**The American Journal of Human Genetics, Volume *101***

## Supplemental Data

# Exonic Mosaic Mutations Contribute Risk for Autism Spectrum Disorder

**Deidre R. Krupp, Rebecca A. Barnard, Yannis Duffourd, Sara A. Evans, Ryan M. Mulqueen, Raphael Bernier, Jean-Baptiste Rivière, Eric Fombonne, and Brian J. O'Roak**

**Supplemental Material and Methods**

**Rare Inherited Variant Simulation**

Variants were required to have an exonic or splicing annotation, population frequency <0.5%, at least 8 reads in all family members, and either 4+ variant reads or 3+ variant reads and allele fraction (AF) >=5% in at least one parent and one child. Variants were excluded if on sex chromosomes, if identified solely by mPUP, or if they had putative mosaic status with AF confidence interval < 40% (*in the parental data only*). This produced a final set of 1,554,918 rare inherited germline SNVs. Indels were treated similarly, then intersected with published calls to produce a final set of 13,479 rare inherited indels.[1] Counts per child are: SNVs-1,103,102 in probands, 825,098 in siblings; Indels-9,782 in probands, 7,197 in siblings.

Variants were divided on their presence in probands or siblings and sampled separately using the R function *sample()* with the Knuth-TAOCP-2002 random number generator. Sampled variants were tested for significant difference from heterozygosity (binomial p <= 0.001 or p <= 0.0001), with lower and higher AF tails evaluated separately, and a count of skewed variants determined for each trial. A total of 10,000 trials were performed for each child. Subsequently, the counts per child were added across trials to obtain distributions of total skewed variants that could be compared to the observed skewing in previously published *de novo* mutations.

**Evaluating Callers with Simulated Data**

These data consisted of 202 synthetic variants in 101 nucleotide single-end Illumina reads generated by simNGS, with variant frequencies ranging from 1-50% and coverage depths (DP) of 30-500 reads.[2] Reads were aligned to the GRCh37-hg19 Broad variant human reference using BWA (0.5.6, 0.7.12)[3] and BWA-mem (0.7.12), and mpileups generated using samtools (1.1).[4] Given that read coverage peaked at variant sites and tapered off over surrounding bases, we only counted bases having at least 90% of the target depth. Callers included: VarScan (2.3.2, 2.3.7)[5], LoFreq (0.4.0, 2.1.1)[6], Atlas2 (1.4.1, 1.4.3)[7], and an in-house mpileup parsing script, referred to as mPUP. For all callers, we required a minimum mapping quality (MAPQ) of 29 and DP >= 8, and disabled samtools base adjusted quality (BAQ). Additional parameters per caller were: VarScan, *--min-var-freq 1x10$^{-15}$ --p-value 0.1*; LoFreq, *--no-default-filter*; mPUP, *-m -c 8 -v 2*. For mPUP calls, a significant difference from the empirical error rate (in simulated data) of 0.005 (binomial p <= 0.005) was required. All caller versions were run on all combinations of variant frequency, coverage depth, and aligner version. Caller performance was evaluated on sensitivity, positive predictive value (PPV), and F-score (beta = 0.5) for each condition.

**Raw Variant Calling**

For all pilot and full cohort analyses, variants were called on individual samples using VarScan 2.3.2, LoFreq 2.1.1, and our in-house script mPUP. Variant calling was performed as described above, with the exception that no error rate test was utilized for mPUP calls in order to maximize sensitivity. Reference and variant allele counts were extracted from mpileups for all family members at all family variant sites using a custom script (*samtools mpileup -B -d 1500 | mPUP -m -q 20 -a count*).

**Initial Variant Filtering: Pilot 24**

To build a systematic PMM calling pipeline, detailed evaluation of the high depth pilot 24 dataset was performed first (Figures S2-S8). The combined annotated raw calls were classified for germline versus mosaic status. Variants with AFs significantly below 50% (binomial p <= 0.001) were considered putative PMMs. For putative transmitted parental PMMs, which also had skewed AFs in child(ren), a significant difference between parent and child AF (Fisher's exact p <= 0.01), with child AF > parental AF was required. Only PMM (child or parental) or GDM calls were considered for validation. For validation sites, we required at least four variant reads with total AF >= 3% or at least three variant reads with AF >= 5% and DP >= 8 in all family members. We removed variants that were: present in the raw calls of more than one of the pilot 24 families, noncoding or non-canonical splicing annotations, or having population frequency >= 0.5% in any reference (Supplemental Note: Model Development). Previously published GDMs[1; 8] were added to the validation set if not identified by our pipeline (19/259 SNVs, 13 of which were called as raw variants but removed by pipeline filters).

**smMIP Design**

Single molecule molecular inversion probes (smMIPs) were designed against candidate variant sites similarly to the method described in O'Roak et al. 2012[9] using MIPGEN[10] (11-25-14 release) with the following parameters: 1) human reference genome GRCh37-hg19 Broad variant, 2) arm length sums 40-44, 3) arm copy product <= 10, 4) min and max capture size 91, 5) three bases degenerate tags on either side of the MIP backbone (total 6Ns), 6) at least five bases flanking target (feature) site, 7) logistic priority score of 0, 8) 60 base maximum overlap between smMIPs, 9) repetitive motifs flagged using Tandem Repeat Finder 4.07b, and 9) smMIPs flagged if arms overlapped a SNP with minor allele frequency >= 0.1% in dbSNP141. A custom picking script was used to select the highest-scoring smMIPs from all designed candidates, with up to four mips covering each validation target and at least one smMIP on each strand where possible. We also required picked smMIPs have at least two base flanking the target site and that smMIP arms be free of recognition motifs for the restriction enzymes StyD41 (CCNGG) and NlaIII (CATG). Probes containing SNPs in targeting arms were accepted only if no others could be designed for the target and provided exome data from the associated family did not contain the problematic SNP; otherwise, SNP MIPs were excluded. If fewer than two smMIPs could be designed for a given site using these parameters, MIPGEN was re-run with the arm copy count first increased to 75. Finally, if probes were still lacking the arm copy count increased to 200 with tandem repeat finder disabled.

Picked smMIPs were divided into pools according to the families they targeted, with roughly equal probe counts in each pool (between 200-1100 probes/pool, Table S3). Pool-specific 20 base PCR adapters were appended to each smMIP arm, with NlaIII and StyD41 recognition sites on the 5' and 3' adapters, respectively. These precursor oligos (total lengths 118-122 nucleotides) were synthesized in bulk by CustomArray, Inc. (Bothell, WA). Probes with logistic scores >= 0.9 were synthesized in a single location. To account for poorer predicted performance and depending on the available synthesis space, probes with logistic scores between 0.7 and 0.9 were replicated 0-5 times and probes with logistic scores <0.7 were replicated between 5-10 times several times (Table S3).

**smMIP Preparation**

Array-synthesized precursor oligos were amplified by pool in a bulk reaction similarly to Boyle et al. 2014[10] with some modifications. Forward PCR primers were biotinylated on the 5' end to permit subsequent strand selection on streptavidin beads (see Table S11 for primer sequences). First, precursor oligos were resuspended at 100 nM in Tris-EDTA and 0.1% Tween (pH 8.0). A 400 µL bulk PCR mix was then prepared using a final concentration of 500 nM for each PCR primer, 1x iProof HF PCR master mix (Biorad, Hercules, CA), 0.2x SYBRGreen (Invitrogen, Carlsbad, CA), and 2.5 nM precursor oligos. This mix was split into eight x 50 µL reactions and amplified with the cycling conditions described in (Table S3). One bulk PCR reaction can be expected to yield ~70 ng of MIP product. Amplified products were combined per pool and purified using the QIAquick PCR purification kit (Qiagen, Hilden, Germany) following the manufacturer's instructions, using 1-2 columns per 400 µL PCR product. Product sizes were verified on a 2% agarose gel and yield quantified with the Qubit High Sensitivity dsDNA Assay Kit (Invitrogen).

Amplified DNA was digested at 37°C overnight in 50 µL of enzyme mix containing 1x CutSmart Buffer and 2 µL (5 U / µL) StyD4I (NEB, Ipswich, MA) to cleave off the 3' PCR adapter. Digested product was verified on a 2% agarose gel, then bound to MyOne Streptavidin C1 beads (Invitrogen) following the manufacturer's protocol, with 10 µL of beads per µg DNA. The bead-bound dsDNA was denatured with 50 µL of 0.125 N NaOH for two minutes (min) at room temperature, followed by supernatant removal, twice. The unbiotinylated antisense strand was washed away using 100 µL of 1x bead wash buffer followed by 100 µL of 1x CutSmart Buffer (NEB), leaving behind only the bead-bound sense smMIP strand.

To remove the remaining forward adapter, pool-specific guide oligos were annealed to the bead-bound 5' adapter sequence to create a double stranded DNA digest substrate. Each guide oligo was designed with two overhanging bases to extend the double-stranded template into the arm sequence of the MIPs. Nucleotide proportions of overhanging bases were proportional to arm composition (a 52/26/22 mixture of NN, GC and GD, respectively - see Table S11). After washing the denatured DNA, beads were resuspended in 50 µL of annealing master mix containing 1x CutSmart Buffer (NEB) and 15uM final concentration of appropriate guide oligo. Annealing was performed in a thermocycler, beginning with a slow ramp (0.1 degree/sec) to 65°C for 4 min and followed by a slow ramp (-0.1 degree/sec) to 37°C. To wash away excess guide oligo, beads were washed with 100 µL of bead wash buffer followed by 100µL of 1x CutSmart Buffer (NEB). Bead-bound DNA

was then resuspended in 50 µL of enzyme mix containing 1x CutSmart Buffer and 1 µL (10 U / µL) of NlaIII (NEB) and incubated for 2 hours (hrs) at 37°C in an Eppendorf ThermoMixerC (Hamburg, Germany) with a speed setting of 800 RPM. To further prevent beads from settling and ensure complete digestion, reactions were lightly vortexed every 30 min throughout the digestion period. Digest product was immobilized on a magnet and the released smMIPs aspirated. smMIPs were purified using the QIAquick column purification kit (Qiagen) following manufacturer's instructions. smMIP size verification was determined by PAGE gel, using a pre-cast 10% TBE-Urea PAGE gel (Invitrogen) and Gel Doc EZ Imager (BioRad). To quantify the amount of probe recovered, a standard curve (5 ng-20 ng) of an 80 bp oligo of known concentration, synthesized by IDT, was also loaded onto the same gel. Probe concentration was determined by relation of band density to DNA concentration derived from our standard curve using ImageLab 4.1's Image Tool (BioRad).

**smMIP Capture and Illumina Sequencing**

DNA prepared from whole blood (WB) and lymphoblastoid cell lines (LCLs) was obtained from the Simons Foundation Autism Research Initiative through the Rutgers University Cell and DNA Repository (Piscataway, NJ). Captures were performed as previously described with some modifications.[11] Hybridization of smMIPs to genomic DNA, gap filling, and ligation were performed in one 25 µL reaction of 1x Ampligase buffer (Epicentre, Madison, WI), with 200 ng of genomic DNA, smMIPs at a ratio of 800-1600 copies to one haploid genome copy [1600:1 for pilot 24, and 800:1 for all others], 0.25 mM dNTPs, 0.32 µL of 5X Hemo KlenTaq DNA polymerase (NEB), and one unit of Ampligase (Epicentre). Reactions were incubated at 95°C for 10 min and at 60°C for 18-42 hrs [18 hrs for pilot 24, 42 hrs for all others]. To degrade un-circularized probe and genomic DNA, 2 µL of exonuclease mix containing 10 units of exonuclease I (Enzymatics, Beverly, MA) and 50 units of exonuclease III (Enzymatics) in 1x Ampligase buffer were added and the reaction was incubated at 37°C for 45 min followed by 95°C for 2 min to inactivate the exonucleases. Subsequently, samples were cooled on ice and stored at 4°C until the time of amplification.

For each capture reaction, 25 µL PCR reactions were prepared [one PCR for pilot 24, two PCRs for other validations] using 5 µL of capture reaction, 0.5 µM forward and reverse barcoded primers (different for each sample), and 1x iProof HF Master Mix (Bio-Rad) at 98°C for 30 seconds (sec); varying cycles of 98°C for 10 sec, 60°C for 30 sec, 72°C for 30 sec; and finally 72°C for 2 min (see Table S3 for cycle number). The optimal number of cycles was determined independently for each pool by observing at what cycle amplification plateaued in a real-time PCR test reaction. Following amplification, a 5 µL aliquot of each sample was run on a 2% agarose gel to confirm correctly sized capture product (~208bp) and to assess relative concentrations of successful captures vs. empty smMIPs and other artifacts.

PCR products were pooled in equal volumes and purified using 0.8x AMPure XP beads (Agencourt-Beckman Coulter, Brea, CA) according to the manufacturer's instructions. Size selection was performed by extraction of correctly sized bands from a 2% agarose gel with the QIAquick Gel Extraction Kit (Qiagen). Pool concentrations were assessed using the Qubit HS dsDNA kit (Invitrogen). The purified PCR pools were then combined into one "megapool" for sequencing. The megapool library (1.8 pmol) was sequenced 2 x 75bp on the NextSeq 500 (Illumina, San Diego, CA) platform, using version 2 chemistry, according to the manufacturer's instructions. We used custom sequencing primers (Table S11) at a final concentration of 0.5 µM.

**PMM Validation Determinations**

Raw paired-end reads were merged using PEAR 0.9.6[12] and mapped to the GRCh37-hg19 Broad variant human reference genome using BWA 0.7.12. Reads which were unmapped (or MAPQ = 0), off-target, soft-clipped, or had insert sizes differing from expected gap-fill size were excluded from analysis. The remainder were collapsed on unique smMIP tags and uniformity of coverage evaluated both per smMIP and per target variant (Figure S3).[9; 11] All validations sets showed similar performance. Variant calls with less than 20-fold Q20 read depth in the family members required to validate a site were excluded from analysis.

Calls without smMIP captured variant reads were classified as false positives if the absence of variant reads was significant given total smMIP depth and expected (exome) AF (i.e. binomial $P(X > 0)$, for $p = AF$, threshold $p <= 0.01$); otherwise, they were considered indeterminate due to insufficient coverage. For calls with observed variant reads, the empirical error rate for that site was determined from all non-target families in the same pool. If smMIP variant AF was not significantly different from the pool error rate (binomial $p <= 0.01$), the variant was considered a sequencing error and thus a false positive.

Calls not excluded as false positives were independently assigned mosaic or germline validation status based on their smMIP data, following the same rubric as exome calling but with less stringent mosaic threshold (binomial p <= 0.01) due to the smaller number of variants being evaluated. Calls were additionally annotated as having either "same" or "different" AF in the target person compared to their exome data (Fisher's exact p <= 0.01). When data from both WB and LCLs was available, the WB validation was given priority. After initial validation assignments were made, two people manually reviewed these data and screenshots of smMIP alignments generated with Integrated Genome Viewer[13] for all validated calls. Variants with adjacent indels, with private SNPs in MIP targeting arms, with highly inconsistent AFs between different MIP probes, located in presumed multicopy regions characterized by multiple segregating mismatches, or having other evidence of problematic alignment were excluded from further analysis.

Resolutions were considered low-confidence if variants had AF <= 10% with only one supporting MIP, if individual MIP AFs differed between mosaic and germline status, or if AF 95% confidence intervals for mosaic validations approached or surpassed 0.5 in either tissue type. High confidence validations were defined based on the reviewers' consensus. Screenshots of exome alignments were generated for all high-confidence mosaic validations and manually reviewed as above, additionally checking for consistent segregation with any nearby SNP haplotypes. Putative mosaic variants were considered confirmed upon passing all review.

**Initial Logistic Regression Model Development**

An initial logistic regression model was trained using the pilot 24 initial resolutions (i.e. prior to analyzing the pilot 400 or full cohort data), using only calls validated as true PMMs or false positives in the smMIP data. Candidate predictors were derived from WES data, e.g. quality-aware total read depth (DP), quality-aware alternative allele read depth (DPALT), sequence context, and which callers identified the variant. Models were built for each candidate predictor using the R function *glm*. Univariate predictors with p <= 0.2 were considered for inclusion in a multivariate model. These terms were ranked in order of most to least significant univariate p-values and successively added into the multivariate model. Any predictor that became nonsignificant (p > 0.05) during this process was excluded. Pairwise interactions were evaluated using the R function *step()*. Finally, any predictors that had become nonsignificant as a result of model adjustments were also excluded, unless the predictor was also present in a significant interacting term. Fit was evaluated for each candidate multivariate model using the Hosmer-Lemeshow test across a range of five group sizes beginning at one greater than the number of model terms, with models rejected at p <= 0.05. Models not rejected were then compared based upon the Akaike information criterion (AIC) and sensitivity (within the dataset) and PPV as determined by 3-fold cross-validation. We selected an initial model that maximized sensitivity and minimized AIC while also maintaining reasonable PPV (Figure S7).

**Initial PMM Filtering and Validation: Pilot 400**

Based on results from the initial pilot 24 dataset, 400 additional pilot quad families were evaluated next (Figures S9-S12). Variant filtering was performed similarly as for the pilot 24 cohort, but calls were could not occur more than five times throughout the entire pilot 400 filtered variant set. For all putative parental transmitted PMMs, more significant skew in parental AF (binomial p <= 0.0001), significant difference between parent and child AF (Fisher's exact p <= 0.01), and child AF > parental AF, having observed that pilot 24 transmitted variants not meeting these criteria largely validated as germline (Figure S8) were required. All putative PMMs were scored using the initial logistic model, and excluded from validations if they scored < 0.2. This threshold was selected to eliminate the majority of false positives but retain high sensitivity and allow further evaluation of model performance. Family 14208 was excluded due to excessive SNV calls. Validation smMIP design, sequencing, analysis, and resolution were performed similarly as for the pilot 24 group, using WB DNA from 78 quad families. All initial validation positive calls, from both pilot sets, were then subjected to an additional manual review of the WES and smMIP alignments to flag potentially problematic calls prior to modeling, e.g. calls with evidence of mismapping, to produce a set of *high-confidence* validation resolutions.

**Refined Logistic Regression Model Development and Evaluation**

Based on manual review, we used only the predictions that were not observed repeatedly in the pilot 400 quad families and removed calls with a median number of mismatches greater than or equal to three in reads with variants. A second improved logistic regression model was trained using all predicted PMMs from this filtered

subset of pilot 400 high-confidence resolutions, including those resolved as germline variants (Table S4). Candidate predictors were as described in initial model development, with the addition of 1) median mismatches in variant reads and 2) variant error rate in a cohort of 400 families not included in either pilot group. Continuous predictors were coded as categorical terms with two or three bins based upon empirical odds ratios from univariate models (Figures S9B-E). A series of bicategorical models was built using successive threshold breakpoints spanning the predictor range, e.g. quartiles or deciles. Values across a range were assigned to the same bin if their odds ratios were similar, with additional thresholds evaluated as needed to identify the most appropriate bin boundaries. After coding continuous variables, univariate and multivariate models were built as previously described. In addition to exclusions already specified, interacting terms were dropped from models if they affected deviance by <10. Model fit and performance were evaluated and the best model selected as previously described.

This model was evaluated using pilot 24 resolutions as a test set and using additional validation data generated after model development (Supplemental Note: Model Development). The refined filtering scheme was retroactively applied to all validations in order to develop a harmonized set of high-confidence resolutions for final model evaluations. Retraining the model on harmonized pilot 400 resolutions did not substantially alter its performance (data not shown). All harmonized resolutions were then scored using the refined model and evaluated sensitivity (defined as the proportion of true variants scoring at or above the filter threshold; at cutoff 0.26) and PPV across those data to select a more stringent score threshold for cohort burden analysis (Figure S12). For cohort burden analysis, the reprocessed pilot 24 WES data was used over the merged pilot 24 WES data used for initial model training.

**Outlier Family Removal**

The 45x joint coverage calls with 5% minimum AFs at refined logistic regression score of >= 0.26 were used to determine if families had an excess of predicted SNVs. To account for coverage differences across families, mutation counts were normalized to reflect the number of calls that would be observed in the full exome (based on 45x joint coverage). Families with individuals that had total coverage adjusted variants above these thresholds were removed: GDMs >= 12, child PMMs >= 10, parental nontransmitted PMMs >= 12, parental transmitted PMMs >= 3, Thresholds were selected based on the distribution of counts in each category across the cohort.

To remove families that did not meet the coverage thresholds stipulated for each variant minimum AF, the total number of jointly sequenced bases within unique autosomal coding regions was calculated for each family at or above the coverage requirement: 45x, 50x, 65x, 85x, and 130x. Families with joint coverage falling below the 5$^{th}$ percentile (45x-85x) or bottom decile (130x) were excluded (Figure S14). Percentile ranking were defined using the whole cohort (quads + trios).

