SFARI GENE Q2/2026 REPORT

Eric Larsen, Senanu Spring-Pearson, Christina Printzis, Girish Nagendra, and Sharmila Banerjee-Basu

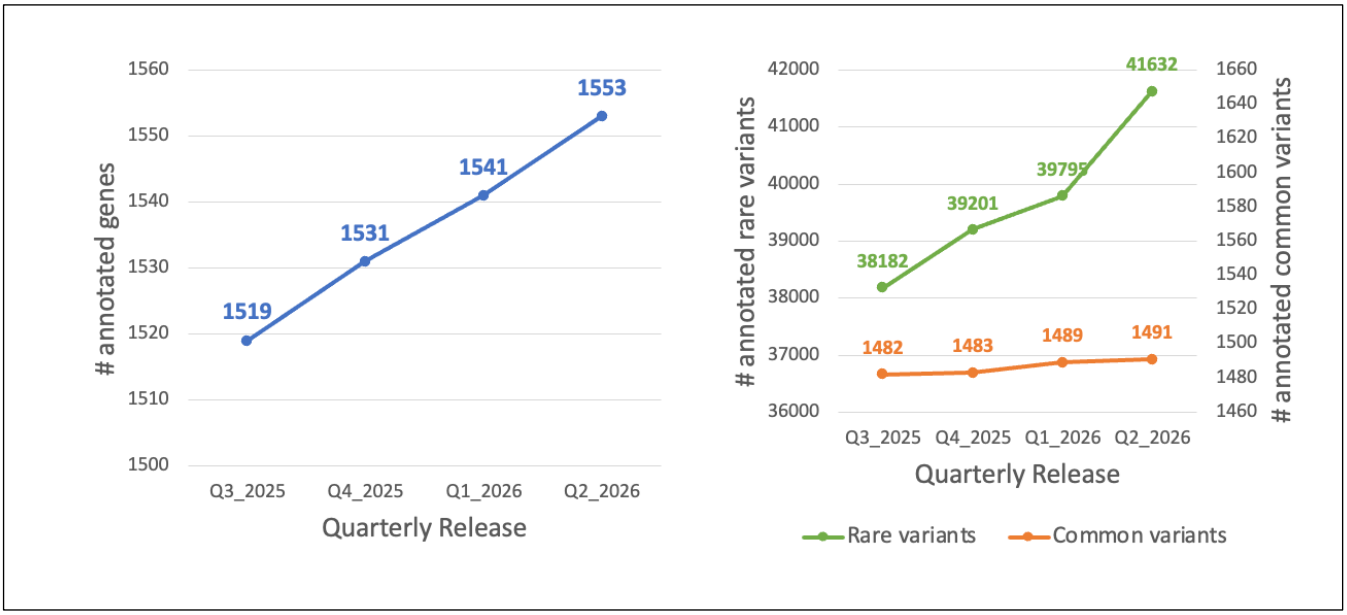
MindSpec Inc.
April – June 2026

Table of Contents

1. Human Gene Module
Quarterly Report, Q2/2026

1.1 Updated Human Gene Dataset

A total of ***twelve*** new genes were added to the Human Gene Module for the Q2/2026 release, bringing the overall number of ASD candidate genes in the module to **1,553** (HG_Figure 1, panel A). In-depth annotation of **1837** rare variants and two common variants were completed in this quarter leading to a total of **41,632** rare and **1,491** common variants, respectively (HG_Figure 1, panel B). Annotation of **122** new references was accomplished for the Human Gene module in Q2/2026, bringing the total number of references to **6,933**.



HG_Figure 1. Number of ASD-linked genes and variants in the Human Gene Module over the last four quarters. (A) The number of genes has grown from 1,519 to 1,553 (B) The number of rare variants has increased from 38,182 to 41,632; number of common variants has increased from 1,482 to 1,491.

