



HAL
open science

Somatic genetic rescue of a germline ribosome assembly defect

Shengjiang Tan, Laëtitia Kermasson, Christine Hilcenko, Vasileios Kargas, David Traynor, Ahmed Boukerrou, Norberto Escudero-Urquijo, Alexandre Faille, Alexis Bertrand, Maxim Rossmann, et al.

► **To cite this version:**

Shengjiang Tan, Laëtitia Kermasson, Christine Hilcenko, Vasileios Kargas, David Traynor, et al.. Somatic genetic rescue of a germline ribosome assembly defect. *Nature Communications*, 2021, 12 (1), 10.1038/s41467-021-24999-5 . hal-03330053v1

HAL Id: hal-03330053

<https://hal.science/hal-03330053v1>

Submitted on 5 Oct 2021 (v1), last revised 31 Aug 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Somatic genetic rescue of a germline ribosome assembly defect

Shengjiang Tan,^{1,2,3,*} Laëtitia Kermasson^{4,*}, Christine Hilcenko^{1,2,3}, Vasileios Kargas^{1,2,3}, David Traynor^{1,2,3}, Ahmed Z Boukerrou^{1,2,3}, Norberto Escudero-Urquijo^{1,2,3}, Alexandre Faille^{1,2,3}, Maxim Rossmann^{1,2,3}, Beatriz Goyenechea^{3,§}, Alexis Bertrand⁴, Jonathan Moreil⁴, Olivier Alibeu⁵, Blandine Beaupain⁶, Christine Bôle-Feysot⁵, Stefano Fumagalli^{7,8}, Sophie Kaltenbach^{9,10}, Jean-Alain Martignoles¹¹, Cécile Masson¹², Patrick Nitschké¹², Mélanie Parisot⁵, Aurore Pouliet⁵, Isabelle Radford-Weiss^{9,10}, Frédéric Tores¹², Jean-Pierre de Villartay⁴, Mohammed Zarhrate⁵, Peter J Bond^{13,14}, Christine Bellanné-Chantelot¹⁵, Isabelle Callebaut^{16,°}, François Delhommeau^{11,°}, Jean Donadieu^{17,°}, Alan J Warren^{1,2,3,@,#}, Patrick Revy^{4,@,#}

¹Cambridge Institute for Medical Research, Cambridge Biomedical Campus Keith Peters Building, Hills Rd, Cambridge CB2 0XY, United Kingdom.

²Wellcome Trust-Medical Research Council Stem Cell Institute, Jeffrey Cheah Biomedical Centre, Puddicombe Way, Cambridge Biomedical Campus, Cambridge, CB2 0AW, UK.

³Department of Haematology, University of Cambridge School of Clinical Medicine, Jeffrey Cheah Biomedical Centre, Puddicombe Way, Cambridge Biomedical Campus, Cambridge, CB2 0AW, UK.

⁴Université de Paris, Imagine Institute, Laboratory of Genome Dynamics in the Immune System, Equipe Labellisée Ligue contre le Cancer, INSERM UMR 1163, F-75015, Paris, France.

⁵INSERM Unité Mixte de Recherche 1163, Structure Fédérative de Recherche Necker INSERM US24/CNRS UMS3633, Genomic Core Facility, Paris Descartes-Sorbonne Paris Cité University, Imagine Institute, Paris, France.

⁶French Neutropenia Registry, Assistance Publique-Hôpitaux de Paris, Trousseau Hospital, Paris, France.

⁷Institut Necker Enfants Malades, Paris, France.

⁸INSERM, U1151, Université Paris Descartes Sorbonne Cité, Paris, France.

⁹Université Paris Descartes, Faculté de Médecine Sorbonne Paris Cité.

¹⁰Service de cytogénétique, Hôpital Necker, Assistance Publique-Hôpitaux de Paris.

¹¹Sorbonne Université, Inserm, Centre de Recherche Saint-Antoine, AP-HP, Hôpital Saint-Antoine, Hématologie Biologique, F-75012 Paris.

¹²INSERM Unité Mixte de Recherche 1163, Bioinformatics Platform, Paris Descartes-Sorbonne Paris Cité University, Imagine Institute, Paris, France.

¹³Bioinformatics Institute (A*STAR), 30 Biopolis Street, 07-01 Matrix, Singapore 138671, Singapore

¹⁴Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Singapore.

¹⁵Dept of Genetics, Pitié-Salpêtrière Hospital, Sorbonne University, Paris, France.

¹⁶Sorbonne Université, Muséum National d'Histoire Naturelle, UMR CNRS 7590, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, IMPMC, 75005 Paris, France.

¹⁷Service d'Hémo-Oncologie Pédiatrique, Assistance Publique-Hôpitaux de Paris Hôpital Trousseau, Registre des neutropénies-Centre de référence des neutropénies chroniques, Paris, France.

[§]Current address: PolyProx Therapeutics, Babraham Research Campus, Cambridge, CB22 3AT, UK.

* Joint first authors

° These authors contributed equally to the study

Corresponding author

@ Co-senior authors

48 **List of supplementary data**

49 Suppl. Figure 1: FISH probes

50 Suppl. Figure 2: Cumulative VAF vs age and cumulative VAF vs mutation count

51 Suppl. Figure 3: Correlation between *EIF6* mutation and blood parameters

52 Suppl. Figure 4: eIF6 Alignment

53 Suppl. Figure 5: Detection of Sbds by immunofluorescence in *Drosophila* mitotic cells

54

55 Suppl. Table S1: List of SDS patients (Excel file)

56 Suppl. Table S2: CADD scores of all *EIF6* SNVs (Excel file)

57 Suppl. Table S3: SNPs /BAF (Excel file)

58 Suppl. Tables S4A, S4B: *Drosophila* genotypes and strains droso

59 Suppl. Table S5: Plasmids (Human)

60 Suppl. Table S6: Oligonucleotides (Human)

61 Suppl. Table S7: Oligonucleotides (*Dictyostelium*)

62 Suppl. Table S8: Strains (yeast)

63 Suppl. Table S9: Oligonucleotides (yeast)

64 Suppl. Table S10: Plasmids (yeast)

65 Suppl. Table S11: Oligonucleotides (*Drosophila*)

66 Suppl. Table S12: Antibodies

67

68

69

70 **Abstract**

71 Indirect somatic genetic rescue (SGR) of a germline mutation is thought to be rare in inherited
72 Mendelian disorders. Here, we establish that acquired mutations in the *EIF6* gene are a frequent
73 mechanism of SGR in Shwachman-Diamond syndrome (SDS), a leukemia predisposition disorder
74 caused by a germline defect in ribosome assembly. Biallelic mutations in the *SBDS* or *EFL1* genes in
75 SDS impair release of the anti-association factor eIF6 from the 60S ribosomal subunit, a key step in
76 the translational activation of ribosomes. We identified diverse mosaic somatic genetic events (point
77 mutations, interstitial deletion, reciprocal chromosomal translocation) in SDS hematopoietic cells that
78 reduce eIF6 expression or disrupt its interaction with the 60S subunit, thereby conferring a selective
79 advantage over non-modified cells. SDS-related somatic *EIF6* missense mutations that reduce eIF6
80 binding to the 60S subunit suppress the defects in ribosome assembly and protein synthesis in multiple
81 *SBDS*-deficient species including yeast, *Dictyostelium* and *Drosophila*. Our data suggest that SGR is a
82 universal phenomenon that may influence the clinical evolution of diverse Mendelian disorders and
83 support eIF6 suppressor mimics as a therapeutic strategy in SDS.

84

85

86

87

88

89

90

91

92 **Introduction**

93 In normal individuals, somatic mutations and chromosomal alterations accumulate with age in
94 cells from diverse tissues, including the hematopoietic system¹⁻⁹. The accumulation of spontaneous
95 genetic variations may contribute to age-related disease, organismal aging, and tumorigenesis^{10,11}.
96 However, more than 40 years ago, Weill and Reynaud proposed that in certain circumstances, somatic
97 mutations might be beneficial to the cell without inducing disease or cellular transformation¹². In
98 inherited Mendelian diseases, this phenomenon, dubbed somatic genetic rescue (SGR)¹, is considered
99 rare and has mainly been observed in hematopoietic disorders, where it may confer a selective
100 advantage and promote recovery of hematopoiesis by counteracting the deleterious effect of the
101 germline mutation¹⁴⁻¹⁶. In most cases, SGR affects the germline mutated gene (direct SGR¹³). In
102 contrast, indirect SGR involves the acquisition of somatic mutations in a distinct gene that participates
103 in the same pathway that is altered by the germline mutation¹³. For instance, indirect SGR has been
104 highlighted in three independent studies on telomeropathies where somatic promoter-activating
105 mutations in *TERT*, the gene encoding the telomerase catalytic subunit that elongates telomeres, were
106 identified in blood cells from patients with germline mutations in genes involved in telomere length
107 regulation, *i.e.* *TERC*, *PARN* and *NHP2*¹⁷⁻¹⁹. To the best of our knowledge, indirect SGR has only been
108 described to date in the telomeropathies.

109 Shwachman-Diamond syndrome (SDS; OMIM #260400) is a rare autosomal recessive disease
110 characterized by bone marrow failure, poor growth, skeletal defects, exocrine pancreatic insufficiency,
111 and predisposition to hematological malignancies²⁰. Biallelic mutations in *SBDS* are the predominant
112 cause of SDS, but biallelic *EFL1* mutations have also been identified². *SBDS* and the GTPase *EFL1*
113 cooperate to evict the anti-association factor eIF6 (yeast Tif6) from the nascent large ribosomal
114 subunit²³⁻²⁵, an essential prerequisite that allows the 60S and 40S subunits to join to form mature
115 actively translating 80S ribosomes. Hence *SBDS* and *EFL1* deficiencies are considered as
116 ribosomopathies since they lead to impaired ribosomal subunit joining and reduced protein synthesis
117 as a consequence of defective eIF6 eviction from the 60S subunit^{20, 23}.

118 Recurrent mosaic acquired interstitial deletions of chromosome 20 (del(20q)) encompassing
119 the *EIF6* gene have been detected in bone marrow cells from some individuals with SDS³. This
120 observation led to the proposal that a reduced dose of eIF6 due to del(20q) might be advantageous to
121 SDS cells by bypassing the defect in ribosomal subunit joining, representing a novel mechanism of
122 indirect SGR^{13, 26-28}. However, the minimal del(20q) region characterized in SDS hematopoietic cells
123 in SDS spanned 2.2 Mb, encompassing 28 genes in addition to *EIF6*²⁸. Furthermore, del(20q) is one of
124 the most common mosaic chromosomal alterations (mCAs) associated with age-related clonal
125 hematopoiesis⁷⁻⁹. Thus, it remains unclear whether *EIF6* haploinsufficiency generated by del(20q)
126 indeed represents a *bona fide* mechanism of indirect SGR in SDS hematopoietic cells.

127 Here, we hypothesize that acquired somatic mutations in the *EIF6* gene might provide a
128 selective advantage for hematopoietic stem and progenitor cells (HSPCs) in SDS that promotes their
129 clonal expansion. To test this hypothesis we performed ultra-deep sequencing of the *EIF6* gene in
130 hematopoietic cells from 40 individuals with SDS carrying biallelic germline *SBDS* mutations. We
131 identified mosaic somatic *EIF6* mutations in 60 % of SDS patients but not in healthy donors. By
132 combining structural modelling and molecular dynamics (MD) simulations with functional studies in
133 yeast, *Dictyostelium discoideum* and *Drosophila melanogaster*, we establish that conserved *EIF6*
134 missense mutations that map to the interface with the 60S subunit bypass SBDS deficiency by
135 reducing the affinity of eIF6 and rescuing the defects in ribosome assembly and global protein
136 synthesis. Our results establish that acquisition of somatic *EIF6* mutations is a frequent mechanism of
137 indirect somatic genetic rescue in hematopoietic cells in SDS, suggesting a strategy for the
138 development of disease-modifying targeted therapeutics in SDS.

139

140

141

142

143 **Results**

144 ***EIF6* mutations as a mechanism of somatic genetic rescue in SDS.**

145 To determine whether acquired mutations in *EIF6* represent a mechanism of SGR in
146 hematopoietic cells in SDS, we performed ultra-deep targeted sequencing of the full genomic *EIF6*
147 gene (introns/exons) after hybridization-based capture with biotinylated ssDNA probes designed and
148 prepared to target a 123 kb chromosomal locus encompassing *EIF6* (chr20:35,256,992-35,380,631
149 according to the GRCh38.p12 assembly of the human reference genome). We analyzed a total of 14
150 SDS patients (hereafter denoted SBDS) carrying biallelic germline mutations in the *SBDS* gene (mean
151 age, 14.7 years; range 1-38.2; DNA extracted from blood n = 8; DNA extracted from bone marrow n =
152 6; **Suppl. Table S1**). We also tested 5 SDS patients who had undergone hematopoietic stem cell
153 transplantation (denoted SBDS post-HSCT; DNA extracted from blood) and fully reconstituted their
154 hematopoietic system as inferred by wild type (WT) *SBDS* sequence in peripheral blood cells (100 %
155 donor, not shown). In addition, we tested 5 patients with neutropenia of uncharacterized genetic origin
156 (denoted Neutro Unkn; in 4, DNA was extracted from blood, in 1 from bone marrow), one SDS-like
157 patient carrying biallelic *SRP54* mutations³⁰ (denoted SRP54; DNA from blood), and 15 healthy age-
158 matched donors (denoted Ctl, DNA from blood). After removing duplicates, ultra-deep *EIF6*
159 sequencing provided a mean depth of 2,807X (ranging from 718X to 7,940X). To accurately identify
160 *EIF6* genetic variants with low rates of somatic mosaicism, we considered all detected genetic variants
161 in the *EIF6* coding sequence with variant allele frequencies (VAF) ≥ 0.5 % as somatic *EIF6*
162 mutations. Using this criterion, we did not detect *EIF6* mutations in the 15 healthy controls, the 5 SDS
163 patients post-HSCT, the 5 patients with neutropenia from unknown molecular origin or the SRP54-
164 deficient patient. In contrast, we detected a total of 10 *EIF6* mutations in 7 out of the 14 SDS patients
165 (50 %) (**Figure 1A**). Nine mutations corresponded to single nucleotide variation (SNVs; 8 missense
166 and 1 nonsense), while one was a 5 bp deletion predicted to cause a frameshift and a premature stop
167 codon (**Figure 1B**). The combined annotation-dependent depletion (CADD) score represents a
168 predictive indicator of the deleterious effect of a genetic variant³¹. Noticeably, the mean CADD score

169 for the 9 *EIF6* SNVs identified in SDS patients was significantly higher than the mean CADD score
170 generated by all possible SNVs in the *EIF6* coding sequence (synonymous, missense, nonsense,
171 start/stop loss; **Figure 1C and Suppl. Table S2**). This observation suggests that clones carrying *EIF6*
172 SNVs predicted to have high deleterious impact were preferentially amplified in blood cells from SDS
173 patients. Moreover, the absence of somatic *EIF6* mutations in normal individuals suggests that they
174 are not favored in cells in normal conditions.

175 The mean VAF of the 10 *EIF6* mutations was 2.15 % (range 0.51-12.32 %). In 3 SDS
176 patients, we detected 2 different *EIF6* mutations (**Figure 1D and Suppl. Table S1**), indicating that
177 distinct *EIF6* mutated clones can emerge independently within the same individual. Strikingly, the
178 same somatic mutation (g.20:33868509A>G; c.317A>G) leading to the eIF6 substitution N106S was
179 detected in four unrelated SDS patients with a VAF ranging from 0.87 to 12.32 %. This suggested to
180 us that N106S might represent a recurrent somatic mutation with a key functional impact in SBDS
181 deficient cells (see below) (**Figure 1D and Suppl. Table S1**).

