The American Journal of Human Genetics, Volume *110*

Supplemental information

Deleterious, protein-altering variants in the transcriptional coregulator *ZMYM3* in 27 individuals with a neurodevelopmental delay phenotype

Susan M. Hiatt, Slavica Trajkova, Matteo Rossi Sebastiano, E. Christopher Partridge, Fatima E. Abidi, Ashlyn Anderson, Muhammad Ansar, Stylianos E. Antonarakis, Azadeh Azadi, Ruxandra Bachmann-Gagescu, Andrea Bartuli, Caroline Benech, Jennifer L. Berkowitz, Michael J. Betti, Alfredo Brusco, Ashley Cannon, Giulia Caron, Yanmin Chen, Meagan E. Cochran, Tanner F. Coleman, Molly M. Crenshaw, Laurence Cuisset, Cynthia J. Curry, Hossein Darvish, Serwet Demirdas, Maria Descartes, Jessica Douglas, David A. Dyment, Houda Zghal Elloumi, Giuseppe Ermondi, Marie Faoucher, Emily G. Farrow, Stephanie A. Felker, Heather Fisher, Anna C.E. Hurst, Pascal Joset, Melissa A. Kelly, Stanislav Kmoch, Benjamin R. Leadem, Michael J. Lyons, Marina Macchiaiolo, Martin Magner, Giorgia Mandrile, Francesca Mattioli, Megan McEown, Sarah K. Meadows, Livija Medne, Naomi J.L. Meeks, Sarah Montgomery, Melanie P. Napier, Marvin Natowicz, Kimberly M. Newberry, Marcello Niceta, Lenka Noskova, Catherine B. Nowak, Amanda G. Noyes, Matthew Osmond, Eloise J. Prijoles, Jada Pugh, Verdiana Pullano, Chloé Quélin, Simin Rahimi-Aliabadi, Anita Rauch, Sylvia Redon, Alexandre Reymond, Caitlin R. Schwager, Elizabeth A. Sellars, Angela E. Scheuerle, Elena Shukarova-Angelovska, Cara Skraban, Elliot Stolerman, Bonnie R. Sullivan, Marco Tartaglia, Isabelle Thiffault, Kevin Uguen, Luis A. Umaña, Yolande van Bever, Saskia N. van der Crabben, Marjon A. van Slegtenhorst, Quinten Waisfisz, Camerun Washington, Lance H. Rodan, Richard M. Myers, and Gregory M. Cooper

Supplemental Note: Case Reports

Individual 1, p.Asp69Asn

Patient was initially referred to the Genetics clinic at 11 months of age for evaluation due to a history of hypoplastic right ventricle following a bidirectional Glenn procedure. He also had a history of hypocalcemia, which had resolved, but consideration for 22q11.2 deletion syndrome was entertained, and a chromosome microarray was obtained which was normal. He was noted at that time to have developmental delays, specifically of gross motor skills as he was unable to crawl, and axial hypotonia. He had no other major findings, and his exam was otherwise mostly normal. Family history was notable for the maternal grandmother and maternal uncle with significant delays in speech acquisition in childhood, although they had no other delays and no current intellectual disabilities. He had a healthy younger full sister and three paternal half-sisters. At the follow-up visit at 19 months of age, he had made little to no developmental progress. Interim brain MRI had revealed abnormalities of the corpus callosum. There was no history of any seizures or loss of previously acquired skills. His developmental delays and findings on MRI were felt to be out of proportion with his history of the complex heart defect and excellent postsurgical progress. Thus, further expanded gene testing was ordered that revealed a maternally inherited *ZMYM3* variant. No additional variants were identified. At three years old, he continues to have gross motors delays and can pull to stand. He is also now noted to have significant speech delay, using babbling and vocalizations but no purposeful words.

Individual 2, p.Glu241Lys

Individual 2 is a 14 year old male with learning difficulties, imbalance, excessive fatigability, heat intolerance, oculomotor dysfunction, and history of growth retardation secondary to insulin-like growth factor 1 (IGF1) deficiency.

This male was born at 35 5/7 weeks gestation by spontaneous vertex vaginal delivery with birth weight 2.61 kg and length 45.7 cm to a G2P1 29 y/o mother and 38 y/o father. The prenatal history was notable for short femurs and humeri and a pericardial effusion. His mother had type I diabetes and hypertension during the pregnancy; the diabetes was under good metabolic control and the hypertension treated with several anti-hypertensive agents. At 20 days of age he was noted to have asymptomatic bilateral pulmonary artery stenoses and a small PFO. Torticollis was noted at 2 months of age. Since early childhood, his clinical phenotype has included dysmorphisms, developmental and neurological issues, and endocrine/growth issues.

A dysmorphology evaluation at 22 months showed head circumference 49 cm (68%; Z = +0.46), weight 8.7 kg (<1%; Z = -3.06) and height 80.3 cm (7%; Z = -1.49). Other notable measurements then included inner canthal distance 2.4 cm (3%), interpupillary distance 4.6 cm (~50%), right hand 9 cm (2%) and right foot 12.2 cm (8%). The anterior fontanelle was fingertip open and there was midfacial hypoplasia, upslanting palpebral fissures, broad nasal root, unilateral single palmar crease, 5th finger clinodactyly, hypermobility of the hips and ankles and mild hypotonia. A dysmorphology evaluation at 9.25 y/o, while receiving treatment with Increlex (recombinant human insulin-like growth factor 1), showed head circumference 55.1 cm (97%; Z = +1.88), weight 27.6 kg (39%; -0.27 SD), height 128.6 cm (21%; -0.81 SD),

ICD 3.0 cm (50%), IPD 5.7 cm (75-97%), palpebral fissure 3.2 cm (>2SD), right hand 14.5 cm (25-50%) and right foot 19.5 cm (3-25%). There was relative macrocephaly and brachycephaly, a prominent forehead with subtle bossing, deep-set eyes, upslanting palpebral fissures, a depressed and broad nasal root, bulbous nasal tip and small chin. Since then, there has been continued height, weight and head growth and the facial features have changed, possibly related to age and/or his medical treatment. Exam at 14 y/o showed head circumference 60.2 cm (>99%; Z = +3.69), weight 71.7 kg (94.5%; Z = 1.60), height 162.5 cm (42.7%; Z = -0.18) and distinctive craniofacial features including marked macrocephaly with prominent forehead, deep-set eyes with upslanting and long palpebral fissures, synophros, increased interpupillary distance and bulbous nasal tip.

Underweight and short stature were noted during infancy; weight and length at 15 months were 7.79 kg (<1%; -3.03 SD) and 70.4 cm (<1%; -2.84 SD). A skeletal survey at 21 months was normal apart from a widely patent anterior fontanelle. Serum IGF1 at 23 months was <25 (NL: 63-279 ng/mL) and repeated on a separate occasion; routine serum chemistries, CBC, serum CRP and WSR and TSH were normal and IGFBP3 low normal. His early childhood clinical history was notable for heat and fasting intolerance and excessive fatigability. He underwent treatment with growth hormone for 7 months beginning at 29 months, without any increase in growth rate or in serum IGF1; GH was discontinued at 3 y/o. Shortly thereafter he was started on Increlex, with good responses in terms of increased serum IGF1 and in growth velocity. He has since been treated with Increlex, with continued good response for linear growth. His height at 13 y/o was 158 cm (57%; Z = +0.18). He had a tendency for hypoglycemia that predated treatment with Increlex and has a longstanding history, still current, of heat intolerance that is associated with irritability, poor attention, and reduced quality of thinking. He also lacks endurance for both gross and fine motor tasks of extended duration compared to age-matched peers. Recent growth parameters are noted above. Recent endocrine data suggest recent endogenous production of IGF1. The basis for the previous short stature is still unexplained.

Developmentally, the proband did not roll over or crawl until one year and walked independently at about 19 months of age; early social, language and fine motor milestones were unremarkable. He received physical therapy since 1 y/o for reduced strength and gross motor delays and wore supramalleolar orthoses from 11 months to 4 y/o. Gross motor delays continued during childhood with poor balance and excessive clumsiness and falls as well as fine motor clumsiness. Gross and fine motor function are currently age-appropriate. He presently has intermittently poor balance and sustains occasional falls, mostly when he is very tired or hot, and has intermittent unexplained left leg collapse. At about 9.5 years, he developed myoclonic jerks of the extremities that occur when ill or very tired, as well as an excessive startle response to unexpected bright light. He has longstanding difficulty with visual tracking that, in turn, causes difficulty with reading. Visual acuity and hearing are unremarkable. There are no behavioral concerns and social skills are normal. He has an individualized educational plan for his schooling. He is in inclusion classes for music, science, social studies, and language arts. His math/calculation skills are not at age level and are felt to be several grades lower; overall, his neuropsychological profile is complex with splintered skills. It has been noted that his cognitive function can vary from day-to-day and that he

sometimes has loss of an established intellectual competency that is regained a few days later.

The family history is notable for a 16 y/o brother who has a diagnosis of high functioning autism, a normal IQ, learning difficulty and a past history of speech delay; he has had little diagnostic testing. The proband's mother, 43 y/o, has type I diabetes since 16 y/o, celiac disease, nephrolithiasis, a bout of lymphocytic colitis at 35 y/o and a past history of a learning disability. She has normal stature and upslanting palpebral fissures. Two of three maternal uncles have normal stature and past histories of speech delays and learning disabilities. A maternal aunt has mild short stature and a history of a learning disability. The maternal grandmother has mild short stature, hypothyroidism and a history of a learning disability. The proband's father repeated one year of elementary school (unknown reason) and his family has largely non-contributory clinical histories except for a sister who is deceased at one day of life (unknown etiology) and a niece with speech delay and scoliosis; he has an ~209 kb deletion of 16p12.2.

Diagnostic testing for the proband includes the following normal studies: MRI pituitary (2.5 y/o), MRI brain (11 y/o), skeletal survey at 21 months (except for widely open anterior fontanelle), EEG (9.5 y/o), ECG, thyroid function tests, serum cortisol, blood lactate and pyruvate, blood ammonia, serum urate, serum CK and aldolase, plasma amino acids, plasma acylcarnitines, and urine organic acids. H19 methylation analysis showed normal methylation at DMR1. Uniparental disomy analysis showed biparental inheritance of chromosome 7. Sequencing of the mitochondrial genome did not reveal any pathogenic mutations or deletions. Exome analysis was unrevealing except, possibly, for maternally inherited hemizygosity for a variant of uncertain significance of ZMYM3, c.721G>A (p.Glu241Lys). No other variants were identified from ES.

A chromosomal microarray analysis showed arr 15q11.2(20,224,751-20,852,202)x1 which was maternally-inherited, and 16p12.2(21,441,805-21,650,621)x1, which was paternally-inherited. Coordinates are based on human genome build 36.3. This indicated a maternally inherited copy number loss of 15q11.2 of about 627 kb and a paternally inherited copy number loss of 16p12.2 of about 209 kb. The loss of 15q11.2 contains 20 genes, 4 of which have OMIM entries (TUBGCP5, CYFIP1, NIPA2, NIPA1). This region is between breakpoint 1 and breakpoint 2 of the Prader Willi/Angelman region; these 4 genes are nonimprinted and conserved. The loss of 16p12.2 contains 3 OMIM gene entries (METTL9, IGSF6, OTOA).

Individual 3, p.Arg302His

Individual 3 is an 8 year old male. He was the product of a normal pregnancy, delivered by cesarean section at 36 weeks and 6 days. Birth weight was 3700 g., APGAR 9/10. The proband was breast fed for 5 months, with normal growth, and teething was observed as normal at this time.

At 4 months old, the proband had a diagnosis of pyelonephritis due to complete bilateral vesicoureteric reflux. The proband sat independently at 9 months of age and walked independently at 18 months of age. The proband exhibited lallation, but otherwise had an absence of language development. He was enrolled in speech therapy and speaks in simple sentences but has persistent pronunciation difficulties.

The proband acquired urinary continence at 5 years old with several episodes of incontinence at 8 years of age. He also has fecal incontinence.

The proband began first grade at age 7 years and has special education classes. He has little relationship with others and shows selective alimentation; he was diagnosed with autism at 2.5 years. The proband exhibits behavioral disturbances, including intolerance to frustration, episodes of aggressivity, and bruxism. He has displayed motor stereotypies since age 5 years, and he has had regular nocturnal awakenings since age 6 years. The proband also has persistent gastric reflux. He has strabismus in his right eye, and lenses were prescribed but he wasn't compliant. Hearing is normal. No MRI or EEG studies have been performed yet.

The proband has a paternally-inherited duplication identified by arrayCGH (arrayCGH: 203-972 Kb duplication in 2q13, paternal arr[GRCh37]2q13(110427255x2,110841715_111044815x3,111399242x2) arr[GRCh37](XY)x). Fragile X testing indicated he has 30 CGG repeats. No other variants were identified by ES other than the *ZMYM3* variant described here.

The proband has two healthy brothers, and no relevant diseases or consanguinity reported in the family.

Individuals 4a, 4b, p.Arg395Ser

Individuals 4a and 4b are full siblings. Individual 4a was born at term and required resuscitation and intensive care after delivery (Apgar Scores of 1,6). In infancy, he was followed up for transient neutropenia and recurrent vomiting. Although he walked independently at one year of age, some motor delay was present that required rehabilitation. Speech was delayed, and communitive speech was not present until three years of age. His behavior was often reported to be stereotypic, marked with aggressivity and throwing things. He was reported to have a lack of need for social interaction at an early age and later was diagnosed with high-functioning autism. His intelligent quotient is low normal (IQ 81). Individual 4a underwent surgery for hypospadias (at 6 years) and bruises easily. He is 24 years old, now. His aggression is controlled by boxing with a punching bag; however, his level of self-care is low.

Individual 4b had autistic features notable in infancy including little eye contact, lack of interaction/affection with his mother, and had poor sleep. Although he was reported to have some single syllable vocalizations at the age of one, loss of the speech occurred after a febrile infection. He has no functional speech. He can say some words, but these appear to represent echolalia. He was also reported to have stereotypic behaviors and suffered from aggression and auto-aggression in the form of biting. Individual 4b has a diagnosis of low-functioning autism and is reported to have severe intellectual disability (IQ ~35). He also bruises easily. Other than the *ZMYM3* variant, no other potentially causal variants were identified by data analyses.

Individual 5, p.Pro398Ser

A male patient was seen in the outpatient Genetics clinic at the age of 9 years old due to global developmental delay. Speech was more delayed than motor skills, but all were behind and noted at 2 years of age. He was also diagnosed with ADD. He required special education classes in school. He was evaluated by a developmental pediatrician who noted he did not have Autism, but he was diagnosed with sensory processing

disorder. Overall health was good. An Autism/ID NGS panel at GeneDx was done revealing a VUS in the X-linked *ZMYM3* gene (c.1192C>T, p.(Pro398Ser)), with no additional variants reported. Mom was found to carry this same genetic variant and has no learning or health problems. This individual has two older brothers, one of whom has ADHD.