**Significance Determination for Burden and Variant Properties Analysis**

To control for type I errors resulting from multiple comparisons, a false discovery rate (FDR) approach utilizing the Benjamini-Yekutieli (BY) procedure was applied.[14] While, less powerful than the Benjamini-Hochberg procedure, BY allows for any dependency structure among the test statistics. We used the R package *Mutoss* implementation, *BY()*, with FDR set to 0.05. For quad data, the paired nonparametric Wilcoxon sign rank test (WSRT) was used. For synonymous variants we used a two-sided test. We used a one-sided test for missense PMMs with the *a priori* assumption that probands would have a higher rate. For full cohort (quad + trio) comparisons the unpaired Wilcoxon rank sum test (WRST) was used.

Families of tests were defined based on the dataset and test statistic used, as follows:

**PMM burden, Probands v. Siblings**

i. Synonymous PMM burden quads two-sided WSRT (5 tests): 1. 15%-45x, 2. 12.5%-50x, 3. 10%-65x, 4. 7.5%-85x, 5. 5%-130x.
ii. 12.5%-50x synonymous PMM burden full/subcohorts, two-sided WRST (5 tests): 1. Full cohort, 2. Has LGD GDM, 3. No LGD GDM, 4. Has NS GDM, 5. No NS GDM.
iii. 15%-45x missense PMM burden full/subcohorts/gene sets, one-sided WRST (15 tests):
    a. subcohorts: 1. All missense full cohort, 2. All missense has LGD GDM, 3. All missense no LGD GDM, 4. All missense has NS GDM, 5. All missense no NS GDM;

b. subcohorts and in essential genes: 6. Full cohort, 7. Has LGD GDM, 8. No LGD GDM, 9. Has NS GDM, 10. No NS GDM;
c. subcohorts and in intolerant genes: 11. Full cohort, 12. Has LGD GDM, 13. No LGD GDM 14. Has NS GDM, 15. No NS GDM .

**Mutation Properties**

iv. AF distribution comparisons, two-sided WRST (7 tests): 1. Probands v. Siblings, 2. Fathers Trans v. Nontrans, 3. Mothers Trans v. Nontrans, 5. Fathers Trans v. Mothers Trans, 6. Fathers Nontrans v. Mothers Nontrans, 7. Children v. Parents Nontrans.
v. Distance to splice site distribution, two-sided WRST (4 tests): 1. Probands v. Siblings, 2. Fathers v. Mothers, 3. Siblings v. Parents, 4. Probands v. Parents.

**Phenotype Information**

We compared 12 subjects LGD PMMs and 45 subjects with missense PMMs whose mutations overlapped genes with GDMs in the SSC. We evaluated developmental history data including: delay in first word use, age of use of first phrases, age at walking, birth weight, gestational age, history of seizures, current body mass index, and head circumference. Standardized head circumference scores (Z-scores) were calculated using norms established by Roche et al. to account for age and gender.[15] We examined measures of autistic symptomatology, including: the Autism Diagnostic Interview-Revised (ADI-R) three domain scores (verbal and non-verbal communication, social interaction and reciprocity, repetitive behaviors), the Autism Diagnostic Observation Schedule (ADOS) calibrated severity scale, the Social Responsiveness Scale (SRS), and total Repetitive Behavior Scale scores. Non-autistic behavioral and emotional problems were examined using the Child behavior Checklist (CBCL). Level of functioning was examined using the Vineland Adaptive Behavior Scales and intellectual quotient (IQ).

When available, the age of parents at blood draw (in years) was retrieved from repository records. If this information was not available, the parental age at blood draw was estimated by adding the proband age at ADOS (months) to the parental age at birth (months) and then rounding to the nearest year. The ADOS was performed near the time of draw. Using these sources, the age of parents at blood draw was estimated for all but two families that passed QC.

**Supplemental Note: Model Development**

Based on the preliminary findings of variants identified using *germline* variant calling pipelines, we sought to perform a systematic analysis of PMMs with methods specifically geared toward mosaic SNV mutations. Several standalone PMM single nucleotide variant (SNV) callers were evaluated and a custom read parser (mPUP) using simulated data containing artificial variants at 202 loci. These loci were simulated at varying AF and depths ranging from 1 to 50% and 30 to 500-fold respectively, allowing a wide evaluation of the possible detection search space (Tables S8 and S9). We found that within the simulated data, caller sensitivity greatly varied at different depths and AFs, but many had high PPV (Table S8). Based on their complementary performances at different depths and AFs, we selected Varscan2, LoFreq, and mPUP for further evaluation.

These three variant callers were applied to the high depth 24 quad families (96 individuals) WES data. This call set included predicted PMM calls from a wide range of AFs (3-50%), at different depths (8x-500x) and support levels (5% at 60x versus 500x). LoFreq showed the best performance as a single caller in terms of correctly validated calls (125/138 LoFreq calls validated true); however, it failed to predict 13/51 validated PMM (Figure S7A). The majority of the PMM calls were validated in both WB and LCL DNA (42/49 with high-confidence dual data).

Using these pilot 24 validation data, an initial logistic regression model was constructed and trained on the validated predicted true/false PMMs, which took into account depth, caller, reference base, and transition vs. transversion changes. A logistic score threshold of >= 0.2, was selected as it performed well in three-way cross validations, but was nevertheless conservative given the limited number of training calls (Figure S7D). Importantly, the initial logistic regression model reduced the raw number of raw PMM calls by 93%.

This initial logistic regression model was then applied as well as additional filters for ambiguous transmitted calls (i.e. binomial p <= 0.0001 and Fisher's exact p <= 0.01) to an independent set of 400 quad families. Validations were then performed. For both pilot 24 and 400 validations, manual inspection of WES and smMIP alignment data was performed for all initially positive validations (based on read count data) and a subset of false positive calls. In doing so, a number of common features associated with poor prediction

outcomes or problematic genomic regions were observed. First, we found that a large number of false positive validations had an excess of multiple mismatches within the variant reads (Figures S6 and S11A). This feature was not present in the vast majority of true germline or mosaic calls. Based on the median number of mismatches we identified <= 3 as a filter threshold that would remove a large number of false positive calls, without dramatically altering sensitivity (Figure S11A). Similarly a number of the pilot 24 calls were detected multiple times in the pilot 400 call set, which had not been processed at the time of selecting pilot 24 validation calls (Figure S11B). Variant calls present in multiple families typically validated as false positives or parental germline. Therefore, all calls with these two features were removed prior to building a refined logistic regression model.

Using the filtered pilot 400 high-confidence validation set, a refined logistic regression model was built on all predicted PMMs (Figure S9). In evaluating the model, calls generally fell within three groups (Figure S12B). First, low scoring and largely false positive calls had low AFs, low read counts, and medium-high empirical error rates. The middle grouping had either low-medium AF, low error rate, and lower variant read counts or low-medium AF, medium-high error rate, and high variant read count. The highest scoring group was largely driven by higher AFs and variant read counts. This group includes the bulk of the true (mosaic and germline) validated calls 87/109 (80%); whereas, the middle grouping contained 15/109 (14%) true validated calls and the low grouping had only 7/109 (6%). Additionally calls validating germline tended to have higher WES AFs and found that the vast majority (99%) of validated PMM calls had upper CI bounds that remained below 0.4, while the majority of true germline calls (76%) fell above this threshold (Figure S10). This threshold was chosen to maximize sensitivity. In addition, a significant fraction of the false positive calls were annotated as SD/TRF calls (Figure S11D). Moving forward SD/TRF calls were removed and re-classified as mosaic versus germline status based on the AF binomial CI.

Pilot 400 family counts for called calls were derived prior to removing outlier families. Re-deriving these counts post outlier removal did not substantially change the call set. Initially, variants that had any population frequency in at least one *but not all three* databases were erroneously omitted from the variant validation sets. Having identified this error, we used this opportunity to generate a third round of validations with which to evaluate our refined model. All pilot 24 and pilot 400 families except 14208 were included in this analysis. Variant filtering was performed similarly to previous iterations, with correction of the population frequency filter and updated filtering rules. Putative PMMs were scored with our refined logistic model and excluded from validations if they scored < 0.26. Validation smMIP design, sequencing, analysis, and resolution were performed similarly as for the pilot groups. Across the test sets (under harmonized filters), both sensitivity and PPV converged at a logistic score of 0.518 (sensitivity 0.83, PPV 0.85) and chose to use this more stringent score threshold (Figures S12E-F). In addition, calls with less than five variant allele reads were removed as these disproportionately contributed to false calls (Figure S11E).

In summation, we identified these parameters as our “best practice calling” and applied this approach to the full cohort to generate our high confidence call set: 1) variant must have at least five reads, 2) AF upper CI must intersect 5%, 3) mismatch <= 3 in variant reads, 4) called by at least two callers, 5) cohort count <= 2, 6) have an AF upper CI < 40%, 7) not be within a known SDTRF loci, 8) refined logistic model score of 0.518. *Specifically for transmitted calls to be considered a putative PMM*, the binomial deviation is more stringent (p <= 0.0001) and the AF between child and parent must be significantly different by Fisher’s exact test (p <= 0.01).

**Supplemental Note: Case Reports**

Reports were generated for a subset of probands with nonsynonymous mutations (both GDMs and PMMs) intersecting the 65 genes meeting an FDR of 0.1 from Sanders et al. (2015)[16] and genes with mosaic and germline LGD mutations. Summaries of patient characteristics—including cognitive ability, presence of comorbid medical and psychiatric disorders, presence of frank dysmorphology, and raw physical measurements (e.g., head circumference)—were culled from the SSC phenotype data distributions (https://sfari.org/resources/sfari-base) and presented in narrative form. Note: *MFRP* was not included because of the presence of a LGD GDM in an unaffected sibling. Individuals with mutations intersecting more than one gene are listed twice.

***BAZ2B* (LGD PMM and GDM)**

**ID: 13694.p1**
*Event: Predicted Mosaic Nonsense*
Patient is a 104 month old non-Hispanic, bi-racial male diagnosed with ASD and Intellectual Disability. Patient is minimally verbal, has a full scale IQ (FSIQ) in the extremely low range (21), and overall adaptive skills in the low range (Vineland ABC = 62). Adaptive skills are uniformly low. Patient does not have a history of seizures, but has a possible history of language regression and has attention difficulties (CBCL Attentional Difficulties T-Score = 74). Patient walked at 12 months of age, but has not yet attained single word use or phrase speech. At time of visit patient's body mass index (BMI) Z-score was -0.80, height Z-score was 0.71, and head circumference Z-score was -0.62.

**ID: 14581.p1**
*Event: Predicted Germline Frameshift Insertion*
Patient is a 64 month old non-Hispanic, white male diagnosed with ASD. Patient is verbally fluent, has a FSIQ in the high average range (113), and overall adaptive skills in moderately low range (Vineland ABC = 80). Adaptive communication falls in the average range (Communication Standard Score = 91), adaptive social skills falls in the average range (Social Standard Score = 86), and daily living skills fall in the moderately low range (DLS Standard Score = 75). Patient does not have a history of seizures, but has a history of word loss. Patient has internalizing (CBCL Internalizing T-score= 76) and externalizing symptoms (CBCL Externalizing T-score= 86) in the clinical range. Patient walked at 12 months of age, used single words at 12 months of age, and used first phrases at 18 months old. At time of visit, patient's BMI Z-score was 1.09, height Z-score was 1.06, and head circumference Z-score was -0.19.

**ID: 11441.p1**
*Event: Predicted Germline Missense*
Patient is a 93 month old non-Hispanic, bi-racial male diagnosed with ASD. Patient is verbally fluent, has a FSIQ in the very high range (125), and overall adaptive skills in the average range (Vineland ABC = 89). However, while adaptive communication and daily living skills fall in average range, social adaptive skills fall in low range (Social Standard Score = 64). Patient does not have a history of seizures or regression. Patient has internalizing symptoms in the borderline clinical range (CBCL Internalizing T-score= 67). Patient walked at 11 months of age, used single words at 11 months of age, and used first phrases at 14 months old.

***UNC79* (LGD PMM and GDM)**

**ID: 14547.p1**
*Event: Predicted Mosaic Nonsense*
Patient is a 99 month old non-Hispanic, Native Hawaiian male diagnosed with ASD. Patient is verbally fluent, has a FSIQ in the very low range (71), with a significant nonverbal (NVIQ = 95) and verbal (VIQ = 60) split. Patient's overall adaptive skills fall in moderately low range (Vineland ABC = 74). Adaptive communication falls in the moderately low range (Communication = 81), adaptive social skills falls in the moderately low range (Social = 76), and daily living skills fall in the low range (DLS = 68). Patient does not have a history of seizures, but had a possible regression. Patient has no elevations in externalizing or externalizing symptoms. Patient walked at 14 months of age and used single words at 15 months of age and first phrases at 26 months old. At time of visit, patient's BMI Z-score was 2.26, height Z-score was 1.72, and head circumference Z-score was 2.26.

**ID: 14530.p1**
*Event: Predicted Germline UNC79 Frameshift Deletion and Predicted Germline GIGYF1 Frameshift Insertion*

Patient is a 49 month old Hispanic male diagnosed with ASD. Patient uses simple phrase speech, has a FSIQ in the low average range (82), and overall adaptive skills in moderately low range (Vineland ABC = 73). Adaptive skills are uniformly in the moderately low range. Patient does not have a history of seizures or regression. Patient has externalizing symptoms in the clinical range (CBCL Externalizing T-score= 74). Patient walked at 12 months of age and had language delays, using single words at 30 months of age and first

phrases at 46 months old. At time of visit, patient's BMI Z-score was 0.45, height Z-score was 0.25, and head circumference Z-score was 1.26.

### *USP15* (LGD PMM and GDM)

**ID: 12025.p1**
*Event: Predicted Mosaic Nonsense*
Patient is an 80 month old non-Hispanic, White male diagnosed with ASD. Patient is minimally verbal, has a FSIQ in the very low range (72), with a significant nonverbal (NVIQ = 96) and verbal (VIQ = 69) split. Patient's overall adaptive skills fall in low range (Vineland ABC = 70). Adaptive communication falls in the moderately low range (Communication = 76), adaptive social skills falls in the low range (Social = 63), and daily living skills fall in the moderately low range (DLS = 77). Patient does not have a history of seizures, but had word loss. Patient has internalizing symptoms in the borderline clinical range (CBCL Internalizing T-score= 65). Patient walked at 10 months of age and used single words at 12 months of age, but had a delay in using phrase speech (first phrases at 48 months old). At time of visit, patient's BMI Z-score was 0.06, height Z-score was -1.42, and head circumference Z-score was -0.22.

**ID: 12521.p1**
*Event: Predicted Germline Frameshift Deletion*
Patient is an 86 month old non-Hispanic, White female diagnosed with ASD. Patient is verbally fluent and has a FSIQ in the very low range (78). Patient's overall adaptive skills fall in moderately low range (Vineland ABC = 78). Adaptive communication falls in the moderately low range (Communication = 84), adaptive social skills falls in the low range (Social = 69), and daily living skills fall in the average range (DLS = 87). Patient does not have a history of seizures, but has a possible regression. Patient has externalizing (CBCL Internalizing T-score= 65) and externalizing (CBCL Externalizing T-score= 66) symptoms in the borderline clinical range. Patient walked at 19 months of age and had language delays, using single words at 36 months of age and first phrases at 48 months old. At time of visit, patient's BMI Z-score was -0.89, height Z-score was -1.22, and head circumference Z-score was 0.65.

### *DIP2A* (ASD 65)

**ID: 13012.p1**
*Event: Predicted Mosaic Frameshift Insertion*
Patient is a 70-month-old Hispanic male diagnosed with ASD and Intellectual Disability. He uses single words, has a FSIQ in the extremely low range (54) with a significant split between nonverbal (NVIQ = 60) and verbal (VIQ = 21) abilities. Patient's overall adaptive skills fall in the low range (Vineland ABC = 54) with uniform deficits across adaptive domains. He has no history of seizures. He has a history of regression. He walked at 10 months, used single words at 11 months of age, and has not developed phrase speech. At time of visit, patient's BMI Z-score was 0.72, height Z-score was -0.13, and head circumference Z-score was 0.63.

**ID: 13106.p1**
*Event: Predicted Germline Nonsense*
Patient is a 198-month-old non-Hispanic White male diagnosed with ASD. He is verbally fluent and has a FSIQ in the average range (100) with a significant split between nonverbal (NVIQ = 79) and verbal (VIQ = 140) abilities. Patient's overall adaptive skills fall in the low range (Vineland ABC = 56) with uniform significant deficits across adaptive domains. Patient has clinically significant internalizing symptoms (CBCL Internalizing T-score = 71) and borderline clinically significant externalizing (CBCL Externalizing T-score = 69) symptoms. He has no history of regression and no history of seizures. He walked at 16 months, used single words at 13 months of age, and used first phrases at 18 months of age. At time of visit, patient's BMI Z-score was 1.59, height Z-score was -1.65, and head circumference Z-score was 0.64.

***GIGYF1* (ASD 65)**

**ID: 11232.p1**
*Event: Predicted Mosaic Frameshift Deletion*
Patient is a 104-month-old non-Hispanic, White male diagnosed with ASD. He is verbally fluent and has a FSIQ in the very low range (74) with a significant split between nonverbal (NVIQ = 68) and verbal (VIQ = 91) abilities. Patient's overall adaptive skills fall in the average range (Vineland ABC = 97) with uniform adaptive functioning across communication, daily living, and social domains. He has no history of seizures. He has no history of regression. He has no elevated clinical symptomatology across internalizing and externalizing disorders. He walked at 12 months, used single words at 11 months of age, and developed phrase speech at 30 months of age. At time of visit, patient's BMI Z-score was 1.43, height Z-score was -0.05, and head circumference Z-score was 0.11.

**ID: 11860.p1**
*Event: Predicted Germline Splicing*
Patient is a 72-month-old Hispanic male diagnosed with ASD. He uses phrase speech and has a FSIQ in the low average range (86) with a significant split between nonverbal (NVIQ = 95) and verbal (VIQ = 75) abilities. Patient's overall adaptive skills fall in the moderately low range (Vineland ABC = 77) with uniform significant deficits across adaptive domains. He has no elevated clinical symptomatology across internalizing and externalizing disorders. He has no history of regression and no history of seizures. He walked at 13 months, used single words at 42 months of age, and used first phrases at 48 months of age. At time of visit, patient's BMI Z-score was 1.94, height Z-score was 0.62, and head circumference Z-score was 0.88.

**ID: 14530.p1**
*Event: Predicted Germline UNC79 Frameshift Deletion and Predicted Germline GIGYF1 Frameshift Insertion*

Patient is a 49 month old Hispanic male diagnosed with ASD. Patient uses simple phrase speech and has a FSIQ in the low average range (82), and overall adaptive skills in moderately low range (Vineland ABC = 73). Adaptive skills are uniformly in the moderately low range. Patient does not have a history of seizures or regression. Patient has externalizing symptoms in the clinical range (CBCL Externalizing T-score= 74). Patient walked at 12 months of age and had language delays, using single words at 30 months of age and first phrases at 46 months old. At time of visit, patient's BMI Z-score was 0.45, height Z-score was 0.25, and head circumference Z-score was 1.26.

***CHD2* (ASD 65)**

**ID: 13073.p1**
*Event: Predicted Mosaic CHD2 Missense and Predicted Germline SYNGAP1 Frameshift Deletion*
Patient is a 58-month-old non-Hispanic, White male diagnosed with ASD. He is minimally verbal and has a FSIQ in the extremely low range (43) with a significant split between nonverbal (NVIQ = 60) and verbal (VIQ = 25) abilities. Patient's overall adaptive skills fall in the low range (Vineland ABC = 57) with uniformly significant deficits across adaptive domains. He has no elevated clinical symptomatology across internalizing and externalizing disorders. He has no history of seizures, but history of a possible regression. In terms of milestones, he walked at 25 months and has not developed single word or phrase speech. At time of visit, patient's BMI Z-score was 1.86 and height Z-score was -1.92.

**ID: 13618.p1**
*Event: Predicted Germline Frameshift Deletion*
Patient is a 159-month-old non-Hispanic White female diagnosed with ASD and Intellectual Disability. She is verbally fluent and has a FSIQ in the extremely low range (44). Patient's overall adaptive skills fall in the low range (Vineland ABC = 57) with uniform deficits across adaptive domains. Patient has clinically significant scores of internalizing (CBCL Internalizing T-score = 75) and borderline externalizing (CBCL Externalizing T-score = 69) symptoms. She has a history of seizures (first grand mal seizure at 11 years of age, with weekly seizures, and reported febrile seizure at 12 years of age), and abnormal EEG (diagnosed at 4 years old). She has no history of regression. She walked at 14 months, used single words at 12 months of age, and used first

phrases at 30 months of age. At time of visit, patient's BMI Z-score was -2.32, height Z-score was -0.34, and head circumference Z-score was -1.65.

**ID: 13614.p1**
*Event: Predicted Germline Nonsense*
Patient is a 113 month non-Hispanic White male diagnosed with ASD. He is verbally fluent and has a FSIQ in the very low range (79). He has moderately low adaptive scores (Vineland ABC = 74) with uniformly low scores in the adaptive subdomains. Patient has clinically significant externalizing symptoms (CBCL Externalizing T-score = 72). Patient has also been diagnosed with Oppositional Defiant Disorder, Attention Deficit Hyperactivity Disorder, and Generalized Anxiety Disorder. He has no history of regression. Patient has had two complex partial seizures. He walked at 13 months, used single words at 30 months or age and phrases at 36 months of age. At time of visit, patient's BMI Z-score was 1.20, height Z-score was 0.20, and head circumference Z-score was -0.22.