182 We next analyzed the B-allele frequency (BAF) across all heterozygous single nucleotide
183 polymorphisms (SNPs) located in the *EIF6* gene. In 9 SDS patients and 10 healthy individuals in
184 whom SNPs were informative, the BAFs were around 0.5 as expected for heterozygous SNPs in
185 diploid cells (**Figure 1E, Suppl. Table S3**)³². In contrast, two SDS patients (SBDS-1 and SBDS-9)
186 exhibited a sharp BAF deviation from 0.5 (**Figure 1E, Suppl. Table S3**), suggesting the existence of a
187 mosaic genetic deletion encompassing the *EIF6* gene. The combination of cytogenetic analysis using
188 specific FISH probes located nearby the *EIF6* locus (**Suppl. Fig. 1**) and array comparative genomic
189 hybridization (CGH) confirmed the presence of an interstitial 20q11.21-q13.2 deletion encompassing
190 *EIF6* in a bone marrow sample from patient SBDS-9 that was estimated to affect 37 % of cells
191 (**Figures 1F, G, and Suppl. Table S1**).

192 Although ultra-deep *EIF6* sequencing did not detect *EIF6* mutations in bone marrow cells
193 from patient SBDS-3, cytogenetic analysis highlighted a reciprocal translocation t(16;20)(q24;q11.2)
194 in 2 out of 20 metaphases (**Suppl. Table S1 and not shown**). Since the *EIF6* gene maps to 20q11.2,

195 we wondered whether the breakpoint in chromosome 20 was located within the *EIF6* gene. A search
196 for chimeric reads from the ultra-deep sequencing containing both the *EIF6* gene and chromosome 16
197 sequences unveiled chimeric sequences in SBDS-3 but not in 4 controls. Analysis of chimeric reads
198 precisely positioned the translocation breakpoints in chromosome 20 within intron 4-5 of *EIF6* and in
199 a non-coding region of chromosome 16 between the *COX4* (9,175 bp at 5' side) and the *IRF8* genes
200 (86,642 bp at 3' side) (**Figure 1H**). We conclude from this analysis that the translocation
201 t(16;20)(q24;q11.2) detected in a mosaic state in bone marrow cells from SBDS-3 disrupted one copy
202 of *EIF6* to cause haploinsufficiency.

203 We conclude that multiple distinct somatic genetic events affecting the *EIF6* gene are frequent
204 in hematopoietic cells in SDS but not in healthy individuals. These *de novo* mosaic genetic
205 modifications consist of chromosomal alterations affecting *EIF6* (interstitial del(20q), reciprocal
206 translocation) or somatic point mutations in the *EIF6* coding sequence (nonsense, missense, and small
207 deletions). These findings support our hypothesis that *EIF6* mutations indeed represent a mechanism
208 of indirect SGR that promotes clonal expansion in the context of a germline ribosome assembly defect
209 in SDS.

210

211 **Spectrum of acquired somatic *EIF6* mutations in SDS**

212 To strengthen this initial genetic analysis and identify further somatic *EIF6* mutations, we
213 performed ultra-deep *EIF6* sequencing of a larger cohort consisting of 26 SDS patients carrying
214 biallelic *SBDS* mutations (mean age: 15.4 years, range 0.47-52.2 years; DNA from blood cells n = 3;
215 DNA from bone marrow n = 23, **Suppl. Table S1**) and 25 age-matched healthy individuals (DNA
216 from blood cells, n = 25). To increase the depth of sequencing with a limited quantity of DNA, we
217 modified the hybridization-based capture strategy by using the *EIF6* cDNA (1,016 bp) as sequence
218 bait. After duplicate removal, this approach yielded a mean depth of 26,873X (range 11,140-47,185X).
219 In this setting we considered all genetic variants in the *EIF6* coding sequence with a VAF \geq 0.25 % as

220 somatic *EIF6* mutations. In total, we identified 56 *EIF6* mutations in 17 of the 26 SDS patients (65.3
221 %), but none in the 25 healthy donors (**Figure 2A**). Up to 8 different *EIF6* mutations were present in
222 the same individual (mean 2.07; range 0-8) (**Figure 2B**). The mean VAF in patients carrying *EIF6*
223 mutations was 1.43 % (range 0.25-27.9 %) (**Figure 2C**). Congruent with the reported accumulation of
224 somatic mutations in hematopoietic cells over time^{5,6}, we found a slight but significant positive linear
225 correlation between the *EIF6* mutation count and age ($r = 0.4105$; $p = 0.0335$; Pearson correlation)
226 (**Figure 2D**). However, the cumulative VAF per patient among SDS patients carrying *EIF6* mutations
227 did not correlate with age or mutation count ($r = 0.04629$; $p = 0.86$ and $r = 0.03589$; $p = 0.8912$,
228 respectively, **Suppl. Figure 2**). Among the 56 *EIF6* mutations, 46 were SNVs (82.1 %) that mainly
229 consisted of C>T transitions (51.1 %), a mutational spectrum that likely reflects the spontaneous
230 deamination of cytosine residues observed in hematopoietic cells from normal individuals^{5,6,33} (**Figure**
231 **2E**). Thirty-one were nucleotide substitutions leading to missense mutations (55.3 %), 20
232 corresponded to nonsense or small indels inducing frameshift and premature stop codons (35.7 %), 4
233 were synonymous (7.1 %) and one corresponded to loss of the start codon (1.8 %; M1L) (**Figure 2F**).
234 The mean CADD score of these 56 SNVs was significantly higher than the mean CADD scores of all
235 possible *EIF6* SNVs (**Figure 2G**). Furthermore, the mutation spectrum among the SNVs highlighted
236 3.4 times more non-synonymous mutations than expected neutrally, as inferred by the ratio of non-
237 synonymous to synonymous variants ($dN/dS = 3.4$; with $dN/dS = 1$ representing neutrality)³⁴.
238 Together, these results further argue that *EIF6* mutations predicted to have a functional impact are
239 positively selected in hematopoietic cells in SDS.

240 Collectively, we identified a total of 66 somatic eIF6 mutations in 24 out of 40 SDS patients
241 (60 %) from two independent genetic analyses, of which 54 (81.8 %) are missense mutations
242 (**Figures 3A, B**) that are distributed throughout the protein (**Figure 3C**). Five SDS patients (12.5 %)
243 exhibited clones with a VAF higher than 5 %. The clones with a VAF > 5 % harbored either nonsense
244 (Q93*, VAF= 6.34 %; Q145*, VAF=10 %) or missense *EIF6* mutations (G69D, VAF=27.9 %; R96W,
245 VAF=7.59 %; N106S, VAF=12.32 %) and 19 SDS patients (47.5 %) exhibited a cumulative VAF > 1
246 % (**Figure 3A and Suppl. Table S1**). Strikingly, 7 amino acids (aa) (N66, G69, R96, N106, D112,

247 L133 and V135) were recurrently targeted by missense mutations (**Figures 3A, B and Suppl. Table**
248 **S1**): 6 patients carried 7 SNVs affecting residue G69, generating distinct missense substitutions
249 (G69A; G69S; G69V; G69D) (**Figure 3B**); 4 patients carried the same R96W substitution; 4 patients
250 carried mutations affecting residue N106 (N106S; N106D), 2 patients had mutations affecting residue
251 N66 (N66H; N66K); 2 patients harbored mutations affecting residue D112 (D112N; D112A); 2
252 patients carried mutations affecting residue L133 (L133P; L133I) and 2 patients harbored the same
253 V135M mutation (**Figure 3B**). Noteworthy, among the somatic missense mutations revealed, G14S
254 and N106S (**Figure 3B**) were previously identified as suppressor mutations that bypassed the
255 ribosome assembly defect in yeast cells lacking the SBDS homolog, Sdo1²⁵. These findings further
256 support the notion that our ultra-deep sequencing had identified mutations that drive positive clonal
257 selection in the context of human SBDS deficiency *in vivo*, likely by increasing fitness at the cellular
258 level.

259 There was no statistical correlation between the presence of *EIF6* mutations (or their VAF)
260 and hemoglobin, platelet or white cell count in SDS individuals at the time of DNA sampling for *EIF6*
261 sequencing (**Suppl. Figure 3 and Suppl. Table S1**).

262 In sum, our genetic analysis demonstrates that clones carrying somatic genetic mutations in
263 the *EIF6* gene are frequent in blood and bone marrow cells from SDS patients, suggesting that they
264 provide a cellular selective advantage in this context. Some of these events, *i.e.* interstitial deletion,
265 reciprocal translocation, nonsense and small indels are predicted to generate *EIF6* null alleles,
266 provoking *EIF6* haploinsufficiency. However, as the majority of *EIF6* mutations detected by ultra-
267 deep sequencing consisted of missense mutations, we set out to assess their impact by structural,
268 biochemical and functional analysis.

269

270

271

272 **Three mutation hotspots in eIF6**

273 We focused on the eIF6 amino acids (N66, G69, R96, N106, D112, L133, and V135) that are
274 recurrently targeted in SDS. These residues are highly conserved across species, with 5 out of the 7
275 amino acids conserved from *Homo sapiens* to the archaeon *Methanopyrus kandleri* (**Suppl. Figure 4**).
276 We used the 2.4 Å cryo-EM structure of human eIF6 bound to the human 60S subunit (PDBID: XXX)
277 to map the eIF6 mutations (**Figure 4A**). As first described for the two homologs in
278 *Methanocaldococcus jannaschii* and *Saccharomyces cerevisiae*³⁷, eIF6 has a pentain fold consisting of
279 five repeated subunits, with 3-stranded β -sheets arranged as blades around a five-fold axis of pseudo-
280 symmetry (**Figure 4A**). The radial arrangement of these subunits is closed by a "velcro" strategy, with
281 the last β -strand of the last blade provided by the N-terminal β -strand, as in β -propeller 3D structures.
282 Five small helices form an inner ring that includes a position invariably occupied by a small amino
283 acid residue (G, A) to allow tight packing (**Figure 4A, Suppl. Figure 4**). Both sides of the pentain
284 fold form flat surfaces, one of which forms the interface with ribosomal proteins uL14 (RPL23), eL24
285 (RPL24), uL3 (RPL3) (using the new nomenclature⁴) and the sarcin-ricin loop (SRL) (**Figure 4A**). We
286 mapped the seven recurrently mutated amino acids to three main hotspots. The first hotspot
287 (highlighted in black in **Figure 4A**), includes residue N106 (blade 3) which is mutated (N106S and
288 N106D) in 6 SDS individuals (**Figure 3B**). The side chain of N106 forms hydrogen (H)-bonds with
289 the main chain oxygen atoms of uL14 residues A133 and A136 (**Figure 4B**). In addition, the backbone
290 nitrogen of N106 forms an intra-molecular H-bond with the backbone oxygen of residue A103. In
291 turn, the backbone nitrogen of A103 forms an H-bond with the backbone oxygen of uL14 residue
292 G137. The side chain and backbone atoms of N106 also form intra-protein H-bonds with the side-
293 chain and backbone atoms of R61 (blade 2) (**Figure 4B**). A network of H-bonding interactions links
294 R61 (blade 2) with the main chain oxygen atoms of G14 (blade 1), I58 and G60 (blade 2) and G149
295 (blade 4) (**Figure 4C**). Interestingly, an R61L mutation was recently identified in a patient with an
296 SDS-like clinical phenotype⁵. The second hotspot (highlighted in cyan in **Figure 4A**) contains 5 aa
297 that cluster at the interface between blade 2 (N66 and G69) and blade 3 (D112, L133 and V135)
298 (**Figure 4D**). Residue N66 forms H-bonds with the main chain oxygen atoms of G69 and L133, while

299 the side chains of L133 and V135 form hydrophobic interactions. At the solvent exposed core of eIF6,
300 D112 forms H-bonds with the backbone nitrogen of R67 and the side chain of N156 (blade 4) as part
301 of a wider network of H-bonds involving residues N21 (blade 1), N111 (blade 3) and D201 (blade 5)
302 (**Figure 4E**). Mutation of any of the five residues lying within the second hotspot is predicted to
303 destabilize the pentain fold as a whole. The third hotspot (highlighted in red in **Figure 4A**) contains
304 residue R96 (at the end of strand β 3 of blade 2), that forms an intra-protein H-bond with the backbone
305 of residue T76 (blade 2) (**Figures 4F**). This interaction may help promote polar interactions between
306 eIF6 residue D78 (blade 2) and eL24 residue K2. The recurrent R96W mutation, identified in 4 SDS
307 patients, likely disrupts both the stability of blade 2 and the interaction of eIF6 with eL24.

308

309 ***EIF6* mutations rescue fitness defect of SBDS-deficient cells *in vivo***

310 We next set out to test the impact of the N66H, G69S, R96W, N106S, D112N, L133P, and
311 V135M mutations on eIF6 protein expression, stability and function. Immunoblotting of extracts from
312 HEK293T cells transfected with equal amounts of WT and mutant FLAG-tagged eIF6-expressing
313 vectors indicated that all but the N106S mutation reduced eIF6 expression, consistent with a reduction
314 in eIF6 stability as predicted by the structural analysis (**Figure 5A**). We further verified that the
315 ectopic expression of the FLAG-eIF6 N106S mutant did not affect the expression and/or stability of
316 the endogenous eIF6 protein (**Figure 5B**). These observations suggest that the selective advantage
317 provided by the N106S mutation (**Figures 5A, B**) is not due to reduced eIF6 dosage, in contrast to the
318 N66H, G69S, R96W, D112N, L133P, and V135M variants.

319 We assessed the ability of the eIF6 N106S mutant to interact with the 60S subunit.
320 Immunoblots of sucrose gradient fractions from HEK293T cells transfected with vectors expressing
321 either WT FLAG-eIF6 or N106S proteins indicated that unlike WT FLAG-eIF6, the N106S mutant
322 did not co-sediment with the 60S subunit (**Figures 5C, D**). We next examined the distribution of WT
323 eIF6 versus the mutants T56K (the most potent gain-of-function mutation identified in yeast⁶) and

324 N106S when expressed in *Dictyostelium discoideum* Ax2 cells lacking the endogenous *EIF6* allele by
325 sucrose gradient fractionation and immunoblotting of cell extracts (**Figure 5E**). Both the endogenous
326 and over-expressed WT eIF6 but not the eIF6-T56K or N106S variants, co-fractionated with the 60S
327 subunit. Furthermore, WT eIF6 but not the T56K or N106S variants, induced a functional defect in
328 ribosomal subunit joining in Ax2 cells.

329 We directly tested the ability of SDS-associated eIF6 missense mutations to rescue the fitness
330 defect of SBDS-deficient cells *in vivo* by engineering a conditional mutation in the yeast SBDS
331 homolog Sdo1 (*sdo1^{ts}*), based on a temperature-sensitive intein which is spliced out to create a
332 functional Sdo1 protein at the permissive (23 °C) but not the restrictive temperatures (30 °C or 37
333 °C)⁷. Compared with empty vector or WT Tif6 controls, expression of the Tif6-G14S, R61L and
334 N106S mutants (but not G69S, R96W, D112E, L133P and V135M), rescued the fitness defect of
335 *sdo1^{ts}* cells at the restrictive temperatures (**Figure 5F**). Immunoblotting revealed that all but the G14S,
336 R61L and N106S mutations reduced the expression of Tif6 compared with the endogenous Tif6
337 protein (**Figure 5G**). These data confirm that SDS-related Tif6 missense mutations that map to the
338 interface with uL14 act as dominant gain-of-function mutations that are able to bypass the fitness
339 defect caused by Sdo1 deficiency and suggest that mutations that destabilise the Tif6 protein confer
340 loss of function. Given the conservation of eIF6 function from human to prokaryotes, these
341 observations collectively support the hypothesis that in SDS, HSPCs positively select somatic
342 mutations that either impair the interaction of eIF6 with the 60S subunit or reduce eIF6 expression.

