [2]Department of Neurology, Massachusetts General Hospital Research Institute and Harvard

10 Medical School, Boston, MA.

11 [3]Program in Medical and Population Genetics and Stanley Center for Psychiatric Research, Broad

12 Institute of Harvard and MIT, Cambridge, MA.

13 [4]Department of Life Sciences and Institute of Genome Sciences, National Yang Ming Chiao Tung

14 University, Taiwan.

15 [5] Bioinformatics and Biostatistics, The Francis Crick Institute, London, UK

16 [6]Mechanisms of Transcription Laboratory, The Francis Crick Institute, London, UK.

17 [7]Department of Cellular and Molecular Medicine, Panum Institute, University of Copenhagen,

18 Copenhagen, Denmark.

19 [8]These authors contributed equally to this work.

20 *Correspondence should be addressed to S.A.S. (slaugenhaupt@mgh.harvard.edu).

21

22 **Abstract**

23   Elongator is a highly conserved protein complex required for transcriptional elongation,

24   intracellular transport and translation. Elongator complex protein 1 (ELP1) is the scaffolding

25   protein of Elongator and is essential for its assembly and stability. Familial dysautonomia (FD), a

26   hereditary sensory and autonomic neuropathy, is caused by a mutation in *ELP1* that lead to a tissue-

27   specific reduction of ELP1 protein. Our work to generate a phenotypic mouse model for FD led to

28   the discovery that homozygous deletion of the mouse *Elp1* gene leads to embryonic lethality prior

29   to mid-gestation. Given that FD is caused by a reduction, not loss, of ELP1, we generated two new

30   mouse models by introducing different copy numbers of the human FD *ELP1* transgene into the

31   *Elp1* knockout mouse (*Elp1$^{-/-}$*) and observed that human *ELP1* expression rescues embryonic

32   development in a dose dependent manner. We then conducted a comprehensive transcriptome

33   analysis in mouse embryos to identify genes and pathways whose expression correlates with the

34   amount of *ELP1*. We found that *ELP1* is essential for the expression of genes responsible for the

35   formation and development of the nervous system. Further, gene length analysis of the

36   differentially expressed genes showed that the loss of *Elp1* mainly impacts the expression of long

37   genes and that by gradually restoring Elongator their expression is progressively rescued. Finally,

38   through evaluation of co-expression modules, we identified gene sets with unique expression

39   patterns that depended on *ELP1* expression. Overall, this study highlights the crucial role of *ELP1*

40   during early embryonic neuronal development and reveals gene networks and biological pathways

41   that are regulated by Elongator.

42

Introduction

43

44 Elongator is a highly conserved multiprotein complex composed of two copies of each of its six

45 subunits, named Elongator complex proteins 1 to 6 (ELP1-6). Elongator subunits are

46 evolutionarily highly conserved from yeast to humans both in their sequence and interaction with

47 other subunits. Conserved function across all species has been clearly demonstrated using a variety

48 of different cross-species rescue experiments (1-3). Deletion of any of the genes encoding the six

49 subunits confers almost identical biochemical phenotypes in yeast (4-6), suggesting that there is a

50 tight functional association between the proteins comprising Elongator complex, and that the

51 functional integrity of Elongator is compromised in the absence of any of its subunits (6). Both

52 yeast and human Elongator have lysine acetyltransferase activity that is mediated by the catalytic

53 subunit Elp3. Elp3 has two identified substrates: histone H3 and α-Tubulin (7-10). While the

54 acetylation of histone H3 is linked to the role of the complex in transcriptional elongation (7, 10,

55 11), cytosolic acetylation of α-Tubulin has been linked to its role in microtubule organization

56 particularly in the context of cell migration (8). Elongator was isolated as a complex that associates

57 with chromatin and interacts with the elongating phosphorylated form of RNA polymerase II

58 (RNAPII) both in yeast and human (7, 10, 12). The catalytic subunit Elp3, by acetylating histone

59 H3, facilitates RNAPII access to actively transcribed genes. In human cells, Elongator is required

60 for the expression of several genes involved in migration and in the expression of HSP70 (13, 14).

61 In addition, accumulating evidence supports the role of this complex in maintaining translational

62 fidelity through tRNA modifications. Specifically, Elongator is essential for the formation of 5-

63 methoxy-carbonylmethyl (mcm5) and 5-car-bamoylmethyl (ncm5) groups on uridine nucleosides

64 present at the wobble position of many tRNAs (15, 16).

65   Several *loss-of- function* studies have demonstrated the key role of Elongator during development.

66   Yeast Elp mutants are hypersensitive to high temperature and osmotic conditions and they showed

67   defects in exocytosis, telomeric gene silencing, DNA damage response and adaption to new growth

68   medium (10, 17, 18) . In *Arabidopsis thaliana*, mutations in Elp subunits resulted in impaired root

69   growth (19) and deletion of Elp3 in *Drosophila melanogaster* was lethal at the larval stage (20).

70   Depletion of Elongator in *Caenorhabditis elegans* led to defects in neurodevelopment (21). In

71   mice, Elp1 knockout results in embryonic lethality due to failed neurulation and vascular system

72   formation (3, 22). Consistent with its crucial role during development, several human

73   neurodevelopmental disorders have been associated with mutations in Elongator complex

74   subunits. Familial dysautonomia (FD) is caused by a splicing mutation in *ELP1* (23-25) that

75   reduces the amount of functional protein in a tissue specific manner, missense mutations in *ELP2*

76   have been found in individuals with severe intellectual disability (ID) (26, 27), variants of *ELP3*

77   have been associated with amyotrophic lateral sclerosis (ALS) (28), *ELP4* variants have been

78   implicated in autism spectrum disorder and Rolandic epilepsy (RE) (29, 30) and a mutation in *Elp6*

79   causes Purkinjie neuron degeneration and ataxia-like phenotypes in mice (31).

80   FD is a neurodevelopmental disorder characterized by widespread sensory and autonomic

81   dysfunction and by central nervous system (CNS) pathology (32-37). The major mutation in FD

82   is a splicing mutation in *ELP1* intron 20 that leads to variable skipping of exon 20 and to a

83   reduction of ELP1 mostly in the nervous system (25, 38). In 2009 we generated a knockout (KO)

84   Elp1 mouse, Elp1$^{-/-}$ and showed that complete ablation of *Elp1* resulted in early embryonic lethality

85   (3, 22). To gain a better understanding how reduction of ELP1 leads to FD, we generated two new

86   mouse models by introducing different copy numbers of the human *ELP1* transgene with the major

87   FD mutation (39), *TgFD1* and *TgFD9,* into the *Elp1$^{-/-}$* mouse. Although the human FD transgene

88 did not rescue embryonic lethality of the $Elp1^{-/-}$ mouse, its expression rescues embryonic

89 development in a dose dependent manner in Tg$FD1$; $Elp1^{-/-}$ and $TgFD9$; $Elp1^{-/-}$ embryos. In order

90 to understand the gene regulatory networks that are dependent on ELP1 expression, we conducted

91 a comprehensive transcriptome analysis in these mouse embryos.

92

93 **Results**

94 **Generation of mice expressing an increasing amount of *ELP1*.**

95 We previously demonstrated that homozygous deletion of the mouse $Elp1$ gene leads to embryonic

96 lethality prior to mid-gestation (3). Detailed phenotypic characterization of $Elp1^{-/-}$ KO embryos at

97 early developmental stages revealed several abnormalities, including a dramatic reduction in size,

98 disruption of the extraembryonic vascular networks, failure of germ layer inversion, and

99 interruption of cephalic neural-tube closure (3, 22). In an effort to understand the molecular

100 mechanisms that characterize FD we have generated several transgenic mouse lines carrying the

101 wild-type (WT) and FD human *ELP1* gene that differ by the copy number of the transgene (39).

102 The murine Elp1 protein is 80% identical to human ELP1 and by introducing the human WT *ELP1*

103 transgene into the $Elp1^{-/-}$ mouse, we completely rescued development and mice were born alive

104 and healthy, confirming ELP1 functional conservation between human and mouse (3). To test

105 whether the abnormalities caused by ablation of mouse $Elp1$ could be improved by the human FD

106 transgene, heterozygote mice carrying different copy numbers of the FD *ELP1* transgene (*TgFD1;*

107 $Elp1^{+/-}$ or *TgFD9;* $Elp1^{+/-}$) were crossed with mice heterozygous for the $Elp1$ knockout allele

108 ($Elp1^{-/+}$). Mice were collected at either E8.5 or P0 and genotyped using genomic DNA from the

109 visceral yolk sac. Although neither *TgFD1* or *TgFD9* rescued embryonic lethality in the $Elp1^{-/-}$

110 mice (Table 1), the development of the FD1/KO (Tg$FD1$; $Elp1^{-/-}$) and FD9/KO (*TgFD9;* $Elp1^{-/-}$)

111 embryos progressed further as human *ELP1* expression increased (Figure 1a). Because familial

112 dysautonomia is caused by a developmental reduction of ELP1 protein, we conducted a

113 comprehensive transcriptome analysis in embryos expressing increasing amounts of ELP1. We

114 collected 29 individual C57BL/6 mouse embryos at E8.5 (n=8 KO, n=7 FD1/ KO, n=6 FD9/ KO,

115 n=8 WT) and total RNA was extracted from each single embryo (Figure 1a). As expected, KO

116 embryos do not express any WT *Elp1* (Figure 1b); whereas FD1/KO and FD9/KO embryos

117 expressed increasing amounts of WT *ELP1* with FD9/KO embryos expressing three times more

118 *ELP1* than FD1/KO embryos (Figure 1c). Normalized gene count comparisons revealed that the

119 expression of the human *ELP1* in FD1/KO embryos is ~6% of WT *Elp1* while in the FD9/KO the

120 human *ELP1* is ~16% WT *Elp1*. Principal component analysis revealed that the four genotypes

121 exhibited distinct expression profiles with PC1 explaining 28% of the total expression variance

122 across samples suggesting that *ELP1* dosage plays a critical role in embryonic transcriptome

123 regulation (Figure 1d).

