nature genetics

Technical Report

https://doi.org/10.1038/s41588-023-01469-w

PhenoScore quantifies phenotypic variation for rare genetic diseases by combining facial analysis with other clinical features using a machine-learning framework

In the format provided by the authors and unedited

Supplementary Information

Contents

1 Supplementary Text **2**

1.1 Permutation test for hypothesis testing 2

1.2 Excluding the presence of a systematic confounder effect 2

Supplemental References **4**

1 Supplementary Text

1.1 Permutation test for hypothesis testing

To provide a p-value for our classification results and to enable the use of our framework for the recognition of specific (sub)groups in genetic syndromes, we developed a permutation test — inspired by the test described by Lopez-Paz & Oquab [1]. For every group of individuals of interest, a group of age-, sex- and ethnicity-matched controls is sampled from our control database. We extract the facial features using VGGFace2 and calculate the HPO similarity using the Resnik score, after which a SVM is trained using cross-validation. A grid search for the optimal hyperparameters is performed, and the Brier score is calculated for this combination of the group of individuals of interest and the matched controls. This process is precisely the same as in our standard analysis. Next, we randomly permute the labels (here, the labels correspond to whether an individual has the syndrome or is a control) 100 times. We ensure that the number of positive and negative classes is the same as in our original distribution of the labels. For each permutation, we repeat the process of training a SVM to obtain a Brier score. We then perform a one-sided Mann-Whitney U test to quantify the probability of the classification results being statistically significantly smaller (since it is the Brier score we are comparing) than the randomly permuted scores.

To further strengthen our permutation test, we repeat the process five times in total, randomly sampling matched controls from our database in each repetition. The five obtained p values were then combined using Fisher's method [2] to gather a definitive p-value for this classification task and, therefore, for this specific group of individuals of interest.

For the subgroup analyses for *SATB1*, *DEAF1*, *SETBP1* and *ADNP*, we do not need to sample controls from our database. However, these datasets are usually imbalanced, sometimes leading to problems for the classifier. We therefore undersample the majority class to the size of the minority class by matching the individuals on sex, ethnicity and age (in that order if possible) and increase the number of permutations of the labels to 1 000 (since we cannot repeatedly sample controls).

Finally, to have a negative control group for this test, we randomly sampled individuals from our control database and calculated p-values for those. We did this for different cohort sizes (n = 3, 5, 10, 20 and 40) for in total 50 trials. Of those 50 trials, two resulted in a p-value smaller than 0.05 - exactly what would be expected by random chance in this number of trials. This shows that our approach leads to the to-be-expected number of false alarms.

1.2 Excluding the presence of a systematic confounder effect

To investigate whether there is a systematic bias in our analysis because of the different origins of the data, since all control individuals were seen at our outpatient clinic at the Radboud university

medical center (RUMC), we performed several analyses.

First, we generated plots using Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP, [3]) of the VGGFace2 feature vectors for all included syndromes (Supplementary Fig. 3). There is no clear distinction between the patients and the controls which could be indicative of a systematic confounder. However, these UMAP plots provide only limited information: indeed, we hope to see some separation between the classes, and while these plots are useful to exclude obvious (large) biases, they do not explain where the difference comes from (i.e. it could be the distinction in syndromes we are looking for, or it could be a bias). Therefore, we performed other analyses investigating this issue as well.

Second, we looked at the largest dataset in this study, the individuals with Koolen-de Vries syndrome (KdVS, caused by pathogenic variants in or a deletion of *KANSL1*). Of the 63 individuals with KdVS, 18 were seen at our outpatient clinic — in the same rooms as the control individuals. We excluded those 18 individuals for this confounder analysis and retrained PhenoScore on the remaining 45, again sampling control individuals for five iterations and averaging the results during cross-validation. The overall performance metrics of this model were comparable to the model when using all KdVS individuals (Brier 0.0917, AUC 0.93, p <0.001).

