

1 **Integrative approach to interpret DYRK1A variants, leading to a frequent**  
2 **neurodevelopmental disorder**

3 Jeremie Courraud<sup>1,2,3,4</sup>, Eric Chater-Diehl<sup>5</sup>, Benjamin Durand<sup>1,2,3,4</sup>, Marie Vincent<sup>6</sup>, Maria del  
4 Mar Muniz Moreno<sup>1,2,3,4</sup>, Imène Boujelbene<sup>1,2,3,4</sup>, Nathalie Drouot<sup>1,2,3,4</sup>, Loréline  
5 Genschik<sup>1,2,3,4</sup>, Elise Schaefer<sup>7</sup>, Mathilde Nizon<sup>6</sup>, Bénédicte Gerard<sup>8</sup>, Marc Abramowicz<sup>9</sup>,  
6 Benjamin Cogné<sup>6</sup>, Lucas Bronicki<sup>10</sup>, Lydie Burglen<sup>11</sup>, Magalie Barth<sup>12</sup>, Perrine Charles<sup>13</sup>,  
7 Estelle Colin<sup>12</sup>, Christine Coubes<sup>14</sup>, Albert David<sup>6</sup>, Bruno Delobel<sup>15</sup>, Florence Demurger<sup>16</sup>,  
8 Sandrine Passemard<sup>17</sup>, Anne-Sophie Denomme<sup>18</sup>, Laurence Faivre<sup>18</sup>, Claire Feger<sup>8</sup>, Mélanie  
9 Fradin<sup>20</sup>, Christine Francannet<sup>21</sup>, David Genevieve<sup>14</sup>, Alice Goldenberg<sup>22</sup>, Anne-Marie  
10 Guerrot<sup>22</sup>, Bertrand Isidor<sup>6</sup>, Katrine M. Johannesen<sup>24,25</sup>, Boris Keren<sup>13</sup>, Maria Kibæk<sup>23</sup>, Paul  
11 Kuentz<sup>18</sup>, Michele Mathieu-Dramard<sup>26</sup>, Bénédicte Demeer<sup>26</sup>, Julia Metreau<sup>27</sup>, Rikke  
12 Steensbjerre Møller<sup>24,25</sup>, Sébastien Moutton<sup>18</sup>, Laurent Pasquier<sup>20</sup>, Kristina Pilekær  
13 Sørensen<sup>23</sup>, Laurence Perrin<sup>28</sup>, Mathilde Renaud<sup>29</sup>, Pascale Saugier<sup>22</sup>, Joane Svane<sup>23</sup>, Julien  
14 Thevenon<sup>30</sup>, Frederic Tran Mau Them<sup>18</sup>, Cathrine Elisabeth Tronhjem<sup>23</sup>, Antonio Vitobello<sup>18</sup>,  
15 Valerie Layet<sup>31</sup>, Marie-Christine Birling<sup>32</sup>, Severine Drunat<sup>33</sup>, Allan Bayat<sup>23</sup>, Christèle  
16 Dubourg<sup>34</sup>, Salima El Chehadeh<sup>7</sup>, Christina Fagerberg<sup>23</sup>, Cyril Mignot<sup>12</sup>, Michel Guipponi<sup>9</sup>,  
17 Thierry Bienvenu<sup>35</sup>, Yann Herault<sup>1,2,3,4</sup>, Julie Thompson<sup>36</sup>, Marjolaine Willems<sup>14</sup>, Jean-Louis  
18 Mandel<sup>1,2,3,4</sup>, Rosanna Weksberg<sup>5</sup>, \*Amélie Piton<sup>1,2,3,4,8,37</sup>

19 *AFFILIATIONS*

20 <sup>1</sup> Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch 67400, France

21 <sup>2</sup> Centre National de la Recherche Scientifique, UMR7104, Illkirch 67400, France

22 <sup>3</sup> Institut National de la Santé et de la Recherche Médicale, U964, Illkirch 67400, France

23 <sup>4</sup> Université de Strasbourg, Illkirch 67400, France

24 <sup>5</sup> Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON M5G 1X8,  
25 Canada

26 <sup>6</sup> Service de Génétique Médicale, CHU de Nantes & Inserm, CNRS, Université de Nantes,  
27 l'institut du thorax, 44000 Nantes, France

28 <sup>7</sup> Service de Génétique Médicale, IGMA, Hôpitaux Universitaires de Strasbourg, Strasbourg,  
29 France

30 <sup>8</sup> Unité de Génétique Moléculaire, IGMA, Hôpitaux Universitaire de Strasbourg, Strasbourg,  
31 France

32 <sup>9</sup> Service of Genetic Medicine, University Hospitals of Geneva, Geneva, Switzerland

33 <sup>10</sup> Department of Genetics, CHEO, Ottawa, ON, Canada.

34 <sup>11</sup> Centre de référence des malformations et maladies congénitales du cervelet et Département  
35 de génétique et embryologie médicale, APHP, Sorbonne Université, Hôpital Armand  
36 Trousseau, 75012 Paris, France

37 <sup>12</sup> Pediatrics & Biochemistry and Genetics, Department, Angers Hospital, Angers, France.

38 <sup>13</sup> Genetic Department, University Hospital Pitié-Salpêtrière, AP-HP, Paris, France

39 <sup>14</sup> Département de Génétique Médicale maladies rares et médecine personnalisée, Centre de  
40 Référence Maladies Rares Anomalies du Développement, Hôpital Arnaud de Villeneuve,  
41 Université Montpellier, France

42 <sup>15</sup> Centre de Génétique Chromosomique, GHICL, Hôpital Saint Vincent de Paul, Lille, France

43 <sup>16</sup> Service de Génétique, CH Bretagne Atlantique- Vannes

44 <sup>17</sup> Département de Génétique, Hôpital Universitaire Robert Debré, APHP, Paris, France.

45 <sup>18</sup> Centre de Génétique et Centre de Référence Anomalies du développement et Syndromes  
46 malformatifs, Hôpital d'Enfants and INSERM UMR1231 GAD, FHU TRANSLAD, CHU de  
47 Dijon, Dijon, France

48 <sup>19</sup> Unité Fonctionnelle d'Innovation en Diagnostique Génomique des Maladies Rares, Pôle de  
49 Biologie, FHU-TRANSLAD, CHU Dijon Bourgogne, F-21000, Dijon, France

50 <sup>20</sup> Centre de Référence Maladies Rares, Unité Fonctionnelle de Génétique Médicale, CHU,  
51 Rennes, France

52 <sup>21</sup> Service de Génétique médicale, CHU de Clermont-Ferrand, Clermont-Ferrand, France

53 <sup>22</sup> Normandie Univ, UNIROUEN, Inserm U1245 and Rouen University Hospital, Department  
54 of Genetics and Reference Center for Developmental Disorders , F 76000, Normandy Center  
55 for Genomic and Personalized Medicine, Rouen, France

56 <sup>23</sup> Department of Clinical Genetics, Odense Denmark Hospital, Odense University Hospital,  
57 Odense, Denmark

58 <sup>24</sup> Department of Epilepsy Genetics and Personalized Treatment, The Danish Epilepsy Centre,  
59 Dianalund, Denmark

60 <sup>25</sup> Institute for Regional Health Services, University of Southern Denmark, Odense Denmark

61 <sup>26</sup> Service de Génétique Clinique, Centre de référence maladies rares, CHU d'Amiens-site Sud,  
62 Amiens, France

63 <sup>27</sup> APHP, Service de neurologie pédiatrique, Hôpital Universitaire Bicetre, Le Kremlin-  
64 Bicetre, France

65 <sup>28</sup> Department of Genetics, Robert Debré Hospital, AP-HP, Paris, France

66 <sup>29</sup> Service de Génétique Clinique et de Neurologie, Hôpital Brabois Enfants, Nancy, France

67 <sup>30</sup> Department of Genetics and Reproduction, Centre Hospitalo-Universitaire Grenoble-Alpes,  
68 Grenoble, France

69 <sup>31</sup> Consultations de génétique, Groupe Hospitalier du Havre, Le Havre, France

70 <sup>32</sup> ICS, Mouse Clinical Institute, Illkirch-Graffenstaden, France

71 <sup>33</sup> Département de Génétique, Hôpital Universitaire Robert Debré, Paris

72 <sup>34</sup> Laboratoire de Génétique Moléculaire, CHU Pontchaillou, UMR 6290 CNRS, IGDR,  
73 Faculté de Médecine, Université de Rennes 1, Rennes, France

74 <sup>35</sup> Molecular Genetics Laboratory, Cochin Hospital, APHP.Centre-Université de Paris, and  
75 INSERM UMR 1266, Institut de Psychiatrie et de Neurosciences de Paris, 75014 Paris,  
76 France

77 <sup>36</sup> Complex Systems and Translational Bioinformatics (CSTB), ICube laboratory - CNRS,  
78 Fédération de Médecine Translationnelle de Strasbourg (FMTS), University of Strasbourg,  
79 Strasbourg, France

80 <sup>37</sup> Institut Universitaire de France

81

82

83 The authors declare no conflict of interest.

84

85 \*Address for correspondence and material request:

86 Amélie Piton, PhD

87 Laboratoire "Mécanismes génétiques des maladies neurodéveloppementales", IGBMC,  
88 Illkirch, France

89 Tel : +33369551652

90 E-mail: [piton@igbmc.fr](mailto:piton@igbmc.fr)

91

92 **Competing interests**

93 None.

94

95

96

97

98 **AbBSTRACT**

99 *DYRK1A*-related intellectual disability (ID) is among the most frequent monogenic form of  
100 ID. We refined the description of this disorder by reporting clinical and molecular data of  
101 forty individuals with ID harboring *DYRK1A* variants. We developed a combination of tools  
102 to interpret missense variants, which remains a major challenge in human genetics: i) a  
103 specific *DYRK1A* clinical score, ii) amino acid conservation data generated from one hundred  
104 of *DYRK1A* sequences across different taxa, iii) *in vitro* overexpression assays to study level,  
105 cellular localization, and kinase activity of *DYRK1A* mutant proteins, and iv) a specific blood  
106 DNA methylation signature. This integrative approach was successful to reclassify several  
107 variants as pathogenic. However, we questioned the involvement of some others, such as  
108 p.Thr588Asn, yet reported as pathogenic, and showed it does not cause obvious phenotype in  
109 mice, emphasizing the need to take care when interpreting variants, even those occurring *de*  
110 *novo*.

111  
112 **Keywords:** *DYRK1A*, clinical score, missense variants, variant of unknown significance  
113 (VUS), functional assays; epismature

114

115

116

117

118

119

120

121

122

123

124

125

## 126 INTRODUCTION

127 Intellectual disability (ID) and autism spectrum disorder (ASD) are two highly  
128 heterogeneous groups of neurodevelopmental disorders (NDD) with substantial genetic  
129 contributions which overlap strongly both at the clinical and genetic levels. Single genetic  
130 events account for about 50% of ID cases (Vissers *et al*, 2016) and for a much smaller  
131 proportion of cases with ASD without ID. More than one thousand genes have been  
132 implicated in monogenic forms of NDD, with an important contribution of autosomal  
133 dominant forms caused by *de novo* mutations (Deciphering Developmental Disorders Study,  
134 2017). One of these genes, *DYRK1A* (*dual specificity tyrosine phosphorylation regulated*  
135 *kinase 1A*)(Gonzalez-Mantilla *et al*, 2016), located on chromosome 21, is among the genes  
136 the most frequently mutated in individuals with ID (Deciphering Developmental Disorders  
137 Study, 2017).