**ID: 13818.p1**
*Event: Predicted Germline Frameshift Insertion*
The patient is a 179 Non-Hispanic, White male. Patient has a diagnosis of ASD as well as Developmental Coordination Disorder, Unspecified Anxiety Disorder, Specific Learning Disorder with impairment in Mathematics, Mild Intellectual Disability, Unspecified Depressive Disorder and Disruptive Mood Dysregulation Disorder. He is verbally fluent and speaks in complex sentences. Patient's cognitive abilities fall in the extremely low range (66) and his adaptive abilities fall in the low range (Vineland ABC = 66). Patient used his first single words at 18 months of age. His first phrases were at 21 months. Patient is color blind, and has a significant visual impairment ("legally blind" without glasses) but wears glasses to correct to normal. Patient has a significant history of chronic constipation, and underwent a testicular hernia repair secondary to constipation. Patient also has a significant history of seizures (grand mal and petit mal reported with age of onset at 2 years of age). Patient has a multidysplastic right kidney. Facial features include horizontal eyebrows, synophrys, horizontal palpebral fissures and a high nasal root. Patient has single palmar crease on right hand, mild 2-3 cutaneous syndactyly of toes, a curved 2$^{nd}$ toe and flat feet. Physical examination reveals one café au lait spot. Patient has a BMI Z-score of -0.92, height Z of 0.6, and head circumference Z of -0.65.

***KMT2C* (ASD 65)**

**ID: 11148.p1**
*Event: Predicted Germline KMT2C Nonsense*
Patient is a 68-month-old non-Hispanic, White male diagnosed with ASD. He uses phrase speech to communicate and has a FSIQ in the low average range (86) with a significant split between nonverbal (NVIQ = 82) and verbal (VIQ = 99) abilities. Patient's overall adaptive skills fall in the moderately low range (Vineland ABC = 81) with adaptive communicative and daily living skills in the average range, but social skills falling in the moderately low range. He has no elevated clinical symptomatology across internalizing and externalizing disorders. He has no history of seizures and no history of regression. He walked at 17 months, used single words at 12 months, and phrase speech at 24 months.

**ID: 11241.p1**
*Event: Predicted Germline KMT2C Missense*
Patient is a 144-month-old non-Hispanic, White male diagnosed with ASD. He is verbally fluent and has a FSIQ in the very low range (76) with similar performance across nonverbal (NVIQ = 77) and verbal (VIQ = 80) domains. Patient's overall adaptive skills fall in the low range (Vineland ABC = 64) with daily living skills in the moderately low range, but social and communication skills falling in the low range. He has no elevated externalizing symptomatology, but clinically elevated internalizing symptoms (CBCL T-score= 70). He has a history of febrile seizures and a possible history of regression. In terms of milestones, he walked at 12 months old, used single words at 9 months and phrase speech at 15 months. Patient has a BMI Z-score of 1.94, height Z of -1.7, and head circumference Z of -0.07.

**ID: 12742.p1**
*Event: Predicted KMT2C Missense*

Patient is a 58-month-old non-Hispanic, White male diagnosed with ASD. He is verbally fluent and has a FSIQ in the average range (105) with similar performance across nonverbal (NVIQ = 103) and verbal (VIQ = 106) domains. Patient's overall adaptive skills fall in the average range (Vineland ABC = 94) with similar functioning across all adaptive subdomains. He has no elevated clinical symptomatology across internalizing and externalizing disorders. He has neither history of seizures nor history of regression. In terms of milestones, he walked at 13 months old, used single words at 24 months and phrase speech at 33 months. Patient has a BMI Z-score of 3.7, height Z-score of -3.88, and head circumference Z of -0.70.

**ID: 13897.p1**
*Event: Predicted Mosaic KMT2C Missense*
Patient is a 127-month-old non-Hispanic, White male diagnosed with ASD. He is verbally fluent and has a FSIQ in the low average range (85) with split performance across nonverbal (NVIQ = 91) and verbal (VIQ = 78) domains. Patient's overall adaptive skills fall in the moderately low range (Vineland ABC = 80) with daily living skills in the average range, but social and communication skills falling in the moderately low range. He has no elevated clinical symptomatology across internalizing and externalizing disorders. He has no history of seizures and no history of regression. In terms of milestones, he walked at 12 months old, used single words at 24 months and phrase speech at 30 months. Patient has a BMI Z-score of 2.0, height Z-score of 3.29, and head circumference Z of 2.95.

***SCN2A* (ASD 65)**

**ID: 13522.p1**
*Event: Predicted Transmitted Mosaic (Germline) Missense*
Patient is a 138-month-old Hispanic male diagnosed with ASD. He is verbally fluent and has a FSIQ in the very low range (79) with split performance across nonverbal (NVIQ = 87) and verbal (VIQ = 70) domains. Patient's overall adaptive skills fall in the moderately low range (Vineland ABC = 72) with similar functioning across adaptive subdomains. He has no elevated clinical symptomatology across internalizing and externalizing disorders. He has no history of seizures and no history of regression. In terms of milestones, he walked at 14 months old, used single words at 12 months and phrase speech at 66 months. Patient has a BMI Z-score of 1.72, height Z-score of -1.2, and head circumference Z-score of 0.04.

**ID: 11892.p1**
*Event: Predicted Germline Nonsense*
Patient is a 12 year old non-Hispanic, White male. Patient has a diagnosis of ASD, Speech Sound Disorder, Mild Intellectual Disability and Developmental Coordination Disorder. Patient is verbally fluent with a FSIQ in the extremely low range (56) and significant nonverbal (NVIQ = 42) and verbal (VIQ = 81) split. His adaptive abilities fall in the low range (Vineland ABC = 62). He first used single words at 16 months and phrase speech at 30 months. He has no history of regression or seizures. Parent report does not indicate any significant internalizing or externalizing behaviors. He has been diagnosed with scoliosis and received corrective surgery for tibeal torsion on both legs at 4 years. Facial features include a broad forehead, a slightly heavy brow that is prominent laterally, slightly high nasal bridge and a thin nose with upturned tip, palpebral fissures at 3.2 cm (+2 SD). Other notable dysmorphology includes scoliosis with a right-to-left curve, multiple nevi scattered on back and chest and hyperreflexia observed in biceps, patellae and Achilles. Patient has a BMI Z-score of -0.35, height Z-score of -0.52, and head circumference Z-score of 0.05.

**ID: 14525.p1**
*Event: Predicted Germline Missense*
Patient is a 142 month old non-Hispanic, White male. Patient has a diagnosis of ASD, Intellectual Disability, and speech delay. He is minimally verbal, uses sign language to communicate and has an estimated verbal mental age of 10 months and a nonverbal mental age of 18 months. His adaptive skills across all domains are in the low range (Vineland ABC = 37). He has clinically significant internalizing symptoms (CBCL Internalizing T-score = 65). In terms of milestones, he walked at 18 months, but never developed language. He has a significant seizure history, starting at 2.5 years of age, with approximately 30 seizures each day, lasting approximately 3-4 months. Seizures were categorized as grand mal, generalized tonic clonic, and atonic and

drop attacks. Patient has a BMI Z-score of -0.94, height Z-score of 0.14, and head circumference Z-score of 0.04.

**ID: 13642.p1**
*Event: Predicted Germline Missense*
Patient is an 111 month old non-Hispanic, White male diagnosed with ASD. He is verbally fluent, with a high average IQ (114) and consistently moderately low adaptive skills (Vineland ABC = 73). Patient has clinically significant internalizing (CBCL Internalizing T-score = 70) and externalizing (CBCL Externalizing T-score = 77) symptoms. He walked at 17 months, used single words at 18 months, and combined words into short sentences at 36 months. Possible loss and regression was reported, but no seizure history. He has a possible hearing problem and corrected vision problems. Patient had chronic diarrhea and suffered severe abdominal pain in early childhood. Patient has recent suspected heart problems (tachycardia). Patient has a BMI Z-score of 0.1, height Z-score of 1.82, and head circumference Z-score of -0.48.

**ID: 11114.p1**
*Event: Predicted Germline Nonsense*
Patient is a 105 month old non-Hispanic, White female diagnosed with ASD and Intellectual Disability. She has several additional diagnoses including: pragmatic language disorder, mixed expressive-receptive language disorder, speech delay, written expression disorder, math disorder, and nonverbal learning disability, attention deficit hyperactivity disorder, and anxiety disorder. She was diagnosed with excessive clumsiness at 2 years, excessive gas at 4 years, and intermittent constipation at 4 months of age. She uses phrase speech and has an IQ in the extremely low range (40). Her adaptive skills are in the low range (Vineland ABC = 67). She has internalizing symptoms in the borderline clinical range (CBCL Internalizing T-score= 65). She has a history of word loss. No history of seizures. Patient has a BMI Z-score of 1.25, height Z-score of 1.56, and head circumference Z-score of 2.91.

**ID: 13544.p1**
*Event: Predicted Germline Missense*
Patient is a 84-month-old non-Hispanic, White male diagnosed with ASD. He is minimally verbal using occasional phrase speech to communicate. He has a FSIQ in the extremely low range (63) with split performance across nonverbal (NVIQ = 77) and verbal (VIQ = 46) domains. Patient's overall adaptive skills fall in the low range (Vineland ABC = 69) with daily living skills in the moderately low range, but social and communication skills falling in the low range. He has no elevated clinical symptomatology across internalizing and externalizing disorders. He has a history of seizures and a possible history of regression. In terms of milestones, he walked at 17 months old, used single words at 12 months and phrase speech at 45 months. Patient has a BMI Z-score of 0.01, height Z-score of 0.14, and head circumference Z-score of -0.73.

**ID: 14280.p1**
*Event: Predicted Germline Missense*
Patient is a 113-month-old non-Hispanic, White male diagnosed with ASD. He is minimally verbal and has a FSIQ in the extremely low range (25). Patient's overall adaptive skills similarly fall in the low range (Vineland ABC = 56) with similar deficits across all subdomains of adaptive functioning. He has no elevated clinical symptomatology across internalizing and externalizing disorders. He has no history of seizures but a possible history of regression. In terms of milestones, he walked at 16 months old but has not developed single word use or phrase speech. Patient has a BMI Z-score of -1.35, height Z-score of -1.74, and head circumference Z-score of -0.13.

***SYNGAP1* (ASD 65)**

**ID: 14001.p1**
*Event: Predicted Mosaic Missense*
Patient is a 91-month-old non-Hispanic, Black male diagnosed with ASD. He is minimally verbal and has a FSIQ in the extremely low range (52) with split performance across nonverbal (NVIQ = 63) and verbal (VIQ = 38) domains. Patient's overall adaptive skills fall in the low range (Vineland ABC = 64) with consistent deficits in the low range across adaptive subdomains. He has no elevated clinical symptomatology across internalizing

and externalizing disorders. He has no history of seizures, but a history of regression with word loss. In terms of milestones, he walked at 12 months old, used single words at 11 months and phrase speech at 54 months. Patient has a BMI Z-score of 0.43 and a height Z-score of 2.16.

**ID: 12804.p1**
*Event: Predicted Germline Missense*
Patient is a 118-month-old non-Hispanic, White male diagnosed with ASD. He is verbally fluent and has a FSIQ in the very low range (77) with split performance across nonverbal (NVIQ = 85) and verbal (VIQ = 69) domains. Patient's overall adaptive skills fall in the moderately low range (Vineland ABC = 77) with similar performance in the moderately low range across adaptive subdomains. He has no elevated clinical symptomatology across internalizing and externalizing disorders. He has no history of seizures and no history of regression. In terms of milestones, he walked at 11 months old, used single words at 18 months and phrase speech at 84 months. Patient has a BMI Z-score of -0.16, height Z-score of 0.49, and head circumference Z-score of 1.26.

**ID: 13073.p1**
*Event: Predicted Mosaic CHD2 Missense and Predicted Germline SYNGAP1 Frameshift Deletion*
Patient is a 58-month-old non-Hispanic, White male diagnosed with ASD. He is minimally verbal and has a FSIQ in the extremely low range (43) with a significant split between nonverbal (NVIQ = 60) and verbal (VIQ = 25) abilities. Patient's overall adaptive skills fall in the low range (Vineland ABC = 57) with uniformly significant deficits across adaptive domains. He has no elevated clinical symptomatology across internalizing and externalizing disorders. He has no history of seizures, but history of a possible regression. In terms of milestones, he walked at 25 months and has not developed single word or phrase speech. At time of visit, patient's BMI Z-score was 1.86, and height Z-score was -1.92.

***KAT2B* (ASD 65)**

**ID: 11592.p1**
*Event: Predicted Mosaic Splicing*
Patient is a 121-month-old non-Hispanic, White male diagnosed with ASD. He is verbally fluent and has a FSIQ in the above average range (115) with split performance across nonverbal (NVIQ = 109) and verbal (VIQ = 122) domains. Patient's overall adaptive skills fall in the average range (Vineland ABC = 92) with communication and daily living skills in the average range, but adaptive social skills falling in the moderately low range. He has symptomatology in the internalizing domain in the borderline clinical range (CBCL Internalizing T-score= 68). He has no history of seizures and no history of regression. In terms of milestones, he walked at 12 months old, used single words at 14 months and phrase speech at 24 months. Patient has a BMI Z-score of -1.96, height Z-score of 1.45, and head circumference Z-score of -0.11.

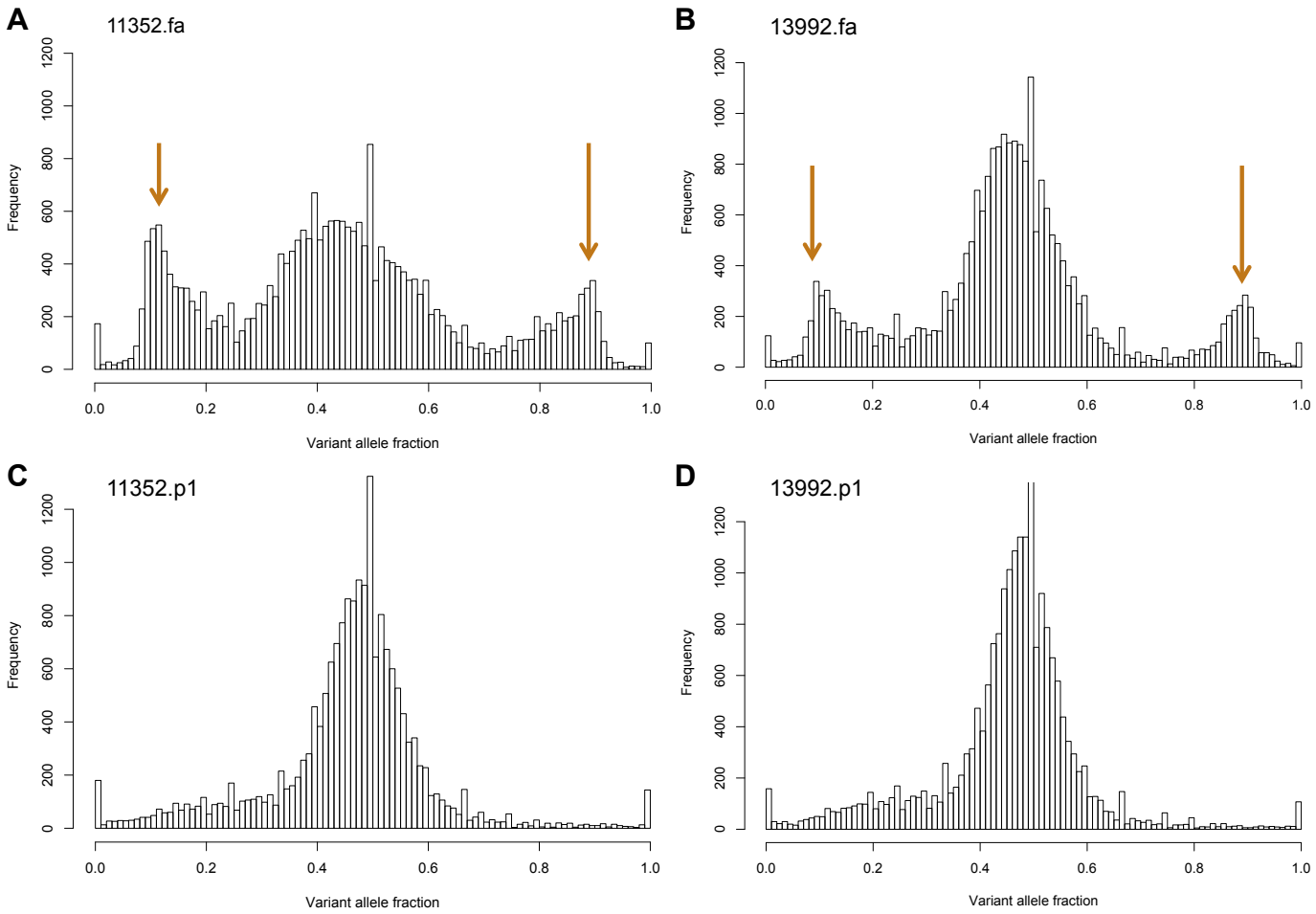



**Figure S1. Representative AF Histograms for Members of Pilot 400 Families Excluded from Model Training Set**

(A) and (B) show individuals identified as having excess SNVs, but no obvious identity or family relationship issues. Secondary peaks suggest sample contamination, indicated by arrows.

(C) and (D) show other members of the same families with typical AF distributions.

Both families were excluded from training of the refined logistic model. Family 11352 was additionally excluded from burden analyses. Family 13992 was included in the burden analyses as more stringent filters ameliorated that family's excess SNVs. Plots use previously published GDMs (Krumm et al. 2015) and exclude calls called homozygous by GATK.

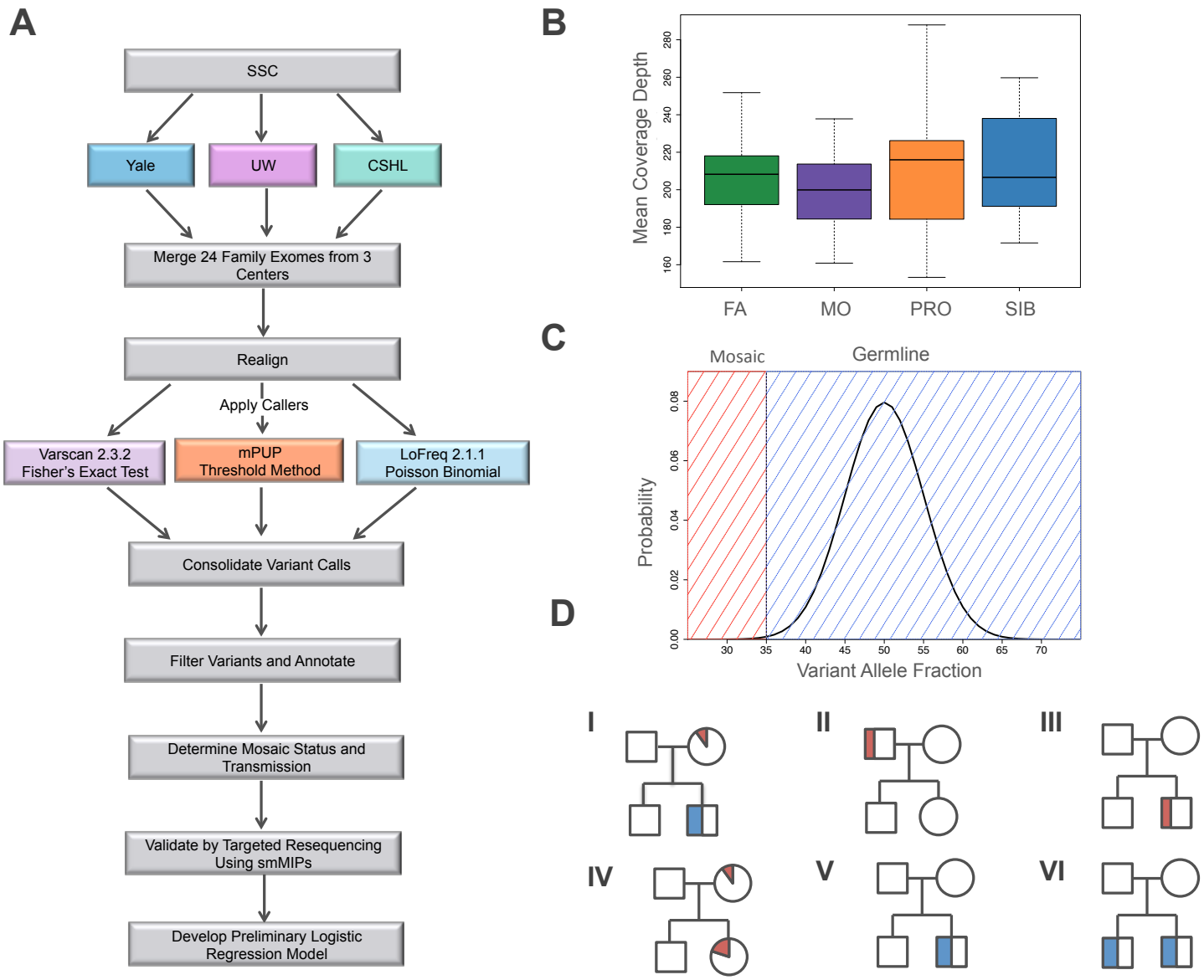



**Figure S2. Analysis Workflow for Pilot 24 PMM Predictions and Validations**

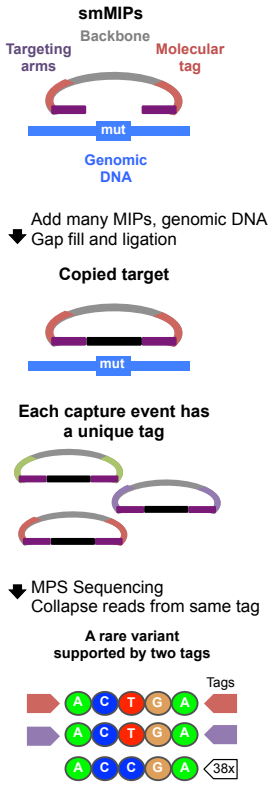
(A) For our first pilot study, we selected 24 families from the SSC collection that had WES performed in parallel by three different sequencing centers (Iossifov et al. 2014). Sequencing data were first merged per sample and then realigned using the method described in Krumm et al. 2015. Variants were called with two established, complementary variant callers (VarScan, LoFreq) and our script mPUP, a read count based method designed to maximize sensitivity. Variants were filtered and annotated as described in methods, then assigned predicted mosaic status and transmission. Candidate variants were validated by targeted resequencing. Results from validation were used as training data to develop a preliminary logistic model for scoring further predictions.