Individual 6, p.Arg441Gln

Individual 6 is a now 10 yo male followed by Genetics due to dysmorphic features, trigonocephaly and autism. He has a complex neurobehavioral history which includes global developmental delays, intellectual impairment, autistic spectrum disorder problems and impulse control. In the past he had problems being a very selective eater with GI difficulties, problems swallowing and feeding which have markedly improved over time. He has a history of exotropia treated with bilateral recessions. His other medical issues include mild thinning of the splenium of the corpus callosum, mild-moderate persistent asthma, eczema, chronically low WBCs and neutropenia, chronic functional constipation and a single renal cyst. The proband's mother had history of ADHD and had an IEP during school years. The proband has a paternal half-sister with ADHD and depression and a maternal half-sister with autism spectrum disorder, ADD, anxiety and possibly dyslexia. There is some distant maternal and paternal family history of autism spectrum disorder and possible ADHD. Other than the *ZMYM3* variant discussed here, no other variants were reported.

Individual 7, p.Arg441Gln

Individual 7 is a 13-year-old boy, a fourth child in a family of European ancestry. He has two healthy sisters. His third sister has severe intellectual deficiency due to a major chromosomal anomaly (maternal isodicentric 15q11.2-q13.1 supernumerary chromosome resulting in tetrasomy). His parents have normal karyotypes. The proband had a normal peri- and postnatal period, but global developmental delay was detected in infancy. The proband exhibits hypotonia, and held his head at 1 year, started to sit at 5 years. He never developed fine motor skills. When put in his hand, he holds a toy but rarely moves it to the other hand. He has a short attention span and is frequently upset but does make eye contact. On clinical examination were also noted hypotrophy of the muscles, and diminished reflexes. The proband never acquired toilet training. His height and weight are deeply below the 3rd percentile. He also has cryptorchidism, and a brain MRI showed enlarged ventricles. Dysmorphic features included coarse and triangular face, widow's peak, long forehead, thick eyebrows, long palpebral fissures, deeply set eyes, blue sclerae, and flashy ears with cupped formed ear lobes. The proband has normally functioning heart and kidneys, and basic biochemical analyses are routinely normal. Genetics testing included FMR1 repeat expansion (Fragile X syndrome) and array-CGH which were negative.

In addition to the ZMYM3 variant discussed here, we also found two compound heterozygous variants (NM_000512.5:c.714del, p.(V239Sfs*80) and c.499T>G, p.(F167V)) in *GALNS* (*612222), a gene associated with Mucopolysaccharidosis IVA (*253000), an autosomal recessive disease. The gnomAD v.2.1.1 database reports at least one p.(F167V) homozygous subject and ClinVar (VCV000321204.9) reports a conflicting interpretation (likely pathogenic/likely benign). However, this variant has been

shown to have reduced enzyme activity (PMID: 1786718). Mucopolysaccharidosis IVA is a severe disease, incompatible with our patient's phenotype; however, we concluded that the biallelic combination of a strong LOF and p.(F167V) in may lead to mild features of Mucopolysaccharidosis IVA, such as the skeletal phenotype of our proband.

Individual 8, p.Arg441Gln

Individual 8 is a now 17y old boy born at 39w3d weighing 2690gr(-1.5 SD); OFC 33.5 cm (-1.5SD); Apgar 7 and 9. Ambiguous genitalia were noted with rugated labioscrotal walls, phallus of 1.4 cm, soft gonads palpable; chromosomes were 46,XY and on ultrasound no Müllerian structures . He was assigned the male gender with scrotal hypospadias. He had a small VSD and ASD type II. Dysmorphisms were fleshy nose, retrognathia, square ears, and simian creases. He was hypotonic with severe delayed development, but with good hearing and vision. He began walking at age 3y. He could not tolerate solid food at age 8 y, had no speech at 15, but was starting to be toilet trained. Dysmorphisms include a hypoplastic midface, protruding ears, thin lips, small hands. At age 15, height was 138.5cm (-5.12 SD), weight 34kg (W/L -1.12 SD). OFC at age 8y was 49cm (-2SD). The proband had a normal multiple congenital anomalies (MCA) sequencing panel, followed by trio exome sequencing which identified the *ZMYM3* variant. The proband's mother has short stature and had a mild learning problem, although she does not need help for her daily functioning.

Individuals 9a, 9b, p.Glu731Asp

Individual 9a and 9b represent two full brothers born from consanguineous parents of Iranian origin. Both affected individuals present with autism and mild intellectual disability. They are reported with normal motor activity. No MRI anomalies have been detected. Both boys have mild facial dysmorphism including long face, high anterior hairline, deep eyes, long philtrum and thin lips. A hemizygous missense variant in *ZMYM3* was identified in both affected brothers. The variant is heterozygous in the mother but absent in the father and the unaffected sister.

Individual 10, p.Ile932Val

Individual 10 was born to a G8P2 mother with dyslexia. The pregnancy was complicated by hypertension, gestational insulin-dependent diabetes, breech presentation, a history of premature births (reportedly caused by low progesterone), and maternal obesity. There were normal ultrasound exams. The proband was delivered at 32 weeks' gestation by c-section delivery due to maternal hypertension and required full resuscitation in the delivery room. At birth he weighed 2020 grams, was 42 cm long, and had an OFT of 31.6 cm. He was discharged home after 2 months in the NICU. At 4 months he was noted to have slightly downslanting palpebral fissures, low set, prominent ears, widely spaced nipples, and mild hypotonia. At three years of age, he has plagiocephaly, torticollis, recurrent acute otitis media, feeding disorder of early childhood, gastro-esophageal reflux disease without esophagitis, constipation, and expressive language delay.

Individual 11, p.Arg1124Gln

Individual 11 is a now 20yo male who first was referred to genetics at age 16 yo due to scoliosis. He was born at term to a 22yo G1 mother, birth weight of 6 pounds, 5 ounces, with no pregnancy or delivery complications. At 3.5 years he was diagnosed with severe delays in receptive and expressive language and poor interaction skills. He was homeschooled after age 11 due to learning and behavior concerns. He left school after the 9th grade and is now attempting to obtain a GED (general educational development) degree. He has oppositional defiant disorder (ODD) and attention deficit hyperactivity disorder (ADHD).

Scoliosis was noted since 9 years of age, which progressed to a left thoracic curve from T2-T7 of 60 degrees, right thoracic curve from T7-L3 of 77 degrees, and left lumbar curve L3-L5 of 38 degrees with kyphosis of 70 degrees. He required spinal fusion surgery at age 16. He needed a tonsillectomy/adenoidectomy at 4 years of age. Brain MRI was normal. He has irregular heartbeats which were evaluated by cardiology with long-term home monitoring, but no treatment was needed.

His physical exam at age 18 years was significant for short stature, long palpebral fissures (3.5 cm (>99%)) which are narrow in height and upslanting, small ears (5.2 cm (1%)) which are posteriorly rotated and low-set, a depressed nasal bridge with wide, thickened and wide nasal alae. He has slight macroglossia with inability to view base of uvula, prominent lips with upturned upper vermillion border, and prognathism.

Prior testing included a microarray which demonstrated a variant of uncertain significance deletion on the X chromosome (Xq26.3-q27.1) which involves *FGF13*. Lysosomal enzyme testing was normal. He was referred for research whole genome sequencing, which revealed the maternally-inherited hemizygous *ZMYM3* variant (c.3371G>A,p.(R1124Q)).

The proband has a maternal aunt with a history of intellectual disability who was in special education classes, but no additional information is known.

Individual 12, p.Tyr1137Asn

Proband 12 is a male born at 38 5/7 weeks gestation by emergency C-section for abnormal fetal positioning after a pregnancy complicated by severe maternal nausea requiring repeated infusions. Measurements at birth were 2780g, 50cm length and 32cm OFC. Initial feeding problems required naso-gastric tube feedings and two weeks hospitalization. Gross motor delays were noted with independent sitting at age 8 months and independent walking at 19 months. Speech development was initially very delayed but progressed more rapidly after tonsillectomy at age 2 years. At the last clinic visit at age 5 years, he was speaking in full sentences albeit with some pronunciation difficulties. He goes to regular Kindergarten where he receives special support. His behavioral profile includes a low tolerance to frustration and some degree of aggressivity but no autism. No health issues besides frequent respiratory infections in infancy and as a toddler. He has no epilepsy and an EEG performed because of a questionable seizure episode during an infection was normal. Family history is unremarkable except for a two-year-old brother with mild speech delay; physical appearance, behavior and overall development of this brother are otherwise unremarkable and very different from the proband's. The brother does not carry the *ZMYM3* variant, which we determined to have occurred *de novo* in the proband's

mother, who is healthy but shares facial features with the proband. No additional variants were reported along with the *ZMYM3* variant described here.

Individual 13, p.Val1202Asp

This individual is a 3-year-old male subject from Nigeria. He is the third-born of four children from healthy non-consanguineous parents. The family history was not contributive except for a record of mother's first-degree cousin with short stature and deafness. The pregnancy was unmonitored. He was born at 35 weeks of gestational age by induced delivery. Swallowing difficulties were reported in the first days of life.

At last clinical evaluation at the age of 3 years and 5 months, he displayed severe intellectual disability with delayed acquisition of all motor milestones such as head control (15 months), sitting position (24 months), and independent walking (36 months). Speech was always absent. He also displayed short stature and rhizo-mesomelia of major upper limbs, microcephaly with trigonocephalic appearance, and some facial dysmorphisms, which include hypertelorism, ptosis major left eyelid, ears with low cup-shaped implantation. Bilateral cryptorchidism was also documented and surgically corrected. Brain MRI revealed global enlargement of the subarachnoid spaces of the posterior cranial fossa. The skeletal X-ray scan documented several skeletal abnormalities. The spine presented with schisis of entire vertebral soma of L5, butterfly vertebrae at T3, median schisis of the posterior arch in T12, and schisis of the posterior arch in L5. Notably, at this level, the hemi lamina right does not merge with the left one but blends with the overlying posterior arch. Humerus and ulna bones were short and dysmorphic, radius was bowed, and Madelung deformity was also noticed bilaterally. Lower limbs presented with dysplastic epiphysis of both proximal fibulae. Biochemical testing and routine laboratory assessments were normal. Karyotype analysis was 46,XY, and a pericentric inversion of the chromosome 9, inv(9)(p11q13) was identified but also considered benign. From ES, the *ZMYM3* variant described here was the only clinically relevant variant reported. No other private/rare (MAF<0.01 in gnomAD) clinically relevant variants involving OMIM genes were identified.

Individual 14, p.Met1213Thr

Individual 14 is a 60 year old gentleman who presented to the Genetics Clinic given a history of cognitive deficits, progressive weakness, and family history of "muscular dystrophy".

His cognitive deficits appeared to represent a mild intellectual disability; however a formal assessment was not available. He attended a "special needs" school. He was unemployed and on social assistance. He describes always having a degree of weakness but that it has become progressively worse over the last few years (distal>proximal) and he was experiencing an increasing number of falls. He previously used a walker and had started to use a scooter. Associated with the weakness, he also experiences a numbness to hands and feet. He does have other health issues including scoliosis, diabetes, hypercholesterolemia, hypertension, asthma, and cataracts. He was estranged from his family for many years, and he does not know where they live or their current health status. Both parents were deceased. He has 4 brothers and 3 sisters. One of his sisters has a diagnosis of "muscular dystrophy" or possibly multiple sclerosis and requires a wheelchair. He is not aware of any genetic testing performed for his

sister. There is no family history of intellectual disability or learning disability in the family.

On examination his OFC was 53cm (-2SD), height was 178cm and weight was 98.1kg. Headshape had a mild plagiocephaly but otherwise normal. He had frontal balding and wore glasses. Eyes were mildly deep-set. Nose, philtrum, mouth, and palate were normally shaped. Ears were normally placed with normal architecture. Strength was mildly reduced (4+/5) to proximal upper extremities and to distal upper extremities (4/5). Hip flexion was diminished 4/5 bilaterally as was knee flexion and extension 4/5, and inversion, eversion, and dorsi- and plantar flexion (4/5).

Brain MRI was normal. Given the family history, he was tested for several genes associated with progressive weakness including myotonic dystrophy 1 and 2, OPMD and FSHD. CK was within normal limits. Given his intellectual disability he was also tested with a microarray (normal) and Fragile X testing (normal). A subsequent comprehensive intellectual disability panel was performed (Fulgent) and was non-diagnostic. However, two variants of uncertain significance and one pathogenic variant were observed. They included *ZMYM3*:c.3638T>C, p.(M1213T); *IGF1R*: NM_000875.4:c.3988G>A, p.(G1330S); and BBS1:NM_024649.4:c.1169T>G, p.(M390R). While this individual was found to be a *carrier of a BBS1 pathogenic variant* he had no retinal dystrophy, no polydactyly or hypogonadism, and no second mutation was observed. He was also found to have a VUS in IGFR1. This variant is seen 15x in gnomAD and the gene is associated with recessive condition. Further, his normal height and that his diabetes is well-controlled and late-onset. Given the lack of diagnosis he was enrolled in the Care4Rare research program. The Care4Rare research program was able to reanalyze the sequence data provided from Fulgent and the research group also identified the variant at *ZMYM3* as a potential candidate.

Individual 15, p.Arg1274Trp

Individual 15 was last evaluated at 16 years 3 months of age. He is a male with a history of global developmental delay, autism, and intellectual disability. Since the start of puberty, the major concern has been with his behavior: he has emotional outbursts, is quick to anger, and has impulsive and aggressive behaviors. He has profound microcephaly and dysmorphic features including angled palpebral fissures, protuberant ears, and a large nose. His extremities are thin with decreased muscle mass, knuckles on his hands are knobby, and he has right concave thoracic scoliosis. He wears glasses to treat myopia. Recently identified mildly dilated aortic root with no cardiac symptoms. General health is fine. No additional variants were reported other than the *ZMYM3* variant described here.

Individual 16, p.Arg1294Cys

The proband was a 26-week estimated gestational age (EGA) fetus of healthy unrelated parents. This was the first pregnancy of the couple. The first trimester ultrasound showed normal nuchal translucency (1.2 mm for a fetal crown-rump length of 60 mm). Estimated risk for trisomy 21 was 1/10,000. The 22-week scan showed a voluminous open dysraphism with significant hydrocephalus with septal rupture and a cardiac defect. Fetal karyotype and CGH-array on amniotic fluid were normal 46, XY.

The parents requested termination of pregnancy, and a male fetus was delivered at 26 weeks. Post-mortem examination showed a male fetus appropriate for gestational age. He had facial dysmorphism with high forehead, synophris, widely spaced eyes, low-set ears and microretrognathia. He had spinal dysraphism with malposition of the lower limbs and feet. External genitalia were abnormal with a median genital tubercle with urethral orifice at the base of the tubercle and unfused genital swelling. Autopsy showed pulmonary abnormal lobulation, thymic hypoplasia and a complex conotruncal heart defect with right aortic arch, overriding aorta, pulmonary stenosis and bicuspidism, atresia of the ductus arteriosus, atrial septal defect, cono-ventricular septal defect, and moderate right ventricular hypertrophy. The brain examination showed ventricular dilatation, right arhinencephaly and neuronal migration abnormalities (heterotopias).

Trio ES identified a *de novo* missense variant in *ZMYM3* and no additional variants.