138

139 The first *DYRK1A* disruptions were identified in individuals with intrauterine growth  
140 restriction (IUGR), primary microcephaly and epilepsy (Møller *et al*, 2008). Few years after,  
141 the first frameshift variant was described in a patient with similar features (Courcet *et al*,  
142 2012). The clinical spectrum associated with *DYRK1A* pathogenic variants (*MRD7 for Mental*  
143 *Retardation 7* in OMIM) was further refined with the publication of additional patients,  
144 presenting suggestive facial dysmorphism, severe speech impairment and feeding difficulty,  
145 while epilepsy and prenatal microcephaly were not always present (Bronicki *et al*, 2015;  
146 Blackburn *et al*, 2019a; van Bon *et al*, 2011, 2016; O’Roak *et al*, 2012; Courcet *et al*, 2012;  
147 Okamoto *et al*, 2015; Iglesias *et al*, 2014; Ruaud *et al*, 2015; Ji *et al*, 2015; Rump *et al*, 2016;  
148 Luco *et al*, 2016; Murray *et al*, 2017; Evers *et al*, 2017; Lee *et al*, 2020a; Dang *et al*, 2018;  
149 Qiao *et al*, 2019; Ernst *et al*, 2020; Tran *et al*, 2020; Møller *et al*, 2008; Fujita *et al*, 2010;  
150 Oegema *et al*, 2010; Yamamoto *et al*, 2011; Valetto *et al*, 2012; Kim *et al*, 2017; Meissner *et*  
151 *al*, 2020; Matsumoto *et al*, 1997). Pathogenic variants were also identified in cohorts of  
152 individuals with ASD (O’Roak *et al*, 2012), but all have ID (Earl *et al*, 2017). The *DYRK1A*  
153 gene encodes a dual tyrosine-serine/threonine (Tyr-Ser/Thr) kinase 763 (NM\_001396.4)  
154 amino acids in length (Becker, 2011) including a DH (DYRK Homology-box) domain (aa  
155 137-154), two nuclear localization signal sequence (NLS)(NLS1 aa 92-104, NLS2 aa 389-  
156 395), a central catalytic domain) (aa 159-479, including Tyrosine 321, involved in the  
157 activation of *DYRK1A* by autophosphorylation (Himpel *et al*, 2001), some Ser/Thr repeats, a  
158 poly-histidine sequence (localization to nuclear speckles) and a PEST domain (aa 525-619)

159 (regulation of DYRK1A level by degradation). *DYRK1A* is ubiquitously expressed during  
160 embryonic development and in adult tissues. Its location is both cytoplasmic and nuclear and  
161 varies by cell type and stage of development (Hämmerle *et al*, 2008). By the number and  
162 diversity of its proposed protein targets, DYRK1A regulates numerous cellular functions  
163 (reviews (Tejedor & Hämmerle, 2011; Duchon & Herault, 2016)), among them the among  
164 them the MAPT (Tau) protein phosphorylated by DYRK1A on its Thr212 position (Woods *et*  
165 *al*, 2001).

166

167 High Throughput Sequencing (HTS) has revolutionized the identification of genetic  
168 variants for diagnostic applications but a major challenge remains in the interpretation of the  
169 vast number of variants, especially for highly heterogeneous disease such ID. A combination  
170 of genetic, clinical and functional approaches, summarized by the American College of  
171 Medical Genetics (ACMG), are commonly used to interpret these variants (Richards *et al*,  
172 2015). A significant proportion of the variants, especially the missense variants, remain  
173 classified as variants of unknown significance (VUS, according to ACMG) after the primary  
174 analysis. For autosomal dominant forms of ID with complete penetrance such as *DYRK1A*-  
175 related ID, the *de novo* occurrence of a variant is a strong argument for pathogenicity,  
176 however: 1) the genotype of the parents is not always available, 2) there is a probability, low  
177 but not negligible, that a variant will occur *de novo* in this gene with no link with the disease.  
178 The clinical argument might also lead to misinterpretation: under-interpretation when the  
179 clinical signs do not correspond to the initial phenotype described, over-interpretation if too  
180 much weight is given to unspecific signs. Many tools have been developed over the past ten  
181 years to predict *in silico* the pathogenicity of missense variants but they remain imperfect and  
182 *in vitro* (or *in vivo*) functional tests are useful to evaluate the true consequences of a variant,  
183 especially in the research setting as these are labor intensive and require the development of  
184 gene/protein-specific assays. DNA methylation (DNAm) is also a powerful tool to test variant  
185 pathogenicity in disorders associated with epigenetic regulatory genes. We discovered that  
186 pathogenic variants in these genes can exhibit disorder-specific DNAm signatures comprised  
187 of consistent, multilocus DNAm alterations in peripheral blood (Choufani *et al*, 2015). More  
188 than 50 DNAm signatures associated with disorders of the epigenetic machinery have now  
189 been established (Choufani *et al*, 2015, 2020; Butcher *et al*, 2017; Aref-Eshghi *et al*, 2019;  
190 Chater-Diehl *et al*, 2019). DNAm signatures are highly sensitive and specific for each  
191 condition, able to discriminate between related disorders, and useful for classifying variants in

192 these genes as pathogenic or benign (Choufani *et al*, 2015; Butcher *et al*, 2017; Aref-Eshghi *et*  
193 *al*, 2019; Chater-Diehl *et al*, 2019). *DYRK1A* has numerous targets, and while it is not well  
194 described as an epigenetic regulator, it has been shown to phosphorylate Histone H3 (Jang *et*  
195 *al*, 2014) as well proteins with acetyltransferase activity such as CBP and p300 (Li *et al*,  
196 2018). Therefore, we hypothesized that pathogenic variants in *DYRK1A* would be associated  
197 with a specific DNAm signature in blood.

198 We reviewed the clinical signs in 32 individuals carrying deletions or clearly  
199 pathogenic variants in *DYRK1A* to refine the clinical spectrum of *DYRK1A*-related ID and we  
200 developed a clinical score, CS<sub>DYRK1A</sub>, to help to recognize affected individuals and to interpret  
201 *DYRK1A* variants (reverse phenotyping). In parallel, we developed *in silico* and *in vitro*  
202 approaches to assess variant effects on *DYRK1A* function. Finally, we defined a DNAm  
203 signature specific to *DYRK1A*-related ID in patient blood. We used this combination of  
204 clinical, *in silico*, *in vitro* and DNAm tools to interpret seventeen variants identified in  
205 *DYRK1A* in patients with ID/NDD and demonstrated the utility of this multifaceted approach  
206 in avoiding misinterpretation/optimizing accurate interpretation of *DYRK1A* variants.

207

208

## 209 **METHODS**

### 210 *Patients and molecular analysis*

211 Variants in *DYRK1A* were identified during genetic analyses carried out in individuals  
212 referred to clinical genetic services for intellectual disability in France, Denmark and  
213 Switzerland: CGH-array, direct Sanger sequencing of *DYRK1A* coding sequences, targeted  
214 next generation sequencing of genes involved in ID (TES)(Redin *et al*, 2014; Carion *et al*,  
215 2020; Nasser *et al*, 2020), trio or simplex clinical or exome sequencing (CES, ES). The  
216 variants reported here were confirmed by an additional method. They are reported according  
217 to standardized nomenclature defined by the reference human genome GRCh37/hg19 and the  
218 *DYRK1A* isoform NM\_001396.4. Predictions of missense variant effects were performed  
219 using *in silico* tools such Combined Annotation Dependent Depletion (CADD)(Kircher *et al*,  
220 2014). Predictions of variant effect on splicing were performed using NnsplICE (Reese *et al*,  
221 1997) and MaxEnt (Eng *et al*, 2004). Fibroblasts established from skin biopsies were  
222 available for Ind #1, #11, #22 and #24 and were cultivated as previously described (Balak *et*



223 *al*, 2019). Paxgene blood samples were collected for Ind #9, #18, #19 and #30. mRNA  
224 extraction, RT-PCR or qPCR were performed as previously described using specific primers  
225 (sequences available on request). For RNA sequencing, libraries, sequencing and analysis  
226 were performed as previously described (Quartier *et al*, 2018).

227

### 228 ***Phenotypic analysis and clinical scoring***

229 A clinical summary, a checklist, and photographs when possible, were provided by the  
230 referring clinicians for the 42 individuals reported here as well as for six French individuals  
231 previously published (individuals Bronicki #2, #3, #8, #9 and #10 and individual Ruaud #  
232 2)(Bronicki *et al*, 2015; Ruaud *et al*, 2015). Based on the most frequent signs and the  
233 morphometric characteristics presented by 32 individuals with truncating variants in *DYRK1A*  
234 (nonsense, frameshift, splice, deletions or translocations affecting *DYRK1A*), a clinical score  
235 out of 20 was established (DYRK1A\_I, n=21 individuals with photographs available **Table**  
236 **S1**). This clinical score was calculated in a second cohort (replication cohort, DYRK1A\_R,  
237 n=13) which includes individuals already described in previous publications carrying a  
238 truncating variant (Bronicki *et al*, 2015; Ruaud *et al*, 2015; van Bon *et al*, 2016). We tested  
239 the score on individuals affected by other frequent monogenic forms of ID, caused by  
240 pathogenic variants in *DDX3X* (n=5), *ANKRD11* (n=5), *ARID1B* (n=8), *KMT2A* (n=6),  
241 *MED13L* (n=5), *SHANK3* (n=6) or *TCF4* (n=6, from (Mary *et al*, 2018)) gene. Score based on  
242 facial features was established from photographs by experimented clinical geneticists.

243

### 244 ***Definition of sets of missense variants and conservation analysis***

245 To evaluate which tools are pertinent to predict effect of missense variants on the  
246 DYRK1A protein, we used different sets of variants : 1) a set of variants presumed to be  
247 benign, i.e missense variants annotated as “benign”/“likely benign” in ClinVar as well as  
248 variants reported more than once in GnomAD (november 2019 release) (negative N-set,  
249 n=115), 2) a set of missense variants reported as “pathogenic”/“likely pathogenic” in Clinvar  
250 (positive P-set, n=16), and 3) all the missense variants reported here, in literature, or as VUS  
251 in Clinvar (test T-set, n=44)(**Table S2**). Orthologs of human DYRK1A were extracted from  
252 the OrthoInspector database version 3.0 (Nevers *et al*, 2019). Using the reference genomes in  
253 the OrthoInspector Eukaryotic database, 123 orthologous sequences were identified and a  
254 multiple sequence alignment (MSA) was constructed using the Clustal Omega software



255 (Sievers *et al*, 2011). The MSA was then manually refined to correct local alignment errors  
256 using the Jalview MSA editor (Waterhouse *et al*, 2009). The refined MSA was then used as  
257 input to the PROBE software (Kress *et al.*, 2019), in order to identify conserved regions in the  
258 sequences. The sequences in the MSA were divided into five separate clades: Vertebrates,  
259 Metazoans, Protists, Plants and Fungi.

260

### 261 ***In vitro analysis of variant effect on DYRK1A protein***

262 *DYRK1A* expression plasmids were generated from the pMH-SFB-*DYRK1A* vector  
263 containing the human *DYRK1A* cDNA sequence (NM\_001396.4) tagged with FLAG peptide  
264 at the N-Terminal side (purchased from addgene #101770; Huen lab). Variant sequences were  
265 obtained by site-directed mutagenesis with the specific primers and confirmed by Sanger  
266 sequencing as described (Quartier *et al*, 2019). HeLa, HEK293 and COS1 cells were  
267 maintained and transfected for 24h with *DYRK1A* plasmids (for immunofluorescence) or  
268 *DYRK1A* plasmids plus pEGFP-N1 plasmid (for Western Blot) as previously done (Quartier  
269 *et al*, 2019). *DYRK1A* proteins were visualized using mouse anti-FLAG antibody (1:1:000;  
270 Sigma Aldrich, #F1804), and their level normalized with EGFP (in-house mouse anti-GFP  
271 antibody). Immunofluorescence experiments were performed in HeLa cells as previously  
272 done (Mattioli *et al.*, 2018), and fluorescence was visualized on an inverted confocal  
273 microscope (SP2UV, Leica, Wetzlar, Germany). For autophosphorylation analysis, proteins  
274 were extracted from HEK293 cells transfected with *DYRK1A* plasmids and  
275 immunoprecipitated with anti-FLAG antibody as described (Mattioli *et al.*, 2018) with the  
276 addition of a phosphatase inhibitor cocktail 2 (Sigma Aldrich). Phosphorylated Tyr321  
277 *DYRK1A* was visualized using rabbit anti-phospho-HIPK2 antibody (1:1000) (Widowati *et*  
278 *al*, 2018) and normalized by the level of total *DYRK1A* protein, visualized using rabbit anti-  
279 *DYRK1A* antibody (1:1000; Cohesion Biosciences, #CPA1357). Kinase activity was  
280 investigated by co-transfecting *DYRK1A* plasmids and MAPT in HEK293 cells  
281 (MAPT\_OHu28029C\_pcDNA3.1(+)-C-HA from geneScript), adapted from what previously  
282 done (Lee *et al*, 2020b). *DYRK1A*, MAPT and pMAPT (Thr212) were visualized using anti-  
283 FLAG antibody, anti-TAU-5 antibody (LAS-12808 Thermofisher) and anti-pTAU-T212  
284 antibody (44-740G Thermofisher) and their level normalized with GAPDH (MAB374  
285 Merck). Interaction with DCAF7/WDR68 were visualized using an anti-WDR68 antibody  
286 (1:2500; abcam, ab138490). Statistical tests used are indicated in Figures's legends.