(B) Boxplots of mean coverages of merged WES data from pilot 24 families split by person type.

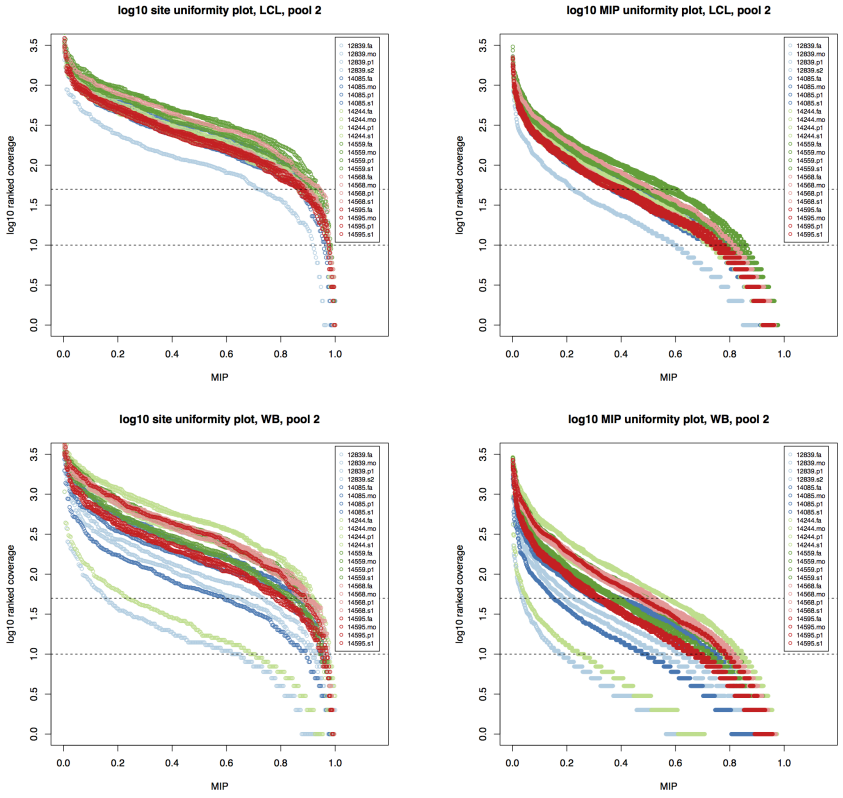
(C) Binomial probability distribution for a theoretical germline variant with 100x sequencing depth. This variant would be considered a putative PMM if fewer than 35 variant reads were observed (binomial p <= 0.001).

(D) Representative pedigrees illustrating variant transmission classifications, with germline variants in blue and PMMs in red. I. transmitted parental mosaic, II. nontransmitted parental mosaic, III. Child mosaic, IV. Possible transmitted parental mosaic (likely false mosaic signal), V. Germline *de novo*. VI. Gonadal mosaic.

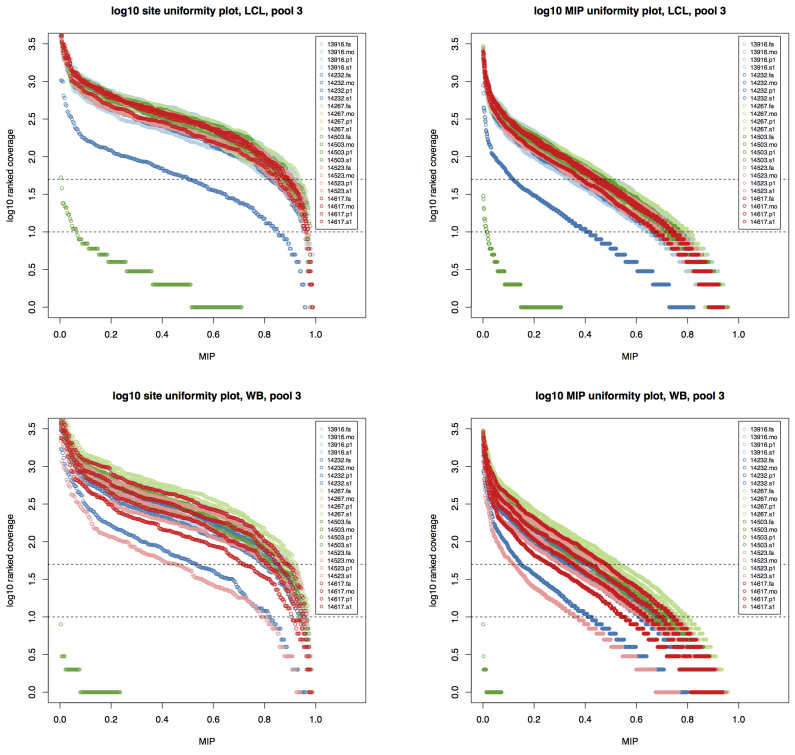
A


smMIPs
Backbone
Targeting arms
Molecular tag
mut
Genomic DNA
Add many MIPs, genomic DNA
Gap fill and ligation
Copied target
mut
Each capture event has a unique tag
MPS Sequencing
Collapse reads from same tag
A rare variant supported by two tags
Tags
A C T G A
A C T G A
A C C G A
38x

B


log10 site uniformity plot, LCL, pool 2
log10 MIP uniformity plot, LCL, pool 2
log10 site uniformity plot, WB, pool 2
log10 MIP uniformity plot, WB, pool 2
log10 ranked coverage
MIP

C


log10 site uniformity plot, LCL, pool 3
log10 MIP uniformity plot, LCL, pool 3
log10 site uniformity plot, WB, pool 3
log10 MIP uniformity plot, WB, pool 3
log10 ranked coverage
MIP

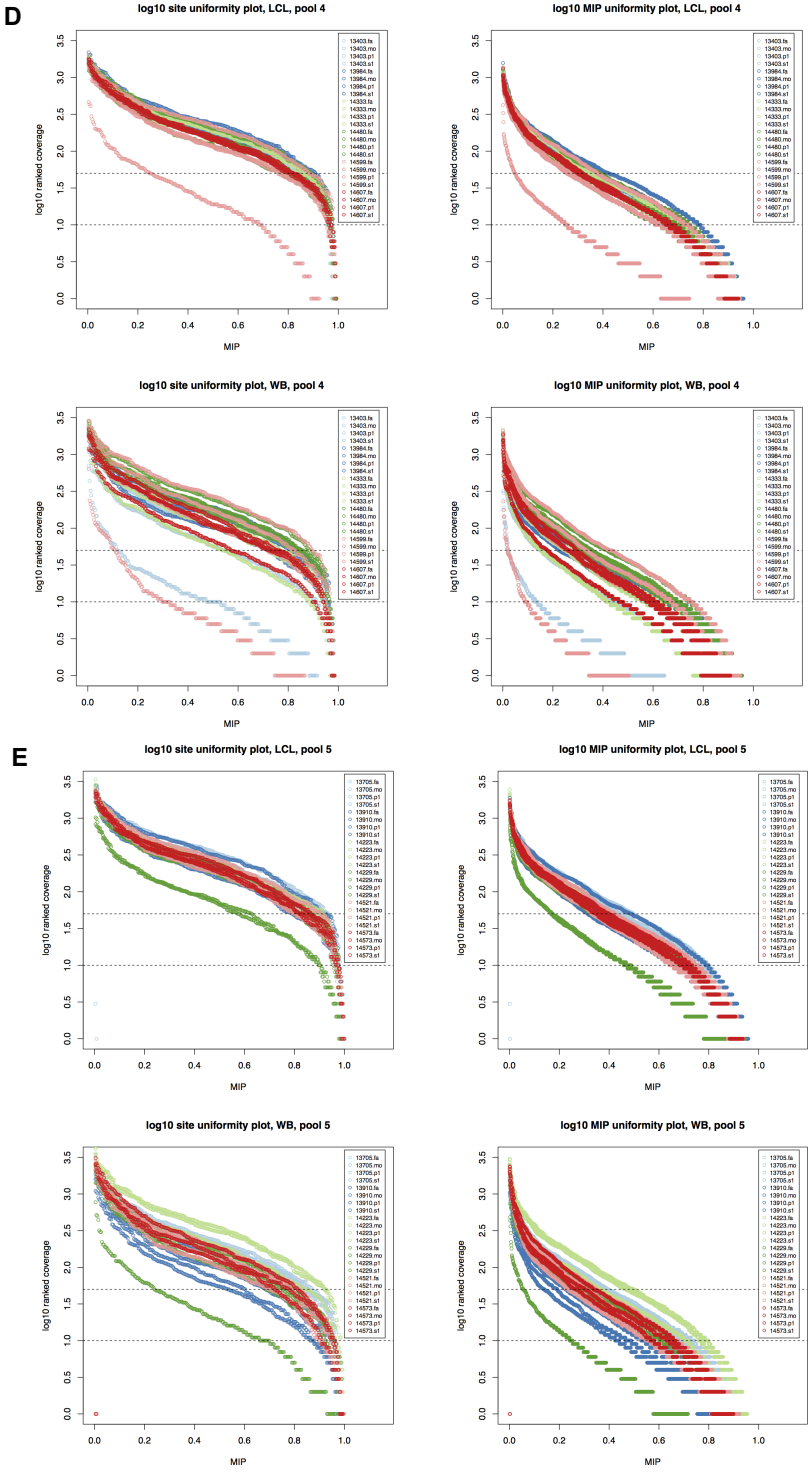



**Figure S3. Coverage Per-Site and Per-MIP Uniformity Plots from Pilot 24 Validation Sequencing**
(A) Schematic of targeted resequencing using smMIPs.
(B-E) Per-site plots show the summed coverage for all MIPs covering each target variant (left) and per-MIP plots show coverage for each MIP (right). Horizontal lines indicate reference thresholds of 10x and 50x coverage; in most pools, approximately 80% of calls achieved at least 50x total read depth. X-axes are scaled to the total number of MIPs or calls per pool for ease of comparison.
(B) Pool 2.
(C) Pool 3.
(D) Pool 4.
(E) Pool 5.

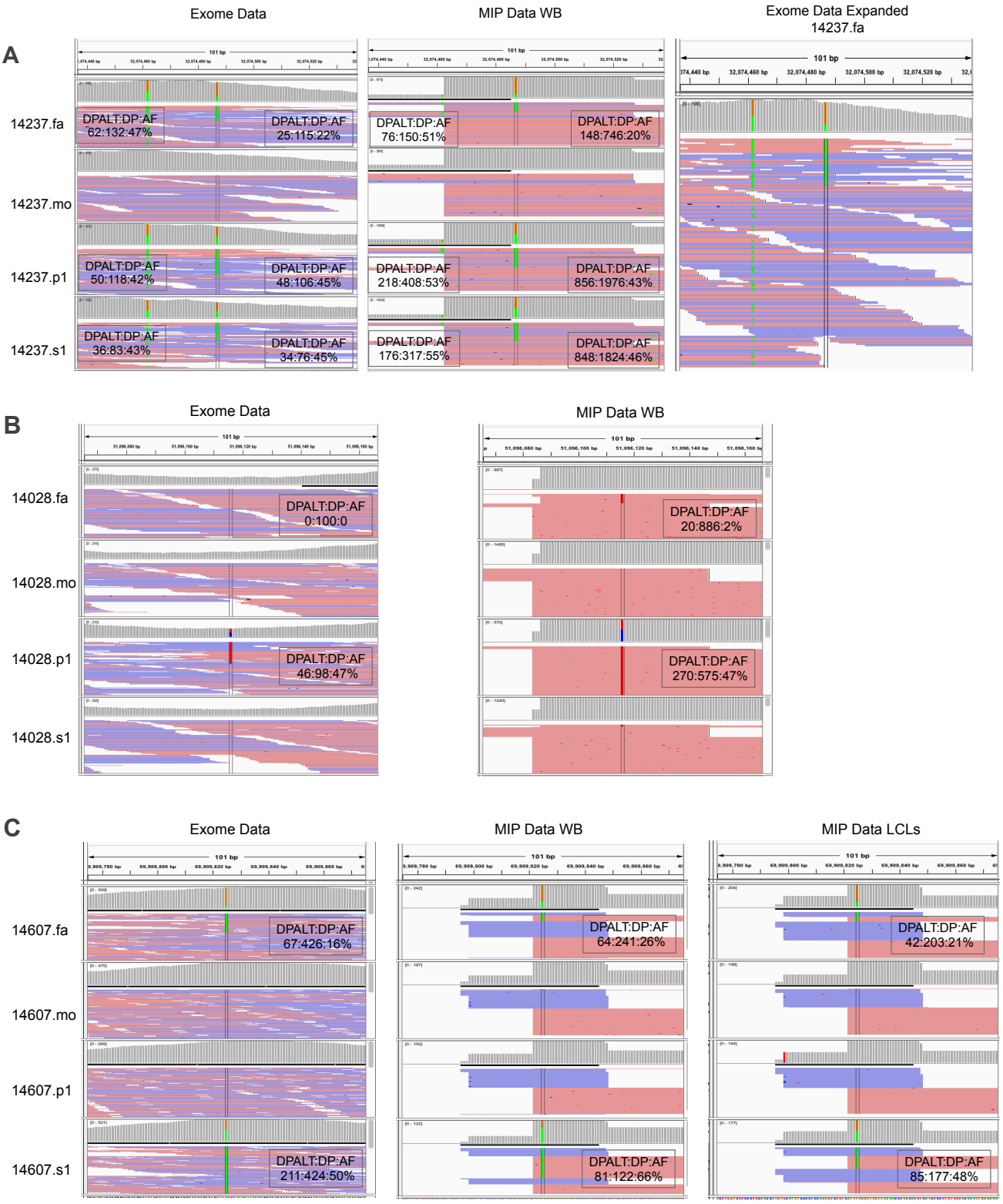



**Figure S4. Representative Read Alignments for Parental Transmitted Mosaic Variants**
(A) Parental PMM and associated germline SNP transmitted to both children.
(B) Example of a putative germline *de novo* call that is actually a cryptic parental mosaic
(C) Transmitted parental mosaic variant supported by exome and validation data. For this particular site (chr10:g.69909825G>A), a second validation was performed with independent probes. In the second validation, the allele counts were consistent in the child, WB: 214/456 (47%) and LCL: 46/98 (47%).
Abbreviations: fa-father, mo-mother, s-sibling, p-proband, WB-whole blood, LCL-lymphoblastoid cell line DPALT-Q20 alternative allele depth, DP-Q20 total site depth, AF-allele fraction.

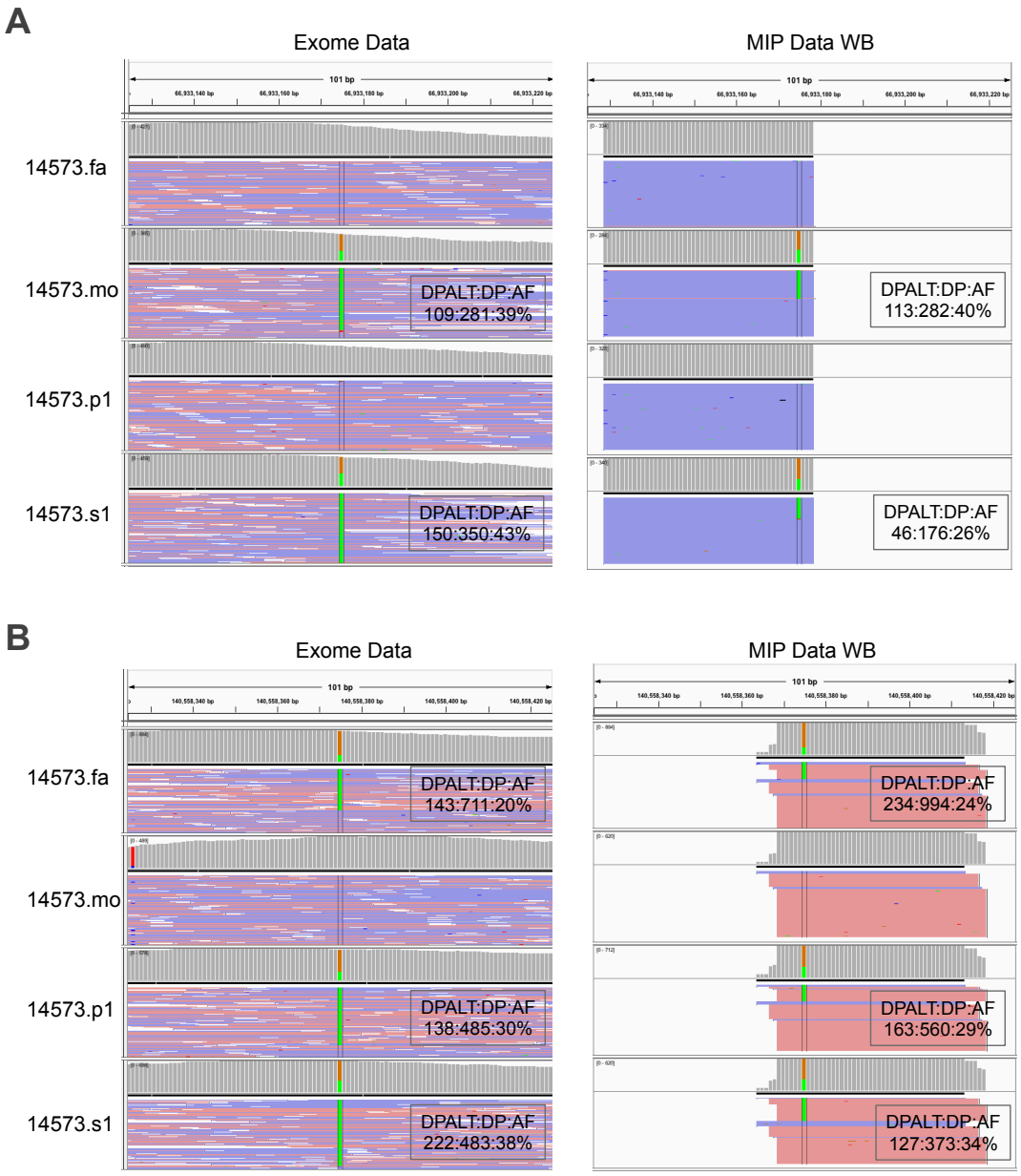



**Figure S5. Representative Read Alignments for Variants Transmitted with Skewed Allele Fractions**
(A) Maternal putative mosaic transmitted to proband with similarly skewed fraction.
(B) Second example of putative mosaic variant also skewed in both proband and sibling.
Abbreviations: fa-father, mo-mother, s-sibling, p-proband, WB-whole blood, LCL-lymphoblastoid cell line, DPALT-Q20 alternative allele depth, DP-Q20 total site depth, AF-allele fraction.