Individual 17, p.Met1343Ile

Individual 17 was the product of a singleton gestation to a then 36 y/o G4O3->4 mom following a pregnancy that was complicated by hyperemesis gravidarum. He was delivered via vaginal delivery at 38 weeks' gestation with normal growth parameters (BW=3.997 kg; L=48.3 cm). He did well in the newborn period. He presented with severe hypoglycemia in the setting of a viral illness at 13 months of age and was subsequently diagnosed with ketotic hypoglycemia by Endocrinology. He still requires continuous GJ-tube feeds during the day due to rapid drop in glucose levels when off feeds. He also has oral aversion and GI dysmotility. At ~2.5 years of age, he was noted to have to have a R-sided facial palsy and his articulation and oromotor function declined at that time. Brain MRI and EMG/NCS were subsequently unremarkable. He has chronic joint pain and swelling for which he takes naproxen and is followed by Rheumatology. He is also followed by Cardiology due to history of dysautonomic symptoms that manifest as dizziness, dehydration, diaphoresis, overheating, headache, and fatigue/decreased endurance. The proband's history is otherwise significant for ocular tracking issues, recurrent otitis media s/p tympanostomy tubes, asthma, frequent UTIs, right distal femur osteochondroma (s/p resection), allergic rhinitis, and non-IgE mediated food allergies; diagnosis of mast cell activation syndrome raised but not confirmed.

Given his joint laxity and concern for possible connective tissue disorder with clinical diagnosis if hypermobile EDS in mother and three maternal half-sisters, he has undergone a Connective Tissue clinic evaluation and did not meet criteria for a primary connective tissues disorder.

He sat at 6-7 mo; walked at 10-11 mo, but never learned to ride a tricycle. He has issues with motor planning and coordination affecting gross and fine motor skills, balance issues; he uses a stroller for distance due to decreased endurance. His first single words and use of short phrases was on-time with some regression of language skills after emergence of R-sided facial palsy. He developed a stutter at ~5 years of age without an inciting event. He has age-level vocabulary but with a pragmatic language disorder and articulation issues (s/p normal palatal evaluation for VPI). He was diagnosed with autism spectrum disorder at 3 yo as well as general anxiety disorder and separation anxiety disorder. At 8 yr of age, he was reported to have age-level math

skills with below grade-level reading and writing skills. Formal neuropsychological testing data not available.

No variants were reported other than the *ZMYM3* variant discussed here.

Individual 18, p.Leu226TrpfTer8

Individual 18 is a female who was born at 38 weeks' gestation to a 37 year old primigravida Caucasian mother and 47 year old African American father. Both are cognitively normal, and their family histories negative for intellectual disabilities, learning disabilities, or fetal loss. At 34 weeks, mild IUGR and short femurs were noted. Delivery was by C/S and Apgars 8/8. Her birth weight was 2450 g (-2.13 SD), length 44.5 cm (-2.61 SD) and OFC 33 cm (-1.90 SD). She had some problems feeding and had symptoms of GERD in the neonatal period. She had early plagiocephaly and torticollis that resolved with helmet treatment. At 2.5 months she had a brief resolved unexplained episode (BRUE) leading to a short ICU admission.

Her growth has been normal; at age 2 her weight was 12.2 kg (0.02 SD), length 86.5 cm (0.25 SD) and OFC 46.5 cm (-0.94 SD). Development has been mildly delayed with sitting at 8 months, crawling at 12 months, and walking at age 2 years. She had delayed receptive language on assessment at 19 months. She has unprecipitated episodes of hand flapping.

The proband has dysmorphic features noted, including deep set eyes, flat nasal bridge with epicanthal folds, full cheeks, a broad nasal tip, long philtrum, small mouth, mild micrognathia and thin vermillion border of the upper lip and a pouting lower lip. Her hands and feet are normal. She has small nipples.

Testing has included a normal SNP microarray, an echocardiogram showing prominence of the papillary muscles, normal sleep and awake EEGs, and a normal swallow study. Exome sequencing studies revealed the maternally-inherited variant in *ZMYM3*. X-inactivation studies at one lab revealed 99% skewing in the child using the AR probe (whole blood sample). Mom's studies were not successful as she is homozygous for the *AR* alleles that were tested. Studies at a second lab using the *RP2* locus revealed skewing with a ratio of 94:6 in both mom and child. Both mom and child have alleles of the same size (366/362), and both individuals appear to be inactivating the same allele (366). No other variants were identified by ES.

Individual 19, p.Tyr752Cys

Individual 19 was a female infant born AGA at 37 3/7 weeks with uncomplicated pregnancy. Her newborn period was complicated by murmur on exam, and she was found to have atrial septal defect that closed spontaneously. Echocardiogram at 7 mo showed small PDA, PFO, and LSVC to coronary sinus, with ASD no longer present. Patient was first evaluated at 17 mo in genetics clinic for hypotonia, poor weight gain, and gross motor delays. History notable for sitting up at 8 mo, walking at 16 mo, and persistent central hypotonia. She initially had delays in speech development, but after therapies, she has now graduated from speech therapy. She has mild textural sensitivities, but overall, there are no concerns for autism spectrum disorder. On exam, patient had apparent telecanthus, bilateral epicanthal folds, flat midface, short nose, soft skin, and central hypotonia. She was observed to be unbalanced while walking and sitting in the tripod position throughout the visit. Growth chart showed her weight in 1st

percentile, length at 18th percentile, and OFC at 6th percentile. Initiated genetic testing with chromosomal microarray, which was negative. Reflexed to the GeneDx Autism/ID Xpanded Panel (trio), which found two variants of uncertain significance: *SCN2A* c.1421T>A, heterozygous, paternally inherited (not thought to be contributing to her features given lack of clinical correlation and unaffected father); and *ZMYM3* c.2255A>G, heterozygous, *de novo*.

Family history is significant for many paternal family members with stretchy skin, joint issues, and hypermobile joints. Cousins of father: one died in infancy of aortic complication, one 6'3" and "lanky", and one girl with low tone and atrial septal defect. Paternal great-great-aunt had twins who both had poor growth and low tone.

Individual 20, p.Arg1294Cys

Individual 20 is a 5 yo girl presenting with global developmental delay. Dysmorphic features include mild microcephaly, plagiocephaly, synophris, thin lips, protruding and badly hemmed ears. Hands and feet are marked by tapered fingers, bilateral single transverse palmar crease, bilateral camptodactyly of third and fourth toes. She sat at 16 month and walked at two years old. Language development is delayed. Other clinical features include volvulus of midgut, pyelectasis, pancreatic cysts. X-inactivation studies using the HUMARA assay indicated 97% skewing in this proband. Parental XCI testing not performed due to the *de novo* nature of her *ZMYM3* variant. No additional variants of interest were identified other than the *ZMYM3* variant discussed here.

Individual 21, p.Arg169Ser

This individual presented for an initial genetic evaluation at 10 years of age for moderate intellectual disability. He was born at 38 weeks gestation. Hypospadias was reported. He walked at 12 months, said his first words around 12 months, spoke in phrases at 3 years, and was toilet trained around 2 years. There was no developmental regression. His mother first noted significant concerns when he started school. He required repeating kindergarten and started attending special-education classes in the 1st grade. At the time of visit, he knew some letters but was not reading and had significant difficulties in all subjects. He was diagnosed with ADHD at 8 years of age. Additional behavior concerns included being easily frustrated and hitting himself. Around age 11, he was diagnosed with autism spectrum. Parents had one daughter born prematurely at 6-1/2 months' gestation; she died at 9 days of age. The child was not known to have any birth defects or other anomalies. No other family members were reported with intellectual disability or birth defects. There was no known history of consanguinity. Genetic testing including chromosomal microarray analysis and testing for Fragile X were normal. An X-linked Intellectual Disability Gene Panel was ordered, which revealed sequence variants in *ZMYM3* and *IL1RAPL1*. The *IL1RAPL1* variant (NM_014271.3:c.1039G>A (p.Val347Ile)) will now be classified as Likely Benign (BS2, BP4). There are 8 hets/2 hemizygotes in gnomAD v2.1.1, in BRAVO there are 13 hets/2 hemizygotes. One report in ClinVar as Likely benign.

Individual 22, p.Arg441Trp

This individual is a now 4 year old male referred to genetics for developmental delay. He had prenatal drug exposure, but there is no other information regarding pregnancy, delivery, or early developmental milestones. He has mild global

developmental delay, mild short stature, a triangular face, prominent forehead, long eyelashes, and single transverse palmar crease. At 4 years old, he is not toiled trained. He only has a few words but he seems to understand more than he can express. He points and has decreased eye contact. He was in a 3K special needs program last year. He does not sleep well. He has one sister with developmental delay, 2 brothers with autism, and one healthy brother. His mother is reported to have bipolar disorder and schizophrenia. There is no known information about his father or other family members. Other than the *ZMYM3* variant, no other variants were reported for this case.

Individual 23, p.Cys454Arg

Individual 23 is a 16-year-old male with global developmental delay, cognitive disorder not otherwise specified, ADHD, short stature, microcephaly, gastrointestinal dysmotility, myopia, and retinopathy. He was born at 36 weeks gestational age. The pregnancy was complicated by intrauterine growth restriction and polyhydramnios. The neonatal period was complicated by jaundice requiring phototherapy.

He has had feeding difficulties since early infancy and has been diagnosed with gastroparesis, ultimately requiring placement of a GJ tube. He underwent a nissen fundoplication for GERD. He also has chronic constipation. His development was globally delayed since birth. He walked at 25 months. Speech was also significantly delayed. He has been diagnosed with cognitive disorder not otherwise specified and ADHD. He also has emotional lability. He has had progressive myopia, and pigmentary abnormalities of the retina concerning for a retinopathy. He has short stature and growth hormone deficiency, on treatment with growth hormone. He also has microcephaly. He has outgoing toes secondary to femoral retroversion/external tibial torsion. He has a history of left-sided vesicoureteric reflux. Investigations include normal audiometry. Brain MRI's have demonstrated mild parenchymal volume loss that has been stable over time (last performed at age 11 years). Electroencephalogram (EEG) performed at 4 years of age was normal. Echocardiograms have been normal.

In terms of family history, he has a sister with articulation difficulties, and a brother with learning disability, dyslexia, and ADHD. There is no other contributory family history. Clinical trio whole exome sequencing reported a *de novo* hemizygous variant in the ZMYM3 gene (c.1360T>C, p.C454R). He also had a heterozygous VUS in the VPS13B gene and a heterozygous known pathogenic variant in the FLG gene.

Individual 24, p.Ser1173Asn

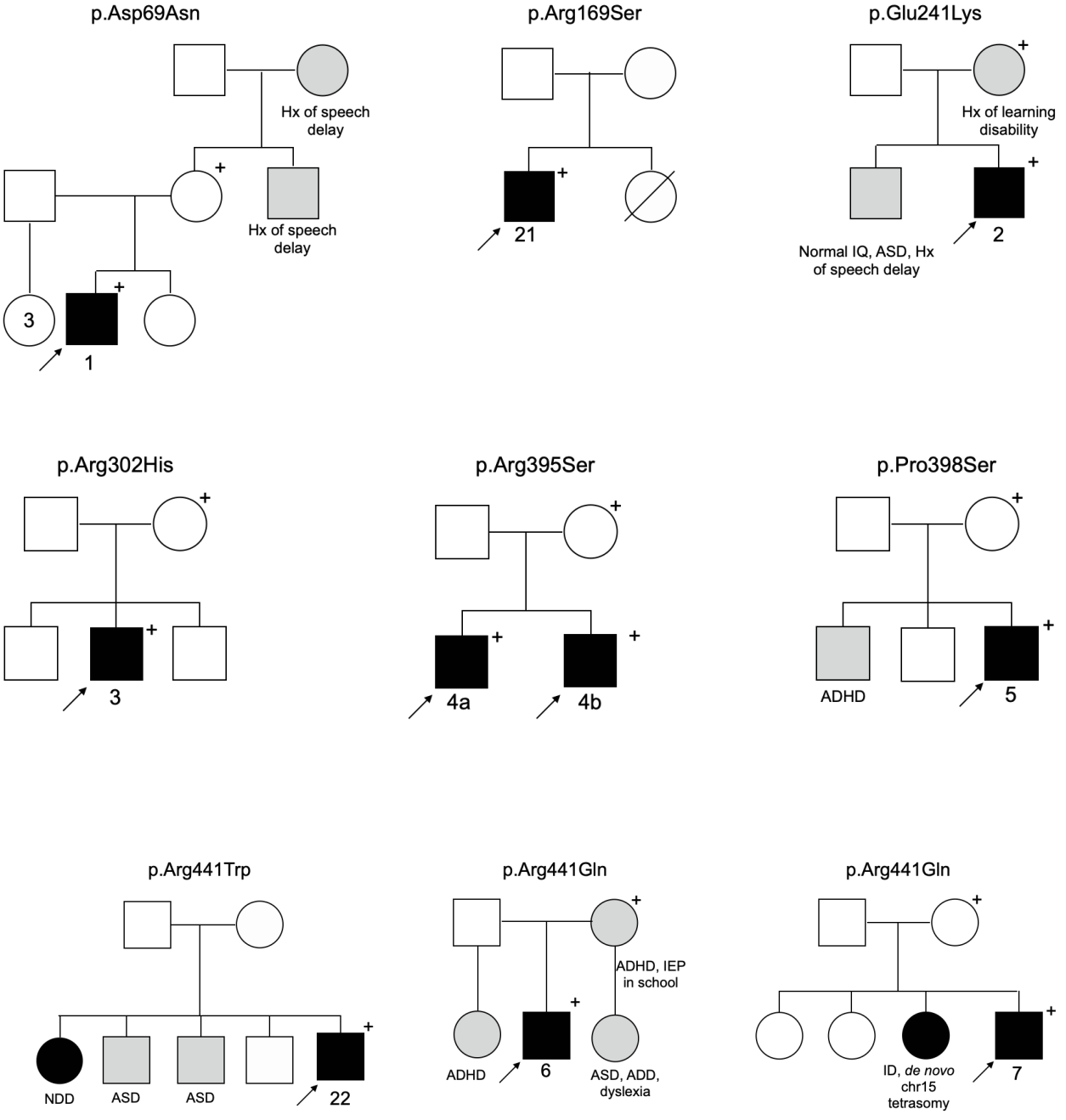
Individual 24 is an 8-year-old male with a history of autism and developmental delay. The proband had normal prenatal screening and was born to a then 17-year-old G1P0 mother via induced vaginal delivery at 40 weeks gestation. Birth weight was 7 pounds and 11 ounces with a birth length of 18 inches. Initial development was delayed, with sitting at 18 months, crawling at 2 years, walking at 2.5 years, and running at 4 years. The individual's first words were at age 2 years, with use of sentences beginning at age 5. He continues to make progress with speech, occupational, physical, and behavioral therapies. Upon examination, the proband was found to have ear size above the 99th percentile. No other dysmorphisms were noted. Family history was significant for ADHD and developmental delays in the proband's mother. A maternal uncle is also reported to

have autism, developmental delays, mild ID, reportedly large ears, and ADHD. Limited paternal history was available, though autism and learning issues were suspected. Initial genetic testing for the proband included chromosomal microarray and Fragile X testing at GeneDx, both of which were negative. Subsequent whole genome sequencing was pursued through the HudsonAlpha Clinical Services Lab, and the *ZMYM3* p.S1173N variant was identified. Sanger testing confirmed maternal inheritance of the variant. This individual is also carrier for a likely pathogenic variant in *ASPM* (NM_018136.5:c.77delG, p.Gly26AlafsTer42) and a likely pathogenic variant in *RARS2* (NM_020320.5:c.35A>G, p.Gln12Arg=). Neither of these are considered a strong phenotypic fit, and no second hits were observed for either gene.