287

### 288 ***DNA methylation signature***

289 Methylation analysis was performed using blood DNA from individuals with *DYRK1A*  
290 LoF (n=16), split into signature discovery (n=10) and validation (n=6) cohorts, based  
291 primarily on whether age at time of blood collection was available, and age- and sex-matched  
292 neurotypical controls (n=24). Whole blood DNA samples were prepared, hybridized to the  
293 Illumina Infinium Human MethylationEPIC BeadChip and analyzed as previously described  
294 (Chater-Diehl *et al*, 2019), a total of n=774,590 probes were analyzed for differential  
295 methylation. Standard quality control metrics showed good data quality for all samples except  
296 Ind#20, which was below the methylated and unmethylated channel median intensity cutoff.  
297 Briefly, *limma* regression with covariates age, sex, and five of the six predicted blood cell  
298 types (i.e. all but neutrophils) identified a DNAm signature with a Benjamini-Hochberg  
299 adjusted  $p$ -value<0.05 and  $|\Delta\beta|>0.10$  (10% methylation difference) comprised of 402 CpG  
300 sites (**Table S3**). Next, we developed a support vector machine (SVM) model with linear  
301 kernel trained on including n=318 non-redundant CpG sites (after CpG sites with >90%  
302 methylation correlation were removed using *caret*) non-redundant CpG sites (Chater-Diehl *et al*,  
303 2019) using the methylation values for the discovery cases vs. controls. The model  
304 generated scores ranging between 0 and 1 (0-100%) for tested samples, classifying samples as  
305 “positive” (score>0.5) or “negative” (score<0.5). Additional neurotypical controls (n=94) and  
306 *DYRK1A* LoF validation samples (n=6) were scored to test model specificity and sensitivity  
307 respectively. EPIC array data from Ind#20 (which failed QC) classified correctly. Also  
308 scored were samples with pathogenic *KMT2A* (n=8) and *ARID1B* (n=4) variants as well as  
309 *DYRK1A* missense (n=10) and distal frameshift (n=1) variants.

### 310 ***Statistics***

311 All statistic tests were performed using the Prism software. The number of replicates is  
312 mentioned in the method section and in Figure’s legends. The statistical tests used, including  
313 the multiple testing correction methods, are indicated in each figure’s legend.

314

## 315 **RESULTS**

### 316 ***Identification of genetic variants in DYRK1A in individuals with ID***

317 We collected molecular and clinical information from 42 individuals with ID carrying  
318 a variant in *DYRK1A* identified in clinical and diagnostic laboratories: structural variants  
319 deleting or interrupting *DYRK1A* and recurrent or novel nonsense, frameshift, splice and  
320 missense variants (**Table 1, Figure S1**). When blood or fibroblast samples were available, we  
321 characterized the consequences of these variants on *DYRK1A* mRNA by RNA-sequencing or  
322 RT-qPCR (**Figure S2, Supplementary Text**). For one variant, c.1978del, occurring in the  
323 last exon of the gene (**Ind #18**), the mutant transcripts escape to nonsense mRNA mediated  
324 decay (NMD) and result in a truncated protein p.Ser660fs (or Ser660Profs\*43) retaining its  
325 entire kinase domain (**Figure S2F**). The variants occurred *de novo* in most of cases (36/42),  
326 one individual had a mosaic father and parental DNA was not available for the others.

327

### 328 *Clinical manifestations in individuals with pathogenic variants in DYRK1A and definition* 329 *of a clinical score*

330 We reviewed the clinical manifestations of the patients with truncating variants, except  
331 p.(Ser660fs)(**Supplementary text**). Recurrent features include, consistently with what was  
332 previously reported (Bronicki *et al*, 2015; Ji *et al*, 2015; Luco *et al*, 2016; van Bon *et al*,  
333 2016; Earl *et al*, 2017): moderate to severe ID, prenatal or postnatal progressive  
334 microcephaly, major speech impairment, feeding difficulties which can be very severe during  
335 infancy, seizures and especially history of febrile seizures, autistic traits and anxiety, delayed  
336 gross motor development with unstable gait, brain MRI abnormalities including dilated  
337 ventricles and corpus callosum hypoplasia and recurrent facial features (**Figure 1A,**  
338 **Supplementary text**). We noted, for the first time, the importance of skin manifestations and  
339 especially atopic dermatitis. We also found some genital abnormalities, as already reported  
340 (Blackburn *et al*, 2019b) (**Supplementary text**). We used recurrent features to establish a  
341 “DYRK1A- clinical score” ( $CS_{DYRK1A}$ ) on 20 points (**Figure 1A**), which aims to reflect  
342 specificity rather than severity of the phenotype. High scores, ranging from 13 to 18.5  
343 (mean=15.5), were obtained for the individuals having a pathogenic variant in *DYRK1A*  
344 described here (*DYRK1A\_I*) or previously (*DYRK1A\_R*) (**Table S1, Figure 1B**). The  
345 threshold of  $CS_{DYRK1A} \geq 13$  appears to be discriminant between individuals with LoF variants  
346 in *DYRK1A* (all  $\geq 13$ ) and individuals suffering from another form of ID (all  $< 13$ ). We then  
347 considered  $CS_{DYRK1A}$  above 13, comprised between 10 and 13 and below 10 as “highly

348 suggestive”, “intermediate” and “poorly evokative” respectively. A clinical score without  
349 photograph could also be calculated but is less discriminative (**Figure S3**).

350

### 351 *In silico analysis of missense variant effects*

352 We evaluated the discriminative power of the CADD score, commonly used in  
353 medical genetics (Kircher *et al*, 2014) to interpret missense variants in *DYRK1A*. If a  
354 significant difference in the CADD score distribution is observed between the variants  
355 presumed to be “benign” (N-set, see **Methods**) and those reported as “pathogenic” by  
356 molecular genetic laboratories (P-set, see **Methods**)(p-value<0.0001), a substantial proportion  
357 of the N-set variants still have a CADD score above the threshold (20 or 25) usually used to  
358 define pathogenicity (**Figure S4A**). This could be explained by the high degree of amino acid  
359 conservation of *DYRK1A* among vertebrates, and this could lead to over-interpretation of  
360 pathogenicity of missense variants. We performed sequence alignment with more than  
361 hundred orthologs of *DYRK1A* from different taxon (**Figure S5**) and confirmed that using  
362 sequences from vertebrate species only is not efficient to classify missense variants, as one  
363 third of the N-set variants affect amino acids conserved in all the vertebrate species (**Figure**  
364 **S4B**, V=100%). Considering conservation parameters going beyond vertebrates (conserved in  
365 100% of vertebrates, at least 90% of metazoan and at least 80% of other animals) appears  
366 more discriminant, keeping most of the variants from P-set (13/16) and only one variant from  
367 N-set (**Figure S4B**).

368

### 369 *In vitro characterization of consequences of missense variants on DYRK1A protein*

370 In order to test the consequences of the missense variants *in vitro*, we overexpressed  
371 wild-type (WT) and mutant *DYRK1A* proteins in three different cell lines (HEK293, HeLa,  
372 COS1) and included a truncating variant Arg413fs and a missense variant Ala341Ser from  
373 GnomAD as pathogenic and benign variants respectively. A significant decrease in *DYRK1A*  
374 protein level, due to a reduction of protein stability (**Figure S6A**), was observed for the  
375 truncating variant Arg413fs but also for the missense variants Asp287Val, Ser311Phe,  
376 Arg467Gln, Gly168Asp and Ile305Arg (**Figure 2A**). None of the variants affect *DYRK1A*  
377 interaction with DCAF7/WDR68 (**Figure S6B**). To be active, *DYRK1A* has to undergo an  
378 autophosphorylation on Tyrosine 321 (Himpel *et al*, 2001). To measure the level of active

379 DYRK1A protein, we detected phospho-DYRK1A (Tyr321) by immunoprecipitation  
380 followed by immunoblot using anti-HIPK2, as previously described (Widowati *et al*,  
381 2018)(**Figure 2B**). We observed no difference in the level of phospho-DYRK1A for the  
382 Ala341Ser variant compared to the WT protein. We confirmed that the three variants  
383 previously tested (Asp287Val, Ser311Phe and Arg467Gln) abolish autophosphorylation  
384 (Widowati *et al*, 2018; Arranz *et al*, 2019), as the Gly168Asp and Ile305Arg variants. The  
385 Ser324Arg DYRK1A variant showed only residual autophosphorylation. No effect on  
386 autophosphorylation was observed for Arg255Gln, Tyr462His, Gly486Asp and Thr588Asn.  
387 No effect was detected neither for the Glu366Asp amino acid change, but the analysis of  
388 patient's blood mRNA showed that the c.1098G>T variant affect splicing and lead to  
389 p.(Ile318\_Glu366del) instead of Glu366Asp (**Figure S2G**). We used this strategy to test  
390 additional variants reported in databases and showed that Arg158His, affecting a highly  
391 conserved amino acid position but reported twice in gnomAD, does not affect DYRK1A  
392 protein. Ala277Pro, reported as pathogenic in ClinVar but affecting a position poorly  
393 conserved beyond vertebrates, as well as Gly171Arg, Leu241Pro and Pro290Arg, reported as  
394 VUS in Clinvar, affect both DYRK1A level and autophosphorylation (**Figure S7A-B, Table**  
395 **S2, Figure S4B**). None of the missense variants appear to affect DYRK1A cellular  
396 localization, contrary to Arg413fs variant or mutations NLS domains (**Supplementary Text,**  
397 **Figure S7C**). However, we observed an aggregation of DYRK1A proteins with the distal  
398 frameshift variant Ser660fs (**Figure 2C**), which prevents us from correctly measuring the  
399 level of mutant protein and its capacity to autophosphorylate.