343

344 **N106S mutation dynamically disrupts the H-bonding interface between eIF6 and uL14**

345 To provide additional insights into the mechanism by which the SDS-related eIF6 missense
346 mutation N106S destabilizes the interaction interface with uL14, we utilized atomic-resolution MD
347 simulations to study the stability of a solvated complex comprising eIF6, uL14, eL24, uL3 and a
348 double stranded helical segment of the 28S ribosomal RNA. Three 500 ns replica simulations were

349 performed for both the WT system and the in silico eIF6 N106S mutant (**Figures 6A-F**). In the WT
350 simulations, the N106 side chain maintained stable H-bond contacts with the backbone carbonyls of
351 uL14 residues A133 and A136, with an average donor-acceptor distance of 2.9 Å (**Figures 6A, C**).
352 The sidechain amide oxygen atom in N106 also retained its native intramolecular contacts with R61
353 (**Figure 6C**), bridging uL14 with the internal network of eIF6 H-bonding interactions spanning blades
354 1-5, as described above. Thus, simulations of the WT complex demonstrated that the key contacts
355 observed in the cryo-EM structure were largely reproduced (**Figure 4B**). In contrast, similar analysis
356 of the eIF6 mutant revealed significant destabilization around S106. The serine sidechain hydroxyl
357 was only able to form weak, intermittent H-bonds with the backbone carbonyl oxygens of uL14
358 residues A133 and A136 (**Figures 6B, C**) or the guanidinium moiety of eIF6 R61 (**Figure 6C**).
359 Supporting the apparently weakened eIF6-uL14 interface, an influx of water molecules was observed
360 after ~150 ns in one of the mutant simulation replicas, satisfying the H-bonding potential of the eIF6
361 S106 sidechain and uL14 A133 and A136 backbone nitrogens (**Figures 6B, D**). These water
362 molecules persisted at the interface throughout the remainder of the simulation, leading to
363 displacement of the eIF6 core relative to uL14, followed by partial solvation of their interaction
364 interface (**Figures 6E, F**). We conclude that comparative MD simulations of the WT and mutant
365 complexes support the hypothesis that the SDS-related eIF6 N106S mutation disrupts the eIF6-uL14
366 interaction interface and ultimately leads to its solvation, due to the lower propensity for the mutant to
367 satisfy the H-bonding network with uL14.

368

369 ***EIF6* mutations rescue larval lethality of *Sbds*-deficient *Drosophila***

370 We sought to test the general concept that somatic *EIF6* mutations can effectively rescue the
371 deleterious effects of a hypomorphic germline *Sbds* mutation in a whole animal context by harnessing
372 *Drosophila* genetics. Based on their strength as suppressors of the fitness defect of *Sdo1*-deleted yeast
373 cells⁶, we expressed three eIF6 missense mutations (eIF6-C56R, eIF6-Y151H and eIF6-V192F) that
374 map to the uL14 interaction interface of eIF6 (**Figure 7A**) in *Sbds*-deficient *Drosophila*.

375 We initially examined the subcellular localization and function of the *Drosophila* SbdS
376 protein. *Drosophila* SbdS localized to the cytoplasm of ovarian follicle cells and in whole larvae
377 (**Figures 7B, C**) but did not colocalize with the mitotic spindle (**Suppl. Figure 5**). In control
378 experiments, SbdS protein expression was selectively lost in the posterior half of the wing disc in cells
379 expressing *SbdS*^{RNAi} (marked with GFP) (**Figure 7D**). We conclude that *Drosophila* SbdS is a
380 cytoplasmic protein, consistent with the localization of its mammalian and *Dictyostelium*
381 counterparts^{7,8}.

382 To examine the consequences of SbdS deficiency in *Drosophila*, we used RNAi to deplete
383 SbdS in the imaginal disc of the developing wing (denoted *SbdS*^{RNAi/+} in **Figures 7E, F**). SbdS depletion
384 reduced the surface area of the adult wing by 10 % compared with control (**Figure 7F**). A
385 corresponding 27 % increase in cell number (as assessed by hair density) indicated a decrease in cell
386 size. We next generated mutant (*SbdS*^{P/P}) animals homozygous for the insertion of a PiggyBac-element
387 transposon (*PBac*{*WH*}*CG8549*⁰¹⁶⁸⁶) within the 5' untranslated region of the *SbdS* (CG8569) gene, 18
388 nucleotides upstream of the start codon, on the third chromosome at cytological position 65C3 (**Figure**
389 **7G**). In addition, we engineered *SbdS*^{P/P} mutants expressing three independent eIF6 missense mutants
390 eIF6-C56R, eIF6-Y151H and eIF6-V192F (marked with a MYC tag). Immunoblotting of cell extracts
391 revealed a marked reduction in SbdS protein expression with a concurrent increase in the amount of
392 endogenous eIF6 in homozygous *SbdS*^{P/P} mutants alone compared with WT or *SbdS*^{P/P} mutants
393 expressing eIF6-C56R-MYC (**Figure 7H**). Phenotypically, compared with WT or *SbdS*^{P/P} mutants
394 expressing eIF6-C56R-MYC, *SbdS*^{P/P} animals alone exhibited a severe growth defect, with only 10 %
395 of third instar larvae surviving to the early pupal stage (**Figure 7I**). Remarkably, all three *EIF6*
396 missense mutant transgenes rescued the viability of the homozygous *SbdS*^{P/P} mutant animals, allowing
397 a significant proportion to hatch as fertile adults (eIF6-C56R, 18.8 %, n = 182; eIF6-Y151H, 71.7 % n
398 = 350; eIF6-V192F, 33.7 %, n = 53) (**Figure 7J**); while overexpression of WT eIF6 induced larval
399 lethality of *SbdS*^{P/P} at the early second instar stage. None of the *EIF6* missense mutant transgenes
400 impaired the viability or fertility of WT *Drosophila*. Importantly, transgenic expression of *Drosophila*
401 or human SBDS rescued the larval lethality of the homozygous *SbdS*^{P/P} mutants (**Figure 7J**),

402 confirming that the mutant phenotype was indeed a consequence of loss of *Sbds* function, attesting to
403 the conservation of SBDS protein function. Immunoblotting of sucrose gradient fractions revealed that
404 expression of the eIF6-C56R mutant rescued the retention of eIF6 on the 60S subunit (**Figure 7K**), the
405 functional impairment of ribosome assembly (**Figure 7K**), the cytoplasmic redistribution of eIF6
406 (**Figures 7L, M**) and the 76 % reduction in global protein synthesis in circulating hemocytes (as
407 assessed by *in vivo* incorporation of O-propargyl-puromycin (OP-Puro)) (**Figure 7N**) observed in
408 *Sbds*^{P/P} mutants compared with WT animals. We conclude that eIF6 missense mutations that map to
409 the interface between eIF6 and uL14 can fully and effectively rescue the deleterious effects of a
410 germline hypomorphic *Sbds* mutation in *Drosophila*.

411

412 **DISCUSSION**

413 In this study, we have identified acquired *EIF6* mutations as a common form of somatic
414 genetic rescue in SDS, a leukemia predisposition disorder caused by a germline defect in ribosome
415 assembly that impairs the release of eIF6 from nascent 60S ribosomal subunits^{2,6-8}. These somatic
416 *EIF6* mutations rescue the primary molecular pathological defect in SDS *in vivo*, either by reducing
417 the dose of eIF6 or by lowering the affinity of eIF6 for the 60S subunit.

418 The development of sensitive and reliable genetic tools has recently enabled the detection of
419 mosaic somatic mutations and spontaneous chromosomal alterations in diverse tissues from normal
420 individuals¹¹. A growing number of studies have demonstrated that such somatic genetic modifications
421 accumulate with age and participate in age-related disease, clonal expansion, and cancer development.
422 However, in the context of Mendelian disease, *de novo* genetic events can counterbalance the
423 deleterious effect of germline mutations, providing the somatically modified cells with a selective
424 advantage compared with their non-modified counterparts. This phenomenon of SGR has been
425 reported in Mendelian hematopoietic disorders where it promotes the clonal expansion of SGR
426 positive cells detectable in blood¹³. In the present study, ultra-deep targeted sequencing has revealed

427 that genetic alterations in the *EIF6* gene that impact the stability or expression of eIF6 or its
428 interaction with the 60S subunit represent a recurrent indirect mechanism of SGR in hematopoietic
429 cells from SDS patients. In agreement with the reported accumulation of somatic genetic alterations
430 over time in hematopoietic cells from normal individuals^{5,6}, we found that the frequency of
431 independent *EIF6* mutations in SDS positively correlates with increasing age. However, the frequency
432 of somatic mutations over time in hematopoietic cells from normal individuals is still a matter of
433 debate¹¹. Strikingly, we detected *EIF6* mutant clones in 4 SDS patients below 10 years of age, one of
434 whom was 3.4 years old. In addition, we detected multiple independent *EIF6* mutant clones (up to 8)
435 in several SDS patients. Together these observations support the idea that the acquisition of somatic
436 mutations in hematopoietic cells is more frequent than previously thought, as they have generally only
437 been unveiled in a context where they provide a selective advantage and promote clonal expansion¹¹.

438 *Sbds* deletion from mesenchymal stem cells in the mouse induces mitochondrial dysfunction,
439 oxidative stress and activation of the DNA damage response (DDR) in HSPCs⁹. These data led to the
440 proposal that mesenchymal inflammation promotes genotoxic stress in SDS HSPCs and drives the
441 evolution to leukemia. However, the mutational signature in our analysis predominantly consists of
442 C>T transitions (**Figure 2E**) that characterize mutations that accumulate with age in normal
443 individuals^{5, 6}, arguing against a strong contribution, if any, of DDR pathways to the promotion of
444 SGR in SDS bone marrow cells. Since somatic mutations accumulate in tissues outwith the
445 hematopoietic system^{4, 11}, it will be interesting to determine whether cellular clones with somatic *EIF6*
446 mutations arise in other organs in SDS, a multi-system disorder caused by a germline ribosome
447 assembly defect.

448 The hematological manifestations in SDS are highly heterogeneous in different individuals
449 who carry identical germline *SBDS* mutations and may even fluctuate within a single individual over
450 time¹⁰. However, we found no correlation between the presence and/or frequency of *EIF6* somatic
451 mutations and the hematological parameters. Longitudinal analysis will be necessary to determine
452 whether clonal expansion promoted by the acquisition of somatic *EIF6* mutations delays or abrogates

453 the emergence of hematological complications such as aplastic anemia, myelodysplastic syndrome
454 (MDS) or acute myeloid leukemia (AML). Clonal hematopoiesis and progression to poor prognosis
455 MDS or AML in SDS is associated with the acquisition of biallelic somatic *TP53* mutations^{11,1242}.
456 Single cell sequencing will be required to determine whether individual clones can carry both *EIF6*
457 and *TP53* somatic mutations or whether these variants are mutually exclusive. Further studies are also
458 warranted to examine the effects of *EIF6* and/or *TP53* mutant clones on disease outcome in SDS.

459 Recently Koh et al. reported an individual with clinical features of SDS in whom a *de novo*
460 heterozygous missense *EIF6* mutation (R61L) was identified by whole exome sequencing of
461 peripheral blood leukocytes⁴³. The authors concluded that this heterozygous *EIF6* mutation caused the
462 disease. Intriguingly, the eIF6-R61L mutation rescued the fitness defect of Sdo1-deficient yeast cells
463 (**Figure 5F**). Furthermore, the hematological abnormalities observed in the patient described by Koh
464 et al. appeared to improve over time. Given our observation that somatic *EIF6* mutations are frequent
465 in blood cells from SDS patients and can promote clonal expansion, we hypothesize that the *EIF6*
466 mutation identified by Koh et al. is an example of SGR counteracting the deleterious effect of an as
467 yet unidentified germline defect in ribosome assembly (we previously reported that pathogenic
468 mutations affecting the *EFL1*-encoding gene that cause SDS may be missed by whole exome
469 sequencing²³). We propose that the selective advantage provided by the somatic *EIF6* mutation
470 promoted expansion of the mutant HSPCs to repopulate the hematopoietic system to a VAF close to
471 50 % in peripheral blood DNA. Similar phenomena have been observed in other Mendelian
472 hematopoietic disorders¹⁴⁻¹⁶. This hypothesis will need to be tested by sequencing of the *EIF6* gene in
473 tissues other than blood.

474 By combining ultra-deep *EIF6* sequencing, cytogenetic, structural, MD simulations and
475 functional analysis, our study provides evidence that distinct genetic *EIF6* alterations can rescue the
476 germline ribosome assembly defect to promote clonal expansion in SDS HSPCs and achieve SGR
477 (**Figure 8**). We confirmed the presence of an interstitial deletion in chromosome 20 that encompasses
478 *EIF6* in hematopoietic cells from some individuals with SDS^{26,28}. However, as the interstitial

479 chromosomal deletion removed additional genes to *EIF6*, we were unable to formally conclude that
480 expansion of del(20q) clones was a specific consequence of *EIF6* haploinsufficiency. The detection in
481 hematopoietic cells from an SDS patient of a reciprocal translocation in which one of the breakpoints
482 disrupted the *EIF6* gene while the other resided within a non-coding region strongly supports the idea
483 that *EIF6* haploinsufficiency does indeed provide a selective advantage and promotes the clonal
484 expansion of SBDS-deficient cells (**Figure 8**). To our knowledge, SGR induced by a reciprocal
485 translocation has not been previously reported¹³. Lastly, our ultra-deep sequencing analysis pinpointed
486 the existence of frequent and distinct point mutations in the coding sequence of *EIF6* that promoted
487 SGR. Interestingly, we detected several mutations that recurrently affected the same conserved
488 residues. We distinguished three categories of *EIF6* point mutations: (1) nonsense and frameshift
489 mutations that led to *EIF6* haploinsufficiency; (2) missense mutations that affected highly conserved
490 amino-acids and strongly reduced eIF6 expression and/or stability; (3) missense mutation that did not
491 impair eIF6 expression but reduced its affinity for the 60S subunit (e.g. N106S, R61L, G14S) (**Figure**
492 **8**). Our MD simulations supported by *in vivo* functional analysis demonstrate that the eIF6 N106S
493 mutant provides a particularly potent selective advantage that is explained by the key structural role of
494 residue N106 in mediating polar interactions between eIF6 and the ribosomal protein uL14 on the
495 intersubunit face of the 60S subunit.

496 In conclusion, our study demonstrates that spontaneous acquired mutations affecting the *EIF6*
497 gene represent a frequent mechanism of indirect SGR of the germline defect in ribosome assembly in
498 SDS. The demonstration that the recurrent missense mutation N106S promotes SGR by reducing the
499 affinity of eIF6 for the 60S subunit provides a compelling *in vivo* rationale for the development of
500 small molecules that mimic the effects of eIF6 suppressor mutations in reducing the affinity of eIF6
501 for the 60S subunit as disease modifying therapeutics in SDS. Lastly, our results support the notion
502 that SGR might represent a universal phenomenon, more frequent than previously suspected, that
503 influences the clinical evolution of diverse Mendelian disorders that not only affect the hematopoietic
504 system. Additionally, the phenomenon of SGR may also be frequent in non-inherited disorders as

505 recently exemplified by chronic liver disease⁴⁵. The continued improvement in sequencing
506 technologies will likely permit the exploration of SGR in many other disorders in the near future.