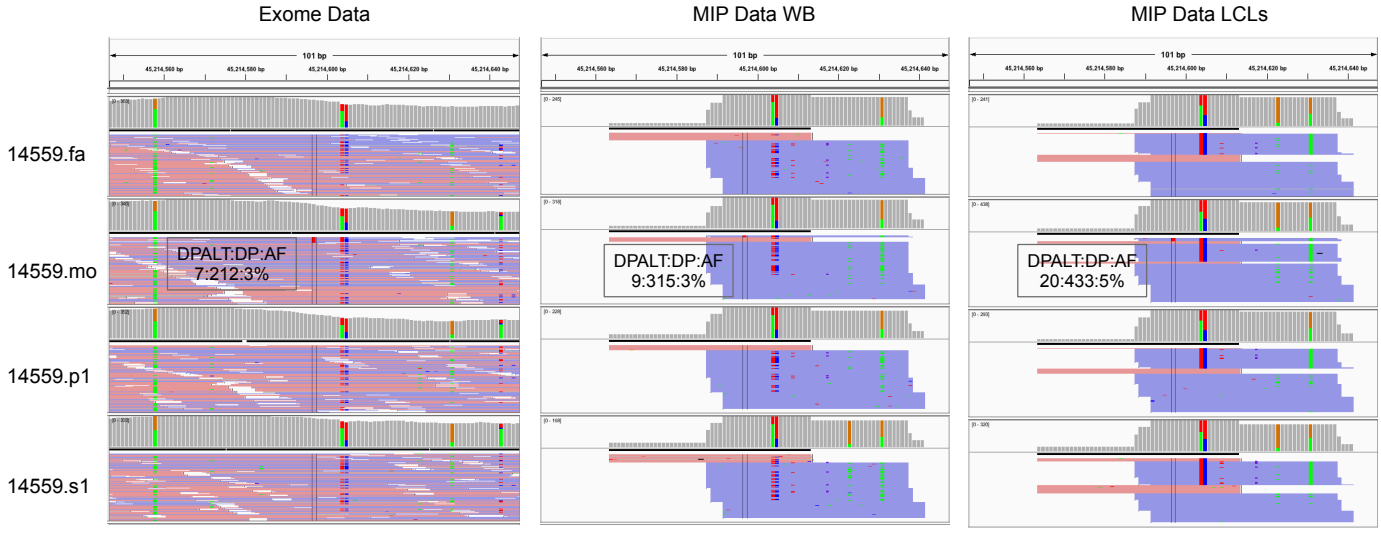
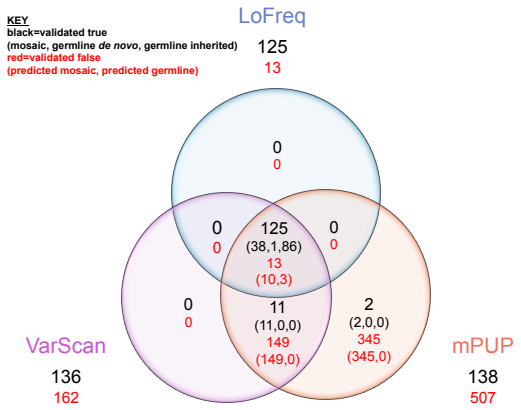



**Figure S6. Representative Read Alignments for Apparently Validated PMMs in Problematic Regions**
Predicted maternal PMM with multiple nearby variants in a segmental duplication.
Abbreviations: fa-father, mo-mother, s-sibling, p-proband, WB-whole blood, LCL-lymphoblastoid cell line DPALT-Q20 alternative allele depth, DP-Q20 total site depth, AF-allele fraction.

**A**




**B**

| Term (Type) | Univariate p-value | Definition |
|---|---|---|
| LOFREQ (T/F) | $7.34 \times 10^{-31}$ | LoFreq call |
| TV (T/F) | $2.65 \times 10^{-22}$ | transversion |
| DP (cont.) | $2.43 \times 10^{-17}$ | variant depth |
| AF (cont.) | $1.76 \times 10^{-12}$ | variant fraction |
| VARSCAN (T/F) | $2.05 \times 10^{-12}$ | VarScan call |
| DPALT (cont.) | $2.23 \times 10^{-11}$ | variant depth |
| SEGDUP (T/F) | $2.39 \times 10^{-7}$ | segmental dup. annotation |
| TRANSMITTED (T/F) | $4.21 \times 10^{-4}$ | transmitted to child(ren) |
| mPUP (T/F) | $1.99 \times 10^{-3}$ | mPUP call |
| GC_7MER (cont.) | $6.56 \times 10^{-3}$ | %GC content of variant and 3 bases either side |
| REF_CG (T/F) | 0.054 | reference allele |
| TRF (T/F) | 0.075 | tandem repeat finder annotation |
| REPEAT (T/F) | 0.178 | RMSK or TRF annotation |
| RMSK (T/F) | 0.342 | repeat masker annotation |
| PHET (cont.) | 0.767 | binomial probability of counts, given 50% expected fraction |

**C**

| Term | Coefficient | OR | P-value |
|---|---|---|---|
| LOFREQ | 0.823 | 2.28 | 0.312 |
| TV | -3.414 | 0.03 | $3.00 \times 10^{-7}$ |
| DP | 0.027 | 1.03 | 0.023 |
| VARSCAN | 2.760 | 15.80 | 0.008 |
| REF_CG | -0.116 | 0.89 | 0.841 |
| TV : DP | 0.010 | 1.01 | 0.001 |
| LOFREQ : REF_CG | 2.966 | 19.42 | 0.002 |
| DP : VARSCAN | -0.028 | 0.97 | 0.021 |

AIC: 188.69

Hoslem-Lemeshow p-value: 0.37

**D**

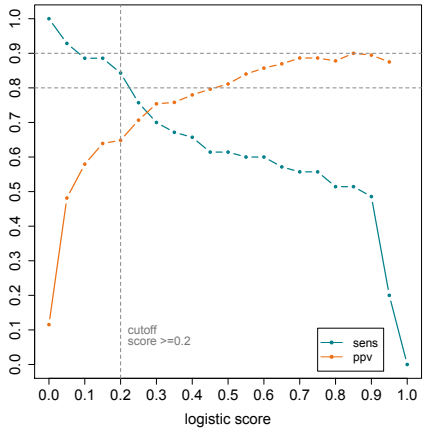



**Figure S7. Evaluation of the Initial Logistic Regression Model**

(A) Performance and Intersection of Variant Callers on Pilot 24 Predicted Mosaic High-Confidence Validation Outcomes

(B) Candidate predictor table with predictors and associated univariate model p-values. Abbreviations: cont-continuous variable, T/F-Boolean variable.

(C) Final model terms and performance metrics. Hoslem-Lemeshow p-value reported for groups = 10.

(D) Sensitivity (sens) and PPV curves from 3-fold cross-validation of model. Briefly, the training data was randomly divided into three groups, with two groups used for training and to score the reserved third. Each group was withheld in turn, with sensitivity and PPV averaged across all three iterations. Sensitivity is defined as the proportion of validated true variants scoring at or above the given value. For score >= 0.2, sensitivity = 0.85 and PPV = 0.67.

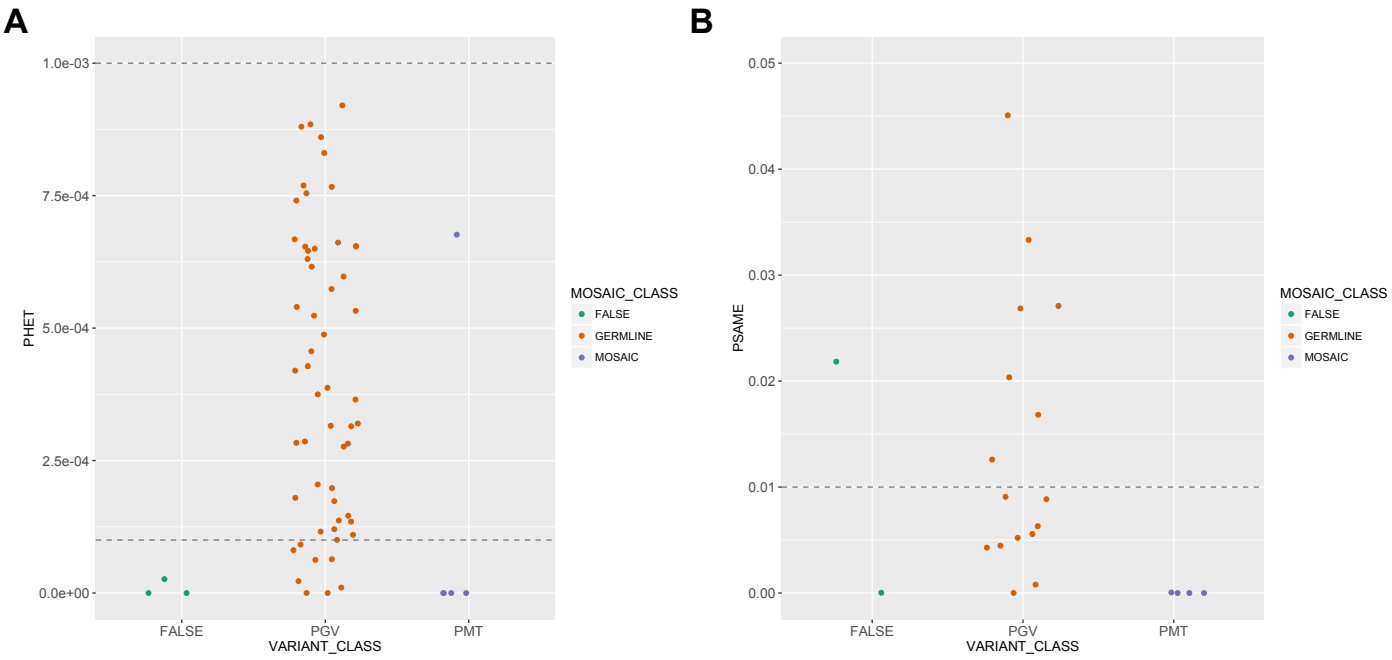
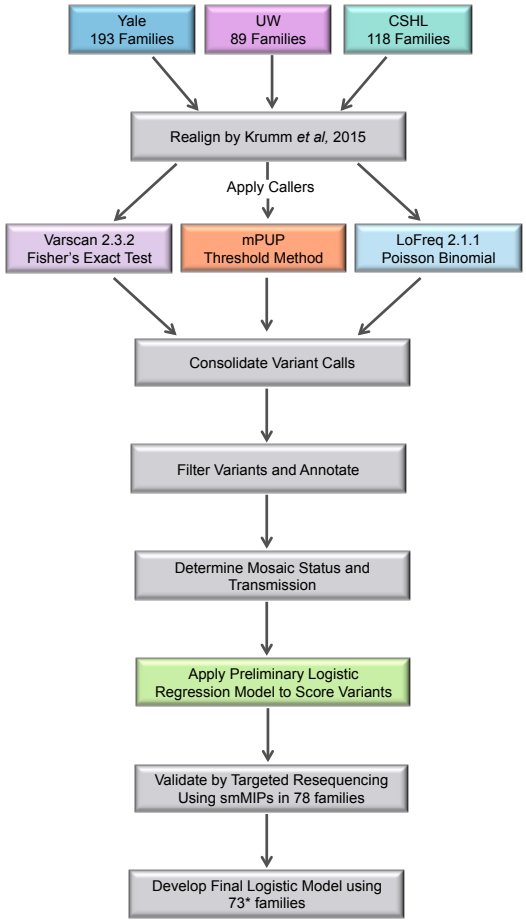



**Figure S8. Filters Applied to Putative Transmitted Variants Subsequent to Pilot 24 Validations**
(A) Binomial probabilities for observed exome read counts of all pilot 24 predicted transmitted PMMs variants with high-confidence resolutions, with original threshold at p <= 0.001 and more stringent cutoff at p <= 0.0001. Nearly all validated PMMs fall well below the stricter threshold. Jitter applied for visibility.
(B) Fisher's exact test probabilities of difference between child and adult allele read counts for the same dataset. All validated PMMs fall well below the threshold of p <= 0.01. Jitter applied for visibility.
Abbreviations: PGV-parental germline transmitted variant, PMT-parental mosaic transmitted.

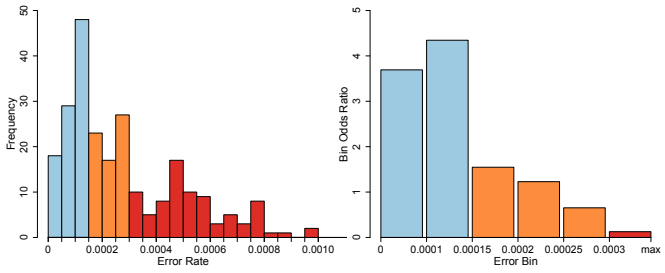
A



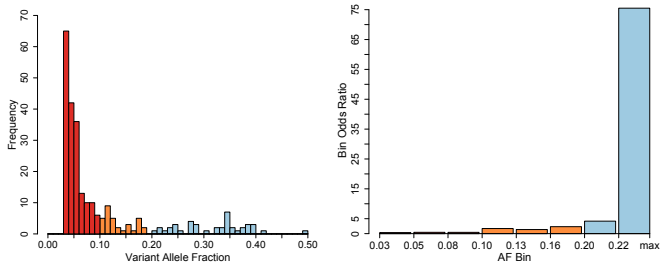

B

| Term (Type) | Univariate p-value | Definition |
|---|---|---|
| DPALT (bin) | $<2.0 \times 10^{-16}$ | variant depth |
| ERROR (bin) | $1.54 \times 10^{-12}$ | site error rate |
| LOFREQ (T/F) | $1.77 \times 10^{-9}$ | LoFreq call |
| AF (bin) | $9.48 \times 10^{-8}$ | variant fraction |
| XM_ALT_1 (T/F) | $2.48 \times 10^{-6}$ | variant read median mismatches |
| DP (bin) | $7.14 \times 10^{-6}$ | total depth |
| REF_CG (T/F) | $8.35 \times 10^{-3}$ | reference allele |
| MAPQ_ALT (bin) | 0.143 | variant read median MAPQ |
| TV (T/F) | 0.161 | transversion |
| TRF (T/F) | 0.215 | tandem repeat finder annotation |
| SEGDUP (T/F) | 0.575 | segmental dup. annotation |
| RMSK (T/F) | 0.916 | repeat masker annotation |
| VARSCAN (T/F) | 0.985 | VarScan call |
| mPUP (T/F) | 0.987 | mPUP call |

C




D




E

| Term | Coefficient | OR | P-value |
|---|---|---|---|
| DPALT | 2.497 | 12.15 | $8.47 \times 10^{-8}$ |
| ERROR - high | -2.421 | 0.09 | $1.59 \times 10^{-5}$ |
| ERROR - med | -0.785 | 0.46 | 0.125 |
| AF - high | 2.830 | 16.94 | $1.79 \times 10^{-3}$ |
| AF - med | 0.584 | 1.79 | 0.329 |
| XM_ALT_1 | 1.496 | 4.46 | $5.64 \times 10^{-3}$ |
| REF_CG | 1.483 | 4.41 | $2.09 \times 10^{-3}$ |

AIC: 171.91

Hoslem-Lemeshow
p-value: 0.56

**Figure S9. Construction Process for the Refined Logistic Model**

(A) For our expanded pilot study, we used existing WES (Krumm et al. 2015) for 400 families from the SSC collection that had WES performed across three sequencing centers. Variants were called with two established, complementary variant callers (VarScan, LoFreq) and our script mPUP, a read count based method designed to maximize sensitivity. Variants were then filtered and annotated as described in methods. Predicted mosaic status and transmission were determined for filtered variants, and predicted PMMs scored using a preliminary logistic regression model trained on the earlier pilot validations. Variants in 78 families with were validated by targeted resequencing using smMIPs. Validation results were then used to develop our refined logistic model. *5 families were excluded as outliers.

(B) Candidate predictor table with predictors and associated univariate model p-values. Abbreviations: bin-binned continuous variable, T/F-Boolean variable, Coef.-term coefficient in model.

(C) Example of binning process showing error rate distribution and associated odds ratio distribution; colors indicate ranges collapsed into categories for final model.

(D) Variant AF distribution and associated odds ratio distribution, similar to (B).

(E) Final model terms and performance metrics. Hoslem-Lemeshow p-value reported for groups = 10.

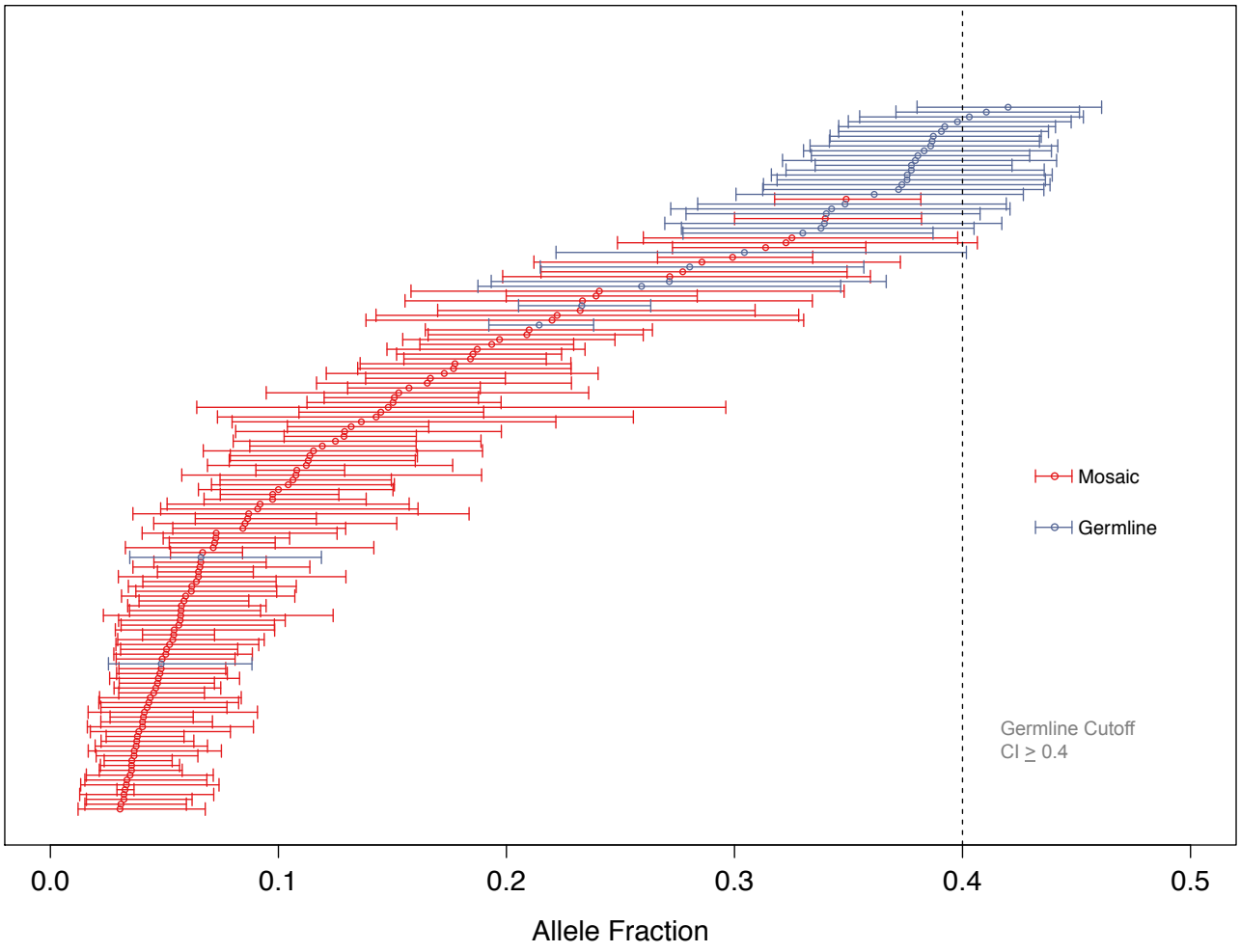



**Figure S10. Distribution of AF Confidence Intervals for Pilot PMMs Validated Mosaic or Germline**
WES AFs and confidence intervals for sites initially predicted mosaic and with validation data for pilot 24 (24 quads) and pilot 400 (78 quads) families. Initial logistic model, pilot 400 singleton, and mismatch filters applied. Reclassifying predicted PMMs with 90% confidence intervals overlapping 0.4 as germline correctly excludes 25/33 (76%) germline resolutions and retains 112/113 (99%) mosaic resolutions. Plot includes validation data for both parents and children. Confidence intervals calculated using Agresti-Coull method.