400

#### 401 *Identification of a DNAm signature associated with DYRK1A pathogenic variants*

402 To determine if *DYRK1A* is associated with specific changes in genome-wide DNAm  
403 in blood, we generated genome-wide methylation profiles using Infinium  
404 HumanMethylationEPIC BeadChip arrays. We compared DNAm in blood for a subset  
405 (discovery) of our cohort carrying pathogenic LoF variants in *DYRK1A* with age- and sex-  
406 matched neurotypical controls and identified n=402 differentially methylated CpG sites (FDR  
407 adjusted  $p < 0.05$ ,  $|\Delta\beta| > 0.10$ ), corresponding to 165 RefSeq genes (**Table S3, Figure 3A-B**).  
408 We trained a support vector machine (SVM) machine learning model on the DNAm ( $\beta$ )  
409 values to generate a score (0-1). We validated its sensitivity and specificity using additional  
410 individuals with *DYRK1A* truncating variants (validation), additional controls as well as  
411 individuals with pathogenic variants in other epigenetic regulatory genes *ARID1B* and

412 *KMT2A* (**Table S4; Figure 3C**). Next, we scored the samples with missense variants in  
413 *DYRK1A* and found that six classified positively: samples with p.Asp287Val, p.Ser311Phe,  
414 p.Arg467Gln, p.Gly168Asp, p.Ile305Arg and p.Ser324Arg, and three negatively: samples  
415 with p.Arg255Gln, p.Tyr462His, p.Thr588Asn (**Figure 3B-C, Table S4, Figure S8**). The  
416 sample with the distal frameshift variant p.Ser660fs also classified as DNAm positive, with a  
417 relatively high score (0.92). The sample with the p.Gly486Asp variant clustered out from both  
418 *DYRK1A* cases and controls using hierarchical clustering and PCA. Specifically, its  
419 methylation profile is opposite to *DYRK1A* LoF cases relative to controls, i.e. reduced DNAm  
420 at the same sites that demonstrate increased methylation in LoF cases, and increased  
421 methylation at sites decreased in LoF cases (**Figure S8**). This pattern is evident when the  $\beta$   
422 values at the top 25 hyper- and hypermethylated signature CpGs are plotted (**Figure S10**). We  
423 found that for one third of the signature sites (134/402) the  $\beta$  value for p.Gly486Asp was  
424 outside the range observed for that of all discovery controls (**Table S3**), suggesting this  
425 variant might have a gain-of-function (GoF) effect. A notable feature of these GoF CpG sites  
426 is that they tended to cluster together, as for instance all seven of the signature sites in the  
427 *HIST1H3E* promoter (**Table S3**).

428

#### 429 ***Integration of the different tools to reclassify variants***

430 We integrated data from the clinical score ( $CS_{DYRK1A}$ ), *in silico* predictions, functional  
431 assays and DNAm model score to evaluate the pathogenicity of the variants and reclassify  
432 them according to ACMG categories (**Figure 4, Table S5**). We found that variants  
433 Gly168Asp, Asp287Val, Ile305Arg, Ser311Phe and Arg467Gln, identified in individuals with  
434 intermediate to high  $CS_{DYRK1A}$  scores, led to reduced protein expression as well as an absence  
435 of autophosphorylation activity, which was previously described for three of them (Widowati  
436 *et al*, 2018; Arranz *et al*, 2019). All classified as DNAm-positive, definitively supporting their  
437 pathogenicity. For the Ser324Arg variant, identified *de novo* in a patient with an intermediate  
438  $CS_{DYRK1A}$  score, we observed only a slight decrease of *DYRK1A* stability and a partial  
439 decrease of its autophosphorylation ability. The binary nature of the DNAm signature,  
440 showing a positive score, definitively supports its pathogenic effect.

441 The Arg255Gln and Tyr462His were identified in individuals with very low  $CS_{DYRK1A}$   
442 score, they had relatively high CADD score (24 and 29.6) but affect amino acids not highly  
443 conserved. They had no effect on protein level, autophosphorylation, and cellular localization



444 of *DYRK1A* and classified DNAm-negative. They were therefore both considered to be likely  
445 benign. Parental DNA was not available to test the inheritance of Arg255Gln, though it is  
446 notable that the variant Tyr462His occurred *de novo*. This individual has an affected brother  
447 who does not carry the variant, and exome sequencing of the whole family failed to identify  
448 additional promising variants. Another *de novo* variant affecting the same position Tyr462Cys  
449 was identified in a girl with mild developmental delay, hypotonia and hypermobility without  
450 facial dysmorphism, who finally obtained another molecular diagnosis, i.e. a *de novo* truncating  
451 variant in a gene that is currently under investigation as a strong candidate gene for syndromic  
452 intellectual disability (personal communication Sander Stegmann, Maastricht University  
453 Medical Center). The Thr588Asn variant, previously reported as likely pathogenic (Bronicki  
454 *et al*, 2015), appears to have no effect on *DYRK1A* mRNA or on *DYRK1A* protein level and  
455 function, consistent with what was described by others (Widowati *et al*, 2018; Arranz *et al*,  
456 2019). To go further, we tested the ability of the mutant Thr588Asn *DYRK1A* to  
457 phosphorylate MAPT on its Thr212 and confirmed it does not affect its kinase activity  
458 (**Figure S9**). Moreover, a knock-in Thr588Asn mouse model was generated and failed to  
459 present any decrease of kinase activity and any obvious behavioral phenotype  
460 (**Supplementary Text, Figure S11**). The patient's DNAm score was negative, and we  
461 therefore reclassified this variant as likely benign. The fact that it occurred *de novo* in a girl  
462 with a high  $CS_{DYRK1A}$  (15.5/20) remains puzzling, while no additional promising variants  
463 were identified in trio-exome sequencing data and no positive classification was found using  
464 ~20 DNAm signatures available which could explain this phenocopy.

465 Only one sample showed a DNAm profile different from controls and individuals with  
466 *DYRK1A*-related ID (**Figure 3A, Figure S8 and S11**). This individual has a very low  
467  $CS_{DYRK1A}$ , presenting relative macrocephaly and ASD without ID and carries a *de novo*  
468 Gly486Asp variant. This variant was previously reported in another individual with NDD  
469 (Dang *et al*, 2018), but it was not possible to obtain DNA or any clinical or inheritance  
470 information. No decrease of protein level, autophosphorylation or modification in kinase  
471 activity was identified (**Figure S9**), consistent with what was reported by others (Arranz *et al*,  
472 2019). Arranz *et al*. has observed on the contrary an increase of *DYRK1A* kinase activity,  
473 which could suggest a potential GoF effect. However, in their study, they reported a  
474 significant increase of kinase activity for five additional variants located all along the protein,  
475 including one reported four times in GnomAD (Arg528Trp), which might raise questions



476 about the sensitivity of the test. We tested the kinase activity of Gly486Asp mutant protein on  
477 MAPT Thr212 but could not observed any difference with WT.

478 We characterized the consequences of a distal frameshift *de novo* variant, Ser660fs. Its  
479 overexpression in HeLa cells leads to cytoplasmic aggregation of DYRK1A, which makes  
480 difficult to quantify the real effect on protein level, autophosphorylation, or kinase activity on  
481 MAPT (**Figure S7 and S10**). However, its DNAm overlaps those of other individuals with  
482 truncating variants located further upstream in the protein, confirming its pathogenic effect  
483 (**Figure S8**). To test if these aggregations could be driven by the novel C-terminal extension  
484 (43 amino acids) added by the frameshift variant, we introduced nonsense variants at the same  
485 positions (Ser660\* and Ser661\*). As no aggregates were detected (**Figure 2C & S8D**), we  
486 concluded that the abnormal C-terminal extension of 43 amino acids is responsible for the  
487 self-aggregation of the mutant DYRK1A protein. Interestingly, the two truncating variants  
488 Ser660\* and Ser661\* did not affect DYRK1A level, localization or autophosphorylation  
489 (**Figure S7A-B**).

490

## 491 **DISCUSSION**

492 Here we report clinical manifestations of 32 patients with clear loss-of-function (LoF)  
493 variants in *DYRK1A*, refining the clinical spectrum associated with *DYRK1A*-related ID. We  
494 used recurrent signs present in more than two-third of the individuals to establish a clinical  
495 score, which may seem outdated in the era of pangenomic approaches, but which it is in fact  
496 very useful to interpret variants of unknown significance identified by these approaches.  
497 Indeed, here we demonstrated that the combination of clinical data together with *in silico* and  
498 *in vitro* observations are essential to interpret variants correctly.

499 Since *DYRK1A* is a highly conserved gene in vertebrates, we assumed that *in silico*  
500 prediction tools using conservation calculated mainly from vertebrates might overestimate the  
501 potential pathogenicity of missense variants. We showed that deeper conservation analyses  
502 using additional taxa are useful to improve the predictions for missense variants. However, *in*  
503 *silico* analyses have their limitations, and functional assays are essential to assess variant  
504 effect conclusively. We therefore tested the effect of 17 variants and showed that half of them  
505 decreased both DYRK1A protein level and DYRK1A autophosphorylation level. For the  
506 remaining variants, no effect on protein function was observed (**Figure 4**). However, the

507 absence of effects observed during a series of functional tests does not totally exclude a  
508 potential effect.

509 Over the past five years, several studies have found patients with specific monogenic  
510 disorders involving genes encoding epigenetic regulatory proteins are associated with DNAm  
511 signatures in blood. The advantage of such signatures is a high rate of clear classification  
512 (positive vs negative) they provide for variants in most cases as pathogenic or benign.  
513 Considering the potential role played by *DYRK1A* in epigenetic regulation (Lepagnol-Bestel  
514 *et al*, 2009; Jang *et al*, 2014; Li *et al*, 2018), we wanted to test whether *DYRK1A* LoF leads to  
515 such a DNAm signature. We identified a DNAm signature associated with pathogenic  
516 variants in *DYRK1A* and demonstrated that it has 100% of sensitivity and specificity (**Figure**  
517 **3**). The combination of clinical score ( $CS_{DYRK1A}$ ), *in silico* predictions, functional assays and  
518 DNAm signature allow to reclassify ten missense variants as pathogenic while three could be  
519 considered as benign: a variant located in the catalytic domain whose inheritance was  
520 unknown, Arg255Gln, and two *de novo* variants located at the end or outside of this domain:  
521 Tyr462His and Thr588Asn. Though the missense variants classified clearly by the DNAm  
522 model, there was a larger range of SVM scores among the positive missense cases than the  
523 validation LoF cases (**Figure 3B**). This could indicate a less severe effect of the missense vs.  
524 LoF variants, however, there was no correlation between the severity of the variant effect *in*  
525 *vitro* and the SVM score, as illustrated by the high score obtained for the variant Ser324Arg  
526 (**Table S4**).

527 Still based on methylation data, we suspected a gain-of-function (GoF) effect for  
528 another *de novo* variant located outside the catalytic domain: Gly486Asp. We have already  
529 shown that DNAm profiles at gene-specific signature sites provide a functional readout of  
530 each variants effect, GoF activity. Indeed, in previous work, we found the same pattern for a  
531 patient with a missense variant in *EZH2*, typically associated with Weaver syndrome. The  
532 patient, presenting undergrowth rather than overgrowth characteristic of Weaver syndrome,  
533 had an opposite DNAm profile to *EZH2* cases relative to controls and carried a Ala738Thr  
534 variant which was demonstrated to increase *EZH2* activity using a luminescence enzymatic  
535 assay (Choufani *et al*, 2020). However, we could not confirm the putative GoF effect of  
536 Gly486Asp with the kinase assay we performed (phosphorylation of MAPT-Thr212).

537 We reported here a distal frameshift *de novo* variant occurring in the last exon of  
538 *DYRK1A* and escaping to NMD. We showed that the novel C-terminal part of the mutant

539 protein leads to DYRK1A aggregation *in vitro*, which needs be confirmed *in vivo* in  
540 physiological context. Interestingly, the two nonsense changes introduced at this position,  
541 Ser660\* and Ser661\*, did not lead to DYRK1A aggregation, and did not affect neither its  
542 expression, ability to autophosphorylate and to phosphorylate MAPT (**Figure 2C, S8C-D,**  
543 **S10**). These results raise questions about the pathogenicity of distal truncating variants which  
544 would escape to NMD. Three additional such variants were reported in individuals with ID or  
545 NDD in Clinvar and in literature (Okamoto et al., 2015)(**Table S7**). Their clinical  
546 interpretation remains ambiguous, especially for c.1726C>T p.Gln576\* and c.2040C>A  
547 p.Tyr680\*, for which inheritance is unknown and clinical manifestations do not really overlap  
548 those of *DYRK1A*-related ID. However, the most distal variant ever reported in *DYRK1A*,  
549 c.2213\_2218delinsAGAG p.Thr738fs, occurred *de novo* in an individual with clinical features  
550 consistent with *DYRK1A*-related ID. It would be interesting to obtain DNA and perform  
551 DNAm analysis in these three individuals to reclassify these variants.