124

125 **Table 1. Genotype ratios of offspring generated by *TgFD; Elp1$^{+/-}$X Elp1$^{+/-}$* pairings.** The time

126 of conception is estimated to be E0.5 day prior to the observation of a vaginal plug. P0, postnatal

127 day zero. Expected mendelian ratio is indicated below the genotype, percentage of embryos of

128 each genotype in parentheses.

TgFD; Elp1$^{+/-}$ X Elp1$^{+/-}$	Time	Number of animals						Total
		Elp1$^{+/+}$ 12.5%	*Elp1$^{+/-}$* 25%	*Elp1$^{-/-}$* 12.5%	*Tg; Elp1$^{+/+}$* 12.5%	*Tg; Elp1$^{+/-}$* 25%	*Tg; Elp1$^{-/-}$* 12.5%	
TgFD1	E8.5	17(11%)	37(25%)	19(13%)	17(11%)	43(29%)	17(11%)	150
TgFD1	P0	43(19%)	88(39%)	0(0%)	29(13%)	63(28%)	0(0%)	223
TgFD9	E8.5	10(7%)	30(21%)	16(11%)	25(17%)	45(31%)	17(11%)	143

TgFD9	P0	24(23%)	42(40%)	0(0%)	13(12%)	27(25%)	0(0%)	106

129

Major transcriptome changes in *Elp1* KO embryos.

Transcriptome profiling in embryos expressing an increasing amount of *ELP1* show that the number of differentially expressed genes (DEGs), (False Discovery Rate or FDR < 0.1; and Fold Change or FC >= 1.5), proportionally declines as *ELP1* expression increases. In KO embryos 2,399 out of 19,619 (12.23%) genes were differentially expressed when compared with WT embryos (Figure 2a, Supplementary Table 1). Strikingly, in FD1/KO embryos the DEGs were only 601 (3.06%), while in FD9/KO embryos there were 494 DEGs (2.52%) (Figure 2a, Supplementary Table 1), demonstrating that a minimal increase in *ELP1* is sufficient to rescue the expression of ~80% of all DEGs. Gene Ontology (GO) analysis of the down-regulated genes (FDR < 0.1) in KO, FD1/KO and FD9/KO embryos highlighted pathways important for nervous system development including synapse formation, neuron projection and axon growth (Figure 2b, Supplementary Table 2). These findings are consistent with the body of work supporting the role of *ELP1* during early development in target tissue innervation and with the fact that neuronal loss in FD is mostly due to failure of innervation (14, 40-45). Notably, of the 71 genes that were significantly downregulated in all three KO genotypes (Supplementary Table 1), 24 of them (~33%) have a critical role in nervous system or brain development (Figure 2c). STRING analysis of these genes revealed an enrichment for protein-protein interactions (PPI enrichment = 2.9E-4 according to STRING v11) (Figure 2c) (46). Among the neuronal genes, *Dbx1* and *Nr2e1* were the two most down-regulated genes across all three KO genotypes (Figure 2d and 2e). *Dbx1*, also known as Developing Brain Homeobox 1, is expressed in a regionally restricted pattern in the developing mouse central nervous system (CNS) and encodes for a transcription factor that plays a pivotal role in interneuron differentiation in the ventral spinal cord (47). In vertebrates, spinal interneurons

152 modulate the motor neuron activity elicited by incoming sensory information and, by relaying the

153 proprioceptive data to the brain, play a critical role in locomotor coordination (48). *Nr2e1* is a

154 transcription factor that regulates the expression of genes essential for retinal development (49).

155 Loss of *Nr2e1* in mice has been shown to cause severe early onset retinal degeneration with death

156 of various retinal cells including retinal ganglion cells (RGCs) (50-52). Interestingly, the list of

157 neuronal genes that were significantly downregulated in all three KO genotypes, KO*,* FD1/KO and

158 FD9/KO, also included the brain-derived neurotrophic factor BDNF-receptor TrkB (*Ntrk2),* the

159 chemorepulsive axon guidance protein draxin, the homeobox protein involved in brain and sensory

160 organ development otx2, and the neuronal adhesion protein involved in neurite growth neurocan

161 (*Ncan*) (Figure 2f-i). GO pathways associated with up-regulated genes in the KO embryos are

162 included in Supplementary Figure S1and Supplementary Table 2.

163

164 **Long neuronal genes require Elongator activity for their expression.**

165 Although Elongator plays a number of roles in the cell, in the nucleus this complex directly

166 interacts with RNAPII and facilitates transcriptional elongation through altering chromatin

167 structure (7, 10, 12, 53). Elongator has histone acetyltransferase activity via its ELP3 subunit and

168 regulates the accessibility of RNAPII to the chromatin. Using Chromatin immunoprecipitation

169 (ChIP) we have previously shown that in the absence of ELP1, histone H3 acetylation was

170 significantly reduced at the 3' ends of genes (3). Moreover, Close et al demonstrated that upon

171 reduction of ELP1 there was progressively lower RNAPII density at the 3' end of target genes than

172 in the promoter region (14), supporting the role of Elongator in transcriptional elongation. To

173 examine if the downregulation in gene expression observed in the KO embryos might be due to a

174 failure in transcriptional elongation, we compared the length distribution of the DEGs among

175   embryos expressing an increasing amount of *ELP1*.   We discovered that long genes, especially

176   those longer than 100 kb, were downregulated significantly more often in the KO embryos than in

177   the FD9/KO embryos (FDR < 0.01, Figure 3a) suggesting that Elongator loss has a more

178   pronounced effect on the expression of longer genes. Importantly, ELP1 restoration efficiently

179   rescued their expression. In contrast, there was no difference in gene length in the upregulated

180   genes between the different genotypes (Figure 3b). Of the 257 long genes (> 100 kb) that were

181   downregulated in KO embryos, 216 (84.05%) were rescued in the FD1/KO embryos and 247

182   (96.11%) were rescued in FD9/KO embryos. GO analysis of these genes highlighted pathways

183   important for synapse formation, neuron projection and axon growth (Figure 3c, Supplementary

184   Table 3). Since long downregulated genes were enriched for pathways important for nervous

185   system development, we compared the average length of neurodevelopmental genes with the

186   average length of all expressed genes (see Material and Methods) and we observed that

187   neurodevelopmental genes (GO:0048666) were significantly longer (Figure 3d). We then

188   investigated whether the ELP1-dependent gene regulation was driven by length or if neuronal

189   genes were more likely to require ELP1 for efficient transcription. We divided all downregulated

190   genes in KO embryos into four gene-length categories and examined the proportion of

191   neurodevelopmental genes (GO:0048666) in each category (Figure 3e *left panel*). If

192   neurodevelopmental genes were more likely to require ELP1 for efficient transcription, we would

193   expect to see more rescue with higher ELP1 dosage. Given that the proportion of

194   neurodevelopmental genes rescued in the FD9/KO embryos in all gene-length categories was

195   similar to those of the downregulated genes in the KO embryos (Figure 3e *right panel*), we

196   concluded that neurodevelopmental genes are more susceptible to ELP1-loss simply because they

197   are longer than the non-neurodevelopmental genes.

198

Identification of genes whose expression correlates with amount of ELP1.

200 Elongator has been linked to transcriptional regulation (13, 14, 54). Our unique mouse models

201 provide, for the first time, the ability to perform a comprehensive transcriptome analysis to identify

202 genes whose expression depends on the amount of *ELP1*. This study is highly relevant to better

203 understanding FD pathogenesis, as the disease is caused by a reduction, not loss, of *ELP1* primarily

204 in the nervous system. We built the gene co-expression network across embryonic RNASeq data

205 for all genotypes (Materials and Methods). We identified thirty-five co-expression distinct

206 Modules Eigengenes (MEs) (Figure 4a). We then postulated that genes whose expression relies on

207 *ELP1* expression would be grouped into three major categories (Figure 4b): (1) genes whose

208 expression changes as a monotonic function of *ELP1*, referred as "dose-responsive genes"; (2)

209 genes whose expression is completely rescued with low *ELP1* expression, referred as "highly

210 responsive genes" and (3) genes whose expression is restored only when *ELP1* is expressed at WT

211 levels, referred as "low responsive genes". Among the thirty-five MEs identified, ME3, ME2 and

212 M12 had the highest positive correlation with these hypothesized gene patterns (Pearson

213 correlation >= 0.85, FDR < 0.1, Figure 4c). ME3 included the dose-responsive genes (Figure 4d);

214 ME2 included the highly responsive genes (Figure 4g) and M12 contained the low responsive

215 genes (Figure 4j). Interestingly, ME4, ME7 and ME10 had the highest negative correlation with

216 our hypothesized patterns (Supplementary Figure S2).