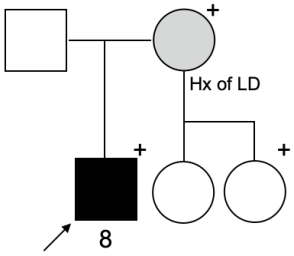
Individual 25, p.Arg1324Trp

This individual is an 8-year-old male first seen by genetics age 3. He was born full term to a 34-year-old female. He was referred to genetics for developmental delay and autism. He had aggressive behaviors as well. He walked at 15 months of age and had very limited language development. He was difficult to examine due to hyperkinesis but did not have major dysmorphic features. His weight was greater than 99th percentile.

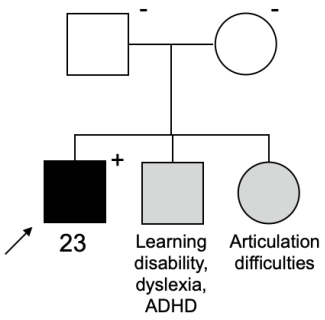
Figure S1. Pedigrees for each family in the study. Variants are shown above the pedigree, with *de novo* variation in red. Affected individuals presented here are indicated with arrows and labeled with IDs matching the text (1-23). Individuals affected with an NDD are shown in black, while mildly-affected individuals, those with a single feature of an NDD, or a history of such are shown in gray. Presence of the variant is indicated with +, while absence of the variant (if tested) is indicated by a -. Numbers in unaffected individuals' shapes indicate the number of unaffected siblings (or half-siblings) when there are more than two.



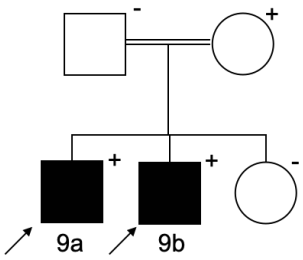
p.Arg441Gln



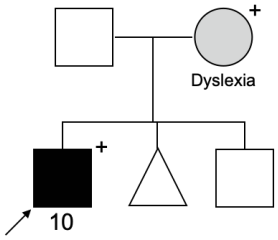
p.Cys454Arg



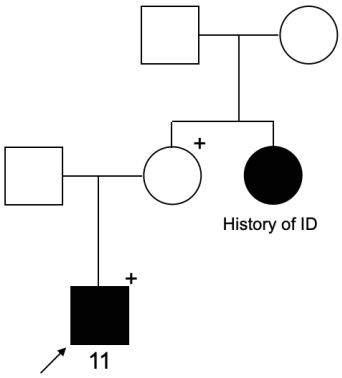
p.Glu731Asp



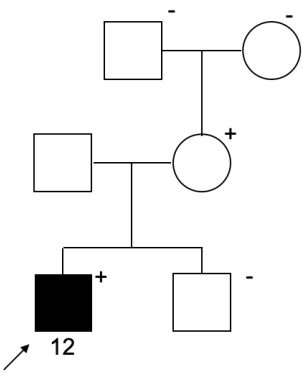
p.Ile932Val



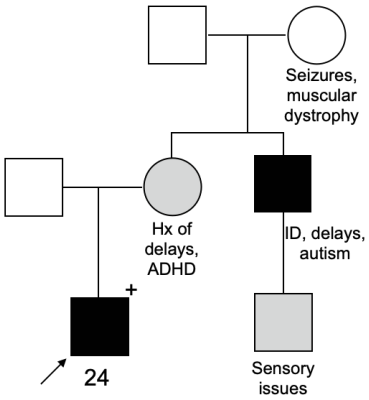
p.Arg1124Gln



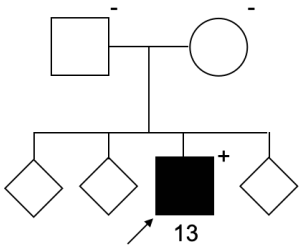
p.Tyr1137Asn



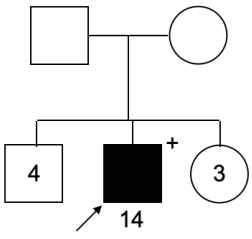
p.Ser1173Asn



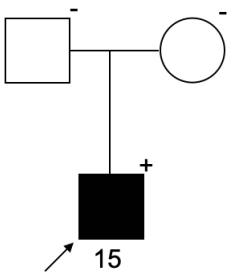
p.Val1202Asp



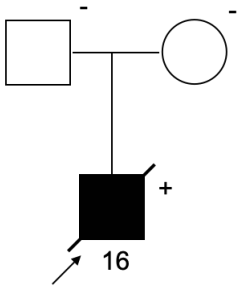
p.Met1213Thr



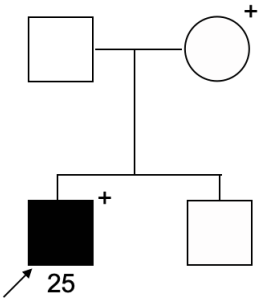
p.Arg1274Trp



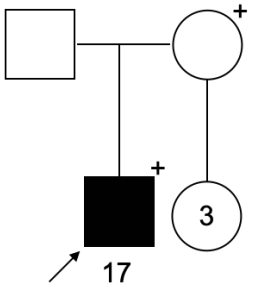
p.Arg1294Cys



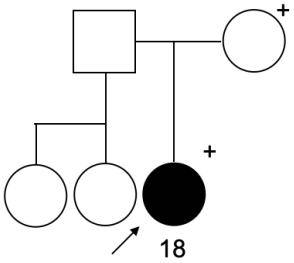
p.Arg1324Trp



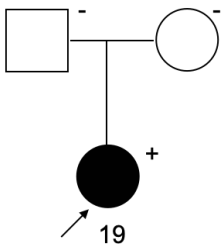
p.Met1343Ile



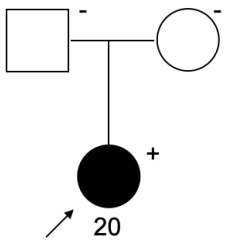
p.Leu226TrpfsTer8



p.Tyr752Cys



p.Arg1294Cys



NDD

Mildly-affected with NDD, single feature of an NDD, or a history of such

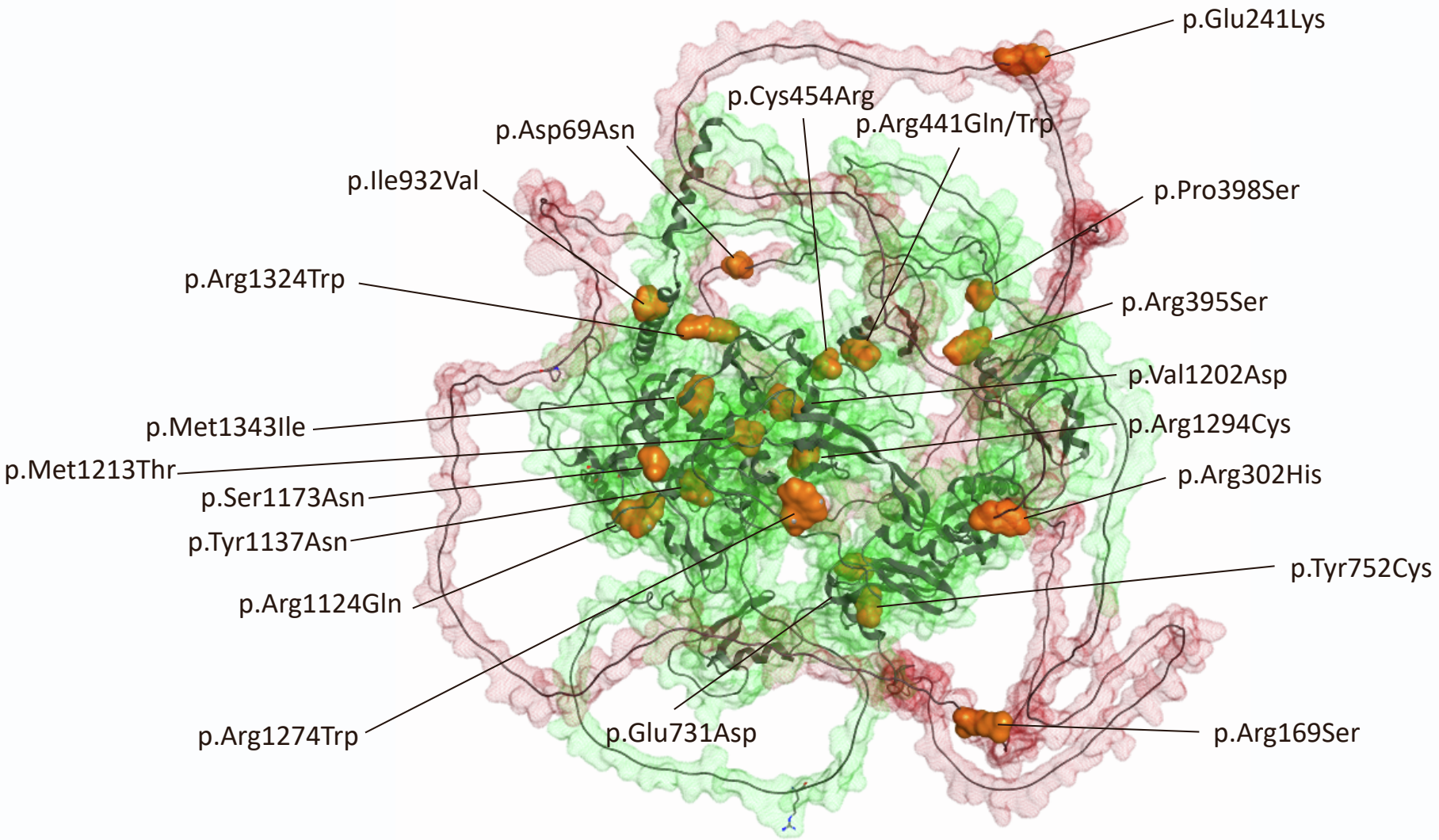


Figure S2. Missense variants in ZMYM3 mainly lie in ordered regions. Visual representation of the variant residues on a 3D model of ZMYM3. Disordered regions are in red, while structured regions are shown in green. Individual variants are shown in orange.

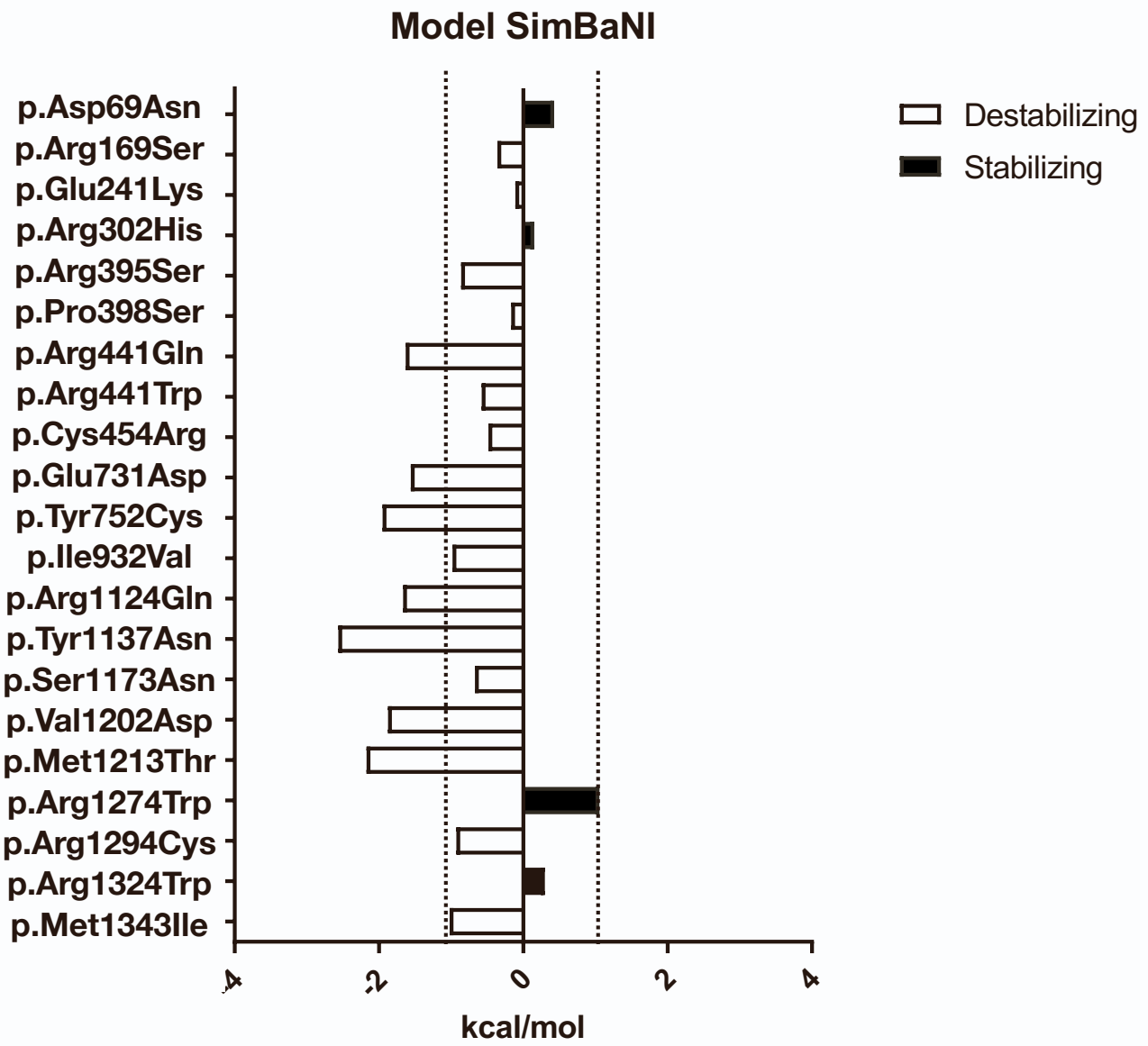


Figure S3. Analysis of the free energy variation upon amino acid substitution suggests a destabilization for the majority of mutants with a major impact of mutants p.Arg441Gln, p.Glu731Asp, p.Tyr752Cys, p.Arg1124Gln, p.Tyr1137Asn, p.Val1202Asp, and p.Met1213Thr . Conversely, p.Arg1274Trp is predicted to stabilize the protein structure. Vertical dotted line indicates threshold of +/- 1 kcal/mol.