552 In conclusion, we developed various tools (clinical score, protein sequence  
553 conservation data, *in vitro* functional assays and a specific DNAm signature) to help to better  
554 diagnose *DYRK1A*-related ID by improving variant interpretation. This combination of tools  
555 was efficient to reclassify variants identified in this gene. We showed that missense variants  
556 located outside but also inside the catalytic domain, even occurring *de novo*, as well as  
557 variants leading to premature stop codon in the last exon of the gene, are not necessarily  
558 pathogenic. These results illustrate that variants in *DYRK1A*, as well as in other NDD  
559 causative genes, must be interpreted with caution, even if they occur *de novo*, to avoid over-  
560 interpretation. In future, we recommend performing DNAm analysis if blood DNA sample is  
561 available or, if not, *in vitro* testing of variant effect on DYRK1A autophosphorylation.

562

## 563 **ACKNOWLEDGMENTS**

564 The authors would like to warmly thank the families for their participation to the study, for  
565 their trust and their support to our research projects. The authors also thank the Agence de  
566 Biomédecine and Fondation APLM for their financial support. We also thank the Centre  
567 National de Génotypage (Jean-François Deleuze, Robert Olasso, Anne Boland and the  
568 technicians and bioinformaticians) for their participation in library preparation and DNA  
569 sequencing. We thank all the people from the GenomEast sequencing platform (Damien  
570 Plassard, Céline Keime, Serge Vicaire, Bernard Jost, Stéphanie Le Gras, Mathieu Jung, etc)

571 for their technical and bioinformatics supports. Also thanks to Paola Rossadillo and Karim  
572 Essabri for their help for the cloning and mutagenesis. They thank people for the diagnostic  
573 laboratory of Hôpitaux Universitaire de Strasbourg (HUS) for performing follow-up of  
574 mutations and giving diagnosis to family (Elsa Nourisson, Céline Cuny, Sylvie Friedman,  
575 Carmen Fruchart) and Anais Philippe and Justine Fraize for their help on clinical files.

576

#### 577 **AUTHOR CONTRIBUTIONS**

578 JC conceived and designed the in vitro experiments, RNA-Seq, RT-PCR experiments ; JC,  
579 BDu, IB, LG performed the experiments, performed statistical analysis and analyzed the data  
580 related to in vitro experiments ; ND provided technical assistance and performed  
581 experiments ; MdM MM, M-C B, YH conceived and created the mouse Thr588Asn knock-in  
582 model, designed and performed behavioral experiments ; ES, MN, BG, MA, BC, LB, LB,  
583 MB, PC, EC, CC, AD, BD, FD, SP, A-SD, LF, CF, MF, CF, DG, AG, A-MG, BI, KM. J,  
584 BK, MK, PK, M MD, BD, JM, RSM, SM, LP, KPS, LPe, MR, PS, JS, JT, FTMT, CET, AV,  
585 VL, SD, AB, CD, SEC, CFa, CM, MG, TB, JLM contributed to clinical and molecular data ;  
586 AP, BDu, MV, MW compiled the molecular and clinical data and established a clinical score ;  
587 Bdu, ES, MZ analyzed and score blindly the facial features ; JT generates multiple sequence  
588 alignment and performed conservation analysis; E C-D &, RW conceived, designed and  
589 performed the DNA methylation analysis; AP conceived, coordinated, supervised the study;  
590 JC & AP wrote the manuscript, with significant contributions of E C-D, RW, BDu, MV, MW  
591 & JLM.

592

#### 593 **CONFLICTS OF INTEREST**

594 None

595

#### 596 **DATA AVAILABILITY**

597 Variants were submitted to ClinVar database. Additional data are available upon request.

598

#### 599 **ETHICS DECLARATION**

600 This study was approved by the local Ethics Committee of the Strasbourg University Hospital  
601 (Comité Consultatif de Protection des Personnes dans la Recherche Biomédicale (CCPPRB)).

602 All patients enrolled in these genetic studies and/or their legal representative have signed  
603 informed consent for genetic testing and authorization for publication.

604

## 605 **WEB RESSOURCES**

606 The URLs for online tools and data presented herein are:

607 Clinvar: <http://www.ncbi.nlm.nih.gov/clinvar/>

608 dbSNP : <http://www.ncbi.nlm.nih.gov/projects/SNP/>

609 Decipher: <https://decipher.sanger.ac.uk/>

610 GnomAD: <http://gnomad.broadinstitute.org/>

611 Integrative Genomics Viewer (IGV): <http://www.broadinstitute.org/igv/>

612 Mutation Nomenclature: <http://www.hgvs.org/mutnomen/recs.html>

613 OMIM: <http://www.omim.org/>

614 UCSC: <http://genome.ucsc.edu/>

615 CADD score: <https://cadd.gs.washington.edu/>

616 OrthoInspector database: <https://www.lbgi.fr/orthoinspectorv3/databases>

617

## 618 **BIBLIOGRAPHY**

619 Aref-Eshghi E, Bend EG, Colaiacovo S, Caudle M, Chakrabarti R, Napier M, Brick L, Brady L, Carere DA,  
620 Levy MA, *et al* (2019) Diagnostic Utility of Genome-wide DNA Methylation Testing in  
621 Genetically Unsolved Individuals with Suspected Hereditary Conditions. *Am J Hum Genet* 104:  
622 685–700

623 Arranz J, Balducci E, Arató K, Sánchez-Elexpuru G, Najas S, Parras A, Rebollo E, Pijuan I, Erb I, Verde G,  
624 *et al* (2019) Impaired development of neocortical circuits contributes to the neurological  
625 alterations in DYRK1A haploinsufficiency syndrome. *Neurobiol Dis* 127: 210–222

626 Balak C, Benard M, Schaefer E, Iqbal S, Ramsey K, Ernoult-Lange M, Mattioli F, Llaci L, Geoffroy V,  
627 Courel M, *et al* (2019) Rare De Novo Missense Variants in RNA Helicase DDX6 Cause  
628 Intellectual Disability and Dysmorphic Features and Lead to P-Body Defects and RNA  
629 Dysregulation. *Am J Hum Genet* 105: 509–525

630 Becker W (2011) Recent insights into the function of DYRK1A. *FEBS J* 278: 222

631 Blackburn ATM, Bekheirnia N, Uma VC, Corkins ME, Xu Y, Rosenfeld JA, Bainbridge MN, Yang Y, Liu P,  
632 Madan-Khetarpal S, *et al* (2019a) DYRK1A-related intellectual disability: a syndrome

- 633 associated with congenital anomalies of the kidney and urinary tract. *Genet Med* 21: 2755–  
634 2764
- 635 Blackburn ATM, Bekheirnia N, Uma VC, Corkins ME, Xu Y, Rosenfeld JA, Bainbridge MN, Yang Y, Liu P,  
636 Madan-Khetarpal S, *et al* (2019b) DYRK1A-related intellectual disability: a syndrome  
637 associated with congenital anomalies of the kidney and urinary tract. *Genet Med* 21: 2755–  
638 2764
- 639 van Bon BWM, Coe BP, Bernier R, Green C, Gerdtts J, Witherspoon K, Kleefstra T, Willemsen MH,  
640 Kumar R, Bosco P, *et al* (2016) Disruptive de novo mutations of DYRK1A lead to a syndromic  
641 form of autism and ID. *Mol Psychiatry* 21: 126–132
- 642 van Bon BWM, Hoischen A, Hehir-Kwa J, de Brouwer APM, Ruivenkamp C, Gijsbers ACJ, Marcelis CL,  
643 de Leeuw N, Veltman JA, Brunner HG, *et al* (2011) Intragenic deletion in DYRK1A leads to  
644 mental retardation and primary microcephaly. *Clin Genet* 79: 296–299
- 645 Bronicki LM, Redin C, Drunat S, Piton A, Lyons M, Passemard S, Baumann C, Faivre L, Thevenon J,  
646 Rivière J-B, *et al* (2015) Ten new cases further delineate the syndromic intellectual disability  
647 phenotype caused by mutations in DYRK1A. *Eur J Hum Genet* 23: 1482–1487
- 648 Butcher DT, Cytrynbaum C, Turinsky AL, Siu MT, Inbar-Feigenberg M, Mendoza-Londono R, Chitayat  
649 D, Walker S, Machado J, Caluseriu O, *et al* (2017) CHARGE and Kabuki Syndromes: Gene-  
650 Specific DNA Methylation Signatures Identify Epigenetic Mechanisms Linking These Clinically  
651 Overlapping Conditions. *Am J Hum Genet* 100: 773–788
- 652 Carion N, Briand A, Cuisset L, Pacot L, Afenjar A & Bienvenu T (2020) Loss of the KH1 domain of FMR1  
653 in humans due to a synonymous variant causes global developmental retardation. *Gene* 753:  
654 144793
- 655 Chater-Diehl E, Ejaz R, Cytrynbaum C, Siu MT, Turinsky A, Choufani S, Goodman SJ, Abdul-Rahman O,  
656 Bedford M, Dorrani N, *et al* (2019) New insights into DNA methylation signatures: SMARCA2  
657 variants in Nicolaides-Baraitser syndrome. *BMC Med Genomics* 12: 105
- 658 Choufani S, Cytrynbaum C, Chung BHY, Turinsky AL, Grafodatskaya D, Chen YA, Cohen ASA, Dupuis L,  
659 Butcher DT, Siu MT, *et al* (2015) NSD1 mutations generate a genome-wide DNA methylation  
660 signature. *Nat Commun* 6: 10207
- 661 Choufani S, Gibson WT, Turinsky AL, Chung BHY, Wang T, Garg K, Vitriolo A, Cohen ASA, Cyrus S,  
662 Goodman S, *et al* (2020) DNA Methylation Signature for EZH2 Functionally Classifies  
663 Sequence Variants in Three PRC2 Complex Genes. *Am J Hum Genet* 106: 596–610
- 664 Courcet J-B, Faivre L, Malzac P, Masurel-Paulet A, Lopez E, Callier P, Lambert L, Lemesle M, Thevenon  
665 J, Gigot N, *et al* (2012) The DYRK1A gene is a cause of syndromic intellectual disability with  
666 severe microcephaly and epilepsy. *J Med Genet* 49: 731–736
- 667 Dang T, Duan WY, Yu B, Tong DL, Cheng C, Zhang YF, Wu W, Ye K, Zhang WX, Wu M, *et al* (2018)  
668 Autism-associated Dyrk1a truncation mutants impair neuronal dendritic and spine growth  
669 and interfere with postnatal cortical development. *Mol Psychiatry* 23: 747–758
- 670 Deciphering Developmental Disorders Study (2017) Prevalence and architecture of de novo  
671 mutations in developmental disorders. *Nature* 542: 433–438