217 Overall, we identified 242 dose-responsive genes whose expression strictly increases as a

218 monotonic function of *ELP1* (Pearson correlation >= 0.85, FDR < 0.1, Figure 4e). The GO

219 analysis of these genes highlighted pathways important for axon and cell projection formation

220 (Figure 4f, Supplementary Table 4). This result supports, once again, the role of ELP1 in the

221 expression of genes important for target tissue innervation and is consistent with the innervation

222 failure observed in FD (14, 40-45). We found 357 highly responsive genes whose expression was

223 completely restored in *FD9/KO* embryos (Pearson correlation >= 0.85, FDR < 0.1, Figure 4h). GO

224 analysis of these genes highlighted pathways associated with transcriptional regulation (Figure 4i,

225 Supplementary Table 4) suggesting that a small increase of functional Elongator is enough to

226 restore normal expression of important transcriptional regulators. This finding might explain the

227 dramatic phenotypic improvement observed in *FD9/KO* embryos when compared with *FD1/KO*

228 and KO. Finally, we identified four low responsive genes whose expression was rescued solely in

229 WT embryos: *Amigo1*, *Snn*, *Unc119b* and *Sh3pxd2b* (Figure 4k) (Pearson correlation >= 0.85,

230 FDR < 0.1).

231 In the negatively correlated patterns ME4, ME7 and ME10 (Pearson correlation >= 0.85, FDR <

232 0.1) we did not find any enrichment for neuronal terms (Supplementary Figure S2, Supplementary

233 Table 5). Genes in ME4 highlighted pathways involved in mitochondrial and respiratory chain

234 activity (Supplementary Figure S3, Supplementary Table 5) while genes in ME7 and ME10 did

235 not shown any significant enrichment for GO terms (FDR < 0.1). These results show that the

236 expression of many genes that are involved in transcriptional regulation and nervous system

237 development positively correlate with *ELP1* expression.

238

239 **Discussion**

240 Variants in Elongator subunits are associated with various human neurodevelopmental disorders,

241 including FD (*ELP1*), ID (*ELP2*), ALS (*ELP3*), autism spectrum disorder and Rolandic epilepsy

242 (*ELP4*) (23-31). To gain a better understanding the pathogenesis of FD, we generated two new

243 mouse models expressing an increasing amount of *ELP1*, FD1/KO and FD9/KO. Although the

244　human FD transgene did not rescue embryonic lethality in the *ELP1* KO mouse, its expression

245　improved embryonic development in a dose dependent manner. To identify genes and pathways

246　whose expression is highly correlated with *ELP1* and that are ultimately essential for embryonic

247　development, we conducted a comprehensive transcriptome analysis in KO, FD1/KO, FD9/KO

248　and WT embryos. We found that even a minimal increase in *ELP1* has a dramatic effect on overall

249　gene expression with the majority of the KO DEGs being completely rescued in the FD1/KO

250　embryos, which only expresses an amount of *ELP1* that is approximately 6% of the *Elp1* amount

251　expressed in the WT embryos. A significant portion of the down-regulated genes across the

252　different genotypes KO, FD1/KO and FD9/KO, have a crucial role in nervous system

253　development. Among these neuronal genes, *Dbx1* and *Nr2e1* were the two most down-regulated

254　genes in all three genotypes with absent or reduced expression of *ELP1*.

255　*Dbx1*, also known as *Developing Brain Homeobox 1*, is expressed in a regionally restricted pattern

256　in the developing mouse central nervous system (CNS) and encodes for a transcription factor that

257　has pivotal role in interneuron differentiation in the ventral spinal cord (47). The spatially restricted

258　expression of *Dbx1* is important to establish the distinction of V0 neurons from the adjacent set of

259　V1 neurons (55). *Dbx1* mutant mice lacking these ventral interneurons exhibit marked changes in

260　motor coordination, supporting the role of Dbx1-dependent interneurons as key components of the

261　spinal locomotor circuits that control stepping movements in mammals (48). Interestingly, one of

262　the most characteristic features of FD is poor locomotor coordination and both patients as well as

263　a phenotypic mouse model of FD showed progressive impairment in coordination that leads to

264　severe gait ataxia (56-58).

265　*Nr2e1*, or nuclear receptor subfamily 2 group E member 1, encodes a highly conserved

266　transcription factor known to be a key stem cell fate determinant in both the developing mouse

267   forebrain and retina (52, 59-63). The mouse *Nr2e1* gene is first detected at embryonic day 8 (E8)

268   in the ventricular zone (VZ) of the neuroepithelium layer and later spreads posteriorly into the

269   diencephalon (E10.5) (63-66). In the developing mouse eye, *Nr2e1* is detected in the optic

270   processes of the developing embryo as early as E9 (32), suggesting a role for this gene in an early

271   phase of retinogenesis (29, 61). During development, *Nr2e1* null mice display an increase in

272   apoptotic levels of RGCs in the ganglion cell layer (GCL), which results in a marked reduction in

273   thickness of the distinct layers in the adult retina and optic nerve dystrophy (51, 52, 66).

274   Intriguingly, degeneration of RGCs is observed in two different *Elp1* conditional knock-out mice

275   (67, 68). In addition, patients with FD show RGCs loss with reduction in the thickness of the retinal

276   nerve fiber layer (RNFL) and progressive vision loss (33, 34, 69). Although further studies will be

277   necessary to determine the link between *ELP1* reduction and downregulation of specific key

278   neurodevelopmental genes, the identified *ELP1*-dependent transcriptome profiles constitute an

279   excellent foundational resource for understanding Elongator biology and also help to shed light

280   into the molecular pathways that underlie diseases caused by disruption of Elongator activity.

281   In the nucleus, Elongator facilitates transcriptional elongation through altering chromatin structure

282   (3, 7, 10, 53). Therefore, we analyzed the length distribution of DEGs among embryos expressing

283   different *ELP1* amounts. Our data clearly showed that long genes are more affected by the loss of

284   functional Elongator compared with shorter genes and, by gradually increasing *ELP1* amount, we

285   were able to progressively restore their expression. Moreover, the observation that neuronal genes

286   are significantly longer than all expressed genes, offers a possible explanation about why the

287   nervous system is the tissue that most relies on functional Elongator during embryonic

288   development. In the future it would be interesting to investigate the role of *ELP1* in transcriptional

289   elongation for each identified target, in order to identify direct regulatory effects.

290   Several genes that require Elongator to be efficiently expressed have been identified using either

291   cell lines or conditional KO mouse tissues with reduced *Elp1* expression (14, 45, 70-73). In

292   addition, RNA microarray analysis in post-mortem FD tissues has shown that a subset of genes

293   involved in myelination require *ELP1* for efficient transcription (74). In the current study we have

294   identified gene patterns whose expression varies as a function of *ELP1* amount: (1) genes whose

295   expression changes as a monotonic function of *ELP1*; (2) genes whose expression is completely

296   restored with low amount of *ELP1* and (3) genes whose expression is restored only when *ELP1* is

297   expressed at WT levels (Figure 4c). It is interesting that distinctive groups of genes display

298   different Elongator dependence. Genes important for axon formation and cell projection responded

299   in a dose-dependent manner to functional Elongator amounts. This is consistent with the

300   observation that, even though Elongator is a ubiquitously expressed protein complex, variants

301   affecting different Elongator subunits all lead to neurodevelopmental diseases (23-31). On the

302   other hand, the expression of important transcriptional regulators is already restored in *FD9/KO*

303   embryos, suggesting that only a small increase of functional Elongator is necessary to rescue their

304   expression. This might underlie the dramatic phenotypic improvement observed in *FD9/KO*

305   embryos when compared with *FD1/KO* and KO. We then identified sets of genes whose expression

306   negatively correlated with *ELP1*. GO analysis of these genes highlighted pathways involved in

307   mitochondrial and respiratory chain activity.

308   In conclusion, our study is the first to assess the *in vivo* dose-dependent effect of *ELP1* in early

309   development using transcriptome analysis. We demonstrated that even a minimal increase in

310   *ELP1* can have a dramatic effect on both mouse embryonic development and global gene

311   expression. Although loss of *ELP1* compromised the expression of a large number of genes, this

312   study shows that neuronal genes are more sensitive to *ELP1* reduction. We recognize that further

313   studies will be necessary to determine which of the individual gene expression changes are direct

314   versus indirect. However, the identification of gene networks and biological pathways that are

315   regulated by *ELP1*-dosage is highly relevant to better understand the pathogenesis of FD, as well

316   as other neurodevelopmental diseases caused by Elongator deficiency. The data presented here

317   will help to identify potential biomarkers for future clinical studies and new targetable pathways

318   for therapy.

319

320   **Materials and Methods**

321   **Generation of *FD1/KO*, *FD9/KO* mouse models and genotyping**

322   The detailed description of the original strategy to generate the *Elp1* knockout mouse line has been

323   previously published (3). A detailed description of the generation the *TgFD* transgenic lines

324   carrying different copy number of the human *ELP1* gene with the IVS20+6T>C mutation can be

325   found in our previous manuscript by Hims et al. (39). To create the *TgFD1; Elp1$^{-/-}$* and *TgFD9;*

326   *Elp1$^{-/-}$* mouse, we crossed the previously generated *TgFD1* or *TgFD9* transgenic mouse line with

327   the mouse line heterozygous for the null allele *Elp1$^{-/+}$*. The resulting progeny was genotyped to

328   detect the presence of the *TgFD1* or *TgFD9* transgene and of the null allele *Elp1^{-}*. As expected,

329   the *TgFD1* and *TgFD9* transgene segregated independently from the null allele; therefore, around

330   one-fourth of the F1 mice carried both the *TgFD* transgene and null *Elp1* alleles (*TgFD1; Elp1$^{+/-}$*

331   or *TgFD9; Elp1$^{+/-}$*). Subsequently, we crossed the *TgFD1; Elp1$^{+/-}$* and *TgFD9; Elp1$^{+/-}$* mice with

332   the mouse line heterozygous for the null Elp1 allele (*Elp1$^{+/-}$*). The resulting progeny was genotyped

333   to detect the presence of the TgFD1 or TgFD9 transgene as well as the null allele in homozygosis

334   (*Elp1$^{-/-}$*). The *Elp1$^{-/-}$*, *TgFD1; Elp1$^{-/-}$*, *TgFD9; Elp1$^{-/-}$* and *Elp1$^{+/+}$* embryos were produced by

335   crossing heterozygote mice carrying different copy numbers of the FD *ELP1* transgene (*TgFD1;*

336     $Elp1^{+/-}$ or $TgFD9; Elp1^{+/-}$) with heterozygote mice ($Elp1^{-/+}$). The day of vaginal-plug discovery

337   was designated E 0.5. The mice used for this study were housed in the animal facility of

338   Massachusetts General Hospital (Boston, MA), provided with constant access to a standard diet of

339   food and water, maintained on a 12-hour light/dark cycle, and all experimental protocols were

340   approved by the Subcommittee on Research Animal Care at the Massachusetts General Hospital.