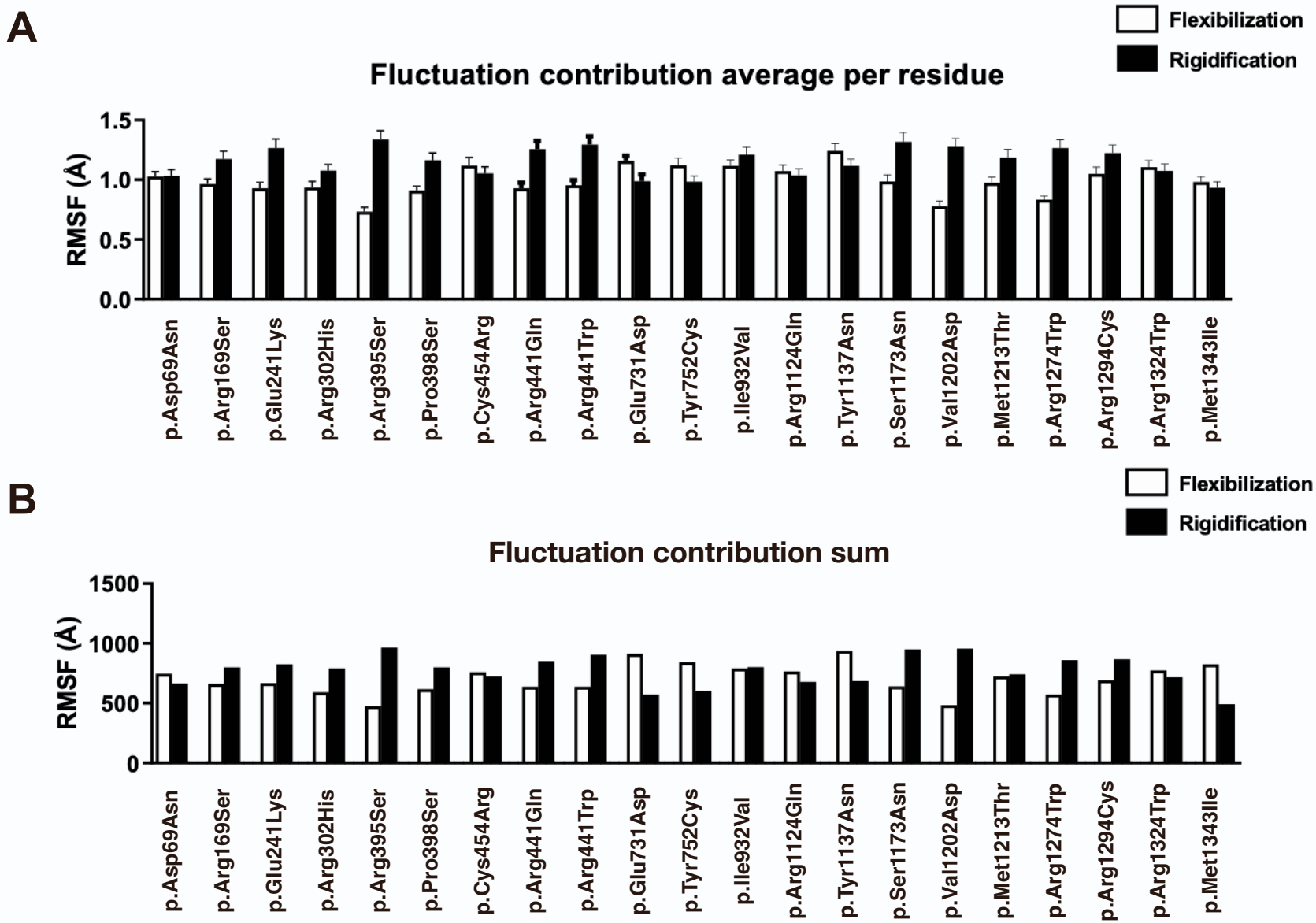


Figure S4: Coarse-Grained molecular dynamics run. Assessment of the flexibility contribution across the entire length of the protein, calculated by comparing the difference of RMSF of the WT minus RMSF of the mutant. Contribution to flexibility is shown in white bars, contribution to rigidity in black bars **A.** Average RMSF per residue for both contributing factors. **B.** Sum of RMSF values for both contributing factors.

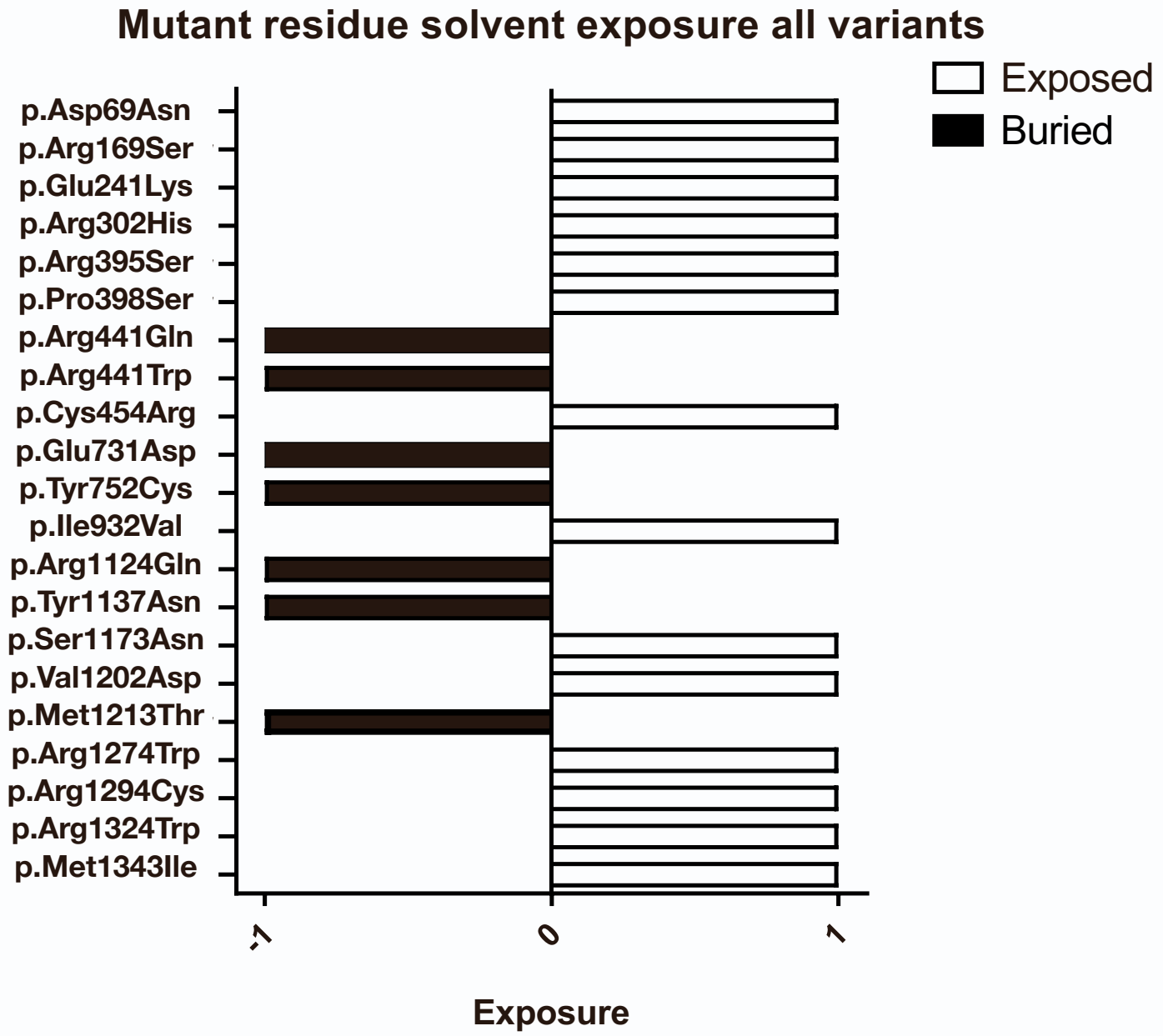


Figure S5: Difference in solvent accessibility of mutant residues. Exposure is expressed as -1 (black bars) for buried and +1 (white bars) for exposed residues. Seven residues are buried, while 14 are exposed.

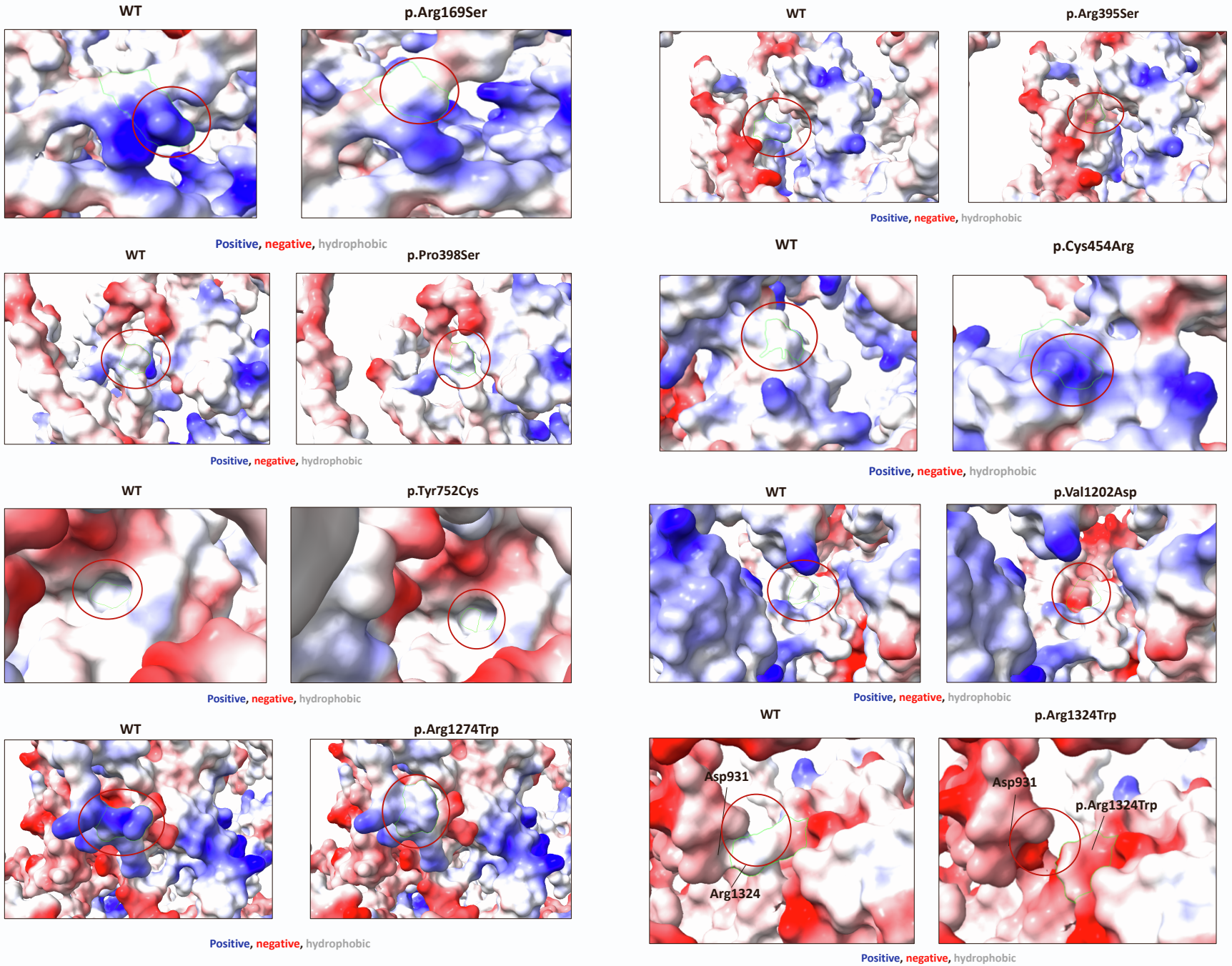


Figure S6. Surface analysis of several selected substitutions.

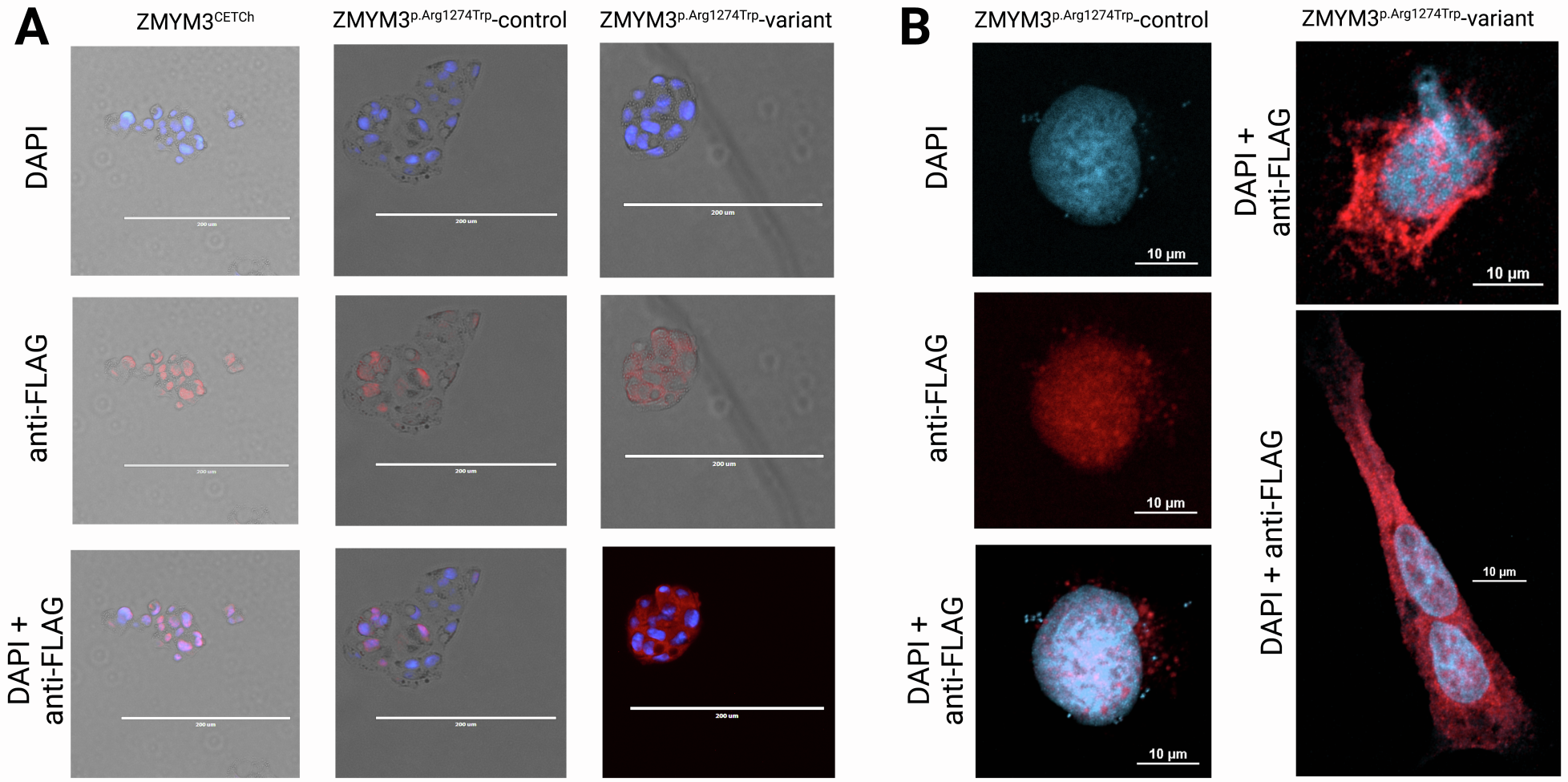


Figure S7. Immunocytochemistry shows localization of ZMYM3CETCh and ZMYM3$^{p.Arg1274Trp}$-control proteins primarily in nuclei of cells, and ZMYM3$^{p.Arg1274Trp}$-variant protein primarily in cytoplasm of cells. A. Images visualized on Evos FL microscope at 20X. Columns of images are labeled with cell line visualized. Separate spectral visualizations of the same field of view are shown in each column; these are (top row) DAPI, showing nuclei; (middle row) secondary antibody spectrum, showing FLAG-tagged protein localization; and (bottom row) both DAPI and anti-FLAG, showing relative localizations. **B.** Images visualized on Nikon AX confocal microscope at 100X. Columns of images are labeled with cell line visualized. First column is organized in labeled rows as in A showing the same field of view. In last column, with ZMYM3$^{p.Arg1274}$Trp-variant only cells, both DAPI and secondary (anti-FLAG) spectra are shown for two separate fields of view; top shows one cell, bottom shows two adjacent cells. Created with www.BioRender.com. See methods for additional details.