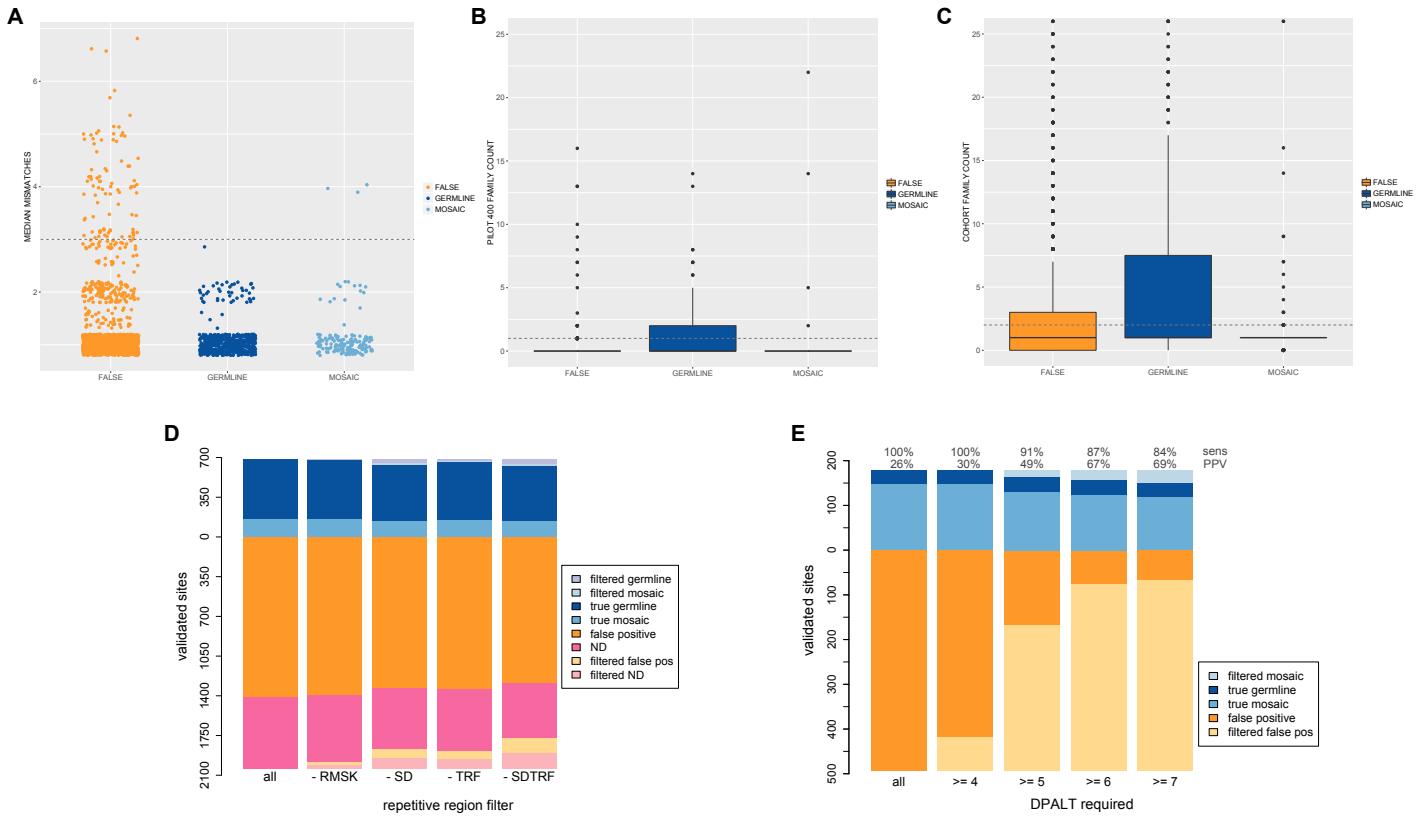



**Figure S11. Development of Additional Filters Based on Validation Outcomes**

(A) Median mismatches in variant reads for pilot 24 and 400 validated calls by validation outcome, with jitter applied to points for visibility. Filter threshold at <= 3 selected to retain all validated germline *de novo* calls.

(B) Occurrence of pilot 24 variants in pilot 400 families, with filter threshold at < 1. Variants in multiple families typically validated as false or parental germline.

(C-E) Evaluation of additional factors driving false calls on pilot 24 and 400 validations after applying refined logistic regression model, variant read mismatch (A), and single pilot 400 (B) filters.

(C) Occurrence of all validated calls across entire cohort, with filter threshold at <= 2. Variants present in more families typically validated as false or parental germline.

(D) Effects on true, false, and indeterminate outcomes of excluding repetitive sequence annotation. Excluding both SD and TRF regions substantially reduced problematic calls and false validations. Abbreviations: RMSK-RepeatMasker, SD-segmental duplication, TRF-Tandem Repeat Finder, ND-indeterminate or low-confidence validations.

(E) Effect of successively more stringent variant read depth (DPALT) filters on sensitivity and PPV for predicted PMMs in all validation groups passing all other filters except logistic score. Threshold of >= 5 variant reads selected to substantially reduce false positives while still passing ~90% of true calls into model scoring. No true germline variants were filtered under any threshold tested. Calls with indeterminate or low-confidence validations were not included.

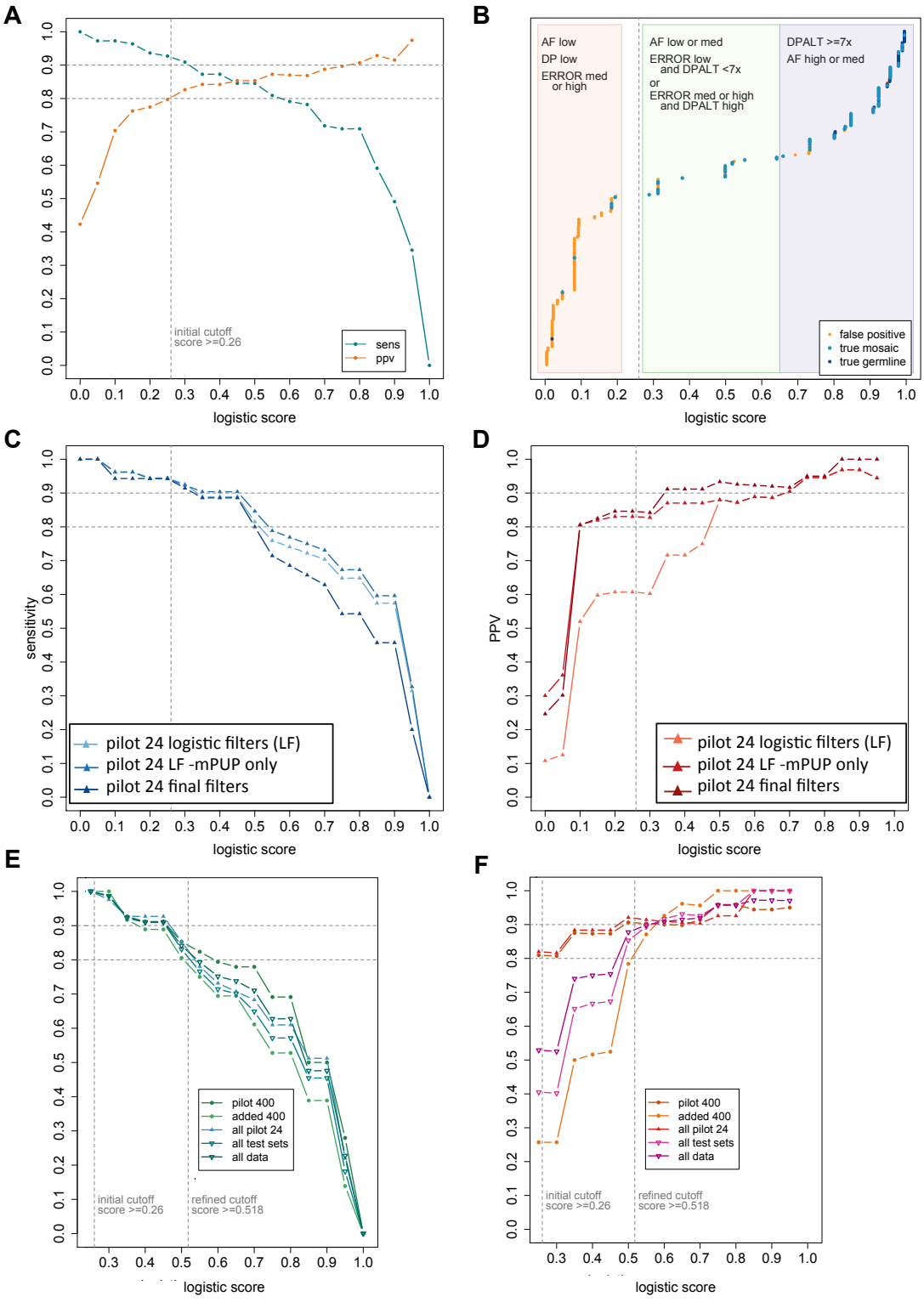

A
initial cutoff
score >=0.26
sens
ppv
logistic score
B
AF low
DP low
ERROR med
or high
AF low or med
ERROR low
and DPALT <7x
or
ERROR med or high
and DPALT high
DPALT >=7x
AF high or med
false positive
true mosaic
true germline
logistic score
C
sensitivity
pilot 24 logistic filters (LF)
pilot 24 LF -mPUP only
pilot 24 final filters
logistic score
D
PPV
pilot 24 logistic filters (LF)
pilot 24 LF -mPUP only
pilot 24 final filters
logistic score
E
pilot 400
added 400
all pilot 24
all test sets
all data
initial cutoff
score >=0.26
refined cutoff
score >=0.518
logistic score
F
pilot 400
added 400
all pilot 24
all test sets
all data
initial cutoff
score >=0.26
refined cutoff
score >=0.518
logistic score

**Figure S12. Evaluation of Refined Logistic Regression Model Performance on Training Set and Pilot 24 Validations**

(A) Sensitivity and PPV curves from 3-fold cross-validation using training set of pilot 400 predicted PMMs with high-confidence resolutions. All validated variants are considered true positives, regardless of germline or mosaic status.

(B) Ranked score plot showing validation outcomes for training set against the characteristic predictors defining score ranges.

(C) Sensitivity curves for successively more stringent filters applied to pilot 24 predicted PMMs with high-confidence resolutions. Sensitivity for each filter set is defined using the set of validated true calls that pass filters regardless of logistic score. At logistic score cutoff 0.26, sensitivity is 0.94 for all filter sets. Logistic filters (LF) are the same filters applied in the pilot 400 dataset for model building. Intermediate line "-mPUP only" removes calls identified solely by the mPUP script. Final filters, adds the additional heuristic established, such as removing mPUP only and SD/TRF calls, updated mosaic predictions based on upper 90% CI, and cohort-wide family count <= 2. Although final filters reduce apparent sensitivity at higher scores, excluded calls were predominantly parental mosaic predictions with germline resolutions (data not shown).

(D) PPV curves for the same filter sets as in C. At cutoff 0.26, PPV values are 0.61 (LF), 0.83 (LF-mPUP), and 0.85 (final filters).

(E-F) Summary of performance of all validation data using refined logistic regression model and final filter heuristics, which are: removing mPUP only and SD/TRF calls, updated mosaic predictions based on upper 90% CI, and cohort-wide family count <= 2, removal of outlier families, logistic score > 0.26, and pilot 400 singletons. Pilot 400 are the training set. Added 400 are new pilot 400 calls tested after model development. All pilot 24 are initial validations and additional calls tested after model development (combined due to low numbers in latter set). All test sets combines the pilot 24 and added pilot 400 calls.

(E) Sensitivity curves for all validation sets. Sensitivity for each set is defined using the set of validated true calls that pass filters regardless of logistic score.

(F) PPV curves for all validation sets.

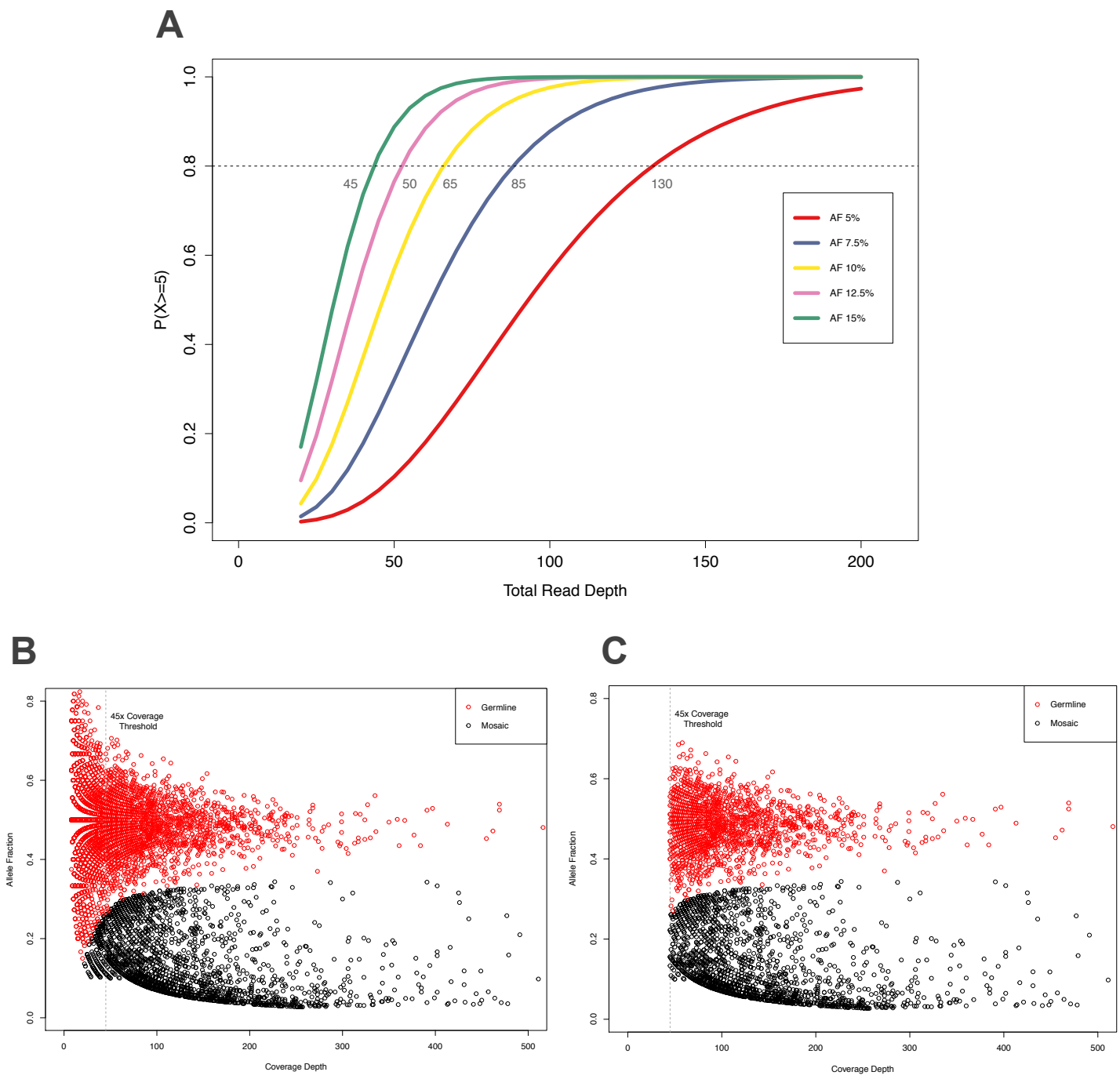



**Figure S13. Defining Coverage Thresholds with Adequate Power to Detect AFs**

(A) Probability of observing at least 5 variant reads across a range of read depths for the given variant allele fractions. Numbers beside lines denote the approximate read depths at which the probability curve crosses 0.8.

(B-C) Comparison of coverage depth to allele fraction of calls within full SSC cohort. Germline variants in red and mosaic in black.

(B) Best practice filters applied but not 5%-45x high confidence threshold.

(C) After 5%-45x threshold applied.

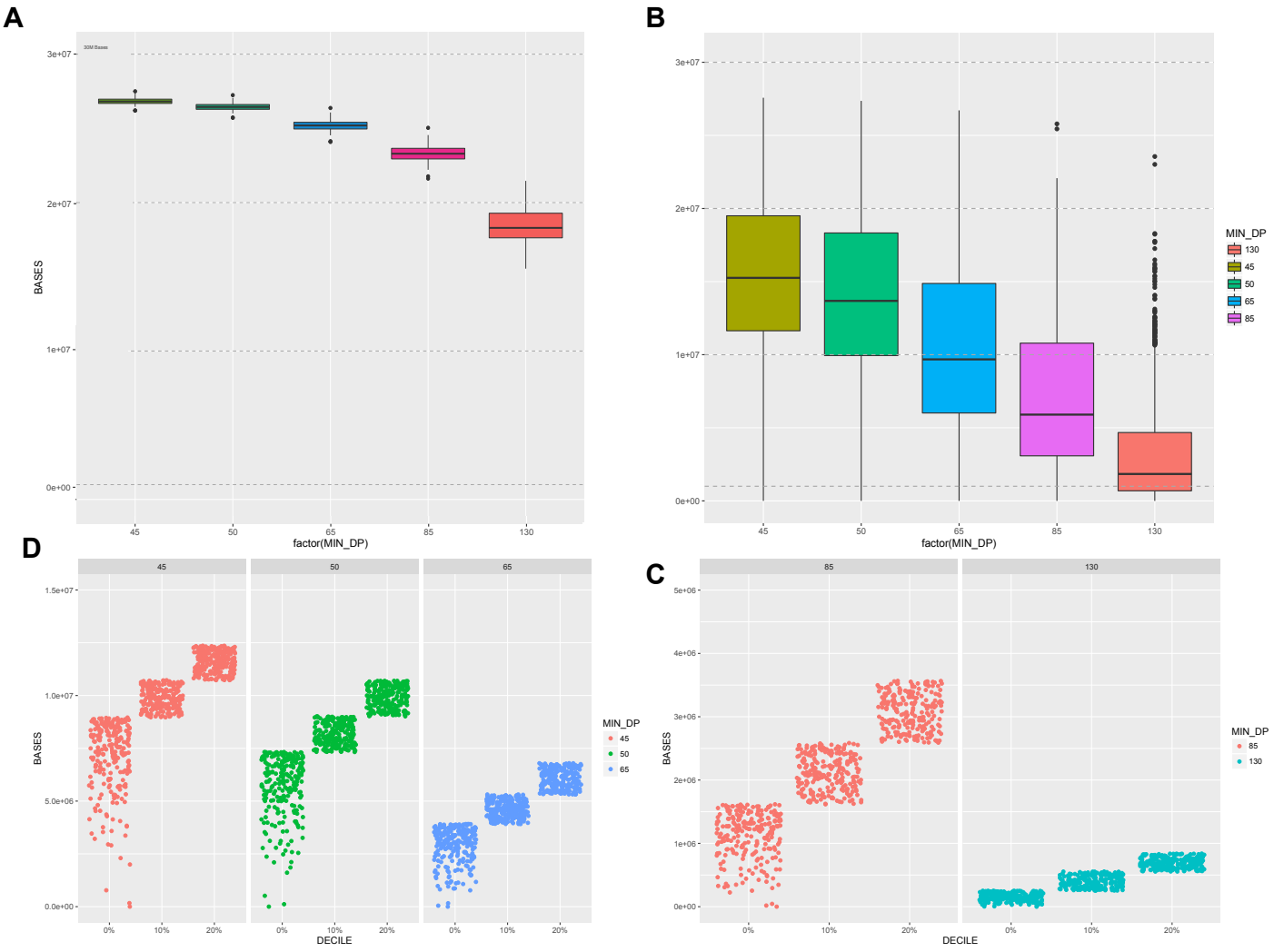



**Figure S14. Coverage Distributions by Burden Analysis Depth Threshold**
(A) Boxplots of total haploid genome bases for merged pilot 24 families at each minimum depth threshold.
(B) Boxplots of total haploid genome bases sequenced across the cohort at each minimum depth threshold.
(C-D) Lowest three coverage deciles for each analysis group, with horizontal jitter applied for visibility of points. Approximately half of the lowest decile shows considerable spread for all coverages except 130x.
Plots include both quad and trio families, and also include families determined to be outliers by SNV counts.
(C) Minimum joint coverage of 45x, 50x, and 65x.
(D) Minimum joint coverage of 85x and 130x.