341   The genotypes of animals and embryos were determined by PCR analysis of genomic DNA from

342   tail biopses and from embryos and/or visceral yolk sacs, respectively. The primer sets used were

343   as follows: for determining the wild-type Elp1 allele, 5'-ACCCTCAGGCAGTTTGATTG-3' and

344   5'-CATGGCTCCATAAAACAAACAC-3';   for   detecting   the   knockout   allele,

345   5'ACCCTCAGGCAGTTTGATTG-3' and 5'-GGCTACCGGCTAAAACTTGA-3'; and

346   for determining the human TgFD transgenes, TgProbe1F 5'-GCCATTGTACTGTTTGCGACT-

347   3' and TgProbe1R 5'-TGAGTGTCACGATTCTTTCTGC-3'.

348

349   **Morphological analysis of embryos**

350   Photographs of visceral yolk sacs and embryos were taken with a digital camera Leica DFC7000

351   T mounted on a Leica M205 FCA dissection microscope. LAS X software (Leica) was used for

352   image processing.

353

354   **RNA Seq experiment**

355   RNA was extracted from 8 $Elp1^{-/-}$, 7 $TgFD1; Elp1^{-/-}$, 6 $TgFD9; Elp1^{-/-}$ and 8 $Elp1^{+/+}$ individual

356   embryos at the embryonic stage of E8.5 using the QIAzol Reagent and following the

357   manufacturer's instructions. RNAseq Libraries were prepared using TruSeq® Stranded mRNA

358   Library Prep kit (Illumina 20020594), using 100 ng total RNA as input. Libraries were evaluated

359   for final concentration and size distribution by Agilent 2200 TapeStation and/or qPCR, using

360   Library Quantification Kit (KK4854, Kapa Biosystems), and multiplexed by pooling equimolar

361   amounts of each library prior to sequencing. Pooled libraries were 50 base pair paired-end

362   sequenced on Illumina HiSeq 2500 across multiple lanes. Real time image analysis and base

363   calling were performed on the HiSeq 2500 instrument using HiSeq Sequencing Control Software

364   (HCS) and FASTQ files demultiplexed using CASAVA software version 1.8.

365   A synthesized transcriptome reference was generated, by artificially adding the sequence of human

366   *ELP1* gene from the Ensembl human transcriptome reference GRCh37.75 to the Ensembl mouse

367   transcriptome reference GRCm38.83 as an independent chromosome. RNASeq reads were

368   mapped to this synthesized transcriptome reference 3 by STAR v2.5.2b, allowing only uniquely

369   mapped reads with 5% mismatch (75).

370

371   **Differential gene expression analysis**

372   Gene counts were generated by HTSeq-count v0.11.2 with "-s reverse" option to be compatible

373   with Illumina TruSeq reads, according to the gene annotations of the synthesized transcriptome

374   reference. Four genotypes were defined as "FD0", "FD1", "FD9" and WT to reflect the amounts

375   of human and mouse ELP1. Gene counts across samples were filtered so that only genes whose

376   median expression amounts were no less than 0.1 counts-per-million-reads in at least one genotype

377   were kept for the downstream analysis. The R (v3.6.1) package "SVA" v3.32.1 was implemented

378   on the filtered gene expression matrix to estimate surrogated variables (SVs) among samples. A

379   generalized linear model was built by the R package "DESeq2" v1.24.0 to correlated gene

380   expression to genotypes reflecting human *ELP1* and mouse *Elp1* amounts, together with all the

381   estimated SVs. During the differential gene expression analysis, the FC and FDR of each genotype

382   was estimated per gene. PCA analysis was implemented based on the top 500 most variable genes

383   across samples, from the filtered gene expression matrix with SVs corrected.

384

385   **GO analysis**

386   GO analysis was done by GOrilla (76) (http://cbl-gorilla.cs.technion.ac.il/). Organism was set as

387   *Mus musculus*. Two unranked lists of genes were used for each GO analysis. The GO analysis for

388   DEGs used the following two lists: 1) the list of DEGs for a genotype and 2) the list of all genes

389   in the filtered gene expression matrix. The GO analysis for co-expression modules used the

390   following two lists: 1) the list of genes highly correlated with a hypothetical pattern in a co-

391   expression module and 2) the list of all genes in the filtered gene expression matrix. In the figures

392   related to the GO analysis for DEGs, ten significant terms in the ontology of Cellular Component

393   were selected to be plotted in Figure 2b. The complete list of DEG GO terms for each genotype

394   can be found in Supplementary Table 2.

395

396   **Mouse neuronal genes**

397   Mouse neuronal genes were extracted and downloaded via AmiGO2 from geneontology.org, using

398   the key word 'neuron' and restricting organism as 'Mus musculus'. Totally 3,442 unique neuronal

399   genes were found.

400

401   **Co-expression modules analysis**

402   R package "WGCNA" (77) v1.68 was implemented to the filtered gene expression matrix with

403   SVs corrected. The soft-thresholding power was determined to be 5. The minimal module size was

404   set to 30. The raw modules were merged using a dis-similarity cutoff of 0.15.

405

406 **Correlation between co-expression modules and hypothetical patterns**

407 Dummy expression data were generated to mimic three hypothetical patterns, namely the

408 monotonic increase pattern, the monotonic decrease pattern, and the saturated pattern. The eigen

409 vector representing each co-expression module was then correlated with each of the hypothetical

410 patterns using Pearson correlation.

411

412 **Statistical Analysis**

413 All raw p values in this study, if multiple tests were involved, were corrected by the Benjamini-

414 Hochberg Procedure and converted to FDR values. Wald test was applied in the differential gene

415 expression analysis. A significant DEG of a genotype, compared to WT, was defined as FDR <

416 0.1 and the FC >= 1.5 for that genotype. Fisher's exact test was used for GO analysis, where a

417 significant enrichment for a GO term was defined as FDR < 0.1. Kolmogorov-Smirnov test (K-S

418 test) was applied to compare the gene length distribution of genes from different groups. A

419 significant difference in length distribution between two groups was defined as FDR < 0.1. A

420 significant correlation throughout this study was defined as the Pearson correlation coefficient >=

421 0.85 and FDR < 0.1. For all box plots, the middle lines inside boxes indicated the medians. The

422 lower and upper hinges corresponded to the first and third quartiles. Each box extended to 1.5

423 times inter-quartile range (IQR) from upper and lower hinges respectively. The symbols *, ** and

424 ***, if appeared in the figures, indicated FDR < 0.1, < 0.01 and < 0.001, respectively.

425

426 **Acknoledgments**

434

Competing interests

436   The authors declare competing financial interests.

437   Funding: Research support from PTC Therapeutics, Inc. (S.A.S.).

438   Personal financial interests: Susan A. Slaugenhaupt is a paid consultant to PTC Therapeutics and

439   is an inventor on several U.S. and foreign patents and patent applications assigned to the

440   Massachusetts General Hospital, including U.S Patents 8,729,025 and 9,265,766, both entitled

441   "Methods for altering mRNA splicing and treating familial dysautonomia by administering

442   benzyladenine," filed on August 31, 2012 and May 19, 2014 and related to use of kinetin; and U.S.

443   Patent 10,675,475 entitled, "Compounds for improving mRNA splicing" filed on July 14, 2017

444   and related to use of BPN-15477. Elisabetta Morini, Dadi Gao, Michael E. Talkowski and Susan

445   A. Slaugenhaupt are inventors on an International Patent Application Number

446   PCT/US2021/012103, assigned to Massachusetts General Hospital and entitled " RNA Splicing

447   Modulation" related to use of BPN-15477 in modulating splicing.

448

Referances

1.   Li F, Lu J, Han Q, Zhang G, Huang B. The Elp3 subunit of human Elongator complex is functionally similar to its counterpart in yeast. Mol Genet Genomics. 2005;273(3):264-72.

2.   Chen Z, Zhang H, Jablonowski D, Zhou X, Ren X, Hong X, et al. Mutations in ABO1/ELO2, a subunit of holo-Elongator, increase abscisic acid sensitivity and drought tolerance in Arabidopsis thaliana. Mol Cell Biol. 2006;26(18):6902-12.

3.   Chen YT, Hims MM, Shetty RS, Mull J, Liu L, Leyne M, et al. Loss of mouse Ikbkap, a subunit of elongator, leads to transcriptional deficits and embryonic lethality that can be rescued by human IKBKAP. Mol Cell Biol. 2009;29(3):736-44.

4.   Fellows J, Erdjument-Bromage H, Tempst P, Svejstrup JQ. The Elp2 subunit of elongator and elongating RNA polymerase II holoenzyme is a WD40 repeat protein. J Biol Chem. 2000;275(17):12896-9.

5.   Winkler GS, Petrakis TG, Ethelberg S, Tokunaga M, Erdjument-Bromage H, Tempst P, et al. RNA polymerase II elongator holoenzyme is composed of two discrete subcomplexes. J Biol Chem. 2001;276(35):32743-9.

6.   Frohloff F, Jablonowski D, Fichtner L, Schaffrath R. Subunit communications crucial for the functional integrity of the yeast RNA polymerase II elongator (gamma-toxin target (TOT)) complex. J Biol Chem. 2003;278(2):956-61.