Then, we classified the 18 individuals (seen at the same institution as the control individuals) using this model and generated prediction probabilities when using HPO data, facial data and the combination of both (a PhenoScore prediction probability), as if these were VUSs. If there was a systematic confounder that would influence the results of PhenoScore not based on the underlying genetic syndrome, but the origin of the data, these individuals should be classified as control individuals. However, this is clearly not the case (Supplementary Table 7), with the median of the PhenoScore prediction being 0.95 and 13 out of 18 individuals correctly classified as having KdVS. One individual is missed and labeled as control, mainly because the phenotype is not matching, while the facial prediction is high (0.86).

Finally, we matched the 18 individuals with KdVS seen at the RUMC with 18 individuals with KdVS seen at other centers on age-, sex- and ethnicity. PhenoScore was then utilized to investigate whether these are two recognizable subgroups, to see if we are able to predict the origin of the data based on the phenotype, since all individuals have the same genetic disorder (KdVS) in this analysis. PhenoScore was not able to detect a phenotypic difference (Brier 0.322, AUC 0.54, p=0.452), supporting the conclusion that there is no systematic confounder in our analyses.

Supplemental references

[1] Lopez-Paz, D. & Oquab, M. Revisiting Classifier Two-Sample Tests. *arXiv* (2016).

[2] Fisher, R. A. *Statistical methods for research workers 12th edition* (Oliver & Body, 1954).

[3] McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv* (2018). 1802.03426.

[4] Rauch, A. *et al.* Systematic assessment of atypical deletions reveals genotype-phenotype correlation in 22q11.2. *J. Med. Genet.* **42**, 871–876 (2005).

[5] Brunet, A. *et al.* Microdeletion and microduplication 22q11.2 screening in 295 patients with clinical features of DiGeorge/Velocardiofacial syndrome. *Am. J. Med. Genet. A* **140**, 2426–2432 (2006).

[6] Shprintzen, R. J. Velo-cardio-facial syndrome: 30 Years of study. *Dev. Disabil. Res. Rev.* **14**, 3–10 (2008).

[7] Lipson, A. H. *et al.* Velocardiofacial (Shprintzen) syndrome: an important syndrome for the dysmorphologist to recognise. *J. Med. Genet.* **28**, 596–604 (1991).

[8] McDonald-McGinn, D. M. & Sullivan, K. E. Chromosome 22q11.2 deletion syndrome (DiGeorge syndrome/velocardiofacial syndrome). *Medicine* **90**, 1–18 (2011).

[9] Grassi, M. S. *et al.* Congenital Heart Disease as a Warning Sign for the Diagnosis of the 22q11.2 Deletion. *Arq. Bras. Cardiol.* **103**, 382–390 (2014).

[10] Meinecke, P., Beemer, F. A., Schinzel, A. & Kushnick, T. The velo-cardio-facial (Shprintzen) syndrome. Clinical variability in eight patients. *Eur. J. Pediatr.* **145**, 539–544 (1986).

[11] Bartsch, O. *et al.* DiGeorge/velocardiofacial syndrome: FISH studies of chromosomes 22q11 and 10p14, and clinical reports on the proximal 22q11 deletion. *Am. J. Med. Genet. A* **117A**, 1–5 (2003).

[12] Marom, R. *et al.* Heterozygous variants in ACTL6A, encoding a component of the BAF complex, are associated with intellectual disability. *Hum. Mutat.* **38**, 1365–1371 (2017).

[13] Alazami, A. M. *et al.* Mutation in ADAT3, encoding adenosine deaminase acting on transfer RNA, causes intellectual disability and strabismus. *J. Med. Genet.* **50**, 425–430 (2013).

[14] El-Hattab, A. W. *et al.* ADAT3-related intellectual disability: Further delineation of the phenotype. *Am. J. Med. Genet. A* **170A**, 1142–1147 (2016).

[15] Coe, B. P. *et al.* Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat. Genet.* **46**, 1063–1071 (2014).