507

508 **Materials and Methods**

509 **Study approval.** Informed and written consent was obtained from donors and patients. The study and
510 protocols comply with the 1975 Declaration of Helsinki as well as with the local legislation and
511 ethical guidelines from the Comité de Protection des Personnes de l'Ile de France II and the French
512 advisory committee on data processing in medical research.

513 **Constructs with human *EIF6*.** Coding sequence of WT or mutant human eIF6 was inserted in the
514 linearized (BglII/NotI) p3X-FLAG-Myc-CMV-26 vector (Sigma) to express FLAG-tagged eIF6
515 protein (**Suppl. Table S5**). The *EIF6* mutations were introduced by hemi-RT-PCR with specific
516 primers (**Suppl. Table S6**). The PCR products and linearized p3X-FLAG-Myc-CMV-26 vector were
517 assembled with NEBuilder® HiFi DNA assembly master mix (New England Biolabs). Nucleotide
518 numbering reflects the cDNA sequence with +1 corresponding to the A of the ATG translation
519 initiation codon in the reference sequence.

520 **Western blotting.** 2×10^6 HEK 293T were transfected with 3 μg of vectors expressing FLAG-eIF6-
521 WT or FLAG-eIF6-mutants by electroporation (Biorad) or lipofectamine 2000 (Invitrogen). 72 hrs
522 post-transfection, cells were scraped, washed in PBS and lysed for 20 min on ice in lysis buffer
523 containing 50 mmol/L Tris (pH 8.0), 2 mmol/L EDTA, 1 % Triton X100, 1 % phosphatase inhibitor
524 cocktails (Sigma) and protease inhibitor (Roche Applied Science, Indianapolis, IN) and centrifuged;
525 supernatant was harvested and protein concentration quantified using the Bradford assay. Whole-cell
526 lysates were analyzed by immunoblotting with appropriate antibodies using the Odyssey® CLx
527 Imaging System (LI-COR Biosciences) for quantification.

528 **Targeted *EIF6* sequencing by NGS (capture by hybridization approach) and genetic analysis.**

529 Genomic DNA was extracted from whole blood cells or bone marrow. Illumina compatible barcoded
530 genomic DNA libraries were constructed according to the manufacturer's sample preparation protocol
531 (Ovation Ultralow V2, Nugen Technologies). Briefly, 400 ng to 3 µg of patient genomic DNA was
532 mechanically fragmented to a median size of 200 bp using a Covaris. 100 ng of double strand
533 fragmented DNA was end-repaired and adaptors containing a Unique Dual Index barcode (IDT) were
534 ligated to the repaired ends (one pair of barcodes per patient). Ligated DNA fragments were PCR
535 amplified to obtain precapture barcoded libraries that are pooled at equimolar concentrations. The
536 capture process was performed using the SureSelect reagents (Agilent), 750 ng of the pool of
537 precapture libraries and home-made biotinylated probes (as previously described in Benyelles et al.⁴⁷
538 and Venot et al.⁴⁸). The biotinylated single stranded DNA probes were designed and prepared to cover
539 a 123 kb chromosomal region including the *EIF6* gene on chromosome 20 (chr20:35,256,992-
540 35,380,631, according to the GRCh38.p12 assembly of the human reference genome) or the *EIF6*
541 cDNA was obtained by PCR amplification with primers located in the 3' and 5' UTR (Sequence (5'-
542 >3') F: CGG GGC CTG AGG GAC GGA GG; R: ACA ACA GAG CAG GTT TTT GC). During the
543 capture process, barcoded library molecules complementary to the biotinylated beads were retained by
544 streptavidin coated magnetic beads on a magnet and PCR amplified to generate a final pool of
545 postcapture libraries covering the targeted genomic regions. Pools of these final libraries were
546 prepared and sequenced either on an Illumina HiSeq2500 or NovaSeq6000 (Paired-End sequencing
547 130+130 on HiSeq, 100+100 bases on NovaSeq, production of ~60 million of clusters per sample).
548 After demultiplexing, sequences were aligned to the reference human genome hg19 using the
549 Burrows-Wheeler Aligner (Li and Durbin, 2010). The mean depth of coverage per sample was
550 $\geq 1,000X$ to enable more accurate Copy Number Variant Analysis. Downstream processing was
551 carried out with the Genome Analysis Toolkit (GATK), SAMtools and Picard, following documented
552 best practices (<http://www.broadinstitute.org/gatk/guide/topic?name=best-practices>). Variant calls
553 were made with the GATK Unified Genotyper. Variants at very low allele frequency were called by
554 freebayes with the option -F 0,0005 (--min_alternate_fraction) (<https://arxiv.org/abs/1207.3907>). The

555 annotation process is based on the latest release of the Ensembl database. Variants were annotated,
556 analyzed and prioritized using the Polyweb/PolyDiag software interface designed by the
557 Bioinformatics platform of University Paris Descartes/Imagine Institute.

558 The sequence analysis dn/ds tool from UCSF (<https://humangenetics.ucsf.edu/sequencing-tool/>) was
559 used to calculate dN/dS.

560 **Cytogenetics and CGH array.** Agilent SurePrint G3 Cancer CGH+SNP 4x180K microarray (Agilent
561 Technologies, Santa Clara, CA) was used for genomic copy number analyses according to
562 manufacturers' recommendations. Genomic positions are relative to the human genome Build
563 NCBI37/hg19. Chromosomal preparation from bone marrow was performed using standard protocols
564 and fluorescence *in situ* hybridization (FISH) was performed using Del (20q) Deletion Probe LPH 020
565 (Cytocell Ltd, Cambridge, UK) according to manufacturers' recommendations.

566 **Determination of the translocation t(16;20)(q24;q11.2) breakpoints with chimeric reads.** To
567 accurately assess the breakpoint location of chromosome 20 / chromosome 16 translocation, we
568 extracted all the reads from chromosome 16 that contain a soft clip in the cigar and determined the
569 position of the last aligned position. We then grouped all those putative break points according to their
570 position to look for clustering. Finally we retained the candidate clusters where mates pointed to
571 chromosome 20 only, and the *EIF6* region in particular, for visual inspection with IGV. The command
572 used was: `samtools view -q 1 sample.bam chr16 | cut -f3,4,6-8 | grep S | awk '{pos=$2;
573 split($3,a,"[IMDSH]"); split($3,b,"[0-9]*"); nb=length(b); for (i=2; i<=nb; i++) if (b[i] ~ /[MD]/)
574 pos=pos+a[i-1]; printf("%s\t%s\t%s\t%s\t%s\t%s\n",$1, pos-1, pos-1, $3, $4, $5)}' | sort -k2,2n |
575 bedtools merge -d 0 -c 5,6 -o distinct,distinct | grep -E '=,chr20|chr20,=' | grep -v -E '=,chr20,|chr20,='
576 | sort -k5,5n. Study of the reads assigned positions of the breakpoint to a position between 85,849,823
577 and 85,849,825 (HG19) on chromosome 16 and to a region ranging from 33,867,599 to 33,867,604 on
578 chromosome 20. The translocation was supported by 10 reads on chromosome 16 in total. The
579 boundary was supported by 6 reads where 3 were inter-chromosomal alignment. On chromosome 20,
580 due to the read-depth greater than 1,800, the situation was less clear. However, we identified 10 inter-`

581 chromosomal alignment reads and 15 more reads supporting the breakpoint region. Similar analysis in
582 4 unrelated controls did not retrieve chimeric reads between chromosome 16 and 20.

583 **Sucrose gradient of human cell extracts.** For ribosome fractionation cytoplasmic extracts from
584 HEK293 cells were prepared as already described¹³. For each sample 1 mg of extract was layered on a
585 10–50 % sucrose gradient containing 20 mM Tris pH 7.6; 80 mM NaCl; 5 mM MgCl₂; 1 mM DTT.
586 The gradients were run in an SW41 Beckman rotor at 220,672 g for 140 min at 4 °C. Following
587 centrifugation gradients were fractionated. Acquisition of the profiles was obtained using the UA6
588 UV/VIS detector from ISCO.

589 **Statistical analyses.** Statistical analyses were performed on Prism (GraphPad Software). Groups were
590 analyzed by Student t-test as indicated and the difference was considered statistically significant for
591 p<0.05. Pearson correlation on Prism (GraphPad Software) was used for correlation determination.

592 **Dictyostelium cell cultivation and transfection.** Ax2 (DBS0235521) cells were grown in filter
593 sterilised HL5 (Formedium #HLE2) containing 200 µg/mL Dihydrostreptomycin (Sigma #D7253) in
594 tissue culture dishes or in shaken suspension at 180 revolutions per minute at 22 °C. For transfection,
595 cells were harvested from tissue culture plates and washed by centrifugation twice in ice-cold H40
596 buffer (40 mM HEPES, 1 mM MgCl₂ pH 7.0). They were resuspended at 4 x 10⁷ cells/mL and 0.1 mL
597 added to a pre-chilled electroporation cuvette (gap width 2 mm, GeneFlow #E6-0062). 1-2 µg of
598 supercoiled or restriction enzyme digested plasmid DNA was added and electroporated with two 350
599 V square wave pulses each of 8 ms duration delivered 1 s apart using a GenePulser Xcell (Bio-Rad)¹⁴.
600 Ax2 cells expressing eIF6 or vector (pDM1203) alone were selected in 10 cm tissue culture dishes
601 using 10 µg/mL G418 (Gibco Geneticin #10131-035). Clonal eIF6 knockout cell lines were selected in
602 96 well tissue culture plates (60 or 600 cells/well) in 0.15 mL of HL5 medium/well containing 10
603 µg/mL blasticidin (InvivoGen #ant-bl-1) and 10 µg/mL G418. After 7-12 days in selection, confluent
604 wells were harvested, the genomic DNA extracted (Quick-DNA™ Miniprep Kit, Zymo research
605 #D3024) and screened by PCR using oligonucleotides DTO16 and DTO18 that bind to regions of the
606 eIF6 genomic locus that are outside that of knockout cassette (**Suppl. Table S7**)¹⁵.

607 **Plasmid construction.** To make knockout vector pDT131 genomic DNA both proximal and distal to
608 the *EIF6* gene were amplified by PCR using primer pairs DTO1/DTO9 and DTO2/DTO3 that
609 introduced restriction enzyme sites for cloning (**Suppl. Table S7**) . The PCR products were digested
610 with ApaI or BamHI/SacII and cloned into pLPBLP either side of the ‘floxable’ bsR cassette¹⁶ and the
611 inserts verified by sequencing. *Dictyostelium* WT or mutant eIF6 expression plasmids were made by
612 PCR amplification of the eIF6 coding sequence (DDB0234038) from Ax2 genomic DNA with the
613 inclusion of BamHI and XbaI restriction sites. The digested PCR product was cloned into the
614 corresponding restriction sites of extrachromosomal vector pDM1203¹⁷. The eIF6 T56K, I58T and
615 N106S point mutations were introduced using PCR mediated site-directed mutagenesis. Primer pairs
616 Max15/Max16 were used for T56K, DTO28/DTO29 for I58T and DTO30/DTO31 for N106S. All
617 mutations were verified by sequencing.

618 **Cell Lysis for ribosome profiles.** Vegetative cells were treated with 100 µg/mL cycloheximide for 5
619 min prior to harvesting. Cells were pelleted by centrifugation and resuspended in buffer KK₂ (16.5
620 mM KH₂PO₄, 3.9 mM K₂HPO₄, 2 mM MgSO₄) plus 100 µg/mL cycloheximide. They were washed
621 twice more in KK₂ , with a final wash in KK₂ containing 100 µg/mL cycloheximide and 1x SigmaFast
622 EDTA-free protease inhibitor cocktail (Sigma #S8830). The cell pellet was resuspended at 2 x10⁸/mL
623 in 50 mM HEPES pH 7.5, 40 mM Mg(CH₃COO)₂, 25 mM KCl, 5 % sucrose, 0.4 % IGEPAL® CA-
624 630 (Sigma #I8896), 100 µg/mL cycloheximide, 1x SigmaFast EDTA-free protease inhibitor cocktail,
625 2 mM PMSF and lysed by passing through a 25 mm diameter Swin-Lok filter holder (GE Healthcare
626 Life Sciences #420200) containing a prefilter (Millipore #AP1002500) together with a 5 µm
627 nucleopore track-etched membrane (Whatman #110613). The lysate was cleared by centrifugation
628 (8,000 g for 30 min at 4 °C) and the supernatant passed through a 33 mm Millex-® GV 0.22 µm
629 PVDF filter unit (Millipore #SLGV033RS). The filtrate was divided into 1.4 mL aliquots after A₂₆₀
630 determination, flash frozen in liquid N₂ and stored at -80 °C. All buffers were at 4 °C.

631 **Sucrose density gradients.** Lysates were loaded onto a 10-40 % (w/v) sucrose gradient in 50 mM
632 Hepes pH 7.5, 25 mM K(CH₃COO)₂, 40mM Mg(CH₃COO)₂ in Polyallomer 14 x 95 mm centrifuge

633 tubes (Beckman). After centrifugation (Beckman SW40Ti rotor) at 260, 900 g for 3 hr at 4 °C,
634 gradients were fractionated at 4 °C using a Gilson Minipuls 3 peristaltic pump with continuous
635 monitoring (A_{254} nm) and polysome profiles recorded using a Gilson N2 data recorder. Proteins were
636 precipitated from 0.5 mL fractions using 20 % (v/v) trichloroacetic acid, separated on SDS-PAGE
637 gels and transferred to nitrocellulose membranes for immunoblotting.

638 **Subcellular fractionation.** Vegetative cells in mid-log phase were harvested, washed in KK2 buffer
639 and resuspended at 2×10^7 cells/mL. One mL of cells was pelleted by centrifugation and lysed in NLB
640 buffer (50 mM Tris-HCl pH 7.4, 5 mM Mg (CH₃COO)₂, 10 % (w/v) sucrose, 2 % (v/v) NP-40 by
641 vortexing for 1 min. Nuclei were pelleted by centrifugation at 2300 g for 5 min at 4 °C and the
642 supernatant saved as the “crude cytoplasmic” fraction. The nuclear pellet, washed once in 1 mL of
643 NLB and resuspended in 100 μ L of NLB, was designated the “nuclear fraction.”

644 **Immunoblotting.** *Dictyostelium* cells were resuspended at 2×10^7 cells/ mL in 1 x NuPAGE® sample
645 buffer (Invitrogen #NP0007) containing 5 % (v/v) 2-mercaptoethanol (Sigma #M6250) and heated at
646 95 °C for 3 min. 2×10^5 cell equivalents were loaded per well of a NuPAGE™ 4-12 % Bis-Tris gel
647 and resolved in 1 x MES SDS running buffer (Life technologies #NP0002). SeeBlue® Plus2
648 (Invitrogen #LC5925) or HiMark™ (ThermoFisher scientific #LC5699) prestained standards were
649 used to calibrate each gel. The iBlot 2 Dry Blotting System (Invitrogen™ #IB21001) was used to
650 transfer the proteins to nitrocellulose membranes (Invitrogen #IB23001). The membranes were
651 blocked for 30 min in block buffer (PBS containing 0.1 % (v/v) TWEEN®20 (Sigma #T2700) and 5
652 % (w/v) dried skimmed milk powder). The primary antibody was diluted in block buffer and
653 incubated with the blocked membrane for 2-4 hr at room temperature or overnight at 4 °C. The
654 membrane was washed for 10 min with gentle agitation in PBS-T buffer (PBS containing 0.1 % (v/v)
655 TWEEN®20) and this was repeated another 3 times with fresh PBS-T. The secondary antibody was
656 diluted in block buffer and incubated with the washed membrane for 1-2 hr at room temperature. The
657 blot was developed in 1.5 mL of Immobilon® Western chemiluminescent HRP substrate (Millipore

658 #WBKLS0500) according to the manufacturer's instructions. The membranes were visualized with the
659 ChemiDoc™ MP imaging system (BIO RAD) using Image Lab software (BIO RAD).