- 672 Duchon A & Herault Y (2016) DYRK1A, a Dosage-Sensitive Gene Involved in Neurodevelopmental  
673 Disorders, Is a Target for Drug Development in Down Syndrome. *Front Behav Neurosci* 10:  
674 104
- 675 Earl RK, Turner TN, Mefford HC, Hudac CM, Gerdtz J, Eichler EE & Bernier RA (2017) Clinical  
676 phenotype of ASD-associated DYRK1A haploinsufficiency. *Mol Autism* 8: 54
- 677 Eng L, Coutinho G, Nahas S, Yeo G, Tanouye R, Babaei M, Dörk T, Burge C & Gatti RA (2004)  
678 Nonclassical splicing mutations in the coding and noncoding regions of the ATM Gene:  
679 maximum entropy estimates of splice junction strengths. *Hum Mutat* 23: 67–76
- 680 Ernst J, Alabek ML, Eldib A, Madan-Khetarpal S, Sebastian J, Bhatia A, Liasis A & Nischal KK (2020)  
681 Ocular findings of albinism in DYRK1A-related intellectual disability syndrome. *Ophthalmic*  
682 *Genet*: 1–6
- 683 Evers JMG, Laskowski RA, Bertolli M, Clayton-Smith J, Deshpande C, Eason J, Elmslie F, Flinter F,  
684 Gardiner C, Hurst JA, *et al* (2017) Structural analysis of pathogenic mutations in the DYRK1A  
685 gene in patients with developmental disorders. *Hum Mol Genet* 26: 519–526
- 686 Fujita H, Torii C, Kosaki R, Yamaguchi S, Kudoh J, Hayashi K, Takahashi T & Kosaki K (2010)  
687 Microdeletion of the Down syndrome critical region at 21q22. *Am J Med Genet A* 152A: 950–  
688 953
- 689 Gonzalez-Mantilla AJ, Moreno-De-Luca A, Ledbetter DH & Martin CL (2016) A Cross-Disorder Method  
690 to Identify Novel Candidate Genes for Developmental Brain Disorders. *JAMA Psychiatry* 73:  
691 275–283
- 692 Hämmerle B, Elizalde C & Tejedor FJ (2008) The spatio-temporal and subcellular expression of the  
693 candidate Down syndrome gene Mnb/Dyrk1A in the developing mouse brain suggests  
694 distinct sequential roles in neuronal development. *Eur J Neurosci* 27: 1061–1074
- 695 Himpel S, Panzer P, Eirnbter K, Czajkowska H, Sayed M, Packman LC, Blundell T, Kentrup H,  
696 Grötzinger J, Joost HG, *et al* (2001) Identification of the autophosphorylation sites and  
697 characterization of their effects in the protein kinase DYRK1A. *Biochem J* 359: 497–505
- 698 Iglesias A, Anyane-Yeboah K, Wynn J, Wilson A, Truitt Cho M, Guzman E, Sisson R, Egan C & Chung WK  
699 (2014) The usefulness of whole-exome sequencing in routine clinical practice. *Genet Med* 16:  
700 922–931
- 701 Jang SM, Azebi S, Soubigou G & Muchardt C (2014) DYRK1A phosphorylates histone H3 to differentially  
702 regulate the binding of HP1 isoforms and antagonize HP1-mediated transcriptional  
703 repression. *EMBO Rep* 15: 686–694
- 704 Ji J, Lee H, Argiropoulos B, Dorrani N, Mann J, Martinez-Agosto JA, Gomez-Ospina N, Gallant N,  
705 Bernstein JA, Hudgins L, *et al* (2015) DYRK1A haploinsufficiency causes a new recognizable  
706 syndrome with microcephaly, intellectual disability, speech impairment, and distinct facies.  
707 *Eur J Hum Genet* 23: 1473–1481
- 708 Kim O-H, Cho H-J, Han E, Hong TI, Ariyasiri K, Choi J-H, Hwang K-S, Jeong Y-M, Yang S-Y, Yu K, *et al*  
709 (2017) Zebrafish knockout of Down syndrome gene, DYRK1A, shows social impairments  
710 relevant to autism. *Mol Autism* 8: 50



- 711 Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM & Shendure J (2014) A general framework for  
712 estimating the relative pathogenicity of human genetic variants. *Nat Genet* 46: 310–315
- 713 Lee K-S, Choi M, Kwon D-W, Kim D, Choi J-M, Kim A-K, Ham Y, Han S-B, Cho S & Cheon CK (2020a) A  
714 novel de novo heterozygous DYRK1A mutation causes complete loss of DYRK1A function and  
715 developmental delay. *Sci Rep* 10: 9849
- 716 Lee K-S, Choi M, Kwon D-W, Kim D, Choi J-M, Kim A-K, Ham Y, Han S-B, Cho S & Cheon CK (2020b) A  
717 novel de novo heterozygous DYRK1A mutation causes complete loss of DYRK1A function and  
718 developmental delay. *Sci Rep* 10: 9849
- 719 Lepagnol-Bestel A-M, Zvara A, Maussion G, Quignon F, Ngimbous B, Ramoz N, Imbeaud S, Loe-Mie Y,  
720 Benihoud K, Agier N, *et al* (2009) DYRK1A interacts with the REST/NRSF-SWI/SNF chromatin  
721 remodelling complex to deregulate gene clusters involved in the neuronal phenotypic traits  
722 of Down syndrome. *Hum Mol Genet* 18: 1405–1414
- 723 Li S, Xu C, Fu Y, Lei P-J, Yao Y, Yang W, Zhang Y, Washburn MP, Florens L, Jaiswal M, *et al* (2018)  
724 DYRK1A interacts with histone acetyl transferase p300 and CBP and localizes to enhancers.  
725 *Nucleic Acids Res* 46: 11202–11213
- 726 Luco SM, Pohl D, Sell E, Wagner JD, Dymont DA & Daoud H (2016) Case report of novel DYRK1A  
727 mutations in 2 individuals with syndromic intellectual disability and a review of the literature.  
728 *BMC Med Genet* 17: 15
- 729 Mary L, Piton A, Schaefer E, Mattioli F, Nourisson E, Feger C, Redin C, Barth M, El Chehadeh S, Colin E,  
730 *et al* (2018) Disease-causing variants in TCF4 are a frequent cause of intellectual disability:  
731 lessons from large-scale sequencing approaches in diagnosis. *Eur J Hum Genet* 26: 996–1006
- 732 Matsumoto N, Ohashi H, Tsukahara M, Kim KC, Soeda E & Niikawa N (1997) Possible narrowed  
733 assignment of the loci of monosomy 21-associated microcephaly and intrauterine growth  
734 retardation to a 1.2-Mb segment at 21q22.2. *Am J Hum Genet* 60: 997–999
- 735 Meissner LE, Macnamara EF, D’Souza P, Yang J, Vezina G, Undiagnosed Diseases Network, Ferreira  
736 CR, Zein WM, Tiffit CJ & Adams DR (2020) DYRK1A pathogenic variants in two patients with  
737 syndromic intellectual disability and a review of the literature. *Mol Genet Genomic Med*:  
738 e1544
- 739 Møller RS, Kübart S, Hoeltzenbein M, Heye B, Vogel I, Hansen CP, Menzel C, Ullmann R, Tommerup N,  
740 Ropers H-H, *et al* (2008) Truncation of the Down syndrome candidate gene DYRK1A in two  
741 unrelated patients with microcephaly. *Am J Hum Genet* 82: 1165–1170
- 742 Murray CR, Abel SN, McClure MB, Foster J, Walke MI, Jayakar P, Bademci G & Tekin M (2017) Novel  
743 Causative Variants in DYRK1A, KARS, and KAT6A Associated with Intellectual Disability and  
744 Additional Phenotypic Features. *J Pediatr Genet* 6: 77–83
- 745 Nasser H, Vera L, Elmaleh-Bergès M, Steindl K, Letard P, Teissier N, Ernault A, Guimiot F, Afenjar A,  
746 Moutard ML, *et al* (2020) CDK5RAP2 primary microcephaly is associated with hypothalamic,  
747 retinal and cochlear developmental defects. *J Med Genet* 57: 389–399
- 748 Nevers Y, Kress A, Defosset A, Ripp R, Linard B, Thompson JD, Poch O & Lecompte O (2019)  
749 Ortholnspector 3.0: open portal for comparative genomics. *Nucleic Acids Res* 47: D411–D418

- 750 Oegema R, de Klein A, Verkerk AJ, Schot R, Dumee B, Douben H, Eussen B, Dubbel L, Poddighe PJ, van  
751 der Laar I, *et al* (2010) Distinctive Phenotypic Abnormalities Associated with Submicroscopic  
752 21q22 Deletion Including DYRK1A. *Mol Syndromol* 1: 113–120
- 753 Okamoto N, Miya F, Tsunoda T, Kato M, Saitoh S, Yamasaki M, Shimizu A, Torii C, Kanemura Y &  
754 Kosaki K (2015) Targeted next-generation sequencing in the diagnosis of  
755 neurodevelopmental disorders. *Clin Genet* 88: 288–292
- 756 O’Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, Carvill G, Kumar A, Lee C, Ankenman K,  
757 *et al* (2012) Multiplex targeted sequencing identifies recurrently mutated genes in autism  
758 spectrum disorders. *Science* 338: 1619–1622
- 759 Qiao F, Shao B, Wang C, Wang Y, Zhou R, Liu G, Meng L, Hu P & Xu Z (2019) A De Novo Mutation in  
760 DYRK1A Causes Syndromic Intellectual Disability: A Chinese Case Report. *Front Genet* 10:  
761 1194
- 762 Quartier A, Chatrousse L, Redin C, Keime C, Haumesser N, Maglott-Roth A, Brino L, Le Gras S,  
763 Benchoua A, Mandel J-L, *et al* (2018) Genes and Pathways Regulated by Androgens in Human  
764 Neural Cells, Potential Candidates for the Male Excess in Autism Spectrum Disorder. *Biol*  
765 *Psychiatry*
- 766 Quartier A, Courraud J, Thi Ha T, McGillivray G, Isidor B, Rose K, Drouot N, Savidan M-A, Feger C,  
767 Jagline H, *et al* (2019) Novel mutations in NLGN3 causing autism spectrum disorder and  
768 cognitive impairment. *Hum Mutat* 40: 2021–2032
- 769 Redin C, Gérard B, Lauer J, Herenger Y, Muller J, Quartier A, Masurel-Paulet A, Willems M, Lesca G, El-  
770 Chehadeh S, *et al* (2014) Efficient strategy for the molecular diagnosis of intellectual disability  
771 using targeted high-throughput sequencing. *J Med Genet* 51: 724–736
- 772 Reese MG, Eeckman FH, Kulp D & Haussler D (1997) Improved splice site detection in Genie. *J*  
773 *Comput Biol* 4: 311–323
- 774 Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, *et al*  
775 (2015) Standards and guidelines for the interpretation of sequence variants: a joint  
776 consensus recommendation of the American College of Medical Genetics and Genomics and  
777 the Association for Molecular Pathology. *Genet Med* 17: 405–424
- 778 Ruaud L, Mignot C, Guët A, Ohl C, Nava C, Héron D, Keren B, Depienne C, Benoit V, Maystadt I, *et al*  
779 (2015) DYRK1A mutations in two unrelated patients. *Eur J Med Genet* 58: 168–174
- 780 Rump P, Jazayeri O, van Dijk-Bos KK, Johansson LF, van Essen AJ, Verheij JBG, Veenstra-Knol HE,  
781 Redeker EJW, Mannens MMAM, Swertz MA, *et al* (2016) Whole-exome sequencing is a  
782 powerful approach for establishing the etiological diagnosis in patients with intellectual  
783 disability and microcephaly. *BMC Med Genomics* 9: 7
- 784 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J,  
785 *et al* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments  
786 using Clustal Omega. *Mol Syst Biol* 7: 539
- 787 Tejedor FJ & Hämmerle B (2011) MNB/DYRK1A as a multiple regulator of neuronal development.  
788 *FEBS J* 278: 223–235

- 789 Tran KT, Le VS, Bui HTP, Do DH, Ly HTT, Nguyen HT, Dao LTM, Nguyen TH, Vu DM, Ha LT, *et al* (2020)  
790 Genetic landscape of autism spectrum disorder in Vietnamese children. *Sci Rep* 10: 5034
- 791 Valetto A, Orsini A, Bertini V, Toschi B, Bonuccelli A, Simi F, Sammartino I, Taddeucci G, Simi P &  
792 Saggese G (2012) Molecular cytogenetic characterization of an interstitial deletion of  
793 chromosome 21 (21q22.13q22.3) in a patient with dysmorphic features, intellectual disability  
794 and severe generalized epilepsy. *Eur J Med Genet* 55: 362–366
- 795 Vissers LELM, Gilissen C & Veltman JA (2016) Genetic studies in intellectual disability and related  
796 disorders. *Nat Rev Genet* 17: 9–18
- 797 Waterhouse AM, Procter JB, Martin DMA, Clamp M & Barton GJ (2009) Jalview Version 2--a multiple  
798 sequence alignment editor and analysis workbench. *Bioinformatics* 25: 1189–1191
- 799 Widowati EW, Ernst S, Hausmann R, Müller-Newen G & Becker W (2018) Functional characterization  
800 of DYRK1A missense variants associated with a syndromic form of intellectual deficiency and  
801 autism. *Biol Open* 7
- 802 Woods YL, Cohen P, Becker W, Jakes R, Goedert M, Wang X & Proud CG (2001) The kinase DYRK  
803 phosphorylates protein-synthesis initiation factor eIF2Bepsilon at Ser539 and the  
804 microtubule-associated protein tau at Thr212: potential role for DYRK as a glycogen synthase  
805 kinase 3-priming kinase. *Biochem J* 355: 609–615
- 806 Yamamoto T, Shimojima K, Nishizawa T, Matsuo M, Ito M & Imai K (2011) Clinical manifestations of  
807 the deletion of Down syndrome critical region including DYRK1A and KCNJ6. *Am J Med Genet*  
808 A 155A: 113–119