[16] Krajewska-Walasek, M. *et al.* Additional data on the clinical phenotype of Helsmoortel-Van der Aa syndrome associated with a novel truncating mutation in ADNP gene. *Am. J. Med. Genet. A* **170**, 1647–1650 (2016).

[17] Van Dijck, A. *et al.* Clinical Presentation of a Complex Neurodevelopmental Disorder Caused by Mutations in ADNP. *Biol. Psychiatry* **85**, 287–297 (2019).

[18] Willemsen, M. H. *et al.* Identification of ANKRD11 and ZNF778 as candidate genes for autism and variable cognitive impairment in the novel 16q24.3 microdeletion syndrome. *Eur. J. Hum. Genet.* **18**, 429–435 (2010).

[19] Youngs, E. L., Hellings, J. A. & Butler, M. G. ANKRD11 gene deletion in a 17-year-old male. *Clin. Dysmorphol.* **20**, 170–171 (2011).

[20] Isrie, M. *et al.* Haploinsufficiency of ANKRD11 causes mild cognitive impairment, short stature and minor dysmorphisms. *Eur. J. Hum. Genet.* **20**, 131–133 (2012).

[21] Sacharow, S., Li, D., Fan, Y. S. & Tekin, M. Familial 16q24.3 microdeletion involving ANKRD11 causes a KBG-like syndrome. *Am. J. Med. Genet. A* **158A**, 547–552 (2012).

[22] Khalifa, M. *et al.* Partial deletion of ANKRD11 results in the KBG phenotype distinct from the 16q24.3 microdeletion syndrome. *Am. J. Med. Genet. A* **161A**, 835–840 (2013).

[23] Ockeloen, C. W. *et al.* Further delineation of the KBG syndrome caused by ANKRD11 aberrations. *Eur. J. Hum. Genet.* **23**, 1270 (2015).

[24] Santen, G. W. E. *et al.* Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. *Hum. Mutat.* **34**, 1519–1528 (2013).

[25] Tsurusaki, Y. *et al.* Coffin-Siris syndrome is a SWI/SNF complex disorder. *Clin. Genet.* **85**, 548–554 (2014).

[26] Miraldi Utz, V., Brightman, D. S., Sandoval, M. A., Hufnagel, R. B. & Saal, H. M. Systemic and ocular manifestations of a patient with mosaic ARID1A-associated Coffin-Siris syndrome and review of select mosaic conditions with ophthalmic manifestations. *Am. J. Med. Genet. C Semin. Med. Genet.* **184**, 644–655 (2020).

[27] van der Sluijs, P. J. *et al.* Discovering a new part of the phenotypic spectrum of Coffin-Siris syndrome in a fetal cohort. *Genet. Med.* **24**, 1753–1760 (2022).

[28] Nagamani, S. C. *et al.* Interstitial deletion of 6q25.2-q25.3: a novel microdeletion syndrome associated with microcephaly, developmental delay, dysmorphic features and hearing loss. *Eur. J. Hum. Genet.* **17**, 573–581 (2009).

[29] Hoyer, J. *et al.* Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. *Am. J. Hum. Genet.* **90**, 565–572 (2012).

[30] Santen, G. W. E. *et al.* Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat. Genet.* **44**, 379–380 (2012).

[31] Michelson, M. *et al.* Delineation of the interstitial 6q25 microdeletion syndrome: refinement of the critical causative region. *Am. J. Med. Genet. A* **158A**, 1395–1399 (2012).

[32] Wieczorek, D. *et al.* A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum. Mol. Genet.* **22**, 5121–5135 (2013).

[33] Vals, M. A. *et al.* Coffin-Siris Syndrome with obesity, macrocephaly, hepatomegaly and hyperinsulinism caused by a mutation in the ARID1B gene. *Eur. J. Hum. Genet.* **22**, 1327–1329 (2014).

[34] Ben-Salem, S. *et al.* Gonadal mosaicism in ARID1B gene causes intellectual disability and dysmorphic features in three siblings. *Am. J. Med. Genet. A* **170A**, 156–161 (2016).