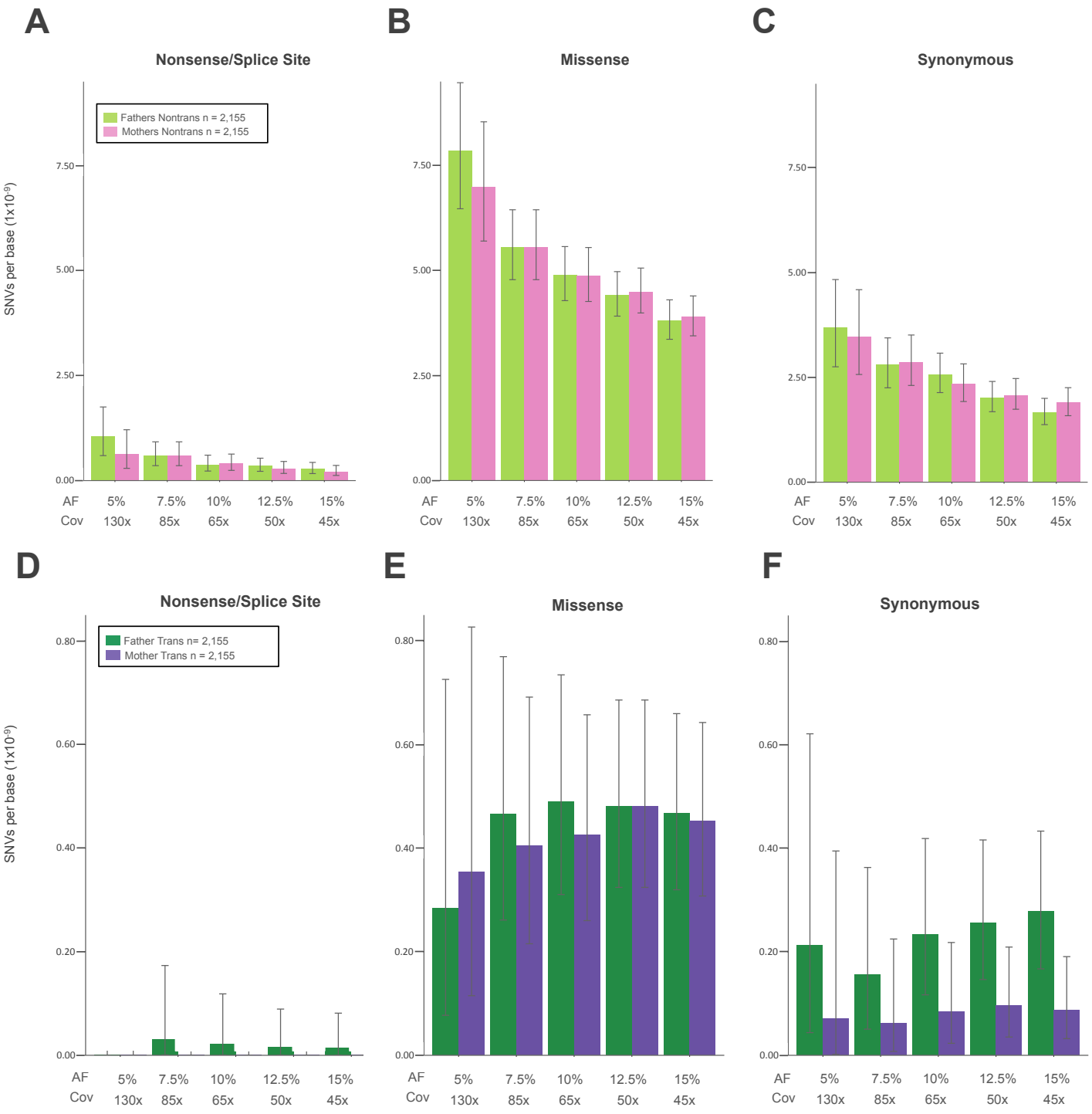



**Figure S15. Rate of Parental PMMs for Different Functional Classes**
Rates and burden analyses of PMMs in full SSC. Mean rates with 95% Poisson CIs (exact method) are shown for parents.
(A) Nonsense/Splice Site Nontransmitted PMMs.
(B) Missense Nontransmitted PMMs.
(C) Synonymous Nontransmitted PMMs.
(D) Nonsense/Splice Site Transmitted PMMs.
(E) Missense Transmitted PMMs.
(F) Synonymous Transmitted PMMs.

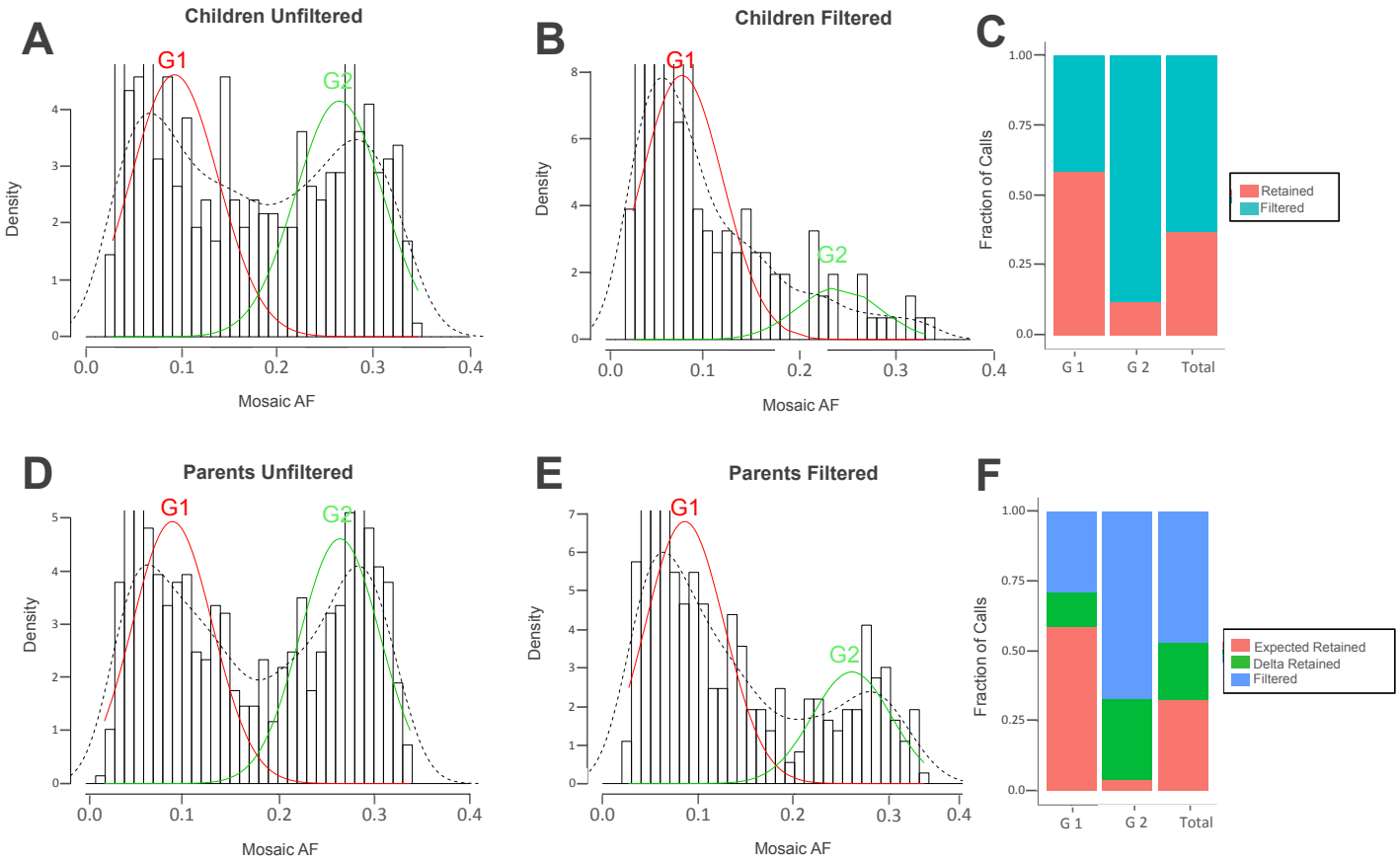



**Figure S16. Distribution of Allele Fractions Before and After Transmission Based Filtering**

To determine the percentage of parental calls that may be due to incomplete filtering from inability to compare to previous generation, we determined the number of mosaic variants within children that were removed due to transmission filters. We took variants from a subset of the harmonized reprocessed cohort (pilot 24 and 400 families) and ignored transmission, but applied model scoring and all other final filters. AF distributions were fitted using a normal mixed model with R package *mixtools,* function *normalmixEM()*. The red distribution represents Gaussian distribution G1 and the green distribution represents G2. Dashed Curve represents the observed AF distribution density.

(A) AF distributions for variants in children (probands and siblings) before applying transmission filters fitted to a mixed model.

(B) AF distributions for variants in children after applying transmission filters fitted to a mixed model.

(C) For G1 (lower AFs), we combined calls within two standard deviations of the estimated mean. For G2, we combined calls more extreme than the mean of G1 plus two standard deviations. We then calculate the fraction of variants remaining after applying transmission filters. In G1, 41% of variants were filtered, 88% of variants in G2, and 63% overall.

(D) AF distributions for variants in parents before applying transmission filters fitted to a mixed model.

(E) AF distributions for variants in parents after applying transmission filters fitted to a mixed model. Plot depicts both nontransmitted and transmitted PMMs. Retained transmitted PMMs required a stricter binomial filter ($p <= 0.0001$)

(F) For parents, 29% of variants in G1 were filtered, 67% of variants in G2, and 47% overall. The number actually retained (observed) is 71% in G1, 33% in G2, and 53% overall. Using the fraction retained for each Gaussian distribution in children, we estimated how many variants in parents we expect to retain if the same transmission data were available. We would expect to only retain 59% in G1, 4% in G2, and 33% overall. The Delta is the difference between the observed calls and expected which is 12% in G1, 29% in G2, and 20% overall. Based on the filter fraction rates from children, we estimate that 20% of the remaining calls in G1, 88% of remaining calls in G2, and 40% of the total remaining calls are likely due to incomplete transmission filtering.

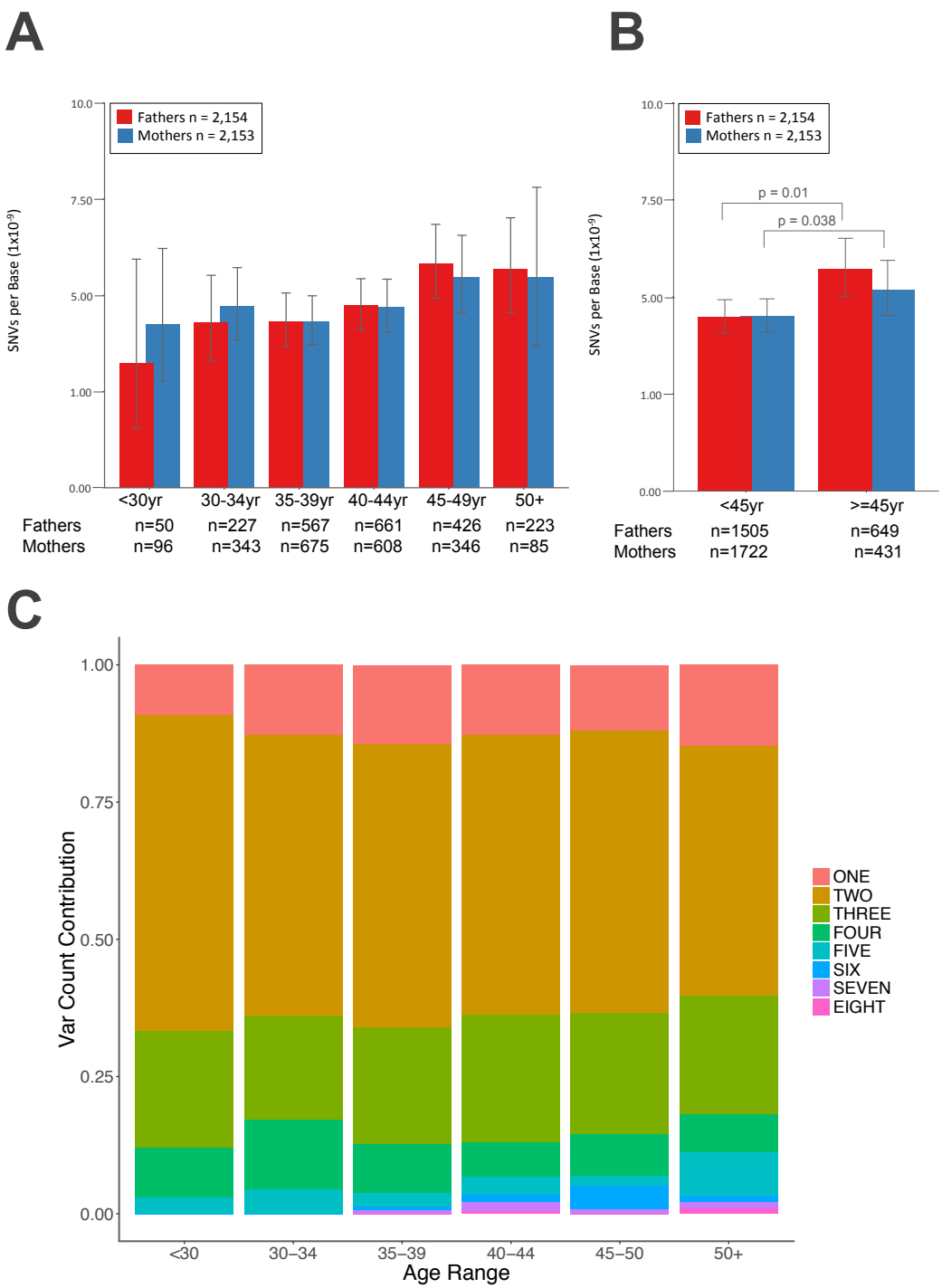



**Figure S17. Rate of Parental Nontransmitted PMMs with Age**

We used age given at time of blood draw. If not available, then we estimated age using age of parent at birth of proband and added age of proband at ADOS, which was conducted near time of blood draw.

(A-B) Rates and burden analyses of PMMs within the 5%-45x set for a given age bin. Mean rates with 95% Poisson CIs (exact method) are shown for parents.

(A) Age of parents divided into six age bins.

(B) To increase power, we divided parents into two age bins to compare mutation burden. Significance determined using Wilcoxon rank sum test, one-sided. We see a significant increase in mutation rate for both mothers and fathers older than 45 yrs.

(C) To adjust for differences in coverage we determined the percentage of the exome covered for each individual and extrapolated the observed variant count to the entirety of the exome. Age of parents was divided into six bins. The fraction of individuals with a given number of coverage adjusted variants within an age bin is shown. Data suggests more individuals appear to accumulate PMMs as they age.

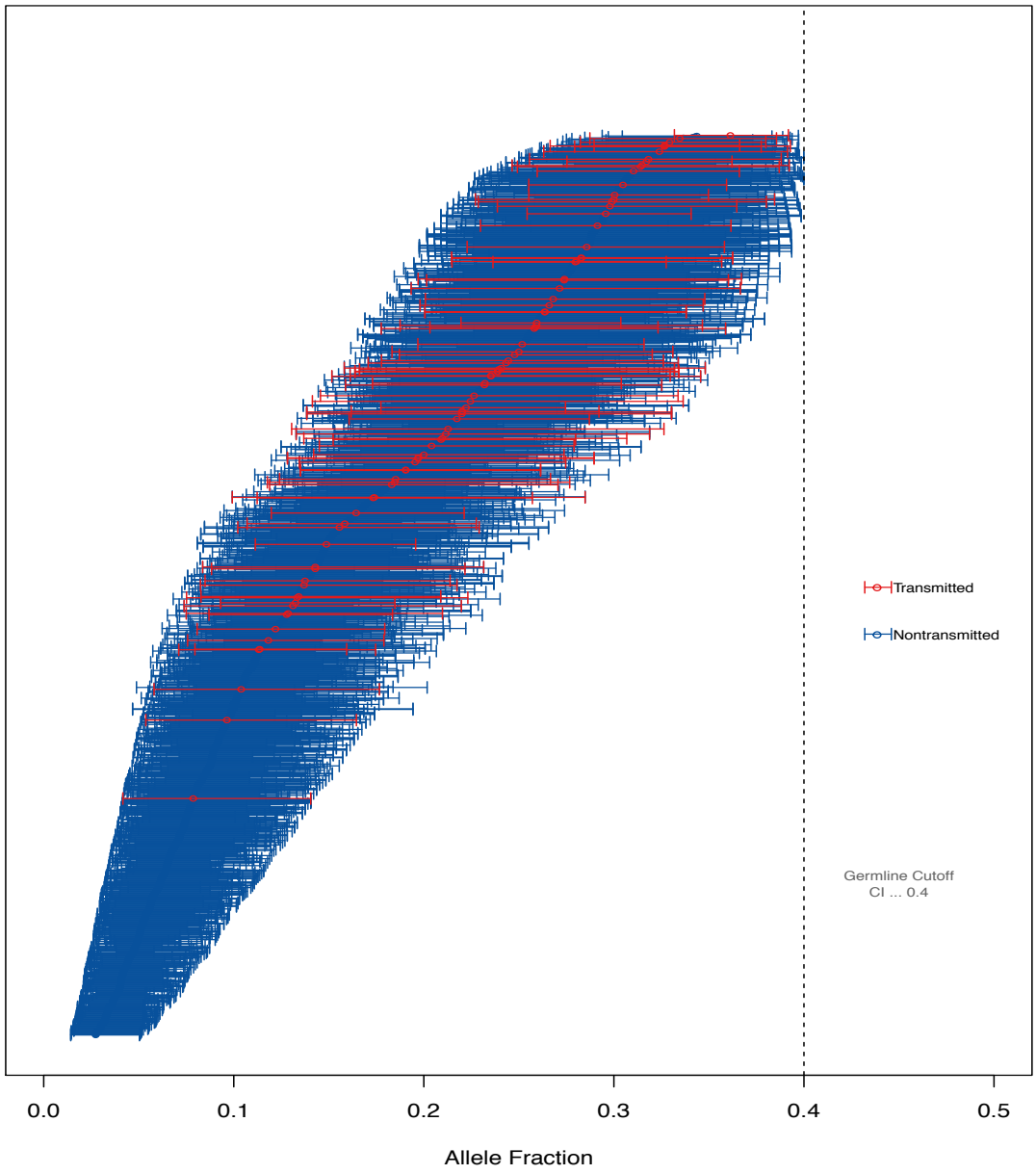



**Figure S18. Distribution of AF Confidence Intervals for Parental PMMs**
WES AFs and confidence intervals for sites validated within the pilot 24 (24 quads) and pilot 400 (78 quads) families. Confidence intervals calculated using Agresti-Coull method. Confidence intervals overlapping 0.4 would be considered germline. Transmitted variants tend to skewer higher in AF.

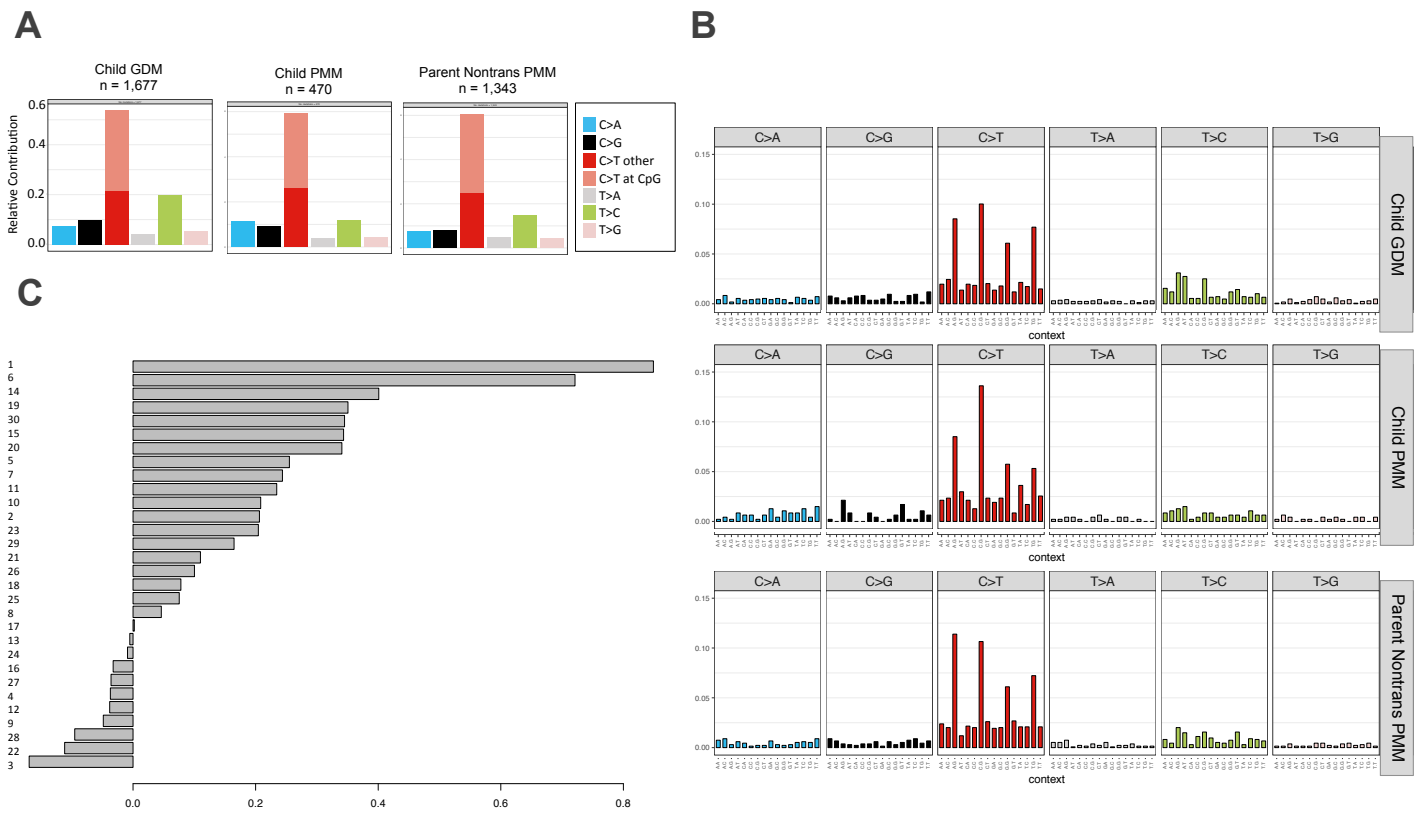



**Figure S19. Mutational Spectrum and Signature**

The R package *MutationalPatterns*[17] was used to extract and plot mutational contexts, as well as calculating their frequency within our high confidence call set.

(A) Mutational spectrum of the six different types of substitutions for child GDMs, child PMMs, and parent nontransmitted PMMs.