1.2 Highlights of Q2/2026 Human Gene Dataset

The identification and evaluation of emerging ASD risk genes in the Q2/2026 release were guided by an integrative evidence-synthesis approach that combines multiple lines of genetic and functional evidence across large-scale and targeted studies. Gene-level assessment incorporated variants from established ASD cohorts including Autism Sequencing Consortium (ASC), Simons Simplex Collection (SSC), SPARK, and MSSNG, alongside data from ancestrally diverse and broad neurodevelopmental cohorts. Notably, recent analyses from the Genetics of Autism in Latin Americans (GALA) Consortium, comprising more than 15,000 individuals including 4,717 participants with ASD, provided an important cohort to evaluate the generalizability of ASD risk genes across populations. Many genome-wide significant genes identified in the Latin American cohort, including PTEN, SHANK3, SCN2A, CHD8, and SYNGAP1, had previously been implicated in predominantly European cohorts, supporting the shared contribution of deleterious coding variation to ASD risk across ancestries. At the

same time, the study identified additional candidate genes, including GSE1, GLS, and OTULIN, which were incorporated into the Q2/2026 dataset.

Each candidate gene was evaluated across variant classes, inheritance patterns, and independent cohorts, with emphasis on replication and convergence of genetic signals. Functional studies, model systems, and detailed clinical phenotyping provide complementary evidence supporting biological relevance. This multi-dimensional evidence synthesis framework enables rigorous evaluation of gene discovery, replication, and biological relevance, supporting the prioritization of both newly emerging candidate genes and established ASD-associated genes within the SFARI Gene database.

The Q2/2026 dataset is described in the following sections:

- **New genes added in this quarter (HG_Table 1)**
- **Summary evidence for new genes (Section 1.3)**
- **New reports added to existing genes (HG_Table 2)**
- **New articles added to the SLOE dataset (HG_Table 3)**

HG_Table 1. New genes added in Q2/2026

Gene Symbol	Rare single gene variant	Syndromic	Functional
MDGA1	√		**1, 2**
UBA7	√		**1**
CDC42BPA	√		
GSE1	√		
GLS	√		
OTULIN	√		
DOP1A	√		**2**
RANBP9	√		**3**
KSR2	√		
NUCLEOLIN	√		
FAM53C	√		**3**
MADD	√	√	

Functional evidence categories:
1.In vitro characterization of disease-associated variants using transfected cell lines or patient-derived fibroblasts.
2.Mouse model recapitulating core ASD-associated phenotypes
•MDGA1: mice with a patient-derived compound heterozygous variant exhibited impaired ultrasonic vocalizations.
•DOP1A: behavioral abnormalities, including impaired discrimination of novel and familiar mice in the 3-chamber social interaction assay, in both knockout mice and mice harboring a patient-derived p.Arg355Cys variant.
3.Interaction with known ASD candidate gene(s)
•MDGA1 inhibited interaction of neuroligin-2 with neurexin.
•RANBP9 interacted with MET and enhanced HGF-MET signaling by recruiting Sos and activating the Ras pathway.
•FAM53C interacted with and inhibited DYRK1A.

1.3 Summary evidence for new genes

MDGA1
Kim et al., 2026 reported two individuals with ASD carrying compound heterozygous missense mutations in this gene [(p.Val116Met/p.Ala688Val) and (p.Tyr635Cys/p.Glu756Gln)]. In utero overexpression of MDGA1 Val116Met/p.Ala688Val in murine altered cortical neuron migration and impaired ultrasound vocalization, while MDGA1 p.Tyr635Cys/p.Glu756Gln disrupted the triangular extracellular structure of MDGA1 and rendered it unable to impact GABAergic synapses in hippocampal CA1 neurons. Furthermore, Kim et al., 2026 found that male Mdga1 knock-in mice harboring the p.Tyr653Cys/p.Glu756Gln mutation exhibited impaired ultrasonic vocalizations and sensorimotor gating, similar to male Mdga1 conditional knockout mice, while no behavioral deficits were seen in female counterparts. Additional de novo and inherited heterozygous variants in the MDGA1 gene were previously identified in ASD probands from the Simons Simplex Collection, the Autism Sequencing Consortium, the SPARK cohort, the MSSNG cohort, and the mAGRE cohort (Lim et al., 2017; Zhou et al., 2022; Fu et al., 2022; Trost et al., 2022; Cirnigliaro et al., 2023). Pettem et al., 2013 demonstrated that MDGA1 inhibited the synapse-promoting activity of neuroligin-2, without altering neuroligin-2 surface trafficking, by inhibiting interaction of neuroligin-2 with neurexin.