[35] Ronzoni, L. *et al.* Interstitial 6q25 microdeletion syndrome: ARID1B is the key gene. *Am. J. Med. Genet. A* **170A**, 1257–1261 (2016).

[36] Gripp, K. W., Baker, L., Telegrafi, A. & Monaghan, K. G. The role of objective facial analysis using FDNA in making diagnoses following whole exome analysis. Report of two patients with mutations in the BAF complex genes. *Am. J. Med. Genet. A* **170**, 1754–1762 (2016).

[37] Zweier, M. *et al.* The HHID syndrome of hypertrichosis, hyperkeratosis, abnormal corpus callosum, intellectual disability, and minor anomalies is caused by mutations in ARID1B. *Am. J. Med. Genet. A* **173**, 1440–1443 (2017).

[38] Määttänen, L., Hietala, M., Ignatius, J. & Arvio, M. A 69-year-old woman with Coffin-Siris syndrome. *Am. J. Med. Genet. A* **176**, 1764–1767 (2018).

[39] van der Sluijs, P. J. *et al.* The ARID1B spectrum in 143 patients: from nonsyndromic intellectual disability to Coffin-Siris syndrome. *Genet. Med.* **21**, 1295–1307 (2019).

[40] Liu, X. *et al.* De Novo ARID1B mutations cause growth delay associated with aberrant Wnt/β-catenin signaling. *Hum. Mutat.* **41**, 1012–1024 (2020).

[41] Lian, S. *et al.* Coffin-Siris Syndrome-1: Report of five cases from Asian populations with truncating mutations in the ARID1B gene. *J. Neurol. Sci.* **414**, 116819 (2020).

[42] Palmer, E. E. *et al.* De Novo Variants Disrupting the HX Repeat Motif of ATN1 Cause a Recognizable Non-Progressive Neurocognitive Syndrome. *Am. J. Hum. Genet.* **104**, 542–552 (2019).

[43] Snijders Blok, L. *et al.* CHD3 helicase domain mutations cause a neurodevelopmental syndrome with macrocephaly and impaired speech and language. *Nat. Commun.* **9**, 4619 (2018).

[44] Drivas, T. G. *et al.* A second cohort of CHD3 patients expands the molecular mechanisms known to cause Snijders Blok-Campeau syndrome. *Eur. J. Hum. Genet.* **28**, 1422–1431 (2020).

[45] O'Roak, B. J. *et al.* Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* **338**, 1619–1622 (2012).

[46] Bernier, R. *et al.* Disruptive CHD8 mutations define a subtype of autism early in development. *Cell* **158**, 263–276 (2014).

[47] Douzgou, S. *et al.* The clinical presentation caused by truncating CHD8 variants. *Clin. Genet.* **96**, 72–84 (2019).

[48] Ostrowski, P. J. *et al.* The CHD8 overgrowth syndrome: A detailed evaluation of an emerging overgrowth phenotype in 27 patients. *Am. J. Med. Genet. C Semin. Med. Genet.* **181**, 557–564 (2019).

[49] Dingemans, A. J. M. *et al.* The phenotypic spectrum and genotype-phenotype correlations in 106 patients with variants in major autism gene CHD8. *Transl. Psychiatry* **12**, 421 (2022).

[50] Nabais Sá, M. J. *et al.* De novo CLTC variants are associated with a variable phenotype from mild to severe intellectual disability, microcephaly, hypoplasia of the corpus callosum, and epilepsy. *Genet. Med.* **22**, 797–802 (2020).

[51] Snijders Blok, L. *et al.* Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. *Am. J. Hum. Genet.* **97**, 343–352 (2015).

[52] Sharma, P., Gambhir, P. S., Phadke, S. R. & Mandal, K. Expanding the phenotype in autosomal dominant mental retardation-24: a novel variation in DEAF1 gene. *Clin. Dysmorphol.* **28**, 94–97 (2019).