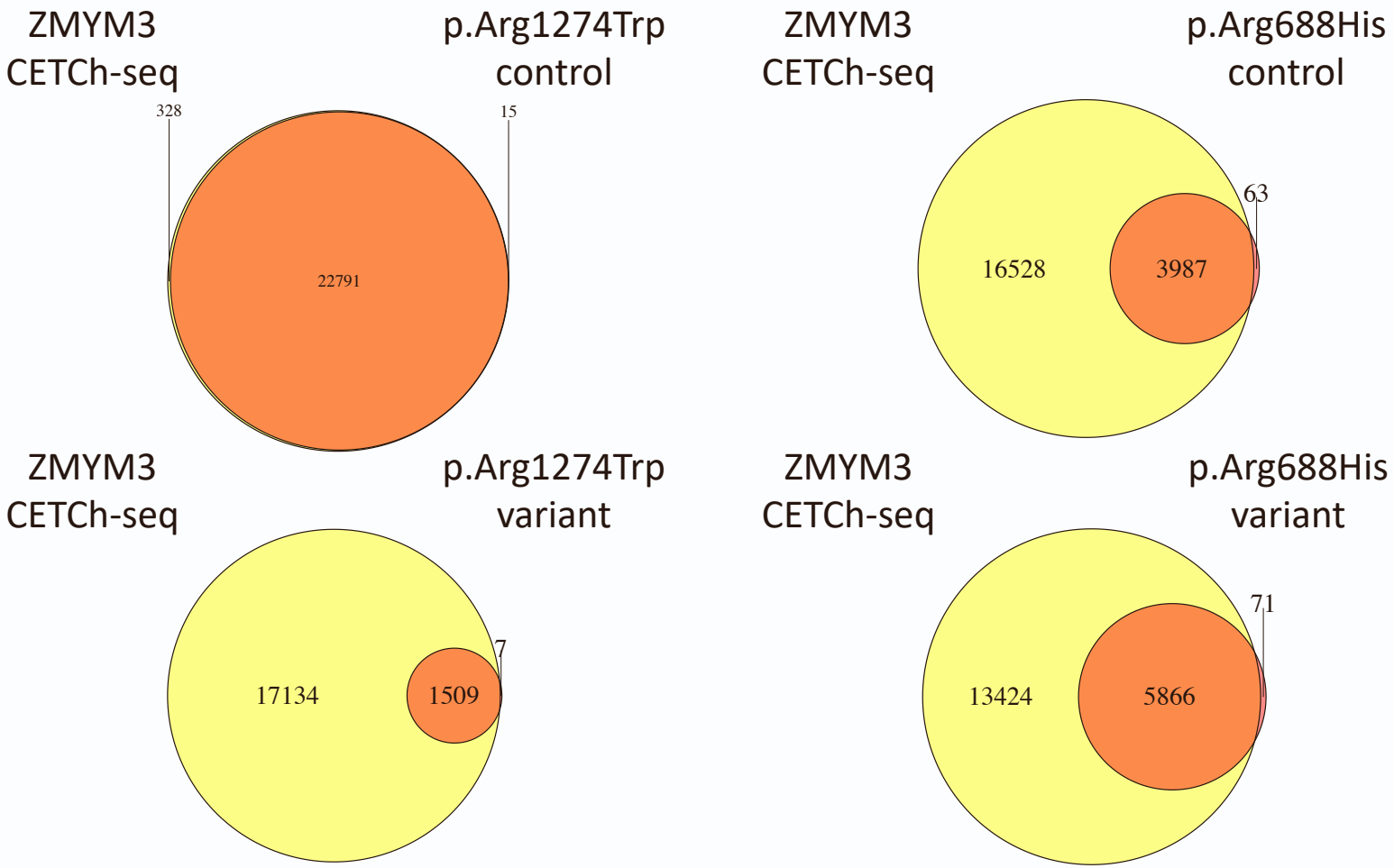


Figure S8. Intersection of csaw analysis and peak calls. For each pairwise comparison, the union of peaks was determined (downsampled to 20 million reads per replicate, IDR 0.05, merged peaks) and intersected with all csaw analyzed regions. Regions with significantly higher reads in the ZMYM3 CETCh-seq experiment are colored in yellow; regions with significantly higher reads in the comparison experiment are colored in red; and regions that are not significantly different are colored in orange (the overlap set; FDR cutoff for significance is 0.05). p.Arg1274Trp control is highly similar to ZMYM3 CETCh-seq, with the vast majority of regions not significantly differential. The other compared experiments are subsets of the ZMYM3 CETCh-seq experiment with very few differentially higher regions.

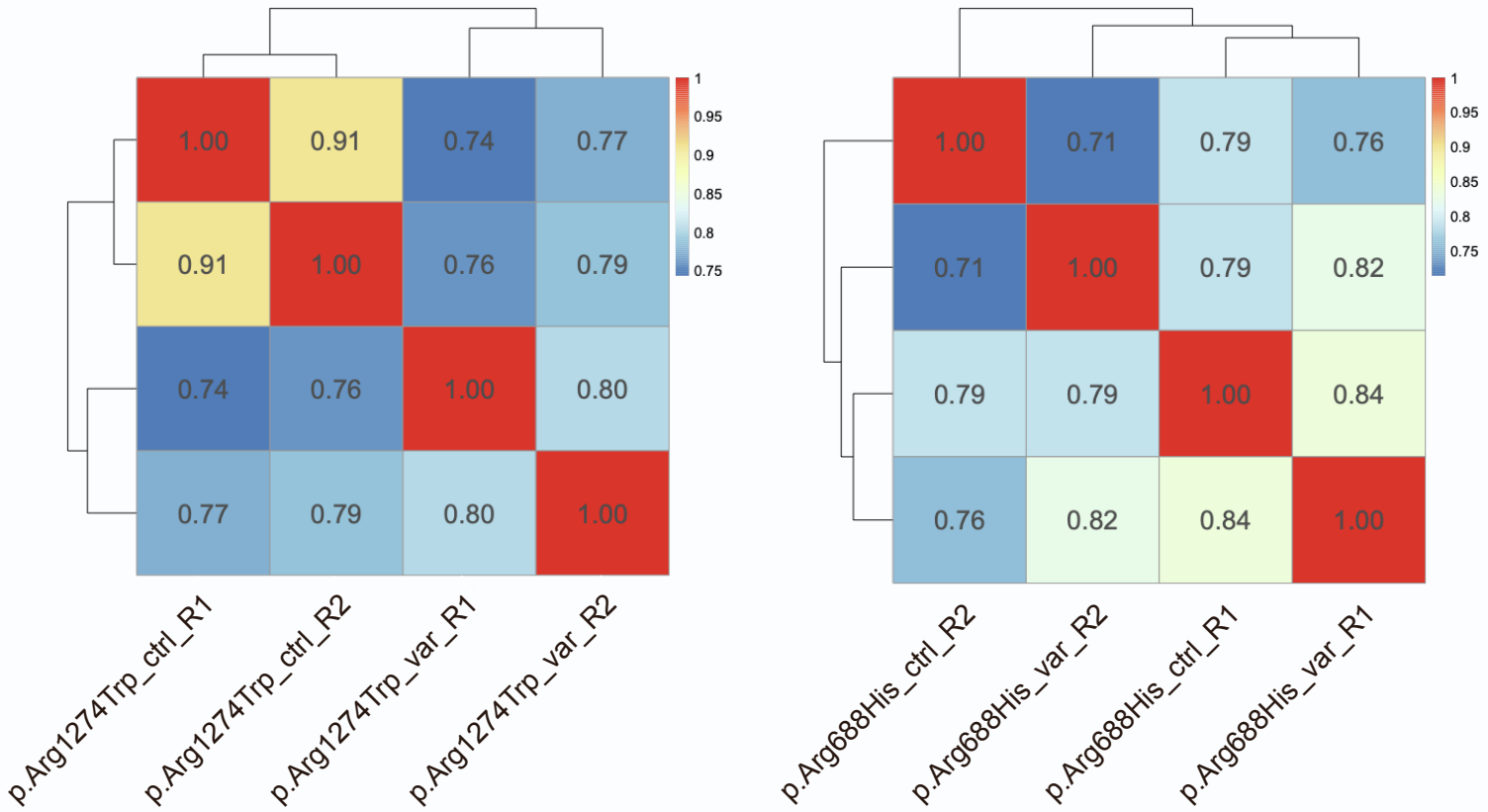


Figure S9. Read count correlations. The union of peaks between all experiments (ZMYM3^{CETCh}, ZMYM3$^{p.Arg1274Trp}$, and ZMYM3$^{p.Arg688His}$) was calculated, and reads from the .bam file for each replicate (downsampled to 20 million reads) were determined at each of these genomic positions. These data were correlated for each pairwise comparison and Pearson correlation coefficient was determined and plotted with clustering by pheatmap v. 1.0.12 in R v. 4.2.1.

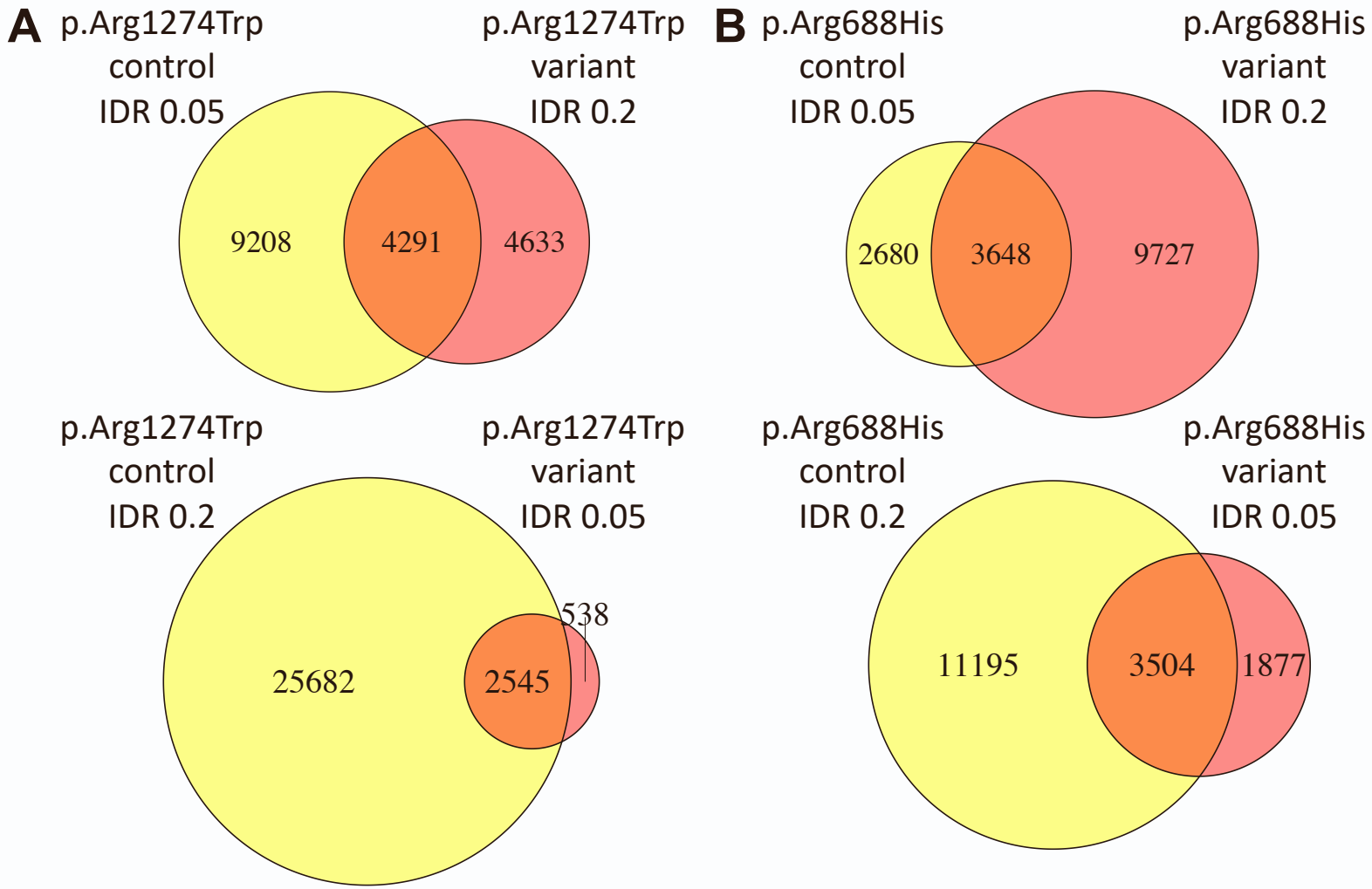


Figure S10. Standard peak overlaps with relaxed peaks. Peaks were called between replicate experiments downsampled to 20 million reads at the standard IDR threshold of 0.05 and also at a relaxed IDR threshold of 0.2. This analysis attempts to find peak overlaps missed by comparing standard peak calls due to enriched regions near peak-calling threshold. For p.Arg1274Trp (A), many regions in the variant experiment called as peaks in control are just under threshold and found by this approach, indicating a general reduced binding by the variant. For p.Arg688His (B), an experiment with greater noise, peaks are gained in both comparisons.

Supplemental Materials and Methods

SEQUENCING

Several probands had either exome sequencing (ES) or the Autism/ID Xpanded Panel (AIDX) through GeneDx, and details are listed in the table below. Using genomic DNA from the proband and parents (if submitted), from the tissues listed below, the exonic regions and flanking splice junctions of the genome were captured using the IDT xGen Exome Research Panel v1.0 (Integrated DNA Technologies, Coralville, IA). Massively parallel (NextGen) sequencing was done on an Illumina system with 100 bp or greater paired-end reads. Reads were aligned to human genome build GRCh37/UCSC hg19 and analyzed for sequence variants using a custom-developed analysis tool. Reported variants were confirmed, if necessary, by an appropriate orthogonal method in the proband and, if submitted, in selected relatives. Additional sequencing technology and variant interpretation protocol has been previously described[1]. The general assertion criteria for variant classification are publicly available on the GeneDx ClinVar submission page (http://www.ncbi.nlm.nih.gov/clinvar/submitters/26957/).