660 **Yeast Strains, plasmids and primers.** *S. cerevisiae* strains used in this study are listed in **Suppl.**
661 **Table S8**, primers are listed in **Suppl. Table S9**, and plasmids in **Suppl. Table S10**. To create the
662 *Sdo1^{int(ts)}* strain, the conditional TS18 intein^{7,18} was amplified by PCR from plasmid pS5DH-G4MINT
663 (gift from N. Perrimon) and inserted between the *SDO1* codons for K73 and C74 by homologous
664 recombination. For the generation of Tif6-GFP mutants, site-directed mutagenesis of the pTIF6-GFP
665 plasmid was performed using the Phusion High-Fidelity PCR kit (NEB) and transformed into XL1-
666 Blue Electroporation-Competent cells (Agilent).

667 **Growth assays.** *sdo1^{ts}* yeast cells were grown in SD –URA liquid medium at 23 °C to stationary
668 phase. 2 OD₆₀₀ of cells were harvested and re-suspended in 500 µL mQ water. 2 µL of serial tenfold
669 dilutions were spotted onto solid SD–URA medium and growth was assessed after 2 d of culture at 30
670 °C, or 3 d at 23 °C or 37 °C.

671 **Immunoblotting.** The *sdo1^{ts}* yeast cells were grown at 23 °C to an OD₆₀₀ of 0.8-1 in SD–URA liquid
672 medium. 1 OD₆₀₀ of cells were harvested, washed and re-suspended in 500 µL of mQ water. 50 µL of
673 1.85 M NaOH was added and the samples incubated on ice for 10 min. Samples were further
674 incubated on ice with 17.5 µL of 100 % (w/v) of TCA and centrifuged for 5 min at 16,000 g. The
675 pellet was washed with 500 µL of 80 % acetone (v/v) and centrifuged for 5 min at 16,000 g. The
676 supernatant was decanted and the resultant pellet air-dried. The pellet was resuspended in 1x NuPAGE
677 LDS sample buffer (Thermo Fisher Scientific) containing 50 mM DTT prior to incubation at 70 °C for
678 10 min. Samples were separated using the NuPAGE 4-12 % Bis-Tris gel (Thermo Fisher Scientific)
679 containing 1x MES buffer (Thermo Fisher Scientific). Proteins were transferred from the gel to the
680 nitrocellulose membrane using the iBlot 2 (Thermo Fisher Scientific) system. The nitrocellulose
681 membrane was blocked with 5 % (w/v) milk dissolved in PBST buffer (137 mM NaCl, 2.7 mM KCl,
682 4.3 mM Na₂HPO₄, 1.47 mM KH₂PO₄ with 0.1 % (v/v) Tween 20) for 30 min. The blot was incubated
683 with 1:1000 dilution of anti-eIF6 antibody (GenTex, #GTX117971) overnight at 4 °C followed by

684 several 5 min washes with PBST buffer. The blot was incubated with 1:5000 dilution of anti-rabbit
685 IgG HRP-linked antibody (Cell Signaling #7074) followed by several 5 min washes with PBST buffer.
686 1 mL of Luminol and 1 mL of Peroxide solution from the Western Chemiluminescent HRP Substrate
687 kit (Immobilon) was incubated with the blot for 1 min. Proteins were visualized using the Bio-Rad
688 Chemidoc MP imaging system.

689 **Genetic complementation.** These were performed as previously described⁶.

690 ***Drosophila melanogaster* strains and genetics.** Flies were maintained using standard culture
691 techniques. All crosses were performed at 25 °C unless otherwise stated. Fly strains and genotypes are
692 described in **Suppl. Table S4**. *CG8549^{f01686}*, *PBac{WH}CG8549[f01686]*, referred to here as *Sbds^P*, is
693 a homozygous lethal piggyBac transposase element insertion in the 5' untranslated region of *CG8549*.

694 **Transgenic *Drosophila* lines.** The coding sequences for WT *Drosophila Sbds* (NM_139800) and
695 *EIF6* (NM_145105) were amplified by PCR from a *Drosophila* embryo cDNA library (gift from
696 Simon Bullock) and cloned into pTWF (The *Drosophila* Gateway vector collection) to generate
697 plasmids pUAS-Sbds-FLAG and pUAS-EIF6-FLAG. *EIF6* suppressor mutations, *EIF6C56R*,
698 *EIF6Y151H* and *EIF6V192F* were generated by PCR site-directed mutagenesis and sub-cloned into
699 vector pPWM (The *Drosophila* Gateway vector collection) using the Gateway system (Invitrogen).
700 Transgenic *pUAS-Sbds-FLAG*, *pUAS-EIF6-FLAG*, *pUAS-EIF6C56R-MYC*, *pUAS-EIF6Y151H-MYC*
701 and *pUAS-EIF6V192F-MYC* flies were generated by P element-mediated germline transformation¹⁹
702 into a *w¹¹¹⁸* strain by Genetic Services Inc. To generate flies expressing human SBDS, the coding
703 sequence for human SBDS (NP_057122) was PCR amplified from a pRSETA-SBDS plasmid⁸ and
704 sub-cloned into plasmid pTWF to generate plasmid pUAS-SBDS-FLAG. Transgenic *pUAS-SBDS-*
705 *FLAG* flies were generated as described above. Primers are listed in **Suppl. Table S11**.

706 **Antibodies.** Antibodies are listed in **Suppl. Table S12**. Rabbit polyclonal antiserum was raised
707 against *Drosophila* Sbds residues 1-252 and affinity purified (Eurogentec).

708 **Protein expression and purification.** Plasmid pSbds-His (encoding *Drosophila* Sbds, amino acids 1-
709 252, fused at the C-terminus to 6 x His residues) was transformed into *E. coli* C41(DE3) cells and
710 Sbds-6xHis protein was purified by Ni-NTA affinity (GE Healthcare) and a Hiload 26/60 Superdex 75
711 column (GE Healthcare). Protein purity was assessed by SDS-PAGE and identity confirmed by mass
712 spectrometry.

713 **Immunofluorescence.** Wing discs dissected from third-instar larvae and ovaries dissected from adult
714 female flies were fixed in 4 % paraformaldehyde in PBS for 30 min at room temperature and
715 processed for immunofluorescence (IF) staining as described^{20,21}. For IF staining of mitotic cells in
716 neuroblasts, *Drosophila* brain squash slides were prepared as described²². Primary antibodies are listed
717 in **Suppl. Table S12**. Alexa 488 (green)- or 563 (red)- or 647 (far red)- conjugated secondary
718 antibodies (Invitrogen) were used at 1:1000 dilution. DNA was stained with DAPI in mounting
719 medium (Vector). Images were collected on a Zeiss LSM780 confocal system, imported to Image J
720 v10.4 (Image J) and Photoshop CS5 (Adobe), and adjusted for brightness and contrast uniformly
721 across entire fields.

722 **Immunoblotting.** *Drosophila* larval extracts were prepared by grinding ten third instar larvae in 150
723 μ L NuPAGE LDS sample buffer (Invitrogen, #NP0007) using a pellet pestle (Eppendorf). Samples
724 were cleared in a microfuge and denatured by heating at 95 °C for 10 min. Third instar larvae cells
725 were fractionated using NE-PER nuclear and cytoplasmic extraction reagents (Thermo Scientific,
726 #78833) according to the manufacturer's instructions. Cell lysates were cleared in a microfuge and
727 normalized for protein concentration using a BCA protein assay kit (Pierce, #23227). Samples were
728 separated using SDS-PAGE for immuno-blotting.

729 **Sucrose gradient sedimentation of *Drosophila* cell extracts.** Ribosomal subunits were separated by
730 sucrose density gradients as previously described⁸. Briefly, *Drosophila* third instar larvae were
731 collected (typically 40 mg), washed with PBS, homogenized in lysis buffer A (10 mM Tris-HCl at pH
732 7.4, 10 mM NaCl, 2.5 mM MgCl₂, 0.5 % (v/v) Triton X-100, 0.5 % (w/v) deoxycholate, 1 % (v/v)
733 Tween 20, 100 mg/mL cycloheximide (Sigma, #C7698) with complete EDTA-free protease inhibitors

734 (Roche) and 0.5 U/mL RNase inhibitor (Promega) and incubated for 10 min on ice. Lysates were
735 cleared in a microcentrifuge. Equal amounts (typically 3-5 A₂₅₄ U) were applied to a 10-50 % (w/v)
736 sucrose gradient in 14 mL of buffer B (10 mM Tris-HCl at pH 7.4, 75 mM KCl, 2.5 mM MgCl₂) and
737 centrifuged (Beckman SW40 rotor) at 284, 600 g for 2 hr at 4 °C). Samples were loaded on a Brandel
738 gradient fractionator, the polysome profiles were detected using a UV monitor (UV-1, Pharmacia) at
739 A₂₅₄, and 0.5 mL fractions collected. The electronic outputs of the UV-1 monitor and fraction collector
740 were fed into a Labjack U3-LV data acquisition device with an LJTick-InAmp preamplifier. Proteins
741 were precipitated from sucrose gradient fractions with 10 % (v/v) trichloroacetic acid (TCA),
742 separated on SDS-PAGE gels and transferred to PVDF membranes for immuno-blotting.

743

744 **Measurement of protein synthesis.** Protein synthesis in *Drosophila* cells was measured using an
745 adapted protocol^{2,23}. Briefly, about thirty third instar larvae were dissected in culture medium (Shields
746 and Sang M3 insect medium (Sigma, #S3652), 10 % fetal bovine serum (Sigma, #F7524) and 1 %
747 Penicillin-Streptomycin (Pen-Strep, Sigma, #P0781)) within 10 min and hemocytes collected into a
748 1.5 mL tube. An equal volume of culture medium with 100 µM of OP-Puro (Invitrogen, #C10456)
749 was added to the hemocytes and incubated at 25 °C for 30 min. Cells were collected by centrifugation
750 at 2, 500 g for 4 min, and washed twice with ice-cold phosphate buffered saline (PBS) (Invitrogen)
751 with 1 % BSA (Sigma) and 100 µg/mL cycloheximide (Sigma, #C7698). Cells were fixed and
752 permeabilized using the Cytotfix/Cytoperm Fixation Permeabilization Kit (BD Biosciences, #554714).
753 Azide-alkyne cycloaddition was performed using the Click-iT Cell Reaction Buffer Kit (Invitrogen,
754 #C10456) with azide conjugated to Alexa Fluor 488 at 5 µM final concentration. Following a 30 min
755 reaction, cells were washed twice in PBS, resuspended in PBS supplemented with 1 % fetal bovine
756 serum and analysed by flow cytometry (Becton Dickinson LSR Fortessa analyzer). Flow cytometry
757 data analysis was performed using FlowJo v10.1 (FlowJo, Ashland, OR). 'Relative protein synthesis
758 rate' was calculated by normalizing OP-Puro signals to control cells after subtracting background
759 fluorescence (cells without OP-Puro incorporation).

760

761 **Molecular dynamics simulations**

762 **System setup.** The atomic model for MD simulations was based on the cryo-EM structure of the
763 human 60S-eIF6 complex at 2.4 Å resolution (PDBID: XXXX). The protein-RNA complex
764 comprised: i) eIF6 residues M1-N225; ii) eL24 residues M1-K60; iii) uL3 residues A45-P82, P206-
765 T223 and H275-R378; iv) uL14 residues S10-A140; and v) 28S rRNA bases A4589-G4639, G4660-
766 U4677 and A4473-U4482. System setup was carried out using the CHARMM-GUI web server²⁴⁻²⁶.
767 Proteins and RNA were inserted into a cubic box (dimension 11.2 nm), allowing a minimum of 1 nm
768 distance from the box edges. Solvation using TIP3P water and sufficient potassium and chloride ions
769 to neutralize the system, to a final, physiologically representative salt concentration of 0.1 M.

770 **Simulation protocol.** All simulations were performed using GROMACS v2019.6²⁷ with the
771 CHARMM36 additive force field²⁸. Energy minimization was performed using the steepest descent
772 algorithm (<5,000 steps) to remove steric clashes, and a 4 ns equilibration phase followed with all
773 protein and RNA atoms were position-restrained with gradually reducing force constants to relax the
774 system, ranging from 400 to 40 kJ mol⁻¹nm⁻². All dihedral angles were restrained during equilibration
775 using a force constant of 4 kJ mol⁻¹ nm⁻². Production simulations were carried out in the NPT ensemble
776 for 500 ns in triplicate for all systems. During production runs, position restraints were applied to uL3
777 (backbone atoms of residues P82, P206, T223 and H275) and the 28S RNA (main chain atoms of
778 bases A4589, G4639, G4660, U4677 and A4473-U4482) to maintain the tertiary structure of uL3 and
779 prevent unfolding of the 28S rRNA. A 2 fs integration time step was used and trajectory frames were
780 written every 20 ps. All covalent bonds hydrogens were constrained using the LINCS algorithm²⁹.
781 Long-range electrostatics were treated with the Particle-Mesh-Ewald algorithm using a real space
782 cutoff of 1.2 nm³⁰. Lennard-Jones interactions were smoothly switched off between 1.0 and 1.2 nm.
783 The Nosé-Hoover thermostat was utilized to maintain the temperature at 303.15 K with a coupling
784 constant of 1 ps^{31,32}. Protein and RNA were coupled separately from the solvent. Isotropic pressure
785 coupling was applied at 1 bar using the Parrinello-Rahman barostat with a coupling constant of 5 ps
786 and compressibility of 4.5x10⁻⁵ bar⁻¹^{33,34}.

787 **Simulation Analysis.** The VMD software was used for trajectory visualization and figure
788 preparation³⁵. All analysis was performed using integrated tools within the GROMACS package²⁷. The
789 Grace plotting tool and the GNU Image Manipulation Program (GIMP) were utilized to visualize the
790 plots.

791 **Acknowledgements.** We thank the patients and their families. P.R. thanks Dr Loélia Babin (Genome
792 dynamics in the Immune system lab, Imagine Institute) for her assistance with graphical representation
793 of mutational landscape and Lolliplot. P.R. thanks Serge Romana and Marc Le Lorch (Hôpital
794 Necker-Enfants malades, Paris France) for the kind gift of the BAC (BAC CTD-2094A15) containing
795 the complete *EIF6* gene.

796 **Author contribution.** L.K., A.B., J.M., and P.R. generated constructs and performed functional
797 experiments in human models. B.B., C.B-C., J.D., J-A.M., and F.D. identified the affected patients and
798 performed related clinical studies. S.K. and I.R-W. performed cytogenetic analysis. S.F. conducted
799 sucrose gradients on human cells. C.B-F., O.A., A.P., M.P., M.Z., performed deep sequencing. P.N.,
800 C.M. and F.T. conducted bioinformatics analysis. I.C. performed structural analysis. A.J.W. conceived
801 and B.G, N.E-U, A.Z.B., A.F., C.H., M.R., D.T. and S.T. performed experiments in yeast,
802 *Dictyostelium* and *Drosophila*. V.K. performed molecular dynamics simulations with input from
803 P.J.B. P.R. conceived the genetic project and did the sequencing analysis. P.R. and A.J.W. wrote the
804 manuscript with editing contributions from J-P.V., I.C., D.T., C.H., V.K. and S.T.