[53] Andoni, T., Ellard, S., Kapadia, J. & Wakeling, E. A novel autosomal recessive DEAF1 nonsense variant: expanding the clinical phenotype. *Clin. Dysmorphol.* **29**, 114–117 (2020).

[54] Sumathipala, D. S. *et al.* A girl with a neurodevelopmental syndrome, adducted thumbs and frequent infections caused by novel homozygous variant in DEAF1. *Clin. Dysmorphol.* **29**, 107–110 (2020).

[55] Møller, R. S. *et al.* Truncation of the Down syndrome candidate gene DYRK1A in two unrelated patients with microcephaly. *Am. J. Hum. Genet.* **82**, 1165–1170 (2008).

[56] van Bon, B. W. *et al.* Intragenic deletion in DYRK1A leads to mental retardation and primary microcephaly. *Clin. Genet.* **79**, 296–299 (2011).

[57] Courcet, J. B. *et al.* The DYRK1A gene is a cause of syndromic intellectual disability with severe microcephaly and epilepsy. *J. Med. Genet.* **49**, 731–736 (2012).

[58] van Bon, B. W. *et al.* Disruptive de novo mutations of DYRK1A lead to a syndromic form of autism and ID. *Mol. Psychiatry* **21**, 126–132 (2016).

[59] Willemsen, M. H. *et al.* Update on Kleefstra Syndrome. *Mol. Syndromol.* **2**, 202–212 (2012).

[60] Lelieveld, S. H. *et al.* Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat. Neurosci.* **19**, 1194–1196 (2016).

[61] Jansen, S. *et al.* De Novo Truncating Mutations in the Last and Penultimate Exons of PPM1D Cause an Intellectual Disability Syndrome. *Am. J. Hum. Genet.* **100**, 650–658 (2017).

[62] Tran Mau-Them, F. *et al.* Expanding the phenotype of IQSEC2 mutations: truncating mutations in severe intellectual disability. *Eur. J. Hum. Genet.* **22**, 289–292 (2014).

[63] Radley, J. A. *et al.* Deep phenotyping of 14 new patients with IQSEC2 variants, including monozygotic twins of discordant phenotype. *Clin. Genet.* **95**, 496–506 (2019).

[64] Koolen, D. A. *et al.* A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat. Genet.* **38**, 999–1001 (2006).

[65] Hoyer, J. *et al.* Molecular karyotyping in patients with mental retardation using 100K single-nucleotide polymorphism arrays. *J. Med. Genet.* **44**, 629–636 (2007).

[66] Koolen, D. A. *et al.* Clinical and molecular delineation of the 17q21.31 microdeletion syndrome. *J. Med. Genet.* **45**, 710–720 (2008).

[67] Dubourg, C. *et al.* Clinical and molecular characterization of 17q21.31 microdeletion syndrome in 14 French patients with mental retardation. *Eur. J. Med. Genet.* **54**, 144–151 (2011).

[68] Diets, I. J. *et al.* De Novo and Inherited Pathogenic Variants in KDM3B Cause Intellectual Disability, Short Stature, and Facial Dysmorphism. *Am. J. Hum. Genet.* **104**, 758–766 (2019).

[69] Clayton-Smith, J. *et al.* Xq28 duplication presenting with intestinal and bladder dysfunction and a distinctive facial appearance. *Eur. J. Hum. Genet.* **17**, 434–443 (2009).

[70] Lugtenberg, D. *et al.* Structural variation in Xq28: MECP2 duplications in 1% of patients with unexplained XLMR and in 2% of male patients with severe encephalopathy. *Eur. J. Hum. Genet.* **17**, 444–453 (2009).

[71] Asadollahi, R. *et al.* Dosage changes of MED13L further delineate its role in congenital heart defects and intellectual disability. *Eur. J. Hum. Genet.* **21**, 1100–1104 (2013).

[72] van Haelst, M. M. *et al.* Further confirmation of the MED13L haploinsufficiency syndrome. *Eur. J. Hum. Genet.* **23**, 135–138 (2015).