809

810

811

812

813

814

815

## 816 **FIGURE LEGENDS**

### 817 **Figure 1. Clinical score for Intellectual Disability associated to *DYRK1A*** 818 **haploinsufficiency**

819 **(A)** Clinical score out of 20 points established according to the most recurrent clinical features  
820 presented by patients (the weight assigned to each symptom being based on its recurrence):

821 clinical symptoms are out of 15 points, while the facial appearance is out of 5 points. EV:  
822 enlarged ventricles; CCA/H: corpus callosum agenesis or hypoplasia, CA: cerebral atrophy,  
823 CeA: cerebellar atrophy (B) Clinical scores calculated for individuals carrying pathogenic  
824 variants in *DYRK1A* reported here and for whom photographs were available (n=21)(initial  
825 cohort, *DYRK1A\_I*, scores 13- 17.5 with a mean of 15.5), the previously published  
826 individuals (replication cohort, *DYRK1A\_R*, scores 13.5-18, mean=15.3) and the individuals  
827 affected with other frequent monogenic forms of ID, associated to mutations in *ANKRD11*,  
828 *MED13L*, *DDX3X*, *ARID1B*, *SHANK3*, *TCF4* or *KMT2A* (scores 3-12.5, mean=7). The  
829 clinical score for the individuals carrying missense or distal frameshift variants are indicated  
830 in yellow (test); The threshold of  $CS_{DYRK1A} \geq 13$  appeared to be discriminant between  
831 individuals with LoF variants in *DYRK1A* (all  $\geq 13$ ) and individuals suffering from another  
832 form of ID (all  $< 13$ ) and a score above this threshold was therefore considered “highly  
833 suggestive”. We classified individuals with  $CS_{DYRK1A} < 10$  as “poorly evokative” and  
834 individuals with a  $CS_{DYRK1A}$  comprised between 10 and 13 as “intermediate”. Brown-  
835 Forsythe and Welsh ANOVA tests with Dunnett’s T 3 multiple comparisons test were  
836 performed. ns: not significant; \*\*  $p < 0.01$ ; \*\*\* $p < 0.001$ , error bars represent SD.

837 **Figure 2. Expression, localization and Tyr321 phosphorylation of DYRK1A mutant**  
838 **proteins**

839 (A) Level of variant DYRK1A proteins expressed in HeLa, HEK293 and COS cells  
840 transiently transfected with DYRK1A constructs. Protein levels were normalized on the level  
841 of GFP proteins (expressed from a cotransfected pEGFP plasmid). Quantifications were  
842 performed on a total of  $n \geq 9$  series of cells ( $n \geq 3$  HeLa cells,  $n \geq 3$  HEK293 and  $n \geq 3$   
843 COS cells). One-way ANOVA with multiple comparison test was performed to compare the  
844 level of variant DYRK1A proteins to the level of wild-type DYRK1A protein, applying  
845 Bonferroni’s correction: ns: not significant; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; error bars  
846 represent SEM, standard error of the mean, (B) DYRK1A’s ability to autophosphorylate on  
847 Tyr321 was tested in HEK293 cells (n=3) by immunoprecipitations with anti-DYRK1A  
848 followed by an immunoblot using an anti-HIPK2 as described in Widowati et al. DYRK1A  
849 phospho-Tyr321 levels were normalized with DYRK1A total level. Variant DYRK1A  
850 phospho-Tyr321 levels were normalized with total DYRK1A protein levels and expressed as  
851 percentage of wild-type level. One-way ANOVA test was performed to compare variants to  
852 wild-type DYRK1A levels. ns: not significant; \*\*\* $p < 0.001$ ; error bars represent SEM,  
853 standard error of the mean (C) Immunofluorescence experiment showing that Ser660fs (alias

854 Ser660Profs\*43) variant leads to DYRK1A protein aggregation when overexpressed in HeLa  
855 cells, using a FLAG-tagged DYRK1A proteins carrying Ser660Profs43. No aggregation was  
856 observed for the Ser660\* variant.

857

858 **Figure 3. DNA methylation signature of DYRK1A loss-of-function functionally classifies**  
859 ***DYRK1A* VUS.**

860 (A) Heatmap showing the hierarchical clustering of discovery *DYRK1A* LoF cases (n=10)  
861 and age- and sex-matched neurotypical discovery controls (n=24) used to identify the 402  
862 differentially methylated signature sites shown. The color gradient represents the normalized  
863 DNA methylation value from -2.0 (blue) to 2.0 (yellow) at each site. DNA methylation at  
864 these sites clearly separate discovery cases (grey) from discovery controls (blue). Euclidian  
865 distance metric is used for the clustering dendrogram. (B) Principal components analysis  
866 (PCA) visualizing the DNAm profiles of the study cohort at the 402 signature sites.  
867 Validation of *DYRK1A* LoF cases (not used to define the signature sites; red) cluster with  
868 discovery cases, while missense (yellow) and distal LoF (green) variants cluster with either  
869 cases or controls. Ind #33 (Gly486Asp) has an opposite DNAm profile to *DYRK1A* LoF cases  
870 at these sites, suggesting a GoF. (C) Support vector machine (SVM) classification model  
871 based on the DNA methylation values in the discovery groups. Each sample is plotted based  
872 on its scoring by the model. All samples are clearly positive (>0.5) or negative (<0.5). All  
873 *DYRK1A* validation cases from our cohort (n=6) classified positively, all control validation  
874 cases (n=94) classified negatively. Missense variants classified clearly positively or  
875 negatively, the distal frameshift variant (Ind #18, c.1978del), analyzed in duplicate, classified  
876 positively. Pathogenic *ARID1B* (Coffin-Siris syndrome) and *KMT2A* (Wiedemann Steiner  
877 syndrome) also classified negatively.

878

879 **Figure 4. Summary of the analysis performed to reclassify variants in *DYRK1A***

880 Representation of the DYRK1A protein (the kinase domain is indicated in red and the  
881 catalytic domain in dark red) with the positions of the different variants tested with the sample  
882 #ID of the individuals indicated inside the circles. Number: number of individuals with ID  
883 reported with the variant; gAD: variant reported in individuals from gnomAD; CS<sub>DYRK1A</sub>  
884 poorly (white), intermediate (grey) or highly (black) evocative, or unknown (-); CADD below

885 25 (white), between 25 and 30 (grey) or above 30 (black); conservation: highly conserved  
886 V=100%, M>90%, O>80% (black), moderately V=100%, M>90%, O<80% (grey) or mildly  
887 V=100% M<90%, O<80% (white); Expression or autophosphorylation being normal (white),  
888 intermediate decreased (grey), strongly decreased (black); Localization was normal (white),  
889 affected (grey) or not tested (-); DNAmethylation positive (black), negative (white),  
890 suggestive of a GoF effect (hashed) or not tested (-). Final classification: Pathogenic (P),  
891 Benign (B), Unknown significance (U).

892

893 **Table 1. List of variants identified in *DYRK1A* in individuals with intellectual disability**

894 del: deletion of the gene ; Trans.: translocation interrupting the gene ; Ns: nonsense ; Fs:  
895 frameshift, Spl: splice ; Ms: missense variants; TES: targeted exome sequencing of ID genes  
896 TES<sup>1</sup> : panel of ID genes from Carion et al. <sup>47</sup>; TES<sup>2</sup> : panel of ID gene adapted from Redin et  
897 al. <sup>46</sup>; TES<sup>3</sup> : panel of 44 ID genes; TES<sup>4</sup>: panel of microcephaly genes from Nasser et al. <sup>48</sup> ;  
898 CES: clinical exome sequencing, ES: exome sequencing, Sanger: Sanger sequencing, CGH-  
899 array: comparative genomic hybridization-array. <sup>a</sup>: the consequences of c.1098G>T is  
900 p.Ile318\_Glu366del instead of p.Glu366Asp ; M : male ; F : female ; DYRK1A\_I : initial  
901 cohort used to establish DYRK1A clinical score (CS<sub>DYRK1A</sub>) ; DYRK1A\_R : replication  
902 cohort used to confirm the relevance of the CS<sub>DYRK1A</sub> ; Disc : discovery cohort used to  
903 establish DNA methylation signature (DNAm) ; Valid: validation cohort used to confirm  
904 DNA methylation signature (DNAm); test : variants tested for pathogenicity using DNAm.

905

906

907

908

909

910

911

Variant					Individual			Reporting		Analyses performed				
GRCh37 (Chr21)	NM_001396.4	NP_001387.2	Type	Method	Ind	Sex	Inherit.	This individual (ClinVar)	Additional ind. (ClinVar)	CS <sub>DYRK1A</sub>	<i>In silico</i>	mRNA	<i>in vitro</i>	DNAm
g:38481804_40190458del (DYRK1A ; >10 other genes)	NA	NA	del.	CGH-array	Ind #1	M	de novo	this report	NA	DYRK1A_↓	-	yes	-	Disc.
g:38722881_39426450del (DYRK1A; KCNJ2)	NA	NA	del.	CGH-array	Ind #2	F	de novo	this report	NA	DYRK1A_↓	-	-	-	-
g:38302140_40041414del (DYRK1A ; >5 other genes)	NA	NA	del.	CGH-array	Ind #40	M	de novo	this report	NA	DYRK1A_↓	-	-	-	-
t(9;21)(p12;q22)	Between exon 2 & 3	NA	transl.	CGH-array	Ind #3	M	de novo	this report	NA	DYRK1A_↓	-	-	-	-
g:38852961C>T	c.349C>T	p.Arg117*	Ns	ES	Ind #4	M	NA	this report	(5x) VCV000373087	DYRK1A_↓	-	-	-	-
				ES	Ind #5	F	father mosaic	this report		DYRK1A_↓	-	-	-	-
				ES	Ind #6	M	de novo	this report		DYRK1A_↓	-	-	-	Disc.
g:38862575C>T	c.763C>T	p.Arg255*	Ns	TES <sup>1</sup>	Ind #7	M	de novo	this report	(5x) VCV000162152	DYRK1A_↓	-	-	-	Valid.
g:38862611C>T	c.799C>T	p.Gln267*	Ns	ES	Ind# 35	M	de novo	this report	-	DYRK1A_↓	-	-	-	-
g:38862748T>A	c.936T>A	p.Cys312*	Ns	CES	Ind #37	F	de novo	this report	-	DYRK1A_↓	-	-	-	-
g:38877655C>T	c.1309C>T	p.Arg437*	Ns	TES <sup>4</sup>	Ind #34	F	NA	this report	(7x) VCV000162158	DYRK1A_↓	-	-	-	-
g:38877745C>T	c.1399C>T	p.Arg467*	Ns	TES <sup>4</sup>	Ind#42	F	de novo	this report	(3x) VCV000204005	DYRK1A_↓	-	-	-	-
g:38850510dup	c.235dup	p.Arg79fs	Fs	TES <sup>4</sup>	Ind #41	F	de novo	this report (=SCV001432470)	-	DYRK1A_↓	-	-	-	-
g:38850565_38850566del	c.290_291 del	p.Ser97fs	Fs	CES	Ind #8	M	de novo	this report	(2x) VCV000418949	DYRK1A_↓	-	-	-	-
g:38850572_38850576del	c.297_301del	p.Leu100fs	Fs	NA	Ind #9	M	de novo	this report (=SCV000485020)	-	DYRK1A_↓	-	-	-	-
g:38853089del	c.477del	p.Tyr159*	Fs	ES	Ind #10	F	de novo	this report	-	DYRK1A_↓	-	-	-	-
g:38862514_38862515del	c.702_703 del	p.Cys235fs	Fs	ES	Ind #11	M	de novo	this report	-	DYRK1A_↓	-	yes	-	Valid.
g:38862594del	c.782del	p.Leu261fs	Fs	TES <sup>2</sup>	Ind #12	F	de novo	this report	-	DYRK1A_↓	-	-	-	Disc.
g:38865371del	c.1004del	p.Gly335fs	Fs	TES <sup>1</sup>	Ind #13	F	absent in mother	this report	-	DYRK1A_↓	-	-	-	Valid.
g:38865375dup	c.1008dup	p.Pro337fs	Fs	ES	Ind #14	F	de novo	this report	-	DYRK1A_↓	-	-	-	-
g:38865400del	c.1033del	p.Trp345fs	Fs	CES	Ind #15	F	de novo	this report	-	DYRK1A_↓	-	-	-	Disc.
g:38877616del	c.1270del	p.His424fs	Fs	ES	Ind #39	F	de novo	this report	-	DYRK1A_↓	-	-	-	-
g:38877679dup	c.1333dup	p.Thr445fs	Fs	ES	Ind #16	F	de novo	this report (=SCV000778256.1)	-	DYRK1A_↓	-	-	-	Disc.
g:38877837del	c.1491delC	p.Ala498fs	Fs	ES	Ind #17	F	de novo	this report	-	DYRK1A_↓	-	-	-	Valid.
g:38884520del	c.1978del	p.Ser660fs= p.Ser660Profs*43	Fs	TES <sup>1</sup>	Ind #18	F	de novo	this report	-	test	-	yes	yes	test
g:38852939G>T	c.328-1G>T	p.?	Spl.	TES <sup>2</sup>	Ind #19	F	de novo	this report	-	DYRK1A_↓	-	yes	-	Disc.
g:38862475A>G	c.665-2A>G	p.?	Spl.	Sanger	Ind #20	M	de novo	this report	SCV000492145	DYRK1A_↓	-	-	-	Valid.
g:38862468_38862472del	c.665-9_665-5del	p.?	Spl.	Sanger	Ind #21	M	de novo	this report	SCV000677027	DYRK1A_↓	-	-	-	Disc.
			Spl.	TES <sup>2</sup>	Ind #36	M	de novo	this report		DYRK1A_↓	-	-	-	-