805 **Funding.** This work has been supported by institutional grants from INSERM, Ligue Nationale contre
806 le Cancer (Equipe Labellisée La Ligue ‘LIGUE 2020’), and CEREDIH (Centre de Référence Déficiets
807 Immunitaires Hérités). This work was supported by State funding from the Agence Nationale de
808 la Recherche under “Investissements d’avenir” program (ANR-10-IAHU-01). This study contributes
809 to the IdEx Université de Paris ANR-18-IDEX-0001. P.R. is a scientist from Centre National de la
810 Recherche Scientifique (CNRS). A.J.W. was supported by a Specialist Programme from Blood Cancer
811 UK (12048, to AJW), the UK Medical Research Council (MR/T012412/1), a Wellcome Trust strategic
812 award to the Cambridge Institute for Medical Research (100140), a core support grant from the

813 Wellcome Trust and MRC to the Wellcome Trust-Medical Research Council Cambridge Stem Cell
814 Institute, the Connor Wright Project, the Cambridge National Institute for Health Research Biomedical
815 Research Centre and the European Cooperation in Science and Technology (COST) Action CA18233
816 “European Network for Innovative Diagnosis and treatment of Chronic Neutropenias, EuNet
817 INNOCHRON”.

818 **Conflict of interest:** The authors declare that there are no competing financial interests in relation to
819 the work described.

820

821

822

823

824

825

826

827

828

829

830

831

832

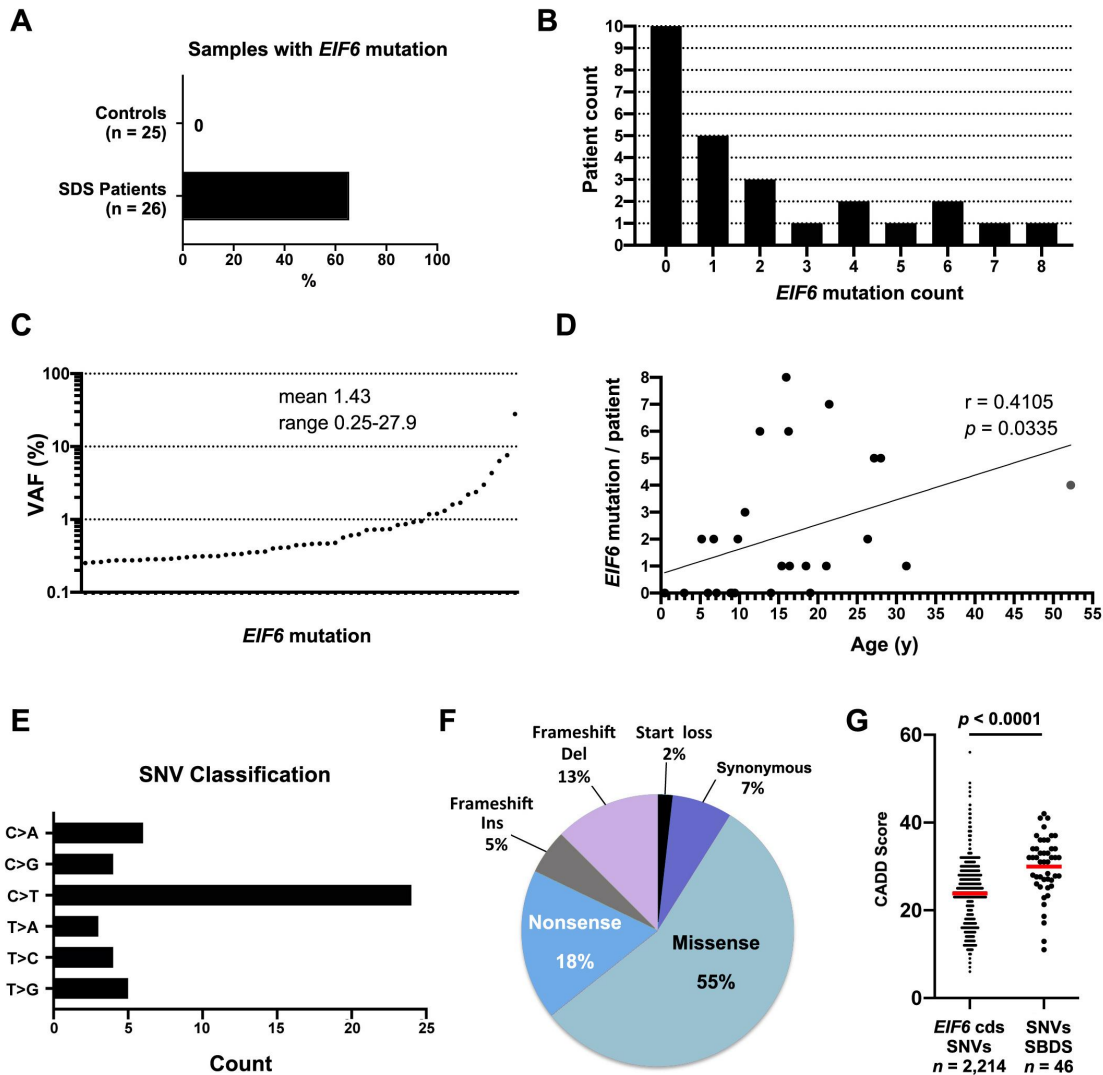
833 **References**

- 834 1. Blokzijl, F. *et al.* Tissue-specific mutation accumulation in human adult stem cells during life.
835 *Nature* **538**, 260-264 (2016).
- 836 2. Martincorena, I. & Campbell, P.J. Somatic mutation in cancer and normal cells. *Science* **349**,
837 1483-1489 (2015).
- 838 3. Garcia-Nieto, P.E., Morrison, A.J. & Fraser, H.B. The somatic mutation landscape of the
839 human body. *Genome biology* **20**, 298 (2019).
- 840 4. Martincorena, I. Somatic mutation and clonal expansions in human tissues. *Genome medicine*
841 **11**, 35 (2019).
- 842 5. Osorio, F.G. *et al.* Somatic Mutations Reveal Lineage Relationships and Age-Related
843 Mutagenesis in Human Hematopoiesis. *Cell Rep* **25**, 2308-2316 e2304 (2018).
- 844 6. Lee-Six, H. *et al.* Population dynamics of normal human blood inferred from somatic
845 mutations. *Nature* **561**, 473-478 (2018).
- 846 7. Terao, C. *et al.* Chromosomal alterations among age-related haematopoietic clones in Japan.
847 *Nature* **584**, 130-135 (2020).
- 848 8. Machiela, M.J. *et al.* Mosaic chromosome 20q deletions are more frequent in the aging
849 population. *Blood advances* **1**, 380-385 (2017).
- 850 9. Loh, P.R. *et al.* Insights into clonal haematopoiesis from 8,342 mosaic chromosomal
851 alterations. *Nature* **559**, 350-355 (2018).
- 852 10. Jaiswal, S. & Ebert, B.L. Clonal hematopoiesis in human aging and disease. *Science* **366**
853 (2019).
- 854 11. Vijg, J. & Dong, X. Pathogenic Mechanisms of Somatic Mutation and Genome Mosaicism in
855 Aging. *Cell* **182**, 12-23 (2020).
- 856 12. Weill, J.C. & Reynaud, C.A. Somatic Darwinism in vivo. *Bio Systems* **12**, 23-25 (1980).
- 857 13. Revy, P., Kannengiesser, C. & Fischer, A. Somatic genetic rescue in Mendelian
858 haematopoietic diseases. *Nat Rev Genet* **20**, 583-598 (2019).
- 859 14. McDermott, D.H. *et al.* Chromothriptic cure of WHIM syndrome. *Cell* **160**, 686-699 (2015).
- 860 15. Le Guen, T. *et al.* An in vivo genetic reversion highlights the crucial role of Myb-Like,
861 SWIRM, and MPN domains 1 (MYSM1) in human hematopoiesis and lymphocyte
862 differentiation. *J Allergy Clin Immunol* **136**, 1619-1626 (2015).
- 863 16. Catto, L.F.B. *et al.* Somatic genetic rescue in hematopoietic cells in GATA2 deficiency. *Blood*
864 **136**, 1002-1005 (2020).
- 865 17. Maryoung, L. *et al.* Somatic mutations in telomerase promoter counterbalance germline loss-
866 of-function mutations. *J Clin Invest* **127**, 982-986 (2017).
- 867 18. Gutierrez-Rodrigues, F. *et al.* Pathogenic TERT promoter variants in telomere diseases. *Genet*
868 *Med* **21**, 1594-1602 (2018).

- 869 19. Benyelles, M. *et al.* NHP2 deficiency impairs rRNA biogenesis and causes pulmonary fibrosis
870 and Hoyeraal-Hreidarsson syndrome. *Hum Mol Genet* **29**, 907-922 (2020).
- 871 20. Warren, A.J. Molecular basis of the human ribosomopathy Shwachman-Diamond syndrome.
872 *Adv Biol Regul* **S2212-4926**, 30153-30157 (2017).
- 873 21. Boocock, G.R. *et al.* Mutations in SBDS are associated with Shwachman-Diamond syndrome.
874 *Nat Genet* **33**, 97-101 (2003).
- 875 22. Stepensky, P. *et al.* Mutations in EFL1, an SBDS partner, are associated with infantile
876 pancytopenia, exocrine pancreatic insufficiency and skeletal anomalies in a Shwachman-
877 Diamond like syndrome. *J Med Genet* **54**, 558-566 (2017).
- 878 23. Tan, S. *et al.* EFL1 mutations impair eIF6 release to cause Shwachman-Diamond syndrome.
879 *Blood* **134**, 270-290(2019).
- 880 24. Finch, A.J. *et al.* Uncoupling of GTP hydrolysis from eIF6 release on the ribosome causes
881 Shwachman-Diamond syndrome. *Genes Dev* **25**, 917-929 (2011).
- 882 25. Menne, T.F. *et al.* The Shwachman-Bodian-Diamond syndrome protein mediates translational
883 activation of ribosomes in yeast. *Nat Genet* **39**, 486-495 (2007).
- 884 26. Pressato, B. *et al.* Deletion of chromosome 20 in bone marrow of patients with Shwachman-
885 Diamond syndrome, loss of the EIF6 gene and benign prognosis. *British journal of*
886 *haematology* **157**, 503-505 (2012).
- 887 27. Valli, R. *et al.* Different loss of material in recurrent chromosome 20 interstitial deletions in
888 Shwachman-Diamond syndrome and in myeloid neoplasms. *Mol Cytogenet* **6**, 56 (2013).
- 889 28. Valli, R. *et al.* Shwachman-Diamond syndrome with clonal interstitial deletion of the long
890 arm of chromosome 20 in bone marrow: haematological features, prognosis and genomic
891 instability. *British journal of haematology* **184**, 974-981 (2019).
- 892 29. Dror, Y. Shwachman-Diamond syndrome. *Pediatric blood & cancer* **45**, 892-901 (2005).
- 893 30. Bellanne-Chantelot, C. *et al.* Mutations in SRP54 gene cause severe congenital neutropenia as
894 well as Shwachman-Diamond-like syndrome. *Blood* **132**, 1318-1331 (2018).
- 895 31. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human
896 genetic variants. *Nat Genet* **46**, 310-315 (2014).
- 897 32. Martincorena, I. *et al.* Tumor evolution. High burden and pervasive positive selection of
898 somatic mutations in normal human skin. *Science* **348**, 880-886 (2015).
- 899 33. Welch, J.S. *et al.* The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150**,
900 264-278 (2012).
- 901 34. Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell*
902 **171**, 1029-1041 e1021 (2017).
- 903 35. Bassler, J. *et al.* Interaction network of the ribosome assembly machinery from a eukaryotic
904 thermophile. *Protein science : a publication of the Protein Society* **26**, 327-342 (2017).
- 905 36. Gartmann, M. *et al.* Mechanism of eIF6-mediated inhibition of ribosomal subunit joining. *The*
906 *Journal of biological chemistry* **285**, 14848-14851 (2010).

- 907 37. Groft, C.M., Beckmann, R., Sali, A. & Burley, S.K. Crystal structures of ribosome anti-
908 association factor IF6. *Nature structural biology* **7**, 1156-1164 (2000).
- 909 38. Morini, J. et al. Radiosensitivity in lymphoblastoid cell lines derived from Shwachman-
910 Diamond syndrome patients. *Radiation protection dosimetry* **166**, 95-100 (2015).
- 911 39. Zambetti, N.A. et al. Mesenchymal Inflammation Drives Genotoxic Stress in Hematopoietic
912 Stem Cells and Predicts Disease Evolution in Human Pre-leukemia. *Cell stem cell* **19**, 613-627
913 (2016).
- 914 40. Donadieu, J., Beaupain, B., Fenneteau, O. & Bellanne-Chantelot, C. Congenital neutropenia in
915 the era of genomics: classification, diagnosis, and natural history. *British journal of*
916 *haematology* **179**, 557-574 (2017).
- 917 41. Brina, D., Miluzio, A., Ricciardi, S. & Biffo, S. eIF6 anti-association activity is required for
918 ribosome biogenesis, translational control and tumor progression. *Biochimica et biophysica*
919 *acta* **1849**, 830-835 (2015).
- 920 42. Xia, J. et al. Somatic mutations and clonal hematopoiesis in congenital neutropenia. *Blood*
921 **131**, 408-416 (2018).
- 922 43. Koh, A.L. et al. Heterozygous missense variant in EIF6 gene: A novel form of Shwachman-
923 Diamond syndrome? *American journal of medical genetics. Part A* (2020).
- 924 44. Miluzio, A. et al. Impairment of cytoplasmic eIF6 activity restricts lymphomagenesis and
925 tumor progression without affecting normal growth. *Cancer cell* **19**, 765-775 (2011).
- 926 45. Zhu, M. et al. Somatic Mutations Increase Hepatic Clonal Fitness and Regeneration in
927 Chronic Liver Disease. *Cell* **177**, 608-621 e612 (2019).
- 928 46. Ziemiecki, A., Muller, R.G., Fu, X.C., Hynes, N.E. & Kozma, S. Oncogenic activation of the
929 human trk proto-oncogene by recombination with the ribosomal large subunit protein L7a.
930 *The EMBO journal* **9**, 191-196 (1990).
- 931 47. Benyelles, M. et al. Impaired telomere integrity and rRNA biogenesis in PARN-deficient
932 patients and knock-out models. *EMBO molecular medicine* **11**, e10201 (2019).
- 933 48. Venot, Q. et al. Targeted therapy in patients with PIK3CA-related overgrowth syndrome.
934 *Nature* **558**, 540-546 (2018).
- 935
- 936
- 937
- 938
- 939
- 940