[73] Cafiero, C. *et al.* Novel de novo heterozygous loss-of-function variants in MED13L and further delineation of the MED13L haploinsufficiency syndrome. *Eur. J. Hum. Genet.* **23**, 1499–1504 (2015).

[74] Asadollahi, R. *et al.* Genotype-phenotype evaluation of MED13L defects in the light of a novel truncating and a recurrent missense mutation. *Eur. J. Med. Genet.* **60**, 451–464 (2017).

[75] Smol, T. *et al.* MED13L-related intellectual disability: involvement of missense variants and delineation of the phenotype. *Neurogenetics* **19**, 93–103 (2018).

[76] Tørring, P. M. *et al.* Is MED13L-related intellectual disability a recognizable syndrome? *Eur. J. Med. Genet.* **62**, 129–136 (2019).

[77] Saunier, C. *et al.* Expanding the Phenotype Associated with NAA10-Related N-Terminal Acetylation Deficiency. *Hum. Mutat.* **37**, 755–764 (2016).

[78] Gupta, A. S., Saif, H. A., Lent, J. M. & Couser, N. L. Ocular Manifestations of the NAA10-Related Syndrome. *Case Rep. Genet.* **2019**, 8492965 (2019).

[79] Cheng, H. *et al.* Phenotypic and biochemical analysis of an international cohort of individuals with variants in NAA10 and NAA15. *Hum. Mol. Genet.* **28**, 2900–2919 (2019).

[80] Bader, I. *et al.* Severe syndromic ID and skewed X-inactivation in a girl with NAA10 dysfunction and a novel heterozygous de novo NAA10 p.(His16Pro) variant - a case report. *BMC Med. Genet.* **21**, 153 (2020).

[81] Gogoll, L. *et al.* Confirmation of Ogden syndrome as an X-linked recessive fatal disorder due to a recurrent NAA10 variant and review of the literature. *Am. J. Med. Genet. A* **185**, 2546–2560 (2021).

[82] Maini, I. *et al.* Clinical Manifestations in a Girl with NAA10-Related Syndrome and Genotype-Phenotype Correlation in Females. *Genes* **12** (2021).

[83] McTiernan, N. *et al.* Biochemical analysis of novel NAA10 variants suggests distinct pathogenic mechanisms involving impaired protein N-terminal acetylation. *Hum. Genet.* **141**, 1355–1369 (2022).

[84] Sandomirsky, K., Marchi, E., Gavin, M., Amble, K. & Lyon, G. J. Phenotypic variability and gastrointestinal manifestations/interventions for growth in NAA10-related neurodevelopmental syndrome. *Am. J. Med. Genet. A* **191**, 1293–1300 (2023).

[85] Homsy, J. *et al.* De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science* **350**, 1262–1266 (2015).

[86] Cheng, H. *et al.* Truncating Variants in NAA15 Are Associated with Variable Levels of Intellectual Disability, Autism Spectrum Disorder, and Congenital Anomalies. *Am. J. Hum. Genet.* **102**, 985–994 (2018).

[87] Stessman, H. A. F. *et al.* Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nat. Genet.* **49**, 515–526 (2017).

[88] Schuurs-Hoeijmakers, J. H. *et al.* Clinical delineation of the PACS1-related syndrome–Report on 19 patients. *Am. J. Med. Genet. A* **170**, 670–675 (2016).

[89] Jansen, S. *et al.* A genotype-first approach identifies an intellectual disability-overweight syndrome caused by PHIP haploinsufficiency. *Eur. J. Hum. Genet.* **26**, 54–63 (2018).

[90] Tanaka, A. J. *et al.* De novo mutations in PURA are associated with hypotonia and developmental delay. *Cold Spring Harb Mol Case Stud* **1**, a000356 (2015).

[91] Reijnders, M. R. F. *et al.* PURA syndrome: clinical delineation and genotype-phenotype study in 32 individuals with review of published literature. *J. Med. Genet.* **55**, 104–113 (2018).