g.38862764G>C	951+1G>C	p.?	Spl.	TES <sup>1</sup>	Ind #38	M	de novo	this report	-	DYRK1A_I	-	-	-	Valid.
g.38862767_38862770del	c.951+4_951+7del	p.?	Spl.	ES	Ind #22	M	de novo	this report (=SCV000435011.5)	SCV000709803	DYRK1A_I	-	yes	-	-
g.38877584A>G	c.1240-2A>G	p.?	Spl.	ES	Ind #23	M	de novo	this report (=SCV000966166)	-	DYRK1A_I	-	-	-	Disc.
g.38877585_38877586insTAA	c.1240-1_1240insTAA	p.?	Spl.	TES <sup>4</sup>	Ind #24	F	de novo	this report	-	DYRK1A_I	-	yes	-	-
g.38853115G>A	c.503G>A	p.Gly168A sp	Mis.	TES <sup>1</sup>	Ind #25	F	de novo	this report	SCV000573105	test	yes	-	yes	test
g.38862576G>A	c.764G>A	p.Arg255Gln	Mis.	TES <sup>1</sup>	Ind #26	F	NA	this report	-	test	yes	-	yes	test
g.38862672A>T	c.860A>T	p.Asp287Val	Mis.	ES	Ind #27	M	de novo	this report (=SCV000598121)	SCV001446739	test	yes	-	yes	test
g.38862726T>G	c.914T>G	p.Ile305Arg	Mis.	Sanger	Ind #28	F	de novo	this report	-	test	yes	-	yes	test
g.38865339T>A	c.972T>A	p.Ser324Arg	Mis.	TES <sup>3</sup>	Ind #29	M	de novo	this report (=SCV000902439)	-	test	yes	-	yes	test
g.38865465G>T	c.1098G>T <sup>3</sup>	p.Glu366A sp <sup>3</sup>	Mis.	ES	Ind #30	F	de novo	this report	-	test	yes	yes	yes	-
g.38877730T>C	c.1384T>C	p.Tyr462His	Mis.	TES <sup>2</sup>	Ind #31	M	de novo	this report (=SCV001437769)	-	test	yes	-	yes	test
g.3887746G>A	c.1400G>A	p.Arg467Gln	Mis.	TES <sup>2</sup>	Ind #32	F	de novo	this report (=SCV001437771)	(3x) VVCV000209150	test	yes	-	yes	test
g.38877803G>A	c.1457G>A	p.Gly486A sp	Mis.	ES	Ind #33	M	de novo	this report (SCV000747759)	-	test	yes	-	yes	test

Variants previously reported :

g.38858865C>T	c.613C>T	p.Arg205*	NS	TES <sup>2</sup>	Bronicki_#2	M	de novo	PMID: 25920557 (SCV000281731=SCV000196058)	(x7) VVCV000162153	DYRK1A_R	-	-	-	-
g.38862656dup	c.844dup	p.Ser282fs	FS	TES <sup>1</sup>	Bronicki_#8	M	De novo	PMID: 25920557 (SCV000196064)	-	DYRK1A_R	-	-	-	-
g.38858873_38858876delinsGAA	c.621_624 delinsGAA	p.Glu208fs	FS	TES <sup>2</sup>	Bronicki_#3	M	de novo	PMID: 25920557 (SCV000281736 = SCV000196059.1)	-	DYRK1A_R	-	-	-	Disc.
g.38868553dup	c.1232dup	p.Arg413fs	FS	TES <sup>2</sup>	Bronicki_#10	F	de novo	PMID: 25920557 (SCV000196066)	-	DYRK1A_R	-	yes	yes	Disc.
g.38862744C>T	c.932C>T	p.Ser311Phe	Mis	NA	Ruud_#2	M	de novo	PMID: 25641759 (SCV000586742)	SCV000520979	test	yes	-	yes	test
g.38884305C>A	c.1763C>A	p.Thr588Asn	Mis	ES	Bronicki_#9	F	de novo	PMID: 25920557 (SCV000965705.1=SCV000196065.1)	-	test	yes	yes	yes	test

**Table 1. List of variants identified in *DYRK1A* in individuals with intellectual disability**

A

Frequent features (% ind. this report - prev. reported)	Points (/20)
Intellectual disability (100% - 100%) <i>(moderate:1; severe:2)</i>	2
Impaired language (100% - 100%) <i>(yes:1; absence or few words:2)</i>	2
Microcephaly (100% - 94%) <i>(yes:1; severe &lt;-3SD:2)</i>	2
Feeding difficulty (neonate or infancy)(93% - 88%)	2
History of seizures (90% - 74%) <i>(yes:1; febrile seizures :2)</i>	2
Autistic traits/Anxiety (93% - 89%) <i>(Anxiety:1; Stereotypies:1; ASD diagnosis:2)</i>	2
Brain anomaly on MRI (83% - 75%) <i>(any:0,5; if EV, CCA/H, CA, CeA:1)</i>	1
Motor delay (83% - 88%)	1
Ataxic walk/hypertonia (62% - 73%)	0.5
Atopic or translucide skin (69% - )	0.5
Dysmorphy:	
<i>Thin hair or erratic hairline</i>	0.5
<i>Upper eyelid edema</i>	1
<i>Deep set eyes</i>	1
<i>Protruding nose or pointed nasal tip</i>	0.5
<i>Dysplastic ears or ear lobe attached</i>	0.5
<i>Thin upper lip</i>	0.5
<i>Widely spaced teeth, protruding upper dental arch</i>	0.5
<i>Retrognathism, smal chin</i>	0.5

B

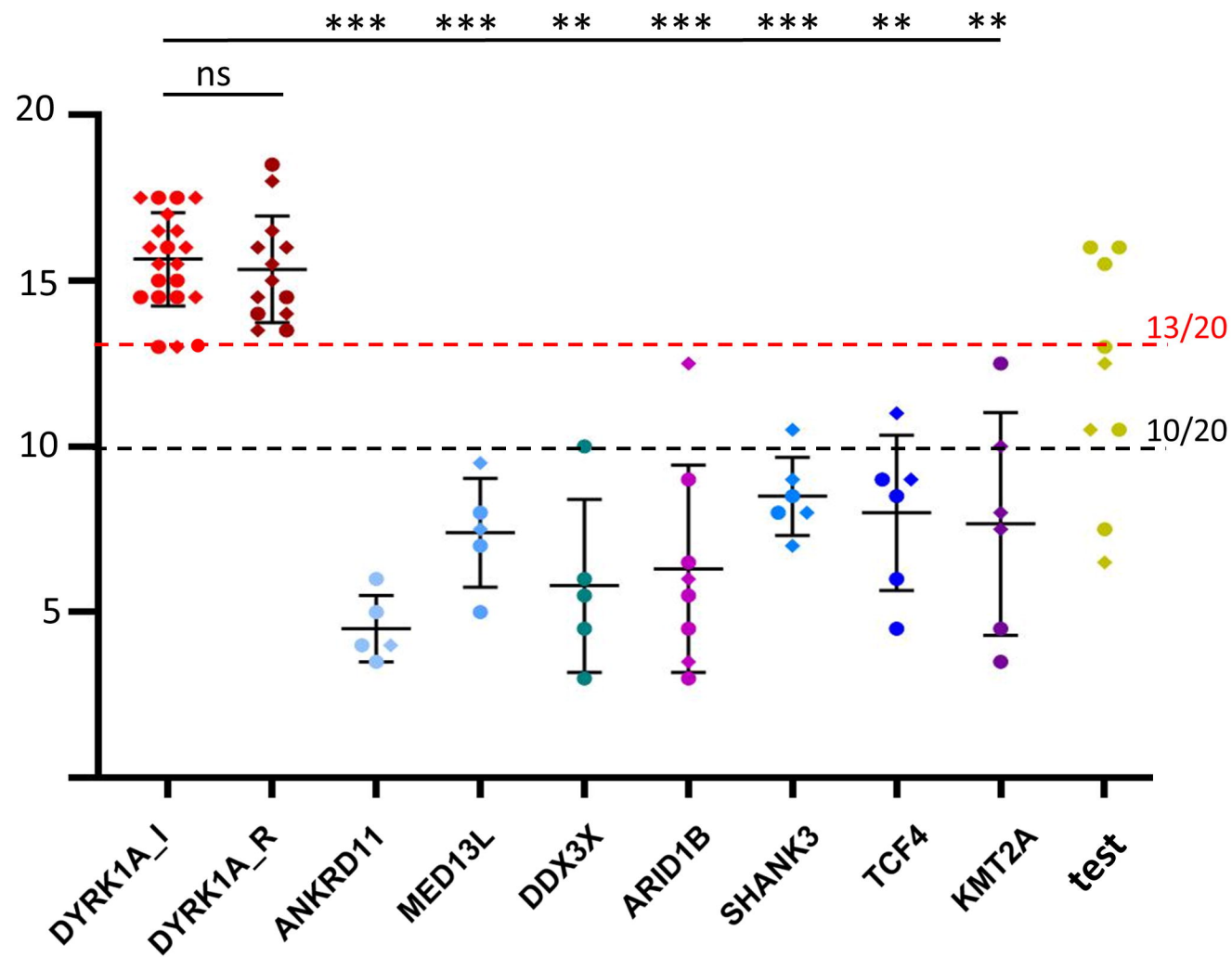


Figure 1