7.   Otero G, Fellows J, Li Y, de Bizemont T, Dirac AM, Gustafsson CM, et al. Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation. Mol Cell. 1999;3(1):109-18.

470   8.   Creppe C, Malinouskaya L, Volvert ML, Gillard M, Close P, Malaise O, et al. Elongator

471   controls the migration and differentiation of cortical neurons through acetylation of alpha-tubulin.

472   Cell. 2009;136(3):551-64.

473   9.   Tran HT, Nimick M, Uhrig RG, Templeton G, Morrice N, Gourlay R, et al. Arabidopsis

474   thaliana histone deacetylase 14 (HDA14) is an alpha-tubulin deacetylase that associates with PP2A

475   and enriches in the microtubule fraction with the putative histone acetyltransferase ELP3. Plant J.

476   2012;71(2):263-72.

477   10.   Wittschieben BO, Otero G, de Bizemont T, Fellows J, Erdjument-Bromage H, Ohba R, et

478   al. A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II

479   holoenzyme. Mol Cell. 1999;4(1):123-8.

480   11.   Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, et al. Genome-wide

481   map of nucleosome acetylation and methylation in yeast. Cell. 2005;122(4):517-27.

482   12.   Kim JH, Lane WS, Reinberg D. Human Elongator facilitates RNA polymerase II

483   transcription through chromatin. Proc Natl Acad Sci U S A. 2002;99(3):1241-6.

484   13.   Han Q, Hou X, Su D, Pan L, Duan J, Cui L, et al. hELP3 subunit of the Elongator complex

485   regulates the transcription of HSP70 gene in human cells. Acta Biochim Biophys Sin (Shanghai).

486   2007;39(6):453-61.

487   14.   Close P, Hawkes N, Cornez I, Creppe C, Lambert CA, Rogister B, et al. Transcription

488   impairment and cell migration defects in elongator-depleted cells: implication for familial

489   dysautonomia. Mol Cell. 2006;22(4):521-31.

490   15.   Huang B, Johansson MJ, Bystrom AS. An early step in wobble uridine tRNA modification

491   requires the Elongator complex. RNA. 2005;11(4):424-36.

492   16.   Fichtner L, Frohloff F, Jablonowski D, Stark MJ, Schaffrath R. Protein interactions within

493   Saccharomyces cerevisiae Elongator, a complex essential for Kluyveromyces lactis zymocicity.

494   Mol Microbiol. 2002;45(3):817-26.

495   17.   Rahl PB, Chen CZ, Collins RN. Elp1p, the yeast homolog of the FD disease syndrome

496   protein, negatively regulates exocytosis independently of transcriptional elongation. Mol Cell.

497   2005;17(6):841-53.

498   18.   Li Q, Fazly AM, Zhou H, Huang S, Zhang Z, Stillman B. The elongator complex interacts

499   with PCNA and modulates transcriptional silencing and sensitivity to DNA damage agents. PLoS

500   Genet. 2009;5(10):e1000684.

501   19.   Nelissen H, Fleury D, Bruno L, Robles P, De Veylder L, Traas J, et al. The elongata

502   mutants identify a functional Elongator complex in plants with a role in cell proliferation during

503   organ growth. Proc Natl Acad Sci U S A. 2005;102(21):7754-9.

504   20.   Walker J, Kwon SY, Badenhorst P, East P, McNeill H, Svejstrup JQ. Role of elongator

505   subunit Elp3 in Drosophila melanogaster larval development and immunity. Genetics.

506   2011;187(4):1067-75.

507   21.   Solinger JA, Paolinelli R, Kloss H, Scorza FB, Marchesi S, Sauder U, et al. The

508   Caenorhabditis elegans Elongator complex regulates neuronal alpha-tubulin acetylation. PLoS

509   Genet. 2010;6(1):e1000820.

510   22.   Dietrich P, Yue J, E S, Dragatsis I. Deletion of exon 20 of the Familial Dysautonomia gene

511   Ikbkap in mice causes developmental delay, cardiovascular defects, and early embryonic lethality.

512   PLoS One. 2011;6(10):e27015.

513   23.   Anderson SL, Coli R, Daly IW, Kichula EA, Rork MJ, Volpi SA, et al. Familial

514   dysautonomia is caused by mutations of the IKAP gene. Am J Hum Genet. 2001;68(3):753-8.

515   24.   Slaugenhaupt SA, Blumenfeld A, Gill SP, Leyne M, Mull J, Cuajungco MP, et al. Tissue-
516   specific expression of a splicing mutation in the IKBKAP gene causes familial dysautonomia. Am
517   J Hum Genet. 2001;68(3):598-605.

518   25.   Cuajungco MP, Leyne M, Mull J, Gill SP, Lu W, Zagzag D, et al. Tissue-specific reduction
519   in splicing efficiency of IKBKAP due to the major mutation associated with familial
520   dysautonomia. Am J Hum Genet. 2003;72(3):749-58.

521   26.   Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, et al. Deep
522   sequencing reveals 50 novel genes for recessive cognitive disorders. Nature. 2011;478(7367):57-
523   63.

524   27.   Cohen JS, Srivastava S, Farwell KD, Lu HM, Zeng W, Lu H, et al. ELP2 is a novel gene
525   implicated in neurodevelopmental disabilities. Am J Med Genet A. 2015;167(6):1391-5.

526   28.   Simpson CL, Lemmens R, Miskiewicz K, Broom WJ, Hansen VK, van Vught PW, et al.
527   Variants of the elongator protein 3 (ELP3) gene are associated with motor neuron degeneration.
528   Hum Mol Genet. 2009;18(3):472-81.

529   29.   Strug LJ, Clarke T, Chiang T, Chien M, Baskurt Z, Li W, et al. Centrotemporal sharp wave
530   EEG trait in rolandic epilepsy maps to Elongator Protein Complex 4 (ELP4). Eur J Hum Genet.
531   2009;17(9):1171-81.

532   30.   Addis L, Ahn JW, Dobson R, Dixit A, Ogilvie CM, Pinto D, et al. Microdeletions of ELP4
533   Are Associated with Language Impairment, Autism Spectrum Disorder, and Mental Retardation.
534   Hum Mutat. 2015;36(9):842-50.

535   31.   Kojic M, Gaik M, Kiska B, Salerno-Kochan A, Hunt S, Tedoldi A, et al. Elongator
536   mutation in mice induces neurodegeneration and ataxia-like behavior. Nature communications.
537   2018;9(1):3195.

538   32.   Axelrod FB, Hilz MJ, Berlin D, Yau PL, Javier D, Sweat V, et al. Neuroimaging supports

539   central pathology in familial dysautonomia. J Neurol.257(2):198-206.

540   33.   Mendoza-Santiesteban CE, Hedges TR, 3rd, Norcliffe-Kaufmann L, Warren F, Reddy S,

541   Axelrod FB, et al. Clinical neuro-ophthalmic findings in familial dysautonomia. J

542   Neuroophthalmol.32(1):23-6.

543   34.   Mendoza-Santiesteban CE, Palma JA, Hedges TR, 3rd, Laver NV, Farhat N, Norcliffe-

544   Kaufmann L, et al. Pathological Confirmation of Optic Neuropathy in Familial Dysautonomia. J

545   Neuropathol Exp Neurol. 2017;76(3):238-44.

546   35.   Ochoa JG. Familial dysautonomia (Riley-Day syndrome) may be associated with epilepsy.

547   Epilepsia. 2003;44(3):472.

548   36.   Mahloudji M, Brunt PW, McKusick VA. Clinical neurological aspects of familial

549   dysautonomia. J Neurol Sci. 1970;11(4):383-95.

550   37.   Pearson J. Familial dysautonomia (a brief review). J Auton Nerv Syst. 1979;1(2):119-26.

551   38.   Slaugenhaupt SA. Genetics of familial dysautonomia. Tissue-specific expression of a

552   splicing mutation in the IKBKAP gene. Clin Auton Res. 2002;12 Suppl 1:I15-9.

553   39.   Hims MM, Shetty RS, Pickel J, Mull J, Leyne M, Liu L, et al. A humanized IKBKAP

554   transgenic mouse models a tissue-specific human splicing defect. Genomics. 2007;90(3):389-96.

555   40.   Ohlen SB, Russell ML, Brownstein MJ, Lefcort F. BGP-15 prevents the death of neurons

556   in a mouse model of familial dysautonomia. Proc Natl Acad Sci U S A. 2017;114(19):5035-40.

557   41.   George L, Chaverra M, Wolfe L, Thorne J, Close-Davis M, Eibs A, et al. Familial

558   dysautonomia model reveals Ikbkap deletion causes apoptosis of Pax3+ progenitors and peripheral

559   neurons. Proc Natl Acad Sci U S A. 2013;110(46):18698-703.

560   42.   Jackson MZ, Gruner KA, Qin C, Tourtellotte WG. A neuron autonomous role for the

561   familial dysautonomia gene ELP1 in sympathetic and sensory target tissue innervation.

562   Development. 2014;141(12):2452-61.

563   43.   Cheishvili D, Maayan C, Cohen-Kupiec R, Lefler S, Weil M, Ast G, et al. IKAP/Elp1

564   involvement in cytoskeleton regulation and implication for familial dysautonomia. Hum Mol

565   Genet. 2011;20(8):1585-94.

566   44.   Johansen LD, Naumanen T, Knudsen A, Westerlund N, Gromova I, Junttila M, et al. IKAP

567   localizes to membrane ruffles with filamin A and regulates actin cytoskeleton organization and

568   cell migration. J Cell Sci. 2008;121(Pt 6):854-64.