[92] Lee, B. H. *et al.* Expanding the neurodevelopmental phenotype of PURA syndrome. *Am. J. Med. Genet. A* **176**, 56–67 (2018).

[93] Al-Mudaffer, M. *et al.* Clinical and radiological findings in Schinzel-Giedion syndrome. *Eur. J. Pediatr.* **167**, 1399–1407 (2008).

[94] Acuna-Hidalgo, R. *et al.* Overlapping SETBP1 gain-of-function mutations in Schinzel-Giedion syndrome and hematologic malignancies. *PLoS Genet.* **13**, e1006683 (2017).

[95] Jansen, N. A. *et al.* Clinical delineation of SETBP1 haploinsufficiency disorder. *Eur. J. Hum. Genet.* **29**, 1198–1205 (2021).

[96] Machol, K. *et al.* Expanding the Spectrum of BAF-Related Disorders: De Novo Variants in SMARCC2 Cause a Syndrome with Intellectual Disability and Developmental Delay. *Am. J. Hum. Genet.* **104**, 164–178 (2019).

[97] Takenouchi, T., Miura, K., Uehara, T., Mizuno, S. & Kosaki, K. Establishing SON in 21q22.11 as a cause a new syndromic form of intellectual disability: Possible contribution to Braddock-Carey syndrome phenotype. *Am. J. Med. Genet. A* **170**, 2587–2590 (2016).

[98] Kim, J. H. *et al.* De Novo Mutations in SON Disrupt RNA Splicing of Genes Essential for Brain Development and Metabolism, Causing an Intellectual-Disability Syndrome. *Am. J. Hum. Genet.* **99**, 711–719 (2016).

[99] Dingemans, A. J. M. *et al.* Establishing the phenotypic spectrum of ZTTK syndrome by analysis of 52 individuals with variants in SON. *Eur. J. Hum. Genet.* **30**, 271–281 (2022).

[100] Kumar, R. *et al.* Severe neurocognitive and growth disorders due to variation in THOC2, an essential component of nuclear mRNA export machinery. *Hum. Mutat.* **39**, 1126–1138 (2018).

[101] Kumar, R. *et al.* Expanding Clinical Presentations Due to Variations in THOC2 mRNA Nuclear Export Factor. *Front. Mol. Neurosci.* **13**, 12 (2020).

[102] Ba, W. *et al.* TRIO loss of function is associated with mild intellectual disability and affects dendritic branching and synapse function. *Hum. Mol. Genet.* **25**, 892–902 (2016).

[103] Pengelly, R. J. *et al.* Mutations specific to the Rac-GEF domain of TRIO cause intellectual disability and microcephaly. *J. Med. Genet.* **53**, 735–742 (2016).

[104] Cogné, B. *et al.* Missense Variants in the Histone Acetyltransferase Complex Component Gene TRRAP Cause Autism and Syndromic Intellectual Disability. *Am. J. Hum. Genet.* **104**, 530–541 (2019).

[105] DeSanto, C. *et al.* WAC loss-of-function mutations cause a recognisable syndrome characterised by dysmorphic features, developmental delay and hypotonia and recapitulate 10p11.23 microdeletion syndrome. *J. Med. Genet.* **52**, 754–761 (2015).

[106] Lugtenberg, D. *et al.* De novo loss-of-function mutations in WAC cause a recognizable intellectual disability syndrome and learning deficits in Drosophila. *Eur. J. Hum. Genet.* **24**, 1145–1153 (2016).

[107] Gabriele, M. *et al.* YY1 Haploinsufficiency Causes an Intellectual Disability Syndrome Featuring Transcriptional and Chromatin Dysfunction. *Am. J. Hum. Genet.* **100**, 907–925 (2017).

[108] Palmer, E. E. *et al.* A Recurrent De Novo Nonsense Variant in ZSWIM6 Results in Severe Intellectual Disability without Frontonasal or Limb Malformations. *Am. J. Hum. Genet.* **101**, 995–1005 (2017).