957 **FIGURE 2**



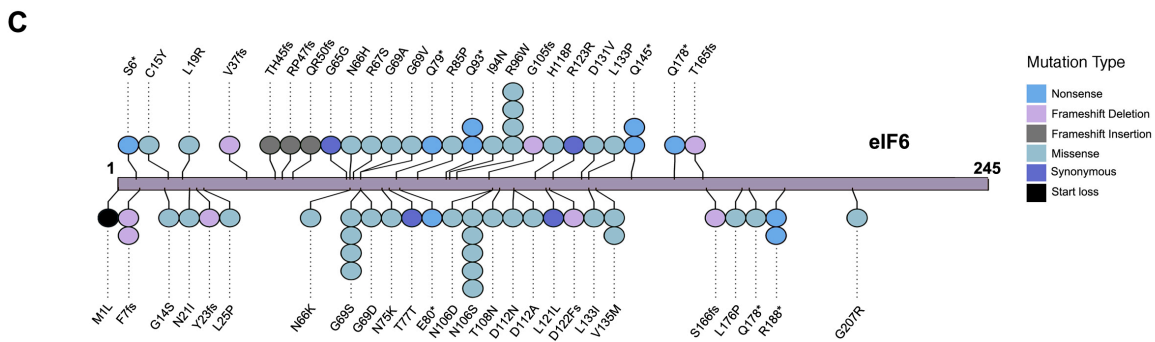
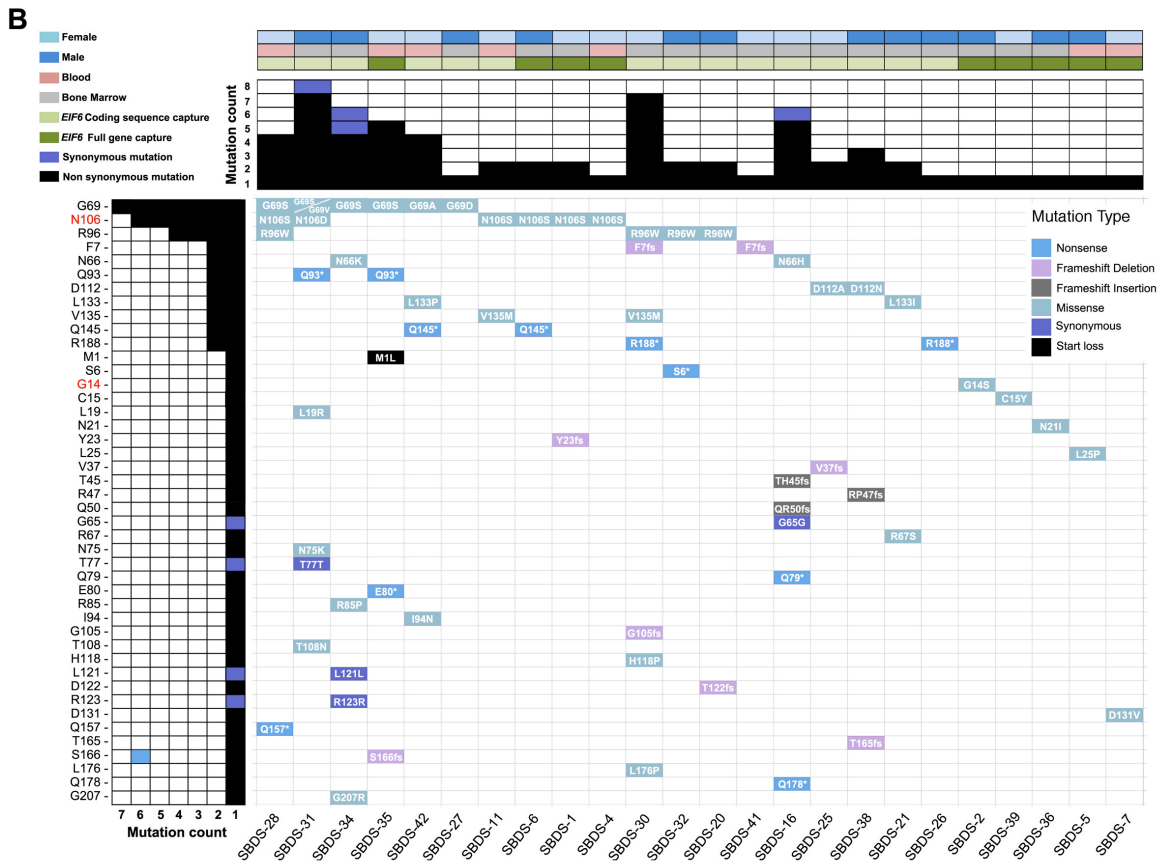
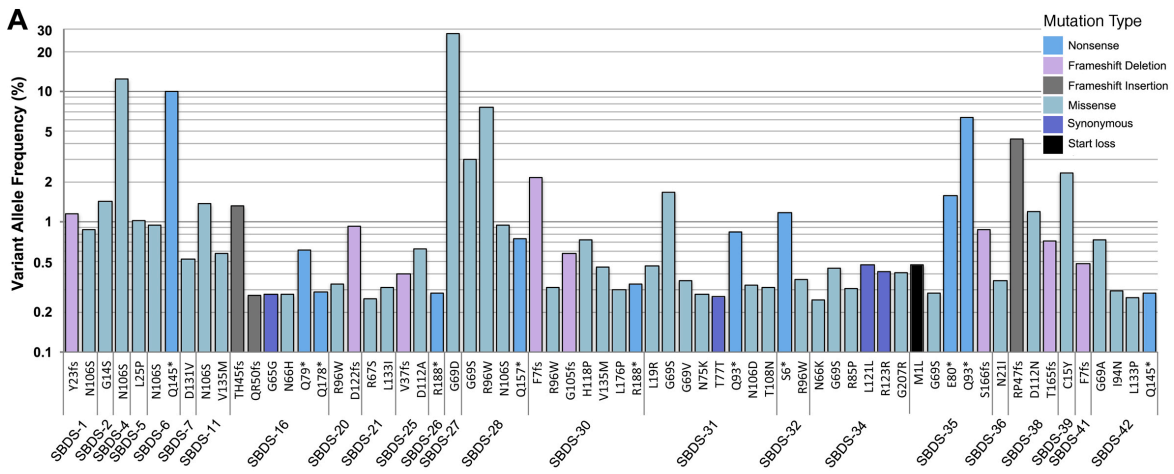
958

959 **Figure 2. Somatic *EIF6* mutations identified in SDS.** (A) Percentage of SDS patients carrying
 960 somatic *EIF6* mutations. (B) *EIF6* mutation count across the 26 SDS patients. (C) VAF distribution of
 961 the 56 identified *EIF6* mutations detected by ultra-deep sequencing. (D) Mutation count in each
 962 individual versus age. (E) Mutational spectrum of the 46 SNVs identified in *EIF6*. P-value and
 963 Pearson correlation are indicated. (F) Classification of the 56 mutations identified in *EIF6*. (G) CADD
 964 scores of all the possible SNVs (n = 2,214; **Suppl. Table S2**) in *EIF6* coding sequences versus the
 965 CADD scores of the 46 SNVs identified in the SDS patients. P-value of unpaired t test is indicated.

966

967

968 **FIGURE 3**



970 **Figure 3. Spectrum of somatic *EIF6* mutations in SDS hematopoietic cells.** (A) Spectrum of 66
971 mutations and their corresponding VAFs identified by ultra-deep sequencing in 24 SDS patients. (B)
972 Waterfall plot of the 66 mutations highlighting the recurrently impacted residues. N106S and G14S
973 (highlighted in red on the left) represent gain-of-function mutations identified in Sdo1-deleted yeast
974 cells²⁵. Gender of patients, origin of DNA, and method of *EIF6* capture for deep-sequencing are
975 indicated. Purple cases represent synonymous mutations. Colors denote type of mutation as listed in
976 the inset (upper right corner). (C) Lollipop showing the distribution of mutations in eIF6.

977

978

979

980

981

982

983

984

985

986

987

988

989

990

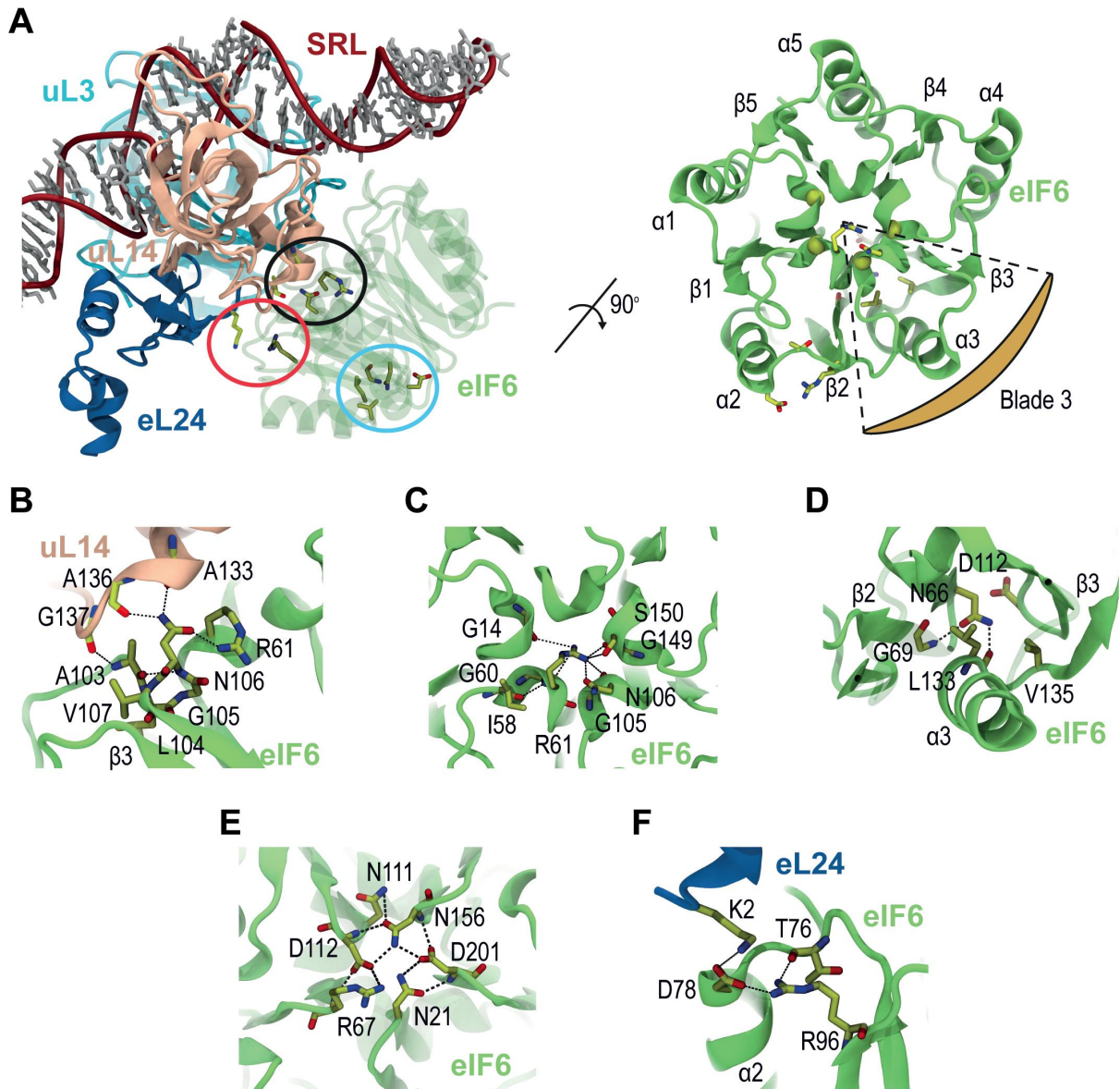
991

992

993

994

995 **FIGURE 4**



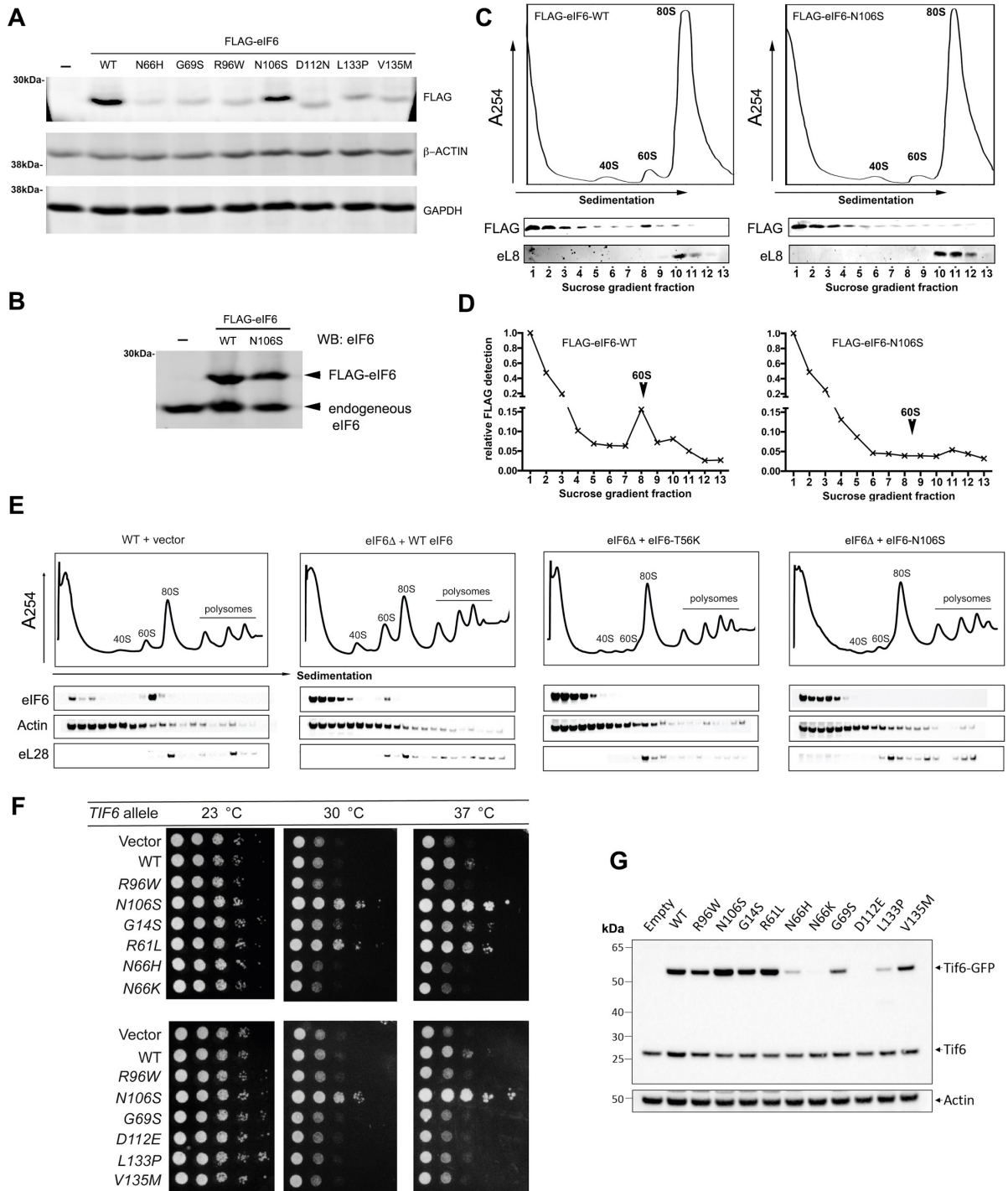
996

997 **Figure 4. SDS-related eIF6 mutations map to three hotspots.** (A) Atomic model (two orthogonal
 998 views) of the interface between human eIF6 and the 60S ribosomal subunit. The eIF6 residues mutated
 999 in SDS cluster in three independent hotspots highlighted in black (interface with uL14), cyan
 1000 (interface between blades 2 and 3) and red (eL24 interface) ellipses. (B-F) Stabilizing interactions
 1001 formed by SDS-related eIF6 residues N106 (B), R61 (C), N66, G69, L133, V135 (D), D112 (E), and
 1002 R96 (F). eL24 is blue; uL14, salmon; eIF6, green. SRL, sarcin-ricin loop.