569   45.   Abashidze A, Gold V, Anavi Y, Greenspan H, Weil M. Involvement of IKAP in peripheral

570   target innervation and in specific JNK and NGF signaling in developing PNS neurons. PLoS One.

571   2014;9(11):e113428.

572   46.   Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11:

573   protein-protein association networks with increased coverage, supporting functional discovery in

574   genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607-D13.

575   47.   Lu S, Shashikant CS, Ruddle FH. Separate cis-acting elements determine the expression

576   of mouse Dbx gene in multiple spatial domains of the central nervous system. Mech Dev.

577   1996;58(1-2):193-202.

578   48.   Lanuza GM, Gosgnach S, Pierani A, Jessell TM, Goulding M. Genetic identification of

579   spinal interneurons that coordinate left-right locomotor activity necessary for walking movements.

580   Neuron. 2004;42(3):375-86.

581   49.   Yu RT, Chiang MY, Tanabe T, Kobayashi M, Yasuda K, Evans RM, et al. The orphan

582   nuclear receptor Tlx regulates Pax2 and is essential for vision. Proc Natl Acad Sci U S A.

583   2000;97(6):2621-5.

584   50.   Abrahams BS, Kwok MC, Trinh E, Budaghzadeh S, Hossain SM, Simpson EM.

585   Pathological aggression in "fierce" mice corrected by human nuclear receptor 2E1. J Neurosci.

586   2005;25(27):6263-70.

587   51.   Zhang CL, Zou Y, Yu RT, Gage FH, Evans RM. Nuclear receptor TLX prevents retinal

588   dystrophy and recruits the corepressor atrophin1. Genes Dev. 2006;20(10):1308-20.

589   52.   Miyawaki T, Uemura A, Dezawa M, Yu RT, Ide C, Nishikawa S, et al. Tlx, an orphan

590   nuclear receptor, regulates cell numbers and astrocyte development in the developing retina. J

591   Neurosci. 2004;24(37):8124-34.

592   53.   Hawkes NA, Otero G, Winkler GS, Marshall N, Dahmus ME, Krappmann D, et al.

593   Purification and characterization of the human elongator complex. J Biol Chem.

594   2002;277(4):3047-52.

595   54.   Li F, Ma J, Ma Y, Hu Y, Tian S, White RE, et al. hElp3 directly modulates the expression

596   of HSP70 gene in HeLa cells via HAT activity. PLoS One. 2011;6(12):e29303.

597   55.   Pierani A, Moran-Rivard L, Sunshine MJ, Littman DR, Goulding M, Jessell TM. Control

598   of interneuron fate in the developing spinal cord by the progenitor homeodomain protein Dbx1.

599   Neuron. 2001;29(2):367-84.

600   56.   Morini E, Gao D, Montgomery CM, Salani M, Mazzasette C, Krussig TA, et al. ELP1

601   Splicing Correction Reverses Proprioceptive Sensory Loss in Familial Dysautonomia. Am J Hum

602   Genet. 2019.

603   57.   Macefield VG, Norcliffe-Kaufmann L, Gutierrez J, Axelrod FB, Kaufmann H. Can loss of

604   muscle spindle afferents explain the ataxic gait in Riley-Day syndrome? Brain. 2011;134(Pt

605   11):3198-208.

606   58.   Macefield VG, Norcliffe-Kaufmann LJ, Axelrod FB, Kaufmann H. Relationship between

607   proprioception at the knee joint and gait ataxia in HSAN III. Mov Disord. 2013;28(6):823-7.

608   59.   Schmouth JF, Banks KG, Mathelier A, Gregory-Evans CY, Castellarin M, Holt RA, et al.

609   Retina restored and brain abnormalities ameliorated by single-copy knock-in of human NR2E1 in

610   null mice. Mol Cell Biol. 2012;32(7):1296-311.

611   60.   Jackson A, Panayiotidis P, Foroni L. The human homologue of the Drosophila tailless gene

612   (TLX): characterization and mapping to a region of common deletion in human lymphoid leukemia

613   on chromosome 6q21. Genomics. 1998;50(1):34-43.

614   61.   Li W, Sun G, Yang S, Qu Q, Nakashima K, Shi Y. Nuclear receptor TLX regulates cell

615   cycle progression in neural stem cells of the developing brain. Mol Endocrinol. 2008;22(1):56-64.

616   62.   Pignoni F, Baldarelli RM, Steingrimsson E, Diaz RJ, Patapoutian A, Merriam JR, et al.

617   The Drosophila gene tailless is expressed at the embryonic termini and is a member of the steroid

618   receptor superfamily. Cell. 1990;62(1):151-63.

619   63.   Yu RT, McKeown M, Evans RM, Umesono K. Relationship between Drosophila gap gene

620   tailless and a vertebrate nuclear receptor Tlx. Nature. 1994;370(6488):375-9.

621   64.   Monaghan AP, Grau E, Bock D, Schutz G. The mouse homolog of the orphan nuclear

622   receptor tailless is expressed in the developing forebrain. Development. 1995;121(3):839-53.

623   65.   Monaghan AP, Bock D, Gass P, Schwager A, Wolfer DP, Lipp HP, et al. Defective limbic

624   system in mice lacking the tailless gene. Nature. 1997;390(6659):515-7.

625   66.   Young KA, Berry ML, Mahaffey CL, Saionz JR, Hawes NL, Chang B, et al. Fierce: a new

626   mouse deletion of Nr2e1; violent behaviour and ocular abnormalities are background-dependent.

627   Behav Brain Res. 2002;132(2):145-58.

628   67.   Ueki Y, Shchepetkina V, Lefcort F. Retina-specific loss of Ikbkap/Elp1 causes

629   mitochondrial dysfunction that leads to selective retinal ganglion cell degeneration in a mouse

630   model of familial dysautonomia. Dis Model Mech. 2018;11(7).

631   68.   Ueki Y, Ramirez G, Salcedo E, Stabio ME, Lefcort F. Loss of Ikbkap Causes Slow,

632   Progressive Retinal Degeneration in a Mouse Model of Familial Dysautonomia. eNeuro.

633   2016;3(5).

634   69.   Mendoza-Santiesteban CE, Hedges Iii TR, Norcliffe-Kaufmann L, Axelrod F, Kaufmann

635   H. Selective retinal ganglion cell loss in familial dysautonomia. J Neurol.261(4):702-9.

636   70.   Goffena J, Lefcort F, Zhang Y, Lehrmann E, Chaverra M, Felig J, et al. Elongator and

637   codon bias regulate protein levels in mammalian peripheral neurons. Nature communications.

638   2018;9(1):889.

639   71.   Zeltner N, Fattahi F, Dubois NC, Saurat N, Lafaille F, Shang L, et al. Capturing the biology

640   of disease severity in a PSC-based model of familial dysautonomia. Nat Med. 2016;22(12):1421-

641   7.

642   72.   Cohen-Kupiec R, Pasmanik-Chor M, Oron-Karni V, Weil M. Effects of IKAP/hELP1

643   deficiency on gene expression in differentiating neuroblastoma cells: implications for familial

644   dysautonomia. PLoS One.6(4):e19147.

645   73.   Boone N, Bergon A, Loriod B, Deveze A, Nguyen C, Axelrod FB, et al. Genome-wide

646   analysis of familial dysautonomia and kinetin target genes with patient olfactory ecto-

647   mesenchymal stem cells. Hum Mutat.33(3):530-40.

648   74.   Cheishvili D, Maayan C, Smith Y, Ast G, Razin A. IKAP/hELP1 deficiency in the

649   cerebrum of familial dysautonomia patients results in down regulation of genes involved in

650   oligodendrocyte differentiation and in myelination. Hum Mol Genet. 2007;16(17):2097-104.

651   75.   Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast

652   universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.

653   76.   Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and

654   visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48.

655   77.   Zhang B, Horvath S. A general framework for weighted gene co-expression network