UBA7
Bandi et al., 2026 reported three unrelated individuals with homozygous truncating variants in the UBA7 gene presenting with a neurodevelopmental disorder characterized by developmental delay, intellectual disability, autism, and nonspecific dysmorphic features; truncating variants resulted in loss of catalytic activity, protein stability, and localization, and patient fibroblasts with the p.Lys709SerfsTer45 variant demonstrated reduced UBA7 transcript, production of a truncated and unstable UBA7 protein, and an inability to induce ISGylation upon interferon beta treatment, indicating a dysfunctional ISGylation system. De novo and inherited heterozygous loss-of-function and missense variants have been identified in ASD probands from the Autism Sequencing Consortium, the SPARK cohort, and the mAGRE cohort (Satterstrom et al., 2020; Trost et al., 2022; Cirnigliaro et al., 2023).

CDC42BPA
A maternally inherited complex genomic rearrangement involving duplication of seven exons and a nested deletion that partially deleted one exon of the CDC42BPA gene was identified by long-read whole-genome sequencing in Mortazavi et al., 2026 in an individual with a diagnosis of ASD and intellectual disability (REACH000529); this rearrangement was predicted to result in protein truncation. Multiple de novo missense variants and a de novo splice-region variant in CDC42BPA have been previously reported in ASD probands from the Simons Simplex Collection, the SPARK cohort, the Autism Sequencing Consortium, and the MSSNG cohort (De Rubeis et al., 2014; Yuen et al., 2017; Satterstrom et al., 2020; Zhou et al., 2022).

GSE1
TADA analysis of >15,000 Latin American individuals from the GALA Consortium (including 4,717 participants with an ASD diagnosis) in Natividad Avila et al., 2026 identified GSE1 as one of 35 genes reaching genome-wide significance (FDR < 0.05); among the Admixed American (AMR) individuals with ASD described in this report were three SPARK probands with de novo loss-of-function variants (one of whom was previously reported in Fu et al., 2022) and a Simons Simplex Collection proband with a de novo synonymous variant originally described in Iossifov et al., 2014. Additional de novo missense variants in this gene were reported in non-AMR ASD probands from the Simons Simplex Collection, the SPARK cohort, and the Sanger_UK10K cohort, as well as in an ASD proband from the MSSNG cohort (De Rubeis et al., 2014; Iossifov et al., 2014; Yuen et al., 2017; Zhou et al., 2022). Inherited loss-of-function and damaging missense (Mis3) variants in GSE1 have also been reported in ASD probands from the Autism Sequencing Consortium and the mAGRE cohort (De Rubeis et al., 2014; Cirnigliaro et al., 2023).

GLS
TADA analysis of >15,000 Latin American individuals from the GALA Consortium (including 4,717 participants with an ASD diagnosis) in Avila et al., 2026 identified GLS as one of 35 genes reaching genome-wide significance (FDR < 0.05); the two Admixed American (AMR) individuals with ASD included in this analysis were SPARK probands originally reported in Zhou et al., 2022 with de novo missense variants with MPC scores > 2. Zhou et al., 2022 also reported two additional de novo GLS variants in SPARK probands that were not included in the Avila et al., 2026 TADA analysis (one loss-of-function variant, one missense variant).

OTULIN
TADA analysis of >15,000 Latin American individuals from the GALA Consortium (including 4,717 participants with an ASD diagnosis) in Avila et al., 2026 identified OTULIN as one of 35 genes reaching genome-wide significance (FDR < 0.05); the two Admixed American (AMR) individuals with ASD included in this analysis were SPARK probands with de novo missense variants with MPC scores > 2. Additional de novo variants in OTULIN have been reported in ASD probands from the Simons Simplex Collection and the mAGRE cohort (Krupp et al., 2017; Zhou et al., 2022; Cirnigliaro et al., 2023).

DOP1A
Ariyama et al., 2026 identified 11 patients from 8 families with either de novo heterozygous or homozygous variants in DOP1A presenting with neurodevelopmental phenotypes, including developmental delay in all individuals and intellectual disability in 8 individuals, as well as autistic features in 4 individuals and stereotypic behavior in two others; in the same report, the authors observed behavioral abnormalities, including impaired discrimination of novel and familiar mice in the 3-chamber social interaction assay, in both Dop1a knockout mice and mice harboring a patient-derived p.Arg355Cys variant. De novo variants in DOP1A, including a de novo loss-of-function variant and four de novo missense variants, have been identified in ASD probands from the SPARK cohort, the Autism Sequencing Consortium, and a Chinese ASD cohort (Zhou et al., 2022; Fu et al., 2022; Trost et al., 2022; Wang et al., 2023; Tan et al., 2025).