Individual	Protein Effect	Testing details	Tissue Tested
1	p.Asp69Asn	AIDX trio	Buccal
2	p.Glu241Lys	ES duo with mom	Blood
5	p.Pro398Ser	AIDX trio	Blood
6	p.Arg441Gln	ES trio	Buccal
15	p.Arg1274Trp	AIDX duo with mom	Blood
18	p.Leu226TrpfsTer8	ES trio	Buccal
19	p.Tyr752Cys	AIDX trio, Sanger confirmation for proband but parents had high quality NGS data and Sanger confirmation was not necessary	Buccal
23	p.Arg454Cys	ES trio	Blood

Individual 3, p.Arg302His and Individual 7, p.Arg441Gln

Exome sequencing was performed on Individuals 3 and 7 as described[2]. Genomic DNA was isolated from blood. The patients were enrolled for trio exome sequencing (ES) as part of the international network of Autism Sequencing Consortium (ASC) (https://asc.broadinstitute.org/). ES was performed at the Broad Institute on Illumina HiSeq sequencers[3]. ES raw data of the trios were processed and analyzed using an in-house implemented pipeline[4,5].

Individuals 4a, 4b, p.Arg395Ser
Genomic DNA extracted from leukocytes of both patients and their parents was used for whole-exome sequencing. Exome enrichment was performed on individually barcoded samples using SeqCap EZ Exome Probes v3.0 (Roche) and sequencing was performed on HiSeq 2500 (Illumina) with 100bp paired-end reads. Reads were aligned to the hg19 reference genome using Novoalign version 3.02.13 (Novocraft) with default parameters. After genome alignment, conversion of SAM format to BAM and duplicate removal was performed using Picard Tools (2.20.8). The Genome Analysis Toolkit, GATK (3.8)[6] was used for local realignment around indels, base recalibration, variant recalibration, and variant calling. Variants were annotated using the GEMINI framework[7] and filtered based on the population frequencies using several public databases and an in-house database of population-specific variants. Identification of candidate variants was performed for autosomal dominant (*de novo* variants) and autosomal recessive inheritance patterns. Variants were further prioritized according to the functional impact and conservation score. Sanger sequencing confirmed presence of the variants in the patients.

Individual 8, p.Arg441Gln
Trio exome sequencing was performed for Proband 8. DNA was enriched using Agilent SureSelect Clinical Research Exome V2 capture and paired-end sequencing using the Illumina platform (outsourced). The aim was to obtain 8.1 Giga base pairs per exome with a mapped fraction of 0.99. The average coverage of the exome was ~50x. Duplicate reads were excluded. Data were demultiplexed with bcl2fastq Conversion Software from Illumina. Reads were mapped to the genome using the BWA-MEM, and variants were called using GATK HaplotypeCaller. Detected variants were filtered and annotated with Cartagenia software and classified with Alamut Visual.

Individuals 9a, 9b, p.Glu731Asp
ES was performed on gDNA of both Individuals 9a and 9b. The exome was captured using the xGen Exome Research Panel v2 (Integrated DNA Technologies) and sequenced using the Illumina HiSeq4000 platform according to the manufacturer's protocols. The overall mean-depth base coverage was 136- and 125-fold, while on average 93% and 92% of the targeted region was covered at least 20-fold, respectively for V.2 and V.3. Read mapping and variant calling were performed as described[8] using the Varapp software[9]. Sanger sequencing confirmed the segregation of the potentially causative variant.

Individual 10, p.Ile932Val
Proband 10 had exome sequencing, and both short-read and long-read genome sequencing as described[10]. The variant was first observed in ES.

Individual 11, p.Arg1124Gln
For Individual 11, patient blood was drawn into an EDTA blood collection tube. Isolation of DNA from whole blood was performed using the QIAsymphony (Qiagen). Sequencing libraries were constructed from patient whole blood genomic DNA using the HudsonAlpha Clinical Services Lab's custom whole genome library preparation protocol.

Patient DNA was sequenced on the Illumina HiSeqX sequencer. DNA library fragments were sequenced from both ends (paired) with a read length of 150 base pairs. Patient genomes were sequenced at an approximate depth of 30X, with at least 80% of base positions reaching 20X coverage. Sequence variants were called using GATK3 and loaded into a custom software analysis application for interpretation. All sequence variants were annotated with relevant information from established data sources to provide support for variant interpretation. Variant pathogenicity was determined using ACMG criteria[11].

Individual 12, p.Tyr1137Asn

For Individual 12, Genomic DNA was extracted from EDTA blood of the patient and his parents. Whole Exome Sequencing (ES) on the patient was performed using the xGen® Exome Research Panel v1.0 (IDT) with paired-end sequencing (HiSeq SBS Kit v4, 125 Fwd-125 Rev, Q30-value: 84) on a HiSeq System (Illumina Inc.). Raw fastQ files were aligned to the hg19 reference genome using NextGene (Softgenetics). The average depth of coverage was 225x and 99.4% of the targeted bases were assessed by ≥20 independent sequence reads. By applying filters for known and candidate ID genes (SYSID and In-House) and Minor Allele Frequency ≤ 2% (gnomAD, ExAC) a total of 37 variants were observed in at least 16% of reads with sufficient quality level. Variants were investigated computationally for deleterious effects, by associations of the affected gene with proband's phenotype and by literature search for functional information. The candidate *ZMYM3* mutation from the ES approach was re-sequenced in the index, his mother and maternal grandparents after PCR amplification by Sanger sequencing using an ABI Genetic Analyzer 3730 (Applied Biosystems, Foster City, California.

Individual 13, p.Val1202Asp

The proband had ES as described[12,13]. Genomic DNA was isolated from circulating leukocytes. The *de novo* status of the *ZMYM3* variant was confirmed by Sanger sequencing.

Individual 14, p.Met1213Thr

A clinical comprehensive intellectual disability panel (555 genes) was ordered through Fulgent Genetics, and testing was done on genomic DNA from blood. This testing resulted in identification of the *ZMYM3* variant described here. Follow-up reanalysis through the Care4Rare project[14,15] also identified this *ZMYM3* variant as the top hit.

Individual 16, p.Arg1294Cys

Individual 16 had trio exome sequencing. DNA was extracted from fetal muscle using the Prepito automate machine. Exome DNA library was prepared with the Agilent Focused Exome preparation kit. High-throughput sequencing was performed on a NextSeq550 sequencer (Illumina) with a 2x75 bp paired-end running method. The BWA-MEM algorithm was used to map the reads on the reference genome (GRCh37/hg19). The variant calling was performed according to GATK and FreeBayes best practices. The ANNOVAR and ALAMUT (Interactive Biosoftware) tools were used for variant annotation.

Individual 17, p.Met1343Ile
For Individual 17, a CHOP Medical Exome was performed on the proband and mother. Genomic DNA was extracted from peripheral blood or other patient tissues following standard DNA extraction protocols. After extraction of genomic DNA, targeted exons are captured with the Agilent SureSelect XT Clinical Research Exome kit (per manufacturer's protocol) and sequenced on the Illumina HiSeq 2000 or 2500 platform with 100bp paired-end reads. Mapping and analysis were based on the human genome build UCSC hg19 reference sequence. Sequencing data is processed using an in-house custom-built bioinformatics pipeline. The bioinformatics protocol utilized for this evaluation is version CWES-2.2. The exome sequencing protocol utilized for this evaluation is version 3.1. Coding exons and splice sites targeted with the exome kit are analyzed and reported. The following pathogenic variants are detectable: single nucleotide variants, small deletions and small insertions.

Individual 20, p.Arg1294Cys
Individual 20 had exome sequencing on DNA extracted from blood. ES was performed using Nimblegen SeqCap Ez MedExome Target Enrichment Kit (Roche Sequencing Solutions, Pleasanton, CA, USA) and an Illumina NextSeq500 (Illumina Inc., CA, USA) as paired-end 150 bp reads. Sequences were analyzed with the SeqOne platform (Montpellier, France). Sanger sequencing was performed for variant confirmation and segregation. X-chromosome inactivation study was performed by methylation analyses using the HUMARA assay[16].

Individual 21, p.Arg169Ser
An Agilent Sure Select Target Enrichment System for the Clinical Research Exome was used to capture the regions of interest using genomic DNA isolated from blood. This method allows for analysis of greater than 98% of the targeted sequence. Analysis of data was performed using NextGENe software (SoftGenetics, State College, PA) along with an in-house bioinformatics pipeline. The data was reviewed with emphasis on novel alterations and those reported in the Human Gene Mutation Database (HGMD). All alterations of potential clinical relevance were confirmed by Sanger sequencing. Routinely observed drop-out and low coverage regions of the NGS data were also Sanger sequenced.

Individual 22, p.Arg441Trp
Exome sequencing for Individual 21 was done on DNA isolated from saliva swab. The Agilent SureSelect^{XT} Clinical Research Exome kit was used to target known disease-associated exonic regions of the genome (coding sequences and splice junctions of known protein-coding genes associated with disease, as well as an exomic backbone) using genomic DNA isolated from the patient. The targeted regions were sequenced using the Illumina NovaSeq^{TM} 6000 System with 150 bp paired-end reads. Using Illumina DRAGEN Bio-IT Platform® software, the DNA sequence was aligned and compared to the human genome build 19 (hg19/NCBI build 37). The average depth of coverage was calculated to be approximately 105X across all targeted regions. The emedgene® software was used to filter and analyze sequence variants identified in the patient.

Individual 24, p.Ser1173Asn
Patient blood was drawn into an EDTA blood collection tube for preparation by the HudsonAlpha Clinical Services Lab. Isolation of DNA from whole blood was performed using the QIAsymphony (Qiagen). Sequencing library was constructed from patient genomic DNA using the Illumina TruSeq PCR-free library preparation protocol and sequenced on the Illumina NovaSeq 6000 sequencer. DNA library fragments were sequenced from both ends (paired) with a read length of 150 base pairs. The patient genome was sequenced to a mean depth of 30X. Raw sequence data was demultiplexed and aligned to reference genome GRCh38 using Sentieon. Sequence variants were called using Sentieon DNAscope and loaded into a custom software analysis application for interpretation. All sequence variants were annotated with relevant information from established data sources to provide support for variant interpretation. Variant pathogenicity was determined using modified ACMG criteria.

Individual 25, p.Arg1324Trp
The Illumina TruSeq Nano DNA Library Prep Kit was used to prepare library for genome sequencing using the genomic DNA isolated from blood. Sequencing was performed on the Illumina NovaSeq™ 6000 System with 150 bp paired-end reads. Using Illumina DRAGEN Bio-IT Platform® software, the DNA sequence was aligned and compared to the human genome build 19 (hg19/NCBI build 37). The average depth of coverage across all genomic regions was calculated to be approximately 40X. The emedgene® software was used to filter and analyze sequence variants identified within the patient's genome sequencing data and to compare variants identified in the patient to the sequences of family members.

p.Arg1294Cys Cases
We note that p.Arg1294Cys has been observed in two individuals here (probands 16 and 20), each confirmed to be *de novo*. Each of these cases were submitted by sites in Europe, and we note that a ClinVar submission (SCV000297052.2) was made by the Children's Hospital of Philadelphia, USA. While this does appear to be a unique case from those described here, we cannot confirm this.

ACMG EVIDENCE CODES
As ACMG evidence codes are only applied to variants in established disease genes, we did not explicitly apply them here. If evidence codes were applied, most variants presented here would likely remain VUSs. All variants presented here are rare (PM2), and most are predicted damaging (PP3). The combination of these two codes would result in an overall status of VUS. For a few variants, additional evidence codes may apply. Functional studies in support of pathogenicity for p.Arg1274Trp (PS3) or in support of a benign status for p.Arg688His (BS3) may also apply. Additionally, while six variants are de novo (PS2), this code can only be applied if "The phenotype in the patient matches the genes' disease association with reasonable specificity."

COMPUTATIONAL MODELING

The wild-type 3D protein structure was downloaded from AlphaFoldDB (https://alphafold.ebi.ac.uk/)[17], which was included with the reference from UniProt (Accession number: Q14202). When not possible online, structures were visualized, colored and the sequence was mutated with Chimera version 1.15, rotamer builder tool[18]. Specifically, structure superposition was obtained in Chimera with the tool Matchmaker. Structure refinement was performed with the Chimera tool Dock Prep with standard settings, as previously described[19]. Depiction of molecular surfaces was defined as VdW surface and colored according to the electrostatic potential.

Wild-Type Protein model analysis: The distinction between organized and disordered regions was based upon the uniprot reference. The pLDDT value for each residue was extracted from the pdb file (opened as text file) from the B-factor field as prescribed on https://alphafold.ebi.ac.uk/about. The PAE matrix was downloaded from the main page of AlphaFoldDB and analyzed with the web-interface built-in tool.

Free folding energy estimation: the ΔΔG was calculated according to the SimBaNI model developed by Caldarau and colleagues[20]. For the SASA term, the RSA obtained with FreeSASA and naccess parameters were considered. Variations >1.0 kcal/mol in module were considered as significant.

Flexibility inspection: The flexibility analysis was performed by submission of pdb structures to CABS-flex2 with no restraints[21]. The results, expressed as RMSF were downloaded and a cutoff of 1 Å was considered as relevant.

Normal mode analysis: The normal mode analysis was performed by submission of the WT pdb file to WebNMA[22] and run in a comparative mode by separately setting the mutation parameters. The deformation energy per-residue was then downloaded and plotted. We applied a cutoff of 2 kcal/mol.

Surfaces calculation: the molecular and solvent accessible surfaces were calculated with two methods. First, both molecular (VdW radii) and solvent-accessible (SASA, probe radius 1.4 Å) total, polar and (by difference) non-polar surfaces were computed to calculate the variations. To this aim, VegaZZ suite was employed[23]. Then, relative-solvent-accessible surface area (RSA) was calculated by submitting the pdb files to FreeSASA with the parameters derived from naccess[24]. Burial-exposure mutant residue classification: the classification into buried/exposed mutant residues was based on the computed RSA values. By comparing the RSAs with the tabulated values[25] and selecting a cutoff of 0.20.

Eukaryotic Linear Motif identification: the UniProt accession (Q14202) was submitted to the online server ELM (http://elm.eu.org/) with standard settings (100 as probability cutoff, species Homo sapiens). The results identified about 107 ELMs. The results were exported as .csv file and the positions were intersected with the mutation sites.

CHIP-SEQ EXPERIMENTS

We edited the genomic DNA at the *ZMYM3* endogenous locus in HepG2 cells to introduce the variant of interest simultaneously with a 3X FLAG tag, 2A self-cleaving peptide, and neomycin resistance gene, using a modified version of the previously published CRISPR epitope tagging ChIP-seq (CETCh-seq) protocol[26]. HepG2 cells were sourced from ATCC (HB-8065) and cultured using the recommended protocol. We identified a CRISPR/Cas9 sgRNA targeting a DNA cleavage site near each variant

(Table S5) using established methods and cloned this sgRNA into pX330-U6-Chimeric_BB-CBh-hSpCas9 (Addgene plasmid #42230)[27]. We designed homology-directed repair donor templates composed of 400 bp of genomic sequence upstream (relative to coding direction of *ZMYM3*) of the cleavage site, exonic sequence surrounding the variant of interest, all exons downstream of the exon harboring the variant without the stop codon, the in-frame FLAG/P2A/NeoR cassette, and 400 bp of genomic sequence downstream of the cleavage site (Table S5). For control experiments, we repaired the genomic DNA with reference sequence at the variant position; for variant experiments, we used the variant nucleotide. For both experiments, an additional mutation was inserted to abolish the PAM site and block cleavage of edited sequence by Cas9. This mutation was a synonymous substitution for PAM sites in coding sequence. The donor templates were synthesized and cloned into the BamHI site of pUC19 by GenScript (Piscataway, NJ, USA). Genomic coordinates for the variant positions, the PAM site mutation positions, and the exons included in the super-exon for each experiment are shown in the table below.