656   analysis. Stat Appl Genet Mol Biol. 2005;4:Article17.
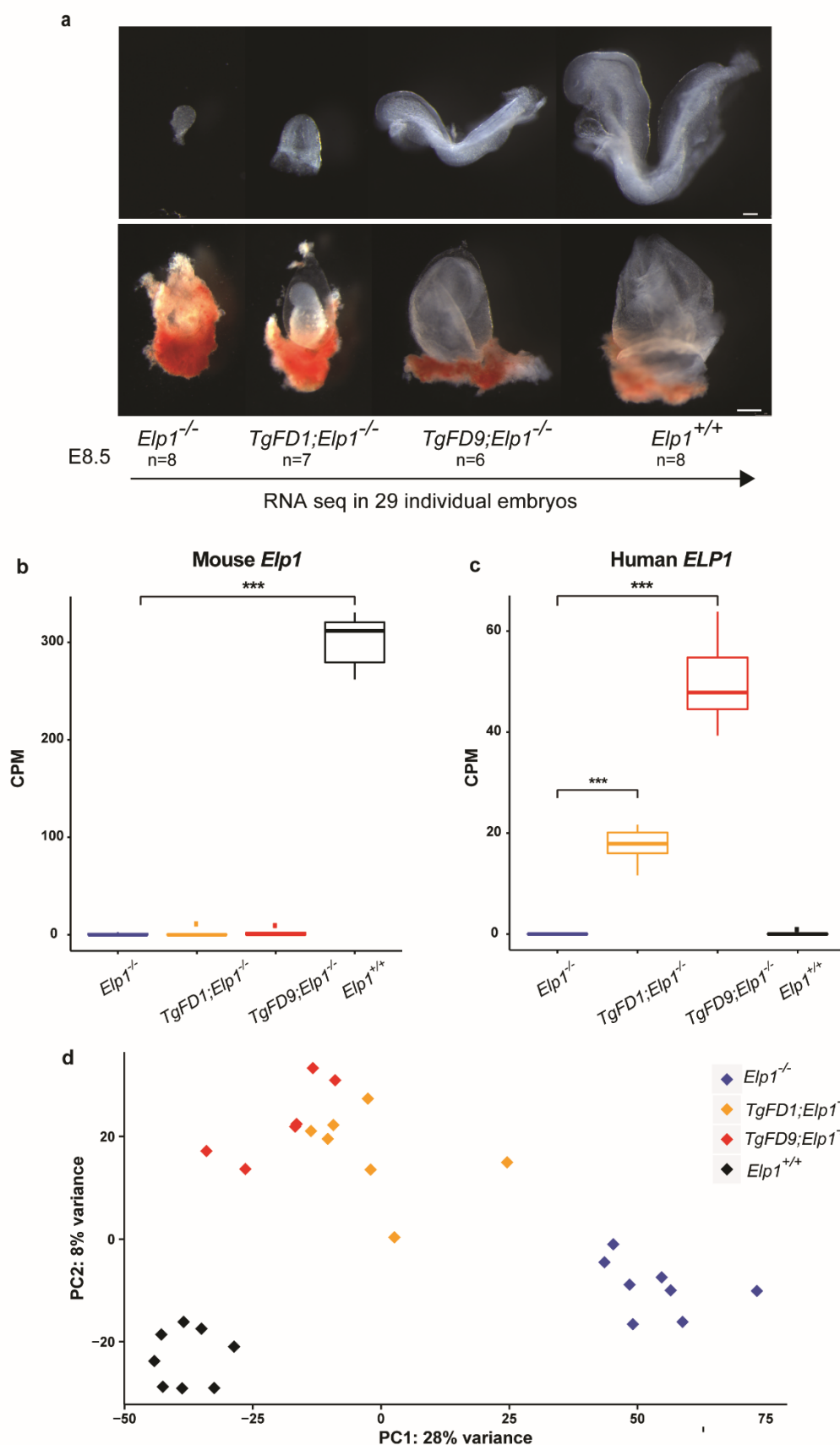
658 **Figure 1. Generating mouse embryos expressing increasing amount of ELP1** (**a**)

659 Morphological phenotypes of *Elp1$^{-/-}$, TgFD1; Elp1$^{-/-}$, TgFD9; Elp1$^{-/-}$* and *Elp1$^{+/+}$* embryos (top,

660 the scale bar represents 100 μm) and extraembryonic components (bottom, the scale bar represents

661 250 μm) at E8.5. RNA-seq experiment was performed using total RNA extracted by individual

662 embryo. (**b**) Expression of the endogenous WT *Elp1* across different embryos. The median for

663 each group is shown. (**c**) Expression of the human WT *ELP1* across different embryos. The median

664 for each group is shown. (**d**) Principal component analysis of all the embryos colored by genotype.

665 In box-and-whisker plots in **b-c**, each box extends to 1.5 times inter-quartile range (IQR) from

666 upper and lower hinges, respectively. Outliers are not shown. Only comparisons with significant

667 difference are marked by stars (two-tailed, unpaired Welch's *t* test with Bonferroni correction). In

668 the figure, *P < 0.05; **P < 0.01; ***P < 0.001.

669

670
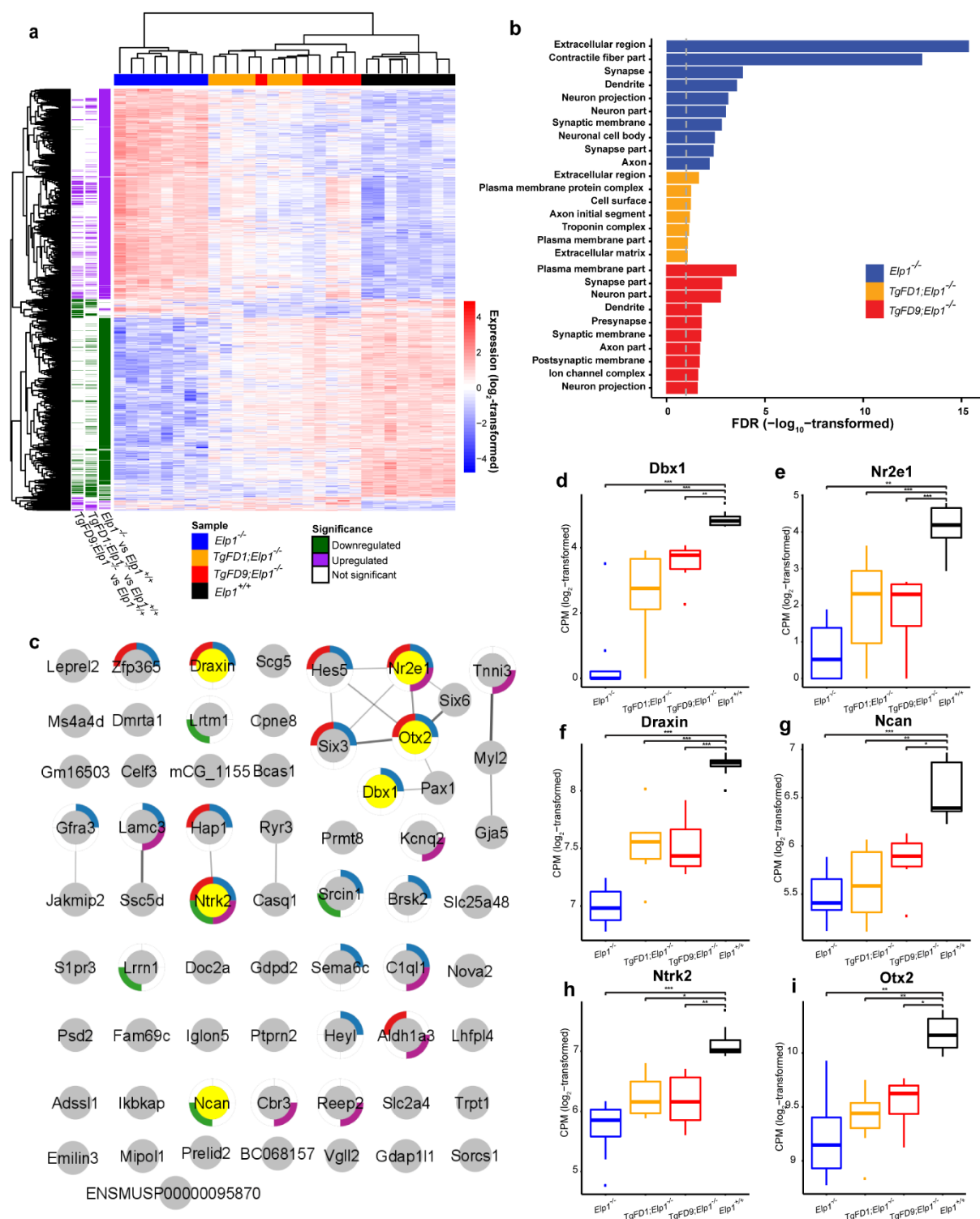
Figure 2. Transcriptome profiling of embryos expressing increasing amount of ELP1. (a)
Heatmap of the 2,619 differentially expressed genes for each genotype compared to WT $Elp1^{+/+}$.
(b) Gene ontology analysis of the downregulated genes in $Elp1^{-/-}$, $TgFD1; Elp1^{-/-}$, $TgFD9; Elp1^{-/-}$

674   embryos. The graph shows FDR values for the most significant specific GO terms (see also

675   Supplementary Table 2). (**c**) STRING interaction map for the 71 genes that were significantly

676   downregulated in all three KO genotypes (see also Supplementary Table 1). Each pie indicates a

677   gene. Note, *Ikbkap* is the alternative name of *Elp1*. The pies surrounded by colored edges indicate

678   genes involved in neurogenesis (GO:0022008, *blue*), brain development (GO:0007420, *red*),

679   regulation of synapse organization (GO:0050807, *green*), and nervous system process

680   (GO:0050877, *purple*), respectively. The pies in yellow indicate the genes whose expression are

681   highlighted in panels d-i. (**d-i**) box-and-whisker plots showing the expression of key genes for

682   nervous system development that are significantly downregulated across all three KO genotypes:

683   Dbx1 (**d**), Nr2e1 (**e**), Draxin (**f**), Ncan (**g**), Ntrk2 (**h**) and Otx2 (**i**). *P <0.05, **P <0.01 and ***P

684   < 0.001, Welch's *t* test.