RANBP9
Whole-exome sequencing of 115 ASD trios from three countries with a high frequency of consanguineous populations (Pakistan, Iran, and Saudi Arabia) in Harripaul et al., 2026 identified a homozygous missense variant in the RANBP9 gene (NM_005493.3:c.1735G>C;p.Gly579Arg) in a female ASD proband from Pakistan. De novo variants in this gene, including a de novo loss-of-function variant, have been previously identified in ASD probands from the Simons Simplex Collection, the SPARK cohort, the Autism Sequencing Consortium, and a Japanese ASD cohort (Iossifov et al., 2014; Lim et al., 2017; Takata et al., 2018; Zhou et al., 2022; Fu et al., 2022). RANBP9 has been shown to interact with the ASD candidate gene MET and enhance HGF-MET signaling by recruiting Sos and activating the Ras pathway (Wang et al., 2022).

KSR2
Whole-exome sequencing of 115 ASD trios from three countries with a high frequency of consanguineous populations (Pakistan, Iran, and Saudi Arabia) in Harripaul et al., 2026 identified a homozygous missense variant in the KSR2 gene (NM_173598.6:c.763C>T;p.Arg255Trp) in a male ASD proband from Iran. De novo variants in this gene, including a de novo loss-of-function variant, have been reported in ASD probands from the SPARK cohort and a Chinese ASD cohort (Zhou et al., 2022; Trost et al., 2022; Wang et al., 2023; Avila et al., 2026).

NUCLEOLIN
Whole-exome sequencing of 115 ASD trios from three countries with a high frequency of consanguineous populations (Pakistan, Iran, and Saudi Arabia) in Harripaul et al., 2026 identified a de novo frameshift variant in the NUCLEOLIN gene (NM_005381.3:c.1991delG;p.Gly664GlufsTer70) in a male ASD proband from Iran. Additional de novo variants in this gene have been previously reported in ASD probands from the Autism Sequencing Consortium, a Brazilian ASD cohort, and a Korean ASD cohort (Montenegro et al., 2020; Fu et al., 2022, Kim et al., 2024).

FAM53C
Whole-exome sequencing of 115 ASD trios from three countries with a high frequency of consanguineous populations (Pakistan, Iran, and Saudi Arabia) in Harripaul et al., 2026 identified a de novo frameshift variant in the FAM53C gene (NM_001135647.2:c.256dupT;p.Ser86PhefsTer14) in a male ASD proband from Pakistan. A de novo missense variant in this gene was previously reported in an ASD proband from the MSSNG cohort in Trost et al., 2022. FAM53C has been shown to interact with and inhibit the ASD candidate gene DYRK1A, and FAM53C knockout human cortical organoids have been shown to display increased cell cycle arrest and growth defects (Miyata et al., 2023; Hammond et al., 2026).

MADD
Whole-exome sequencing of 115 ASD trios from three countries with a high frequency of consanguineous populations (Pakistan, Iran, and Saudi Arabia) in Harripaul et al., 2026 identified a homozygous missense variant in the MADD gene (NM_001135943.2:c.1883A>G;p.Tyr628Cys) in a male ASD proband from Pakistan. Anazi et al., 2017 had previously reported two unrelated individuals with biallelic MADD variants presenting with developmental delay and poor eye contact; one of these individuals was also reported to have a diagnosis of autism spectrum disorder. Schneeberger et al., 2020 reported that, of the 14 patients with Group 1 MADD-associated disorder (DEEAH syndrome), one presented with autism, while of the nine individuals with Group 2 MADD-associated disorder (neurodevelopmental disorder with dysmorphic facies, impaired speech and hypotonia), one presented with ASD and another with stereotypic movements. De novo heterozygous missense variants in MADD have been previously reported in ASD probands from the Simons Simplex Collection and the SPARK cohort (Iossifov et al., 2014; Feliciano et al., 2019; Satterstrom et al., 2020; Zhou et al., 2022; Avila et al., 2026), while inherited loss-of-function variants in this gene were observed in multiple individuals with ASD from the iHART and mAGRE cohorts (Ruzzo et al., 2019; Cirnigliaro et al., 2023).