1003

1004

1005 **FIGURE 5**



1006
1007

1008 **Figure 5. Functional consequences of SDS-related eIF6 mutations.** (A, B) The eIF6-N106S
1009 mutation does not alter eIF6 protein stability in human cells. Cell extracts from HEK293T cells were
1010 immunoblotted to detect the indicated FLAG-eIF6 variants compared with (A) GAPDH, β -ACTIN or
1011 (B) endogenous eIF6. Representative of three independent experiments. (C) The N106S mutation

1012 reduces eIF6 affinity for the 60S subunit in human cells. Cell extracts from HEK293T cells transfected
1013 with FLAG-eIF6-WT or FLAG-eIF6-N106 were fractionated by sucrose gradient sedimentation and
1014 immunoblotted to visualize eIF6 or eL8. Representative of two independent experiments. **(D)**
1015 Quantification of FLAG-eIF6 expression in the experiments depicted in **(C)**. **(E)** The eIF6-N106S
1016 mutant has lower affinity for the 60S subunit in *Dictyostelium* cells. Extracts from eIF6-deleted
1017 (*EIF6Δ*) *D. discoideum* Ax2 cells transformed with plasmids expressing eIF6-N106S or eIF6-T56K
1018 variants versus WT cells transformed with vector alone were fractionated by sucrose gradient
1019 sedimentation and immunoblotted to visualize the indicated proteins. **(F)** SDS-related Tif6 missense
1020 variants rescue the fitness defect of Sdo1-deficient cells. Tenfold serial dilutions (from left to right) of
1021 conditional Sdo1-deficient (*sdo1^{ts}*) cells complemented with plasmids expressing empty vector
1022 (pRS316), WT Tif6 or the indicated Tif6 variants were spotted onto SD-URA medium at the
1023 permissive (23 °C, 3 days) or restrictive (30 °C, 2 days; 37 °C, 3 days) temperatures. **(G)** SDS-related
1024 Tif6 missense mutations that map to the uL14-binding interface do not alter protein stability. Cell
1025 extracts from *sdo1^{ts}* cells expressing empty vector, WT or mutant Tif6-GFP were immunoblotted to
1026 detect Tif6 or actin loading control.

1027

1028

1029

1030

1031

1032

1033

1034

1035

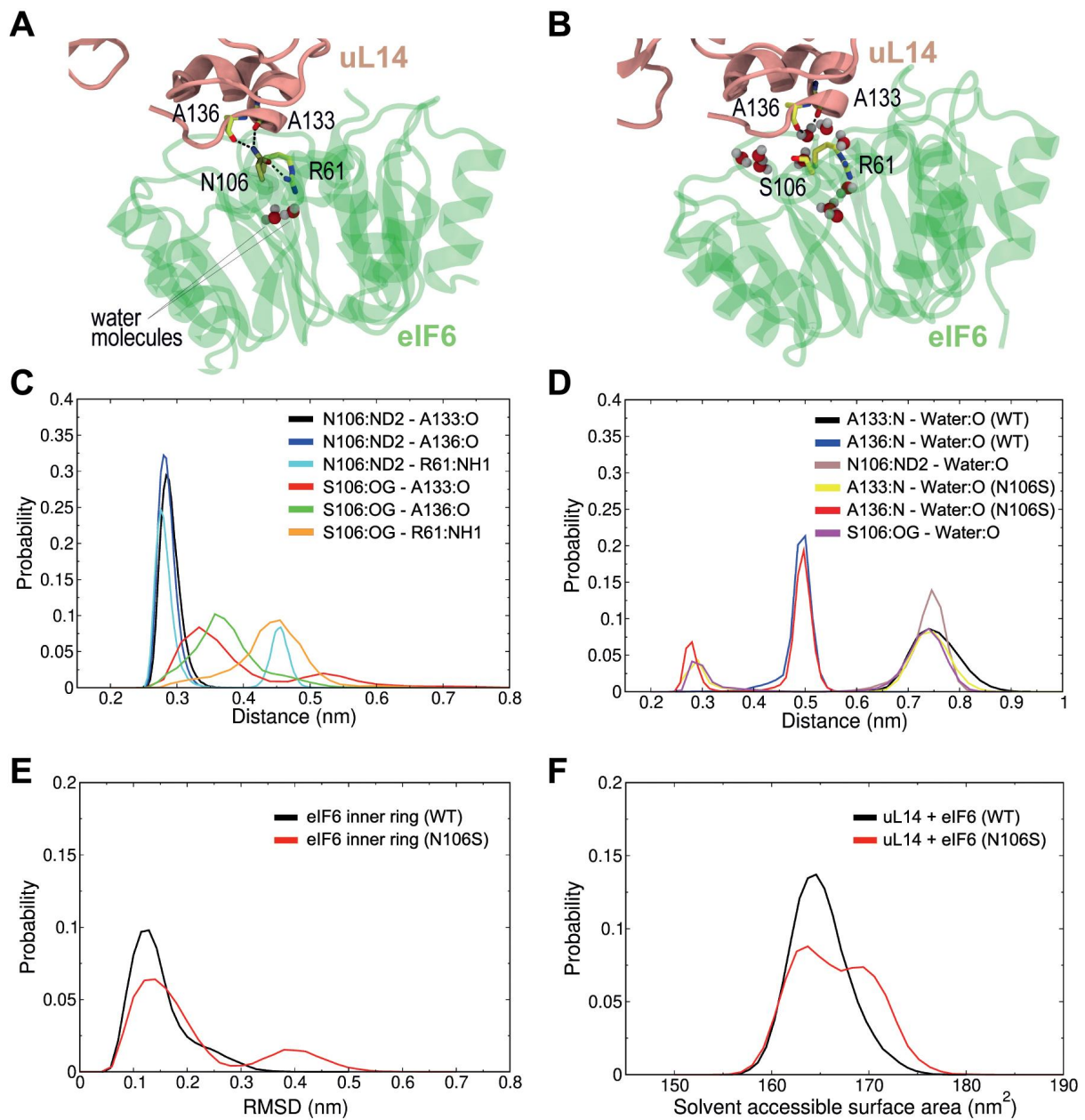
1036

1037

1038

1039

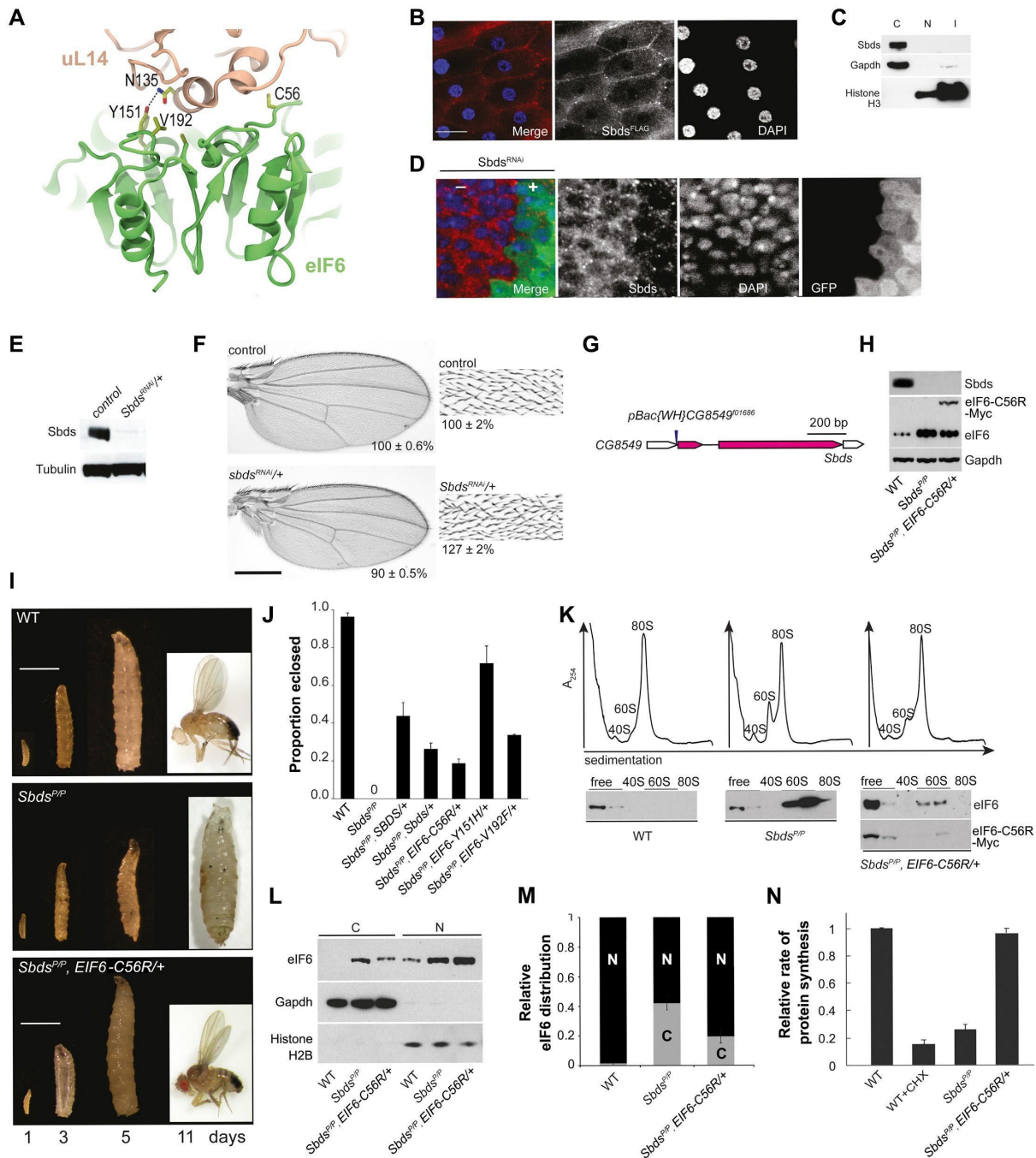
1040 **FIGURE 6**



1041

1042 **Figure 6. N106S mutation disrupts the H-bonding capacity of the eIF6-uL14 interaction**
 1043 **interface.** (A, B) Representative snapshots of the interaction interface between eIF6 N106 WT or
 1044 S106 mutant (green) and uL14 (salmon) after 500 ns of simulation. Key water molecules are indicated
 1045 in CPK format. (C, D) Distances (nm) between the indicated atoms of eIF6 WT and mutant (residues
 1046 N106, S106 and R61), and either uL14 (residues A133, A136) (C) or water (D). (E) Root mean square
 1047 deviation (RMSD) of the distance (nm) between the WT or mutant eIF6 inner ring and uL14. (F)
 1048 Solvent accessible surface area of the WT or mutant eIF6-uL14 complex. Curves in each plot include
 1049 data from all 3 replicas per system.

1050 **FIGURE 7**



1051
1052

1053 **Figure 7.** eIF6 missense mutations fully rescue the larval lethality of Sbd-deficient *D. melanogaster*.
 1054 (A) eIF6 residues C56, V192 and Y151 lie at the interface with uL16. eIF6 is colored green; uL14,
 1055 salmon. (PDBID: XXX) (B-D) Cytoplasmic localization of *Drosophila* Sbd revealed by (B)
 1056 immunostaining of FLAG-tagged Sbd (red) in ovarian follicle cells, nucleus shown in blue (DAPI);
 1057 (C) immunoblotting of subcellular fractions from *Drosophila* third instar larvae cells labeled (C)

1058 cytoplasmic, (N) soluble nuclear and (I) insoluble nuclear fraction containing nucleoli and chromatin;
1059 (D) indirect immunofluorescence of third instar larval wing disc cells. Sbds (red) was depleted by
1060 RNAi in posterior wing disc cells (marked with GFP); nucleus is blue, (DAPI). (E) RNAi depletion of
1061 Sbds in third instar larval extracts revealed by immunoblotting. (F) *Sbds* is required for cellular
1062 growth. RNAi depletion of *Sbds* in developing wings versus control. Wing size (n = 15, p < 0.0001,
1063 left) and bristle density (n = 10, p < 0.0001, right) expressed as a percentage (\pm SE) of control. Scale
1064 bar, 200 μ m. (G) Genomic organisation of the *Drosophila Sbds* (CG8549) locus with PiggyBac-
1065 element insertion site indicated in blue. White boxes, untranslated regions of *Sbds* mRNA; magenta,
1066 *Sbds* coding region; grey line, intron. (H) Sbds depletion in homozygous *Sbds*^{P/P} and *Sbds*^{P/P}, *EIF6*-
1067 *C56R*/+ mutants revealed by immunoblotting. (I) Three independent eIF6 missense mutations rescue
1068 the larval lethality of *Sbds*-deficient flies. Indicated genotypes shown at indicated time-points after
1069 fertilization. Scale bar, 1 mm. (J) Genetic complementation of *Sbds*-deficient flies. Proportion of
1070 eclosed flies with the indicated genotypes is quantified. (K) *EIF6-C56R* rescues the ribosome
1071 assembly defect in *Sbds*-deficient flies. Polysome profiles from the indicated fly genotypes are shown.
1072 The indicated proteins are visualised by immunoblotting. (L) *EIF6-C56R* rescues the cytoplasmic
1073 redistribution of eIF6 in *Sbds*-deficient flies. Subcellular fractions of third instar larvae cells with the
1074 indicated genotypes were immunoblotted to visualize the indicated proteins. (M) Subcellular
1075 distribution of endogenous eIF6 in the indicated genotypes quantified by densitometry of (L). Error
1076 bars, mean \pm SE, n=6, from three replicates. (N) *EIF6-C56R* rescues the protein synthesis defect in
1077 *Sbds*-deficient flies. Relative rates of *ex-vivo* OP-Puro incorporation in the indicated genotypes are
1078 shown. *Drosophila* genotypes are listed in **Suppl. Table S4**.

1079

1080

1081

1082

1083

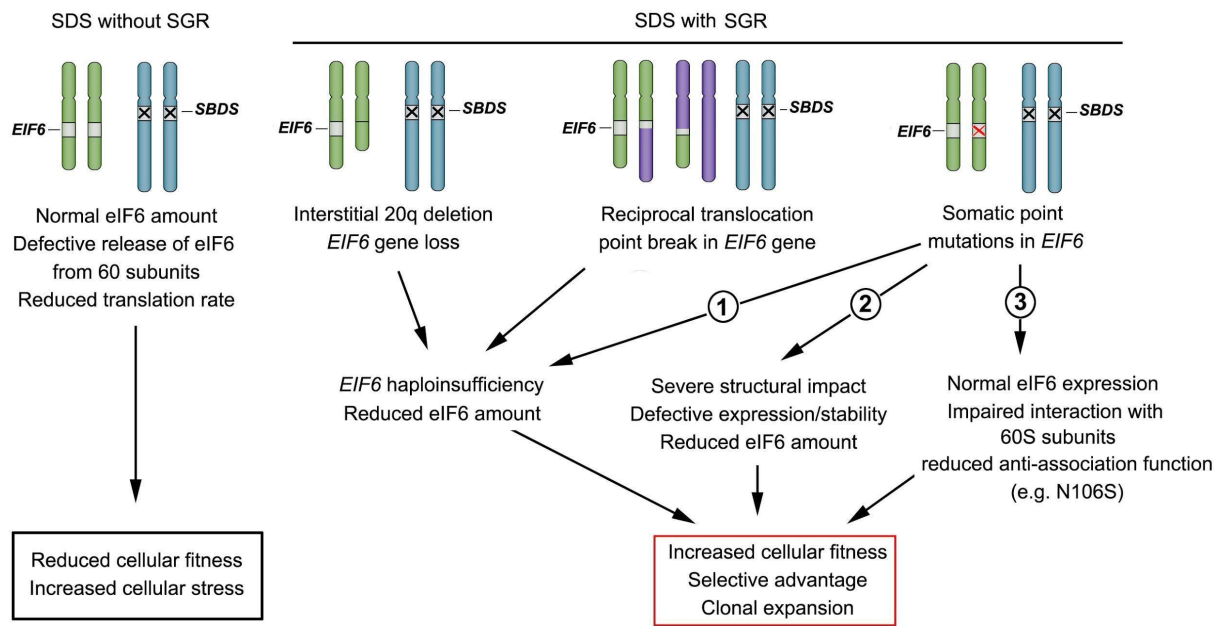
1084

1085

1086

1087

1088 **FIGURE 8**



1089

1090 **Figure 8.** Schematic representation of *EIF6* somatic genetic rescue mechanisms in SDS.

1091

1092

1093