Variant	**hg38 substitution**	**super-exon**	**PAM mutation**
p.Arg441Trp	chrX:71,249,610 G>A	chrX:71,249,461-71,249,679 chrX:71,249,020-71,249,169 chrX:71,248,686-71,248,799 chrX:71,248,438-71,248,524 chrX:71,248,163-71,248,312 chrX:71,247,734-71,247,904 chrX:71,247,346-71,247,510 chrX:71,246,595-71,246,690 chrX:71,246,354-71,246,512 chrX:71,245,986-71,246,098 chrX:71,245,669-71,245,842 chrX:71,245,340-71,245,483 chrX:71,244,790-71,244,891 chrX:71,244,303-71,244,470 chrX:71,243,829-71,243,978 chrX:71,242,971-71,243,084 chrX:71,242,171-71,242,422 chrX:71,241,229-71,241,342 chrX:71,240,919-71,241,107	chrX:71,249,611 G>A
p.Arg688His	chrX:71,247,819 C>T	chrX:71,247,734-71,247,904 chrX:71,247,346-71,247,510 chrX:71,246,595-71,246,690 chrX:71,246,354-71,246,512 chrX:71,245,986-71,246,098 chrX:71,245,669-71,245,842 chrX:71,245,340-71,245,483 chrX:71,244,790-71,244,891 chrX:71,244,303-71,244,470	chrX:71,247,824 G>A

		chrX:71,243,829-71,243,978 chrX:71,242,971-71,243,084 chrX:71,242,171-71,242,422 chrX:71,241,229-71,241,342 chrX:71,240,919-71,241,107	
p.Arg1274Trp	chrX:71,241,327 G>A	chrX:71,241,229-71,241,342 chrX:71,240,919-71,241,107	chrX:71,241,351 G>A

For each of the control and variant experiments, we nucleofected two million HepG2 cells with 5 ug total plasmid DNA (2.5 ug sgRNA/Cas9 plasmid and 2.5 ug respective donor plasmid) using a Lonza Nucleofector Kit V with an Amaxa Nucleofector 2. Immediately after nucleofection, each experiment was split into two wells of a 6-well plate, and these two replicates were recovered and grown separately. Two days post-nucleofection, we began selection with Geneticin (Invitrogen 10131) at 300 ug/mL. Cells were grown under selection for two weeks, and afterwards continued to expand in non-selective media for another three to four weeks. The p.Arg441Trp experiment did not survive selection, and no further work was performed on this variant. Genomic DNA was purified from cells and used as template for PCR and Sanger sequence validation of edits using a Qiagen DNeasy Blood & Tissue Kit (Table S5). Cells (20 million for each replicate) were crosslinked and harvested, immunoprecipitation with M2 FLAG monocolonal antibody (Sigma F1804) was performed, and sequencing libraries were constructed, all as previously described[28]. The libraries were pooled with other CETCh-seq libraries and sequenced on a NovaSeq S2 flowcell yielding total aligned read counts as shown in Table S4. We performed peak calling using SPP[29] and Irreproducible Discovery Rate (IDR)[30] using ENCODE-standardized pipelines for analysis and quality-control[31]. We also downsampled all replicate bam files to 20 million reads and performed peak calling, using either the standard IDR cutoff of 0.05 or a relaxed IDR cutoff of 0.2. We performed additional differential binding analyses using the R package csaw v1.28.0[32], using window widths of 10 nucleotides and background bins of 10,000 nucleotides, and using downsampled (to 20 million reads) bam files for all replicates. We generated Activity-by-Contact (ABC) v0.2 loop calls[33] using hg38 reference sequence and gene coordinates downloaded from the UCSC genome table browser, the ENCODE datasets ENCSR149XIL (DNase-seq), ENCSR000AMO (H3K27Ac), and ENCFF356LFX (blacklist), RNA-seq for HepG2 downloaded from Expression Atlas, and processed HepG2 Hi-C data from 4D Nucleome (4DNESC2DEQIJ).

The overlaps of peaks called from downsampled bam files (20 million reads) were 67.8% between ZMYM3$^{p.Arg1274Trp}$-control and ZMYM3$^{p.Arg1274Trp}$-variant, and 46.9% between ZMYM3$^{p.Arg688His}$-control and ZMYM3$^{p.Arg688His}$-variant. Knowing that peak overlaps suffer from missed calls at regions near threshold in experiments, we expanded the peak overlap analysis to use peaks called at the standard IDR cutoff of 0.05 in one experiment and peaks called at a relaxed IDR cutoff of 0.2 in the other experiment. This increased overlaps to 82.5% and 65.1%, respectively (Figure S10). To examine experiment similarity agnostic to peak calls, we performed read count correlations using each separate replicate for each experiment. Rather than performing read count correlations across the entire genome (an analysis that suffers from the

majority of regions being near background level of read counts), we filtered the analysis space to ZMYM3-specific regions. We determined the union of all peaks called in ZMYM3^{CETCh}, ZMYM3$^{p.Arg688His}$, and ZMYM3$^{p.Arg1274Trp}$, and used this set of regions for read counts. Reads from each bam file were determined at each genomic location and the Pearson correlation coefficient was calculated for each pairwise comparison (Figure S9). For p.Arg1274Trp, replicate experiments were highly correlated (r=0.80-0.91), and the correlation between control and variant experiments was high (r=0.74-0.79). For p.Arg688His, replicate experiments correlated similarly to correlations between control and variant (r=0.71-0.84). These results are consistent with the p.Arg688His experiments being of somewhat lower quality than p.Arg1274Trp experiments, and with no distinguishable difference between p.Arg688His control and variant beyond the noise of the assay.

IMMUNOCYTOCHEMISTRY EXPERIMENTS

We grew ZMYM3^{CETCh}, ZMYM3$^{p.Arg1274Trp}$-control, and ZMYM3$^{p.Arg1274Trp}$-variant cells in 6-well plates, seeding 50,000 cells per well, and in Millicell EZ Slides (Millipore), seeding 2,000 cells per chamber. After 24 hours, we performed fixation and permeabilization of cells on plates and slides using the Image-iT kit (Invitrogen R37602) following the manufacturer's protocol. For primary antibody, we used Sigma F1804 mouse anti-FLAG. For secondary (labeled) antibodies, we used Invitrogen T6390 goat anti-mouse IgG Texas Red-X, Abcam ab150116 goat anti-mouse IgG Alexa Fluor 594, Abcam ab6787 goat anti-mouse IgG Texas Red, or Invitrogen A21236 goat anti-mouse IgG Alexa Fluor 647. Each cell line/secondary antibody combination was performed in duplicate, and the complete experiment (from seeding to visualization) was performed twice. We imaged 6-well plates on an Evos FL microscope (ThermoFisher), and Millicell EZ Slides on a Nikon Eclipse Ti2 AX confocal microscope. In both replicates and both experiments, >95% of cells showed the localization patterns in Figure S7: ZMYM3^{CETCh} and ZMYM3$^{p.Arg1274Trp}$-control cells showed primarily nuclear localization of protein, and ZMYM3$^{p.Arg1274Trp}$-variant cells showed primarily cytoplasmic localization of protein.

SUPPLEMENTAL REFERENCES

1. Retterer, K., Juusola, J., Cho, M.T., Vitazka, P., Millan, F., Gibellini, F., Vertino-Bell, A., Smaoui, N., Neidich, J., Monaghan, K.G., et al. (2016). Clinical application of whole-exome sequencing across clinical indications. Genet Med *18*, 696–704.
2. Pavinato, L., Trajkova, S., Grosso, E., Giorgio, E., Bruselles, A., Radio, F.C., Pippucci, T., Dimartino, P., Tartaglia, M., Petlichkovski, A., et al. (2021). Expanding the clinical phenotype of the ultra-rare Skraban-Deardorff syndrome: Two novel individuals with WDR26 loss-of-function variants and a literature review. Am. J. Med. Genet. A *185*, 1712–1720.
3. Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., De Rubeis, S., An, J.Y., Peng, M., Collins, R., Grove, J., Klei, L., et al. (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell *180*, 568-584.e23.
4. Bauer, C.K., Calligari, P., Radio, F.C., Caputo, V., Dentici, M.L., Falah, N., High, F., Pantaleoni, F., Barresi, S., Ciolfi, A., et al. (2018). Mutations in KCNK4 that Affect Gating Cause a Recognizable Neurodevelopmental Syndrome. Am. J. Hum. Genet. *103*, 621–630.
5. Flex, E., Martinelli, S., Van Dijck, A., Ciolfi, A., Cecchetti, S., Coluzzi, E., Pannone, L., Andreoli, C., Radio, F.C., Pizzi, S., et al. (2019). Aberrant Function of the C-Terminal Tail of HIST1H1E Accelerates Cellular Senescence and Causes Premature Aging. Am. J. Hum. Genet. *105*, 493–508.
6. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res *20*, 1297–1303.
7. Paila, U., Chapman, B.A., Kirchner, R., and Quinlan, A.R. (2013). GEMINI: integrative exploration of genetic variation and genome annotations. PLoS Comput Biol *9*, e1003153.
8. Alfaiz, A.A., Micale, L., Mandriani, B., Augello, B., Pellico, M.T., Chrast, J., Xenarios, I., Zelante, L., Merla, G., and Reymond, A. (2014). TBC1D7 Mutations are Associated with Intellectual Disability, Macrocrania, Patellar Dislocation, and Celiac Disease. Hum. Mutat. *35*, 447–451.
9. Delafontaine, J., Masselot, A., Liechti, R., Kuznetsov, D., Xenarios, I., and Pradervand, S. (2016). Varapp: A reactive web-application for variants filtering. BioRxiv 060806.
10. Cohen, A.S.A., Farrow, E.G., Abdelmoity, A.T., Alaimo, J.T., Amudhavalli, S.M., Anderson, J.T., Bansal, L., Bartik, L., Baybayan, P., Belden, B., et al. (2022). Genomic answers for children: Dynamic analyses of >1000 pediatric rare disease genomes. Genet. Med. *24*, 1336–1348.
11. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med *17*, 405–424.
12. Motta, M., Fasano, G., Gredy, S., Brinkmann, J., Bonnard, A.A., Simsek-Kiper, P.O.,

Gulec, E.Y., Essaddam, L., Utine, G.E., Guarnetti Prandi, I., et al. (2021). SPRED2 loss-of-function causes a recessive Noonan syndrome-like phenotype. Am. J. Hum. Genet. *108*, 2112–2129.
13. Lin, Y.C., Niceta, M., Muto, V., Vona, B., Pagnamenta, A.T., Maroofian, R., Beetz, C., van Duyvenvoorde, H., Dentici, M.L., Lauffer, P., et al. (2021). SCUBE3 loss-of-function causes a recognizable recessive developmental disorder due to defective bone morphogenetic protein signaling. Am. J. Hum. Genet. *108*, 115–133.
14. Beaulieu, C.L., Majewski, J., Schwartzentruber, J., Samuels, M.E., Fernandez, B.A., Bernier, F.P., Brudno, M., Knoppers, B., Marcadier, J., Dyment, D., et al. (2014). FORGE Canada consortium: Outcomes of a 2-year national rare-disease gene-discovery project. Am. J. Hum. Genet. *94*, 809–817.
15. Hamilton, A., Tétreault, M., Dyment, D.A., Zou, R., Kernohan, K., Geraghty, M.T., FORGE Canada Consortium, Care4Rare Canada Consortium, Hartley, T., and Boycott, K.M. (2016). Concordance between whole-exome sequencing and clinical Sanger sequencing: implications for patient care. Mol. Genet. Genomic Med. *4*, 504–512.
16. Bertelsen, B., Tümer, Z., and Ravn, K. (2011). Three new loci for determining x chromosome inactivation patterns. J. Mol. Diagn. *13*, 537–540.
17. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589.
18. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.
19. Rossi Sebastiano, M., Ermondi, G., Hadano, S., and Caron, G. (2022). AI-based protein structure databases have the potential to accelerate rare diseases research: AlphaFoldDB and the case of IAHSP/Alsin. Drug Discov. Today *27*, 1652–1660.
20. Caldararu, O., Blundell, T.L., and Kepp, K.P. (2021). Three Simple Properties Explain Protein Stability Change upon Mutation. J. Chem. Inf. Model. *61*, 1981–1988.
21. Kuriata, A., Gierut, A.M., Oleniecki, T., Ciemny, M.P., Kolinski, A., Kurcinski, M., and Kmiecik, S. (2018). CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. Nucleic Acids Res. *46*, W338–W343.
22. Tiwari, S.P., Fuglebakk, E., Hollup, S.M., Skjærven, L., Cragnolini, T., Grindhaug, S.H., Tekle, K.M., and Reuter, N. (2014). WEBnm@ v2.0: Web server and services for comparing protein flexibility. BMC Bioinformatics *15*,.
23. Pedretti, A., Mazzolari, A., Gervasoni, S., Fumagalli, L., and Vistoli, G. (2021). The VEGA suite of programs: an versatile platform for cheminformatics and drug design projects. Bioinformatics *37*, 1174–1175.
24. Mitternacht, S. (2016). FreeSASA: An open source C library for solvent accessible surface area calculations. F1000Research *5*,.
25. Creighton, T.E. (2013). Proteins : structures and molecular properties (New York, NY, USA: Freeman).
26. Savic, D., Partridge, E.C., Newberry, K.M., Smith, S.B., Meadows, S.K., Roberts, B.S., Mackiewicz, M., Mendenhall, E.M., and Myers, R.M. (2015). CETCh-seq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins. Genome Res. *25*, 1581.
27. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using

CRISPR/Cas systems. Science *339*, 819–823.
28. Meadows, S.K., Brandsmeier, L.A., Newberry, K.M., Betti, M.J., Nesmith, A.S., Mackiewicz, M., Partridge, E.C., Mendenhall, E.M., and Myers, R.M. (2020). Epitope tagging ChIP-seq of DNA binding proteins using CETCh-seq. Methods Mol. Biol. *2117*, 3–34.
29. Kharchenko, P. V., Tolstorukov, M.Y., and Park, P.J. (2008). Design and analysis of ChIP-seq experiments for DNA-binding proteins. Nat. Biotechnol. 2008 2612 *26*, 1351–1359.
30. Li, Q., Brown, J.B., Huang, H., and Bickel, P.J. (2011). Measuring reproducibility of high-throughput experiments. Https://Doi.Org/10.1214/11-AOAS466 *5*, 1752–1779.
31. Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., et al. (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res. *22*, 1813–1831.
32. Lun, A.T.L., and Smyth, G.K. (2016). csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. Nucleic Acids Res. *44*, e45.
33. Nasser, J., Bergman, D.T., Fulco, C.P., Guckelberger, P., Doughty, B.R., Patwardhan, T.A., Jones, T.R., Nguyen, T.H., Ulirsch, J.C., Lekschas, F., et al. (2021). Genome-wide enhancer maps link risk variants to disease genes. Nature *593*, 238–243.