1.4 New evidence supporting existing genes in Q2/2026

RNU4-2 and RNU2-2
In Q2, we incorporated multiple studies that substantially expand the genetic architecture and clinical spectrum of RNU4-2 and RNU2-2, two spliceosomal small nuclear RNA genes that have recently emerged as important contributors to neurodevelopmental disorders with ASD as a frequent clinical feature. These studies added newly characterized pathogenic variants, expanded patient cohorts, refined genotype–phenotype relationships, and provided additional functional evidence, considerably strengthening support for both genes.

Conceptually, these studies establish that both RNU4-2 and RNU2-2 exhibit distinct autosomal dominant and autosomal recessive disease mechanisms. For RNU4-2, ASD has been reported in approximately 48% of individuals with the dominant ReNU syndrome (Chen *et al*., 2024; Nava *et al*., 2025), compared with approximately 29% of individuals with the recessive disorder (Rius *et al*., 2026). Similarly, recurrent de novo variants in RNU2-2 define an autosomal dominant neurodevelopmental syndrome with autistic behavior as a prominent clinical feature (Greene *et al*., 2025), while three independent international studies from the United Kingdom, France, and collaborative global cohorts (Greene *et al*., 2026; Jackson *et al*., 2026; Leitao *et al*., 2026) established biallelic RNU2-2 variants as a frequent cause of recessive developmental and epileptic encephalopathy, with approximately 52% of affected individuals presenting with ASD. Together, these studies substantially strengthen the evidence supporting RNU2-2 through independent replication, expanded patient cohorts, transcriptomic characterization, and refined genotype–phenotype correlations. Finally, Turner *et al*. (2026) demonstrated a highly significant enrichment of de novo RNU2-2 variants in males with ASD from the Simons Simplex Collection and SPARK cohorts ($P < 2.2 \times 10^{-16}$), with de novo variants identified in approximately 0.2% of individuals with ASD compared with 0.05% of unaffected individuals. Collectively, these findings considerably strengthen the evidence supporting RNU4-2 and RNU2-2 as ASD-associated non-coding genes and highlight the growing contribution of spliceosomal RNAs to autism genetics.

In parallel, the Q2/2026 update incorporated evidence from recent studies that strengthen the characterization of established ASD-associated genes. These additions include large multi-individual cohort expansions, detailed case series, and integrative functional studies spanning in vitro systems, organoids, and in vivo animal models. Newly curated clinical reports further refine phenotypic profiles and inheritance patterns, including variable expressivity and emerging genotype–phenotype correlations. Selected examples of new evidence added to existing genes are shown below.

ERBB4
Rare deletions affecting the ERBB4 gene have been identified with autism (Pinto et al., 2010) as well as schizophrenia (Walsh et al., 2008). More recently, **Mademont-Soler et al., 2026** described 24 new patients from 13 unrelated families with predicted null variants in the ERBB4 gene and clinical manifestations characterized by intellectual disability/neurodevelopmental delay (71%), speech delay (70%), ADHD (47%), ASD (39%), and/or epilepsy/seizures (22%).

PTCHD1
A recent study from **Ko et al., 2026** identified exon 3 of PTCHD1 as a high-risk locus through mapping of clinically reported loss-of-function mutations in human patients; subsequent characterization of a Ptchd1 knockout mouse model targeting Ptchd1 exon 3 (Ptchd1Δexon3) demonstrated loss of both Ptchd1_a (full-length) and Ptchd1_c (shorter) transcripts accompanied by dysregulation of social and communication behaviors, increased repetitive behavior, and motor and learning impairments.

CEP41
Functional assessment of the ASD-associated CEP41 p.Arg242His variant in human cortical organoids indicated this variant affected the development of projection neurons and interneurons (**Hasenpusch-Theil et al., 2026**).

PTCHD1-AS
Bradley et al., 2026 examined whole-genome sequencing data in ASD cases (9,349) and controls (8,332) and identified 27 male individuals with ASD carrying microdeletions within PTCHD1-AS exons that were associated with an increased risk of ASD (odds ratio 2.56, P=0.01). In the same report, the authors found that two mouse models disrupting the canonical exon 3 of Ptchd1-as showed ASD-like features in male mice, including increased repetitive behaviors and impaired social interaction and communication without cognitive comorbidities or ADHD-like behaviors.