685

686
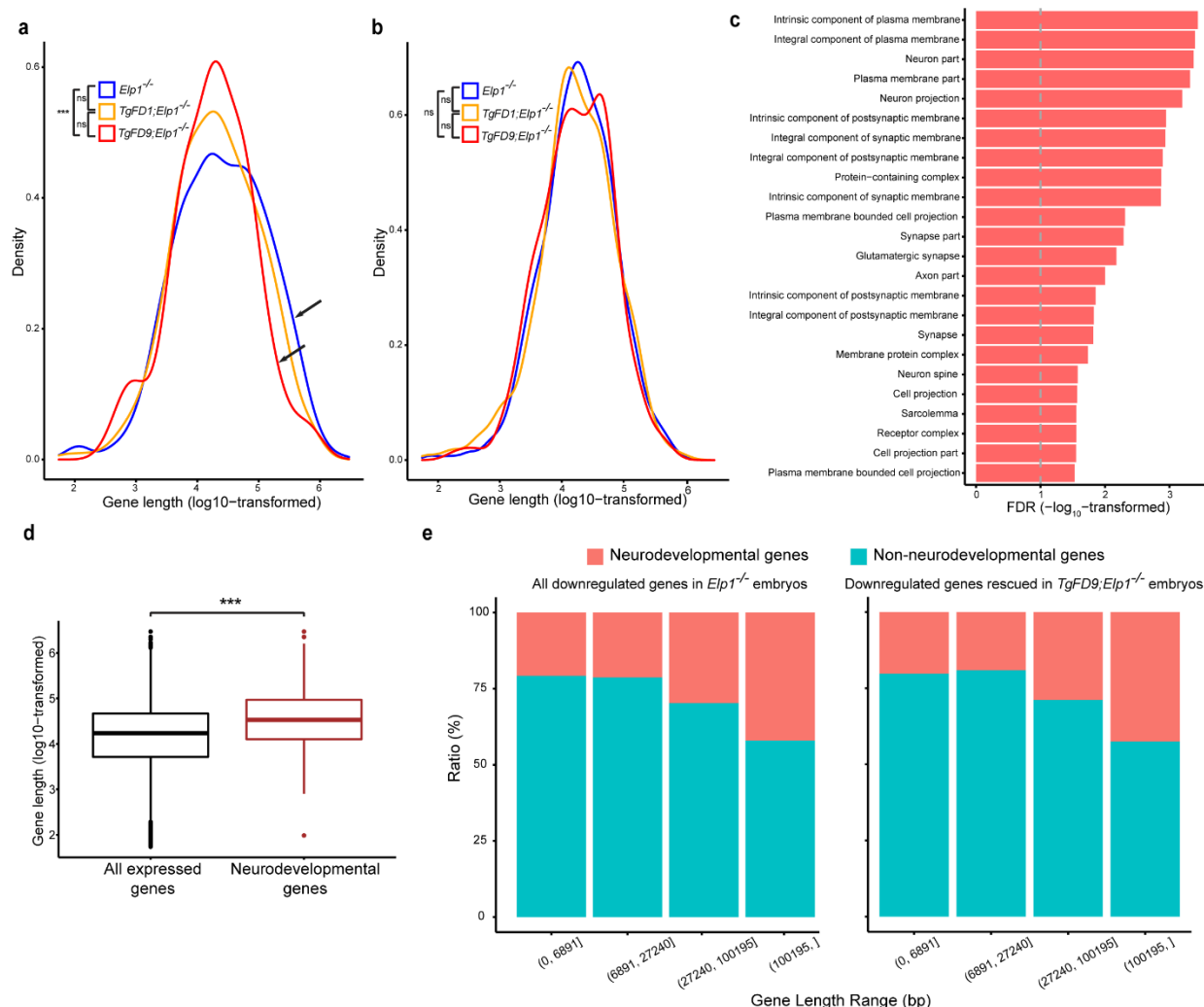
Figure 3. Gene length distribution in embryos expressing different ELP1 amounts. (a) Length distribution of the down-regulated genes in *Elp1⁻/⁻, TgFD1; Elp1⁻/⁻* and *TgFD9; Elp1⁻/⁻* embryos. Arrows indicate the comparison between *Elp1⁻/⁻* and *TgFD9; Elp1⁻/⁻* distribution. **(b)** Length distribution of the up-regulated genes in *Elp1⁻/⁻, TgFD1; Elp1⁻/⁻* and *TgFD9; Elp1⁻/⁻* embryos. In both **(a)** and **(b)**, K-S test was applied. **(c)** Gene ontology analysis of the 247 long genes (> 100 kb) downregulated in *Elp1⁻/⁻* embryos and rescued at *TgFD9; Elp1⁻/⁻*. The graph shows FDR values for each specific GO term (see also Supplementary Table 3). **(d)** Comparison between length distributions of all expressed genes in RNASeq with the length distribution of neuronal genes (GO:0048666 in Gene Ontology database) via two-tailed, unpaired Welch's *t* test. **(e)** Ratio of

696   neurodevelopmental and non-neurodevelopmental genes in different gene-length ranges. The *left*

697   panel represents all downregulated genes in *Elp1^-/-^* embryos, while the *right* panel represents genes

698   which expression was rescued in *TgFD9; Elp1^-/-^* embryos. The gene-length ranges were

699   determined in order to have an equal number of genes in each range. In the figure, ns, not

700   significant and *P < 0.05; **P < 0.01; ***P < 0.001.

701
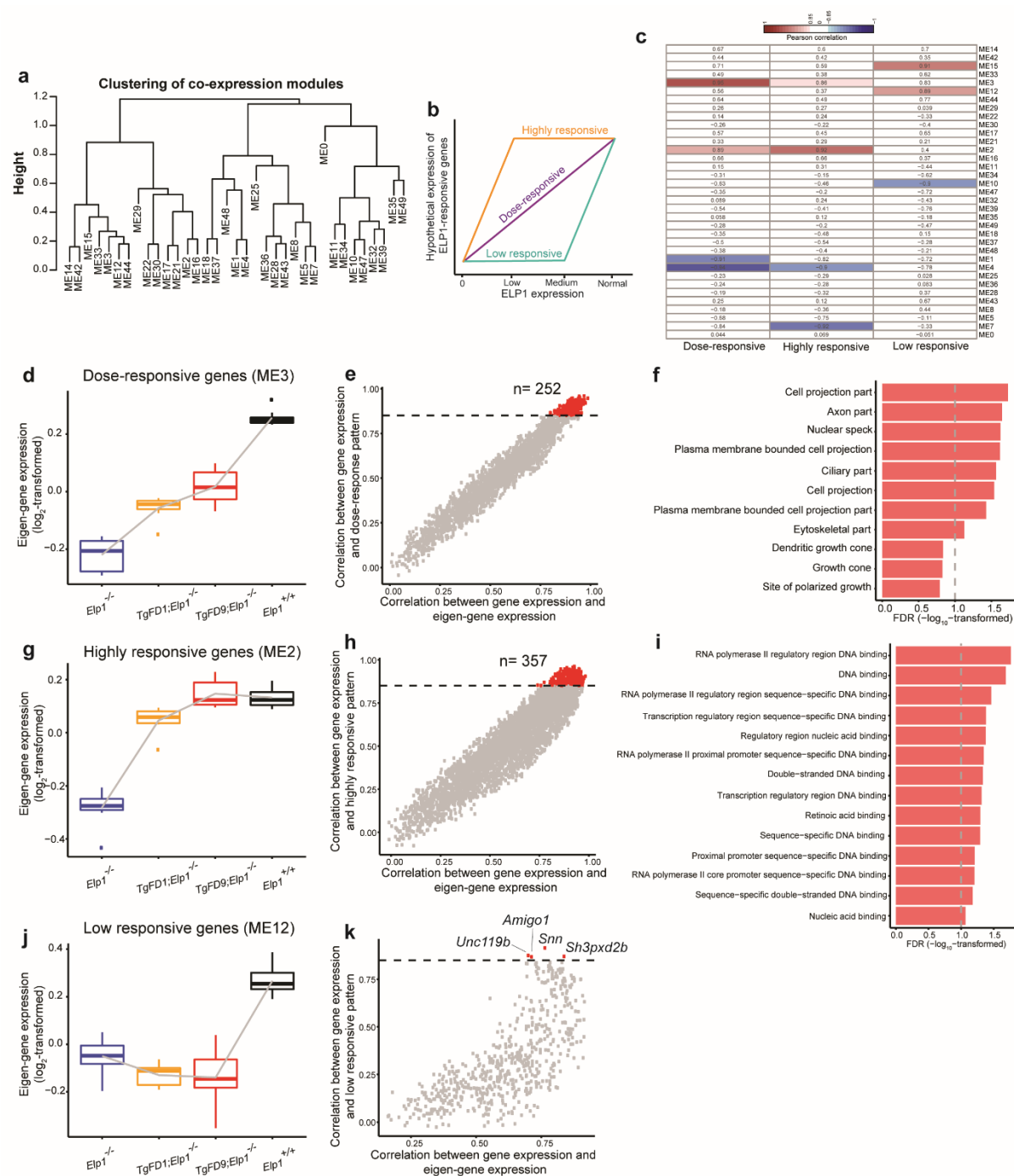
703 **Figure 4. Identification of *ELP1*-responsive gene patterns.** (**a**) Modules of distinct Eigen (ME)

704 analysis across embryos expressing increasing amount of *ELP1* identified thirty-five distinct MEs

705 (see also Materials and Methods). (**b**) Hypothetical expression trajectories of *ELP1*-responsive

706 genes. (**c**) Heatmap displays the Pearson correlation between the eigen values of each co-

707 expression module and the eigen values of each hypothetical expression trajectory. The blue

708 domain indicates a negative correlation while the red domain indicates a positive correlation. The

709 number in each grid demonstrates the Pearson correlation coefficient. (**d**) Boxplot displays the

710 eigen values of gene expression of ME3 at each genotype. (**e**) Module membership of each gene

711 in ME3. Each dot represents a gene. The x-axis demonstrates the Pearson correlation between gene

712 expression and module eigen values of ME3 while the y-axis demonstrates the Pearson correlation

713 between gene expression and the eigen values of each hypothetical expression trajectory. The

714 horizontal dashed line shows a correlation coefficient of 0.85. (**f**) Gene ontology analysis of the

715 252 dose-responsive genes which expression strictly increase as monotonic function of *ELP1*. The

716 graph shows FDR values for each specific GO term (see also Supplementary Table 4). (**g**) Boxplot

717 displays the eigen values of gene expression of ME2 at each genotype. (**h**) Module membership

718 of each gene in ME2. (**i**) Gene ontology analysis of the 357 highly responsive genes whose

719 expression is completely restored in *FD9/KO* embryos. The graph shows FDR values for each

720 specific GO term (see also Supplementary Table 4). (**j**) Boxplot displays the eigen values of gene

721 expression of ME12 at each genotype. (**k**) Module membership of each gene in ME12. *$P < 0.05$;

722 **$P < 0.01$; ***$P < 0.001$, Welch's *t* test.