KDM5B
El Hayak et al., 2026 assembled a cohort of 24 additional individuals from 21 families with either monoallelic or biallelic KDM5A variants who presented with severe speech impairment and intellectual disability, often alongside autism spectrum disorder and other neurodevelopmental features; the authors also demonstrated reduced KDM5A protein levels in cell lines derived from a subset of affected individuals.

MSL2
Multiple de novo variants in the MSL2 gene, including two de novo loss-of-function variants and several de novo missense variants, have been identified in ASD probands (Iossifov et al., 2014; Zhou et al., 2022; Trost et al., 2022; Costa et al., 2023). Two de novo loss-of-function variants in this gene were also identified in Chinese NDD probands (Zhang et al., 2021). More recently, **Luo et al., 2026** demonstrated using a nervous system-specific Msl2 conditional knockout mouse model that Msl2 deficiency led to impaired social novelty recognition, learning deficits, and spatial memory impairments, and that these behavioral abnormalities were characterized by disrupted neocortical lamination, compromised proliferation capacity and differentiative ability of neural progenitor cells, and reduced neuronal migration. Furthermore, Luo et al., 2026 reported that an ASD-associated missense variant located in the MSL2 DNA-binding domain (p.Gly465Val) exhibited impaired FMR1 promoter binding.

HG_Table 2. Selected examples of new studies added to existing genes in Q2/2026

Gene	Title	First Author	Journal	Year
FOXP1	Resolving non-coding splice-altering variants using an integrative genomic and transcriptomic workflow: application to FOXP1	Planté-Bordeneuve P	Hum Genomics	2026
POGZ	Reciprocal regulation between autism risk gene POGZ and circadian clock	Wu T	JCI Insight	2026
TANC2	De novo TANC2 stop-loss variant associated with developmental impairment and drug-resistant epilepsy	Hintermayer MA	Epileptic Disord	2026
ZBTB18	Adaptive evolution of gene regulatory networks in mammalian neocortex	Li Z	Nature	2026
CHD2	Uncoupling memory impairments from autism-associated behaviors in Chd2 deficient mice	Yoon SH	Mol Psychiatry	2026
RNU2-2	Biallelic variants in RNU2-2 cause the most prevalent known recessive neurodevelopmental disorder	Greene D	Nat Genet	2026
RNU2-2	Biallelic variants in RNU2-2 cause a remarkably frequent developmental and epileptic encephalopathy	Jackson A	Nat Genet	2026
RNU2-2	Systematic analysis of snRNA genes reveals frequent RNU2-2 variants in dominant and recessive developmental and epileptic encephalopathies	Leitão E	Nat Genet	2026

1.5 Human Gene Standardization

We continue to standardize the allele change and residue change data fields to the terminology developed by the Human Genome Variation Society (HGVS) for the Human Gene dataset. All new variant annotations were performed with standardized terminology and include genomic coordinates in GRCh38 genome build for allele change, residue change and correct genome build.

1.6 Development of a Candidate Gene Pool

The Single Line of Evidence (SLOE) dataset comprises genes and variants reported in the scientific literature that lack sufficient evidence to meet the established inclusion criteria for the database. This dataset serves as a repository for genetic findings that may have potential relevance but have not been fully validated or sufficiently supported by robust data. Regular queries of the SLOE dataset play a crucial role in gene selection process by facilitating cross-referencing and validating newly identified ASD candidate genes. By systematically examining the SLOE dataset, we compare recently discovered variants against existing evidence, allowing for a thorough assessment of their relevance and significance. In Q2/2026, we updated SLOE dataset by adding three new articles, detailed in HG_Table 3.

HG_Table 3: New articles added to the SLOE dataset.

Title	First Author	Journal/Book	Publication Year
Autism spectrum disorder trios from consanguineous populations are enriched for rare homozygous variants, identifying 32 new candidate genes	Harripaul R	Sci Rep	2026
Genetic variants identification through whole-genome sequencing based on dried blood spots in 92 Chinese children with Autism	Wang L	Hum Genomics	2026
Deleterious coding variation associated with autism is shared across ancestries	Natividad Avila M	Nat Med	2026