(B) Mutational signature of the relative frequency of mutations (Y-axis) within trinucleotides (context) for child GDMs, child PMMs, and parent nontransmitted PMMs.

(C) We determined the correlation by Pearson method of the trinucleotide frequencies with the 30 different cancer signatures observed in Alexandrov et al. 2013 (see Web Resources for download).[18] We found child GDMs, child PMMs, and parent nontransmitted PMMs all are most correlated with cancer signature 1 and all have similar correlation profiles. Shown is the correlation profile of child PMMs and cancer signatures as a representative profile.

**Table S7. Results of Rare Inherited Variant Simulations**

| | | | AF < 0.5 (left tail) | | | | | AF > 0.5 (right tail) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Region | Total Mut # | Exp | Obs | E-Frac | O-Frac | p-value | Exp | Obs | E-Frac | O-Frac | p-value |
| Probands | | | | | | | | | | | | |
| p <= 0.001 True | | | | | | | | | | | | |
| SNVS | Unique CDS | 2662 | 33 | 250 | 0.01 | 0.09 | < 0.0001 | 6 | 7 | 0.002 | 0.003 | 0.399 |
| | SD/TRF | 231 | 42 | 55 | 0.18 | 0.24 | 0.017 | 4 | 2 | 0.017 | 0.009 | 0.87 |
| | Total | 2893 | 78 | 305 | 0.03 | 0.11 | < 0.0001 | 10 | 9 | 0.003 | 0.003 | 0.667 |
| Indels | Unique CDS | 250 | 15 | 50 | 0.06 | 0.20 | < 0.0001 | NA | NA | NA | NA | NA |
| | SD/TRF | 18 | 1 | 7 | 0.06 | 0.39 | 0.0003 | NA | NA | NA | NA | NA |
| | Total | 268 | 16 | 57 | 0.06 | 0.21 | < 0.0001 | NA | NA | NA | NA | NA |
| p <= 0.0001 True | | | | | | | | | | | | |
| SNVS | Unique CDS | 2662 | 19 | 200 | 0.007 | 0.08 | < 0.0001 | 2 | 2 | 0.001 | 0.001 | 0.493 |
| | SD/TRF | 231 | 33 | 51 | 0.14 | 0.22 | 0.0007 | 3 | 1 | 0.013 | 0.004 | 0.943 |
| | Total | 2893 | 56 | 251 | 0.02 | 0.09 | < 0.0001 | 5 | 3 | 0.002 | 0.001 | 0.849 |
| Indels | Unique CDS | 250 | 6 | 35 | 0.02 | 0.14 | < 0.0001 | NA | NA | NA | NA | NA |
| | SD/TRF | 18 | <1 | 5 | 0.00 | 0.28 | < 0.0001 | NA | NA | NA | NA | NA |
| | Total | 268 | 7 | 40 | 0.03 | 0.15 | < 0.0001 | NA | NA | NA | NA | NA |
| Siblings | | | | | | | | | | | | |
| p <= 0.001 True | | | | | | | | | | | | |
| SNVS | Unique CDS | 1849 | 24 | 163 | 0.02 | 0.09 | < 0.0001 | 4 | 2 | 0.002 | 0.001 | 0.902 |
| | SD/TRF | 144 | 27 | 28 | 0.19 | 0.19 | 0.4 | 3 | 3 | 0.021 | 0.021 | 0.391 |
| | Total | 1993 | 47 | 191 | 0.03 | 0.10 | < 0.0001 | 7 | 5 | 0.004 | 0.003 | 0.8 |
| Indels | Unique CDS | 124 | 8 | 39 | 0.06 | 0.31 | < 0.0001 | NA | NA | NA | NA | NA |
| | SD/TRF | 16 | 1 | 5 | 0.06 | 0.31 | < 0.0001 | NA | NA | NA | NA | NA |
| | Total | 140 | 10 | 48 | 0.07 | 0.34 | < 0.0001 | NA | NA | NA | NA | NA |
| p <= 0.0001 True | | | | | | | | | | | | |
| SNVS | Unique CDS | 1849 | 15 | 136 | 0.008 | 0.07 | < 0.0001 | 1 | 1 | 0.001 | 0.001 | 0.623 |
| | SD/TRF | 144 | 22 | 22 | 0.15 | 0.15 | 0.49 | 2 | 2 | 0.014 | 0.014 | 0.516 |
| | Total | 1993 | 41 | 158 | 0.02 | 0.08 | < 0.0001 | 3 | 3 | 0.002 | 0.002 | 0.606 |
| Indels | Unique CDS | 124 | 4 | 25 | 0.03 | 0.20 | < 0.0001 | NA | NA | NA | NA | NA |
| | SD/TRF | 16 | <1 | 8 | 0.00 | 0.50 | < 0.0001 | NA | NA | NA | NA | NA |
| | Total | 140 | 4 | 33 | 0.03 | 0.24 | < 0.0001 | NA | NA | NA | NA | NA |
| Combined | | | | | | | | | | | | |
| p <= 0.001 True | | | | | | | | | | | | |
| SNVS | Unique CDS | 4511 | 57 | 413 | 0.01 | 0.09 | < 0.0001 | 10 | 9 | 0.002 | 0.002 | 0.665 |
| | SD/TRF | 375 | 68 | 83 | 0.18 | 0.22 | 0.03 | 6 | 5 | 0.016 | 0.01 | 0.682 |
| | Total | 4886 | 136 | 496 | 0.03 | 0.10 | < 0.0001 | 17 | 14 | 0.003 | 0.003 | 0.78 |
| Indels | Unique CDS | 374 | 23 | 89 | 0.06 | 0.24 | < 0.0001 | NA | NA | NA | NA | NA |
| | SD/TRF | 34 | 3 | 16 | 0.09 | 0.47 | < 0.0001 | NA | NA | NA | NA | NA |
| | Total | 408 | 25 | 105 | 0.06 | 0.26 | < 0.0001 | NA | NA | NA | NA | NA |
| p <= 0.0001 True | | | | | | | | | | | | |
| SNVS | Unique CDS | 4511 | 33 | 336 | 0.007 | 0.07 | < 0.0001 | 3 | 3 | 0.001 | 0.001 | 0.485 |
| | SD/TRF | 375 | 55 | 73 | 0.15 | 0.19 | 0.006 | 4 | 3 | 0.011 | 0.008 | 0.826 |
| | Total | 4886 | 97 | 409 | 0.02 | 0.08 | < 0.0001 | 8 | 6 | 0.002 | 0.001 | 0.796 |
| Indels | Unique CDS | 374 | 10 | 60 | 0.03 | 0.16 | < 0.0001 | NA | NA | NA | NA | NA |
| | SD/TRF | 34 | 2 | 12 | 0.06 | 0.35 | < 0.0001 | NA | NA | NA | NA | NA |
| | Total | 408 | 11 | 77 | 0.03 | 0.19 | < 0.0001 | NA | NA | NA | NA | NA |

Total mutation # is the total number of mutations analyzed within each set. Exp column shows the expected number of variants with AFs exceeding the given threshold. Expected derived from the mean number of rare variants meeting the indicated binomial p-value threshold simulated over 10,000 trials. Observed are the counts of *de novo* variants meeting the indicated binomial p-value threshold and characterized as potential PMMs. The simulated p-value was calculated from the number of trials that met or exceeded our observed over 10,000 trials. Note: variants in sex chromosomes were excluded for this analysis and no observed indels met any >0.5 threshold (listed as NAs). Abbreviations: AF-allele fraction, Exp-expected, Obs-observed, E-frac-the expected number of variants flagged as PMMs within a set (e.g. unique CDS) divided by the total, O-frac-the observed number of variants within a set (e.g. unique CDS) flagged as PMMs divided by the total, CDS-coding sequence, SD/TRF-coding sequence overlapping segmental duplication or tandem repeat finder tracks.

**Table S8. Summary of Top Performing Callers on Simulated Data at Varying Depth and Coverage**

| DEPTH | AF | BEST SENS | SENS | BEST PPV | PPV | BEST F0.5 | F0.5 |
|---|---|---|---|---|---|---|---|
| 30 | 0.01 | --- | --- | --- | --- | --- | --- |
| 30 | 0.05 | --- | --- | --- | --- | --- | --- |
| 30 | 0.10 | mPUP | 0.762 | LoFreq 2.1.1, mPUP | 1.000 | mPUP | 0.941 |
| 30 | 0.25 | mPUP | 0.856 | LoFreq 0.4.0/2.1.1, Varscan 2.3.2/2.3.7 | 1.000 | Varscan 2.3.2 | 0.965 |
| 30 | 0.50 | LoFreq 0.4.0/2.1.1 | 0.901 | LoFreq 0.4.0/2.1.1 | 1.000 | LoFreq 0.4.0/2.1.1 | 0.978 |
| 60 | 0.01 | --- | --- | --- | --- | --- | --- |
| 60 | 0.05 | mPUP | 0.755 | LoFreq 2.1.1 | 1.000 | mPUP | 0.899 |
| 60 | 0.10 | mPUP | 0.847 | LoFreq 2.1.1, Varscan 2.3.2/2.3.7 | 1.000 | Varscan 2.3.2/2.3.7 | 0.954 |
| 60 | 0.25 | LoFreq 0.4.0 | 0.900 | LoFreq 0.4.0/2.1.1, Varscan 2.3.2/2.3.7 | 1.000 | LoFreq 0.4.0 | 0.978 |
| 60 | 0.50 | LoFreq 0.4.0 | 0.915 | LoFreq 0.4.0/2.1.1, Varscan 2.3.2/2.3.7 | 1.000 | LoFreq 0.4.0 | 0.982 |
| 100 | 0.01 | mPUP | 0.015 | mPUP | 0.300 | mPUP | 0.062 |
| 100 | 0.05 | mPUP | 0.801 | LoFreq 2.1.1, Varscan 2.3.2/2.3.7 | 1.000 | mPUP | 0.922 |
| 100 | 0.10 | Varscan 2.3.2 | 0.871 | LoFreq 0.4.0/2.1.1, Varscan 2.3.2/2.3.7 | 1.000 | Varscan 2.3.2 | 0.971 |
| 100 | 0.25 | LoFreq 0.4.0 | 0.906 | LoFreq 0.4.0/2.1.1, Varscan 2.3.2/2.3.7 | 1.000 | LoFreq 0.4.0 | 0.980 |
| 100 | 0.50 | LoFreq 0.4.0/2.1.1 | 0.891 | LoFreq 0.4.0/2.1.1, mPUP, Varscan 2.3.2/2.3.7 | 1.000 | LoFreq 0.4.0/2.1.1 | 0.976 |
| 250 | 0.01 | mPUP | 0.010 | mPUP | 0.500 | mPUP | 0.046 |
| 250 | 0.05 | Varscan 2.3.2 | 0.891 | LoFreq 0.4.0/2.1.1, Varscan 2.3.2/2.3.7 | 1.000 | Varscan 2.3.2 | 0.976 |
| 250 | 0.10 | LoFreq 0.4.0, mPUP | 0.891 | LoFreq 0.4.0, Varscan 2.3.2/2.3.7 | 1.000 | LoFreq 0.4.0 | 0.976 |
| 250 | 0.25 | LoFreq 0.4.0 | 0.905 | LoFreq 0.4.0/2.1.1, Varscan 2.3.2/2.3.7 | 1.000 | LoFreq 0.4.0 | 0.980 |
| 250 | 0.50 | LoFreq 0.4.0/2.1.1, mPUP | 0.905 | LoFreq 0.4.0/2.1.1, mPUP, Varscan 2.3.2/2.3.7 | 1.000 | LoFreq 0.4.0/2.1.1, mPUP | 0.980 |
| 500 | 0.01 | Varscan 2.3.2/2.3.7 | 0.557 | mPUP | 1.000 | Varscan 2.3.2/2.3.7 | 0.858 |
| 500 | 0.05 | LoFreq 0.4.0, mPUP, Varscan 2.3.2/2.3.7 | 0.891 | LoFreq 0.4.0, mPUP, Varscan 2.3.2/2.3.7 | 1.000 | LoFreq 0.4.0, mPUP, Varscan 2.3.2/2.3.7 | 0.976 |
| 500 | 0.10 | LoFreq 0.4.0 | 0.906 | LoFreq 0.4.0, mPUP, Varscan 2.3.2/2.3.7 | 1.000 | LoFreq 0.4.0 | 0.980 |
| 500 | 0.25 | LoFreq 0.4.0, mPUP | 0.901 | LoFreq 0.4.0, mPUP, Varscan 2.3.2/2.3.7 | 1.000 | LoFreq 0.4.0, mPUP | 0.978 |
| 500 | 0.50 | LoFreq 0.4.0/2.1.1, mPUP | 0.906 | LoFreq 0.4.0/2.1.1, mPUP, Varscan 2.3.2/2.3.7 | 1.000 | LoFreq 0.4.0/2.1.1, mPUP | 0.980 |

Abbreviations: AF-allele fraction, SENS-sensitivity, PPV-positive predictive value, F0.5-F-score with 0.5 beta value.

**Table S10. Rank Enrichments for Genomewide ASD Predictions**

| Missense | | | ASD Association Rank | LGD Rank | LGD&RVIS Avg. Rank |
|---|---|---|---|---|---|
| | Count Pro | Count Sib | p-value | p-value | p-value |
| Whole Cohort | 184 | 134 | 0.6808 | 0.7358 | 0.5445 |
| Pro Has LGD GDM | 25 | 32 | 0.7993 | 0.4172 | 0.3246 |
| Pro No LGD GDM | 159 | 102 | 0.5408 | 0.7709 | 0.5551 |
| Pro Has NS GDM | 114 | 91 | 0.9056 | 0.2011 | 0.7252 |
| Pro No NS GDM | 70 | 43 | 0.1595 | 0.3234 | 0.1524 |
| | | | | | |
| Synonymous | | | ASD Association Rank | LGD Rank | LGD&RVIS Avg. Rank |
| | Count Pro | Count Sib | p-value | p-value | p-value |
| Whole Cohort | 80 | 42 | 0.1855 | 0.346 | 0.4358 |
| Pro Has LGD GDM | 20 | 11 | 0.04931 | 0.5165 | 0.849 |
| Pro No LGD GDM | 60 | 31 | 0.5217 | 0.3047 | 0.2555 |
| Pro Has NS GDM | 52 | 31 | 0.07623 | 0.5431 | 0.8687 |
| Pro No NS GDM | 28 | 11 | 0.6266 | 0.2176 | 0.02911* |
| | | | | | |
| Essential Missense | | | ASD Association Rank | LGD Rank | LGD&RVIS Avg. Rank |
| | Count Pro | Count Sib | p-value | p-value | p-value |
| Whole Cohort | 41 | 24 | 0.9697 | 0.7183 | 0.2527 |
| Pro Has LGD GDM | 5 | 6 | 0.7316 | 0.3961 | 0.1645 |
| Pro No LGD GDM | 36 | 18 | 0.9625 | 0.8458 | 0.35 |
| Pro Has NS GDM | 27 | 16 | 0.9285 | 0.7055 | 0.4359 |
| Pro No NS GDM | 14 | 8 | 0.8175 | 0.7589 | 0.07252* |
| | | | | | |
| Intolerant Missense | | | ASD Association Rank | LGD Rank | LGD&RVIS Avg. Rank |
| | Count Pro | Count Sib | p-value | p-value | p-value |
| Whole Cohort | 59 | 34 | 0.8538 | 0.593 | 0.3839 |
| Pro Has LGD GDM | 7 | 7 | 0.9869 | 0.1914 | 0.5 |
| Pro No LGD GDM | 52 | 27 | 0.467 | 0.8506 | 0.4547 |
| Pro Has NS GDM | 36 | 21 | 0.9676 | 0.2233 | 0.5359 |
| Pro No NS GDM | 23 | 13 | 0.2146 | 0.9192 | 0.2446 |

Analysis performed on high confidence call set (5%-45x). Significance determined using unpaired Wilcoxon rank sum test, one-sided for missense and two-sided for synonymous. ASD Association rank obtained from per gene ASD association scores in Krishnan et al. 2016.[19] LGD rank and LGD&RIVIS Avg. rank obtained from per gene ranks derived in Iossifov et al. 2015.[20] *Nominally significant values called out in text. Abbreviations: Pro-probands (Quads + Trios), Sib-siblings, LGD-likely gene disrupting, NS-nonsynonymous GDM-germline *de novo* mutation.

**Table S11. Primer and Guide Sequences Used in smMIP Preparation and Sequencing**

| PROBE SET | PRIMER | SEQUENCE | GUIDE OLIGO | GUIDE SEQUENCE |
|---|---|---|---|---|
| Set 02 | ArrayMIP_02_FWD | /5BiosG/GCCGGTCAACAAACTCGCATG | Guide_02_NlaIII_2N | NNCATGCGAGTTTGTTGACCGGC |
| | ArrayMIP_02_REV | TGCGCAGTGCCATCATCCTGG | Guide_02_NlaIII_GC | CGCATGCGAGTTTGTTGACCGGC |
| | | | Guide_02_NlaIII_GD | DGCATGCGAGTTTGTTGACCGGC |
| Set 03 | ArrayMIP_03_FWD | /5BiosG/CCATAGCCGAGTCCACACATG | Guide_03_NlaIII_2N | NNCATGTGTGGACTCGGCTATGG |
| | ArrayMIP_03_REV | GCCAGACGCTGTCATTCCTGG | Guide_03_NlaIII_GC | CGCATGTGTGGACTCGGCTATGG |
| | | | Guide_03_NlaIII_GD | DGCATGTGTGGACTCGGCTATGG |
| Set 04 | ArrayMIP_04_FWD | /5BiosG/CCCTTCACGCGTTCTTCCATG | Guide_04_NlaIII_2N | NNCATGGAAGAACGCGTGAAGGG |
| | ArrayMIP_04_REV | ATGCTATGGAGCGTCACCTGG | Guide_04_NlaIII_GC | CGCATGGAAGAACGCGTGAAGGG |
| | | | Guide_04_NlaIII_GD | DGCATGGAAGAACGCGTGAAGGG |
| Set 05 | ArrayMIP_05_FWD | /5BiosG/GTCCGGCTCTCCTCAGTCATG | Guide_05_NlaIII_2N | NNCATGACTGAGGAGAGCCGGAC |
| | ArrayMIP_05_REV | AACCTATGACCTCACGCCTGG | Guide_05_NlaIII_GC | CGCATGACTGAGGAGAGCCGGAC |
| | | | Guide_05_NlaIII_GD | DGCATGACTGAGGAGAGCCGGAC |
| Set 06 | ArrayMIP_06_FWD | /5BiosG/CTGAATAGCAGCTACCGCATG | Guide_06_NlaIII_2N | NNCATGCGGTAGCTGCTATTCAG |
| | ArrayMIP_06_REV | CTCGGTCACTATGTGCCCTGG | Guide_06_NlaIII_GC | CGCATGCGGTAGCTGCTATTCAG |
| | | | Guide_06_NlaIII_GD | DGCATGCGGTAGCTGCTATTCAG |
| Set 07 | ArrayMIP_07_FWD | /5BiosG/GAACACGTACCAATCCGCATG | Guide_07_NlaIII_2N | NNCATGCGGATTGGTACGTGTTC |
| | ArrayMIP_07_REV | AAAGATACCAGTCGTGCCTGG | Guide_07_NlaIII_GC | CGCATGCGGATTGGTACGTGTTC |
| | | | Guide_07_NlaIII_GD | DGCATGCGGATTGGTACGTGTTC |
| Set 08 | ArrayMIP_08_FWD | /5BiosG/TCGCAAGTCTTGAACCGCATG | Guide_08_NlaIII_2N | NNCATGCGGTTCAAGACTTGCGA |
| | ArrayMIP_08_REV | GTTCAGTGATCTCGTGCCTGG | Guide_08_NlaIII_GC | CGCATGCGGTTCAAGACTTGCGA |
| | | | Guide_08_NlaIII_GD | DGCATGCGGTTCAAGACTTGCGA |
| Set 09 | ArrayMIP_09_FWD | /5BiosG/TACAGGTCCGTGCCATTCATG | Guide_09_NlaIII_2N | NNCATGAATGGCACGGACCTGTA |
| | ArrayMIP_09_REV | TCGTGTGGCTAGATTCCCTGG | Guide_09_NlaIII_GC | CGCATGAATGGCACGGACCTGTA |
| | | | Guide_09_NlaIII_GD | DGCATGAATGGCACGGACCTGTA |
| Set 10 | ArrayMIP_10_FWD | /5BiosG/CACTGTCCCCTTGCTTCCATG | Guide_10_NlaIII_2N | NNCATGGAAGCAAGGGGACAGTG |
| | ArrayMIP_10_REV | GATTCGATAGGCTGACCCTGG | Guide_10_NlaIII_GC | CGCATGGAAGCAAGGGGACAGTG |
| | | | Guide_10_NlaIII_GD | DGCATGGAAGCAAGGGGACAGTG |
| Set 11 | ArrayMIP_11_FWD | /5BiosG/TCGTCGCACTACTCTGACATG | Guide_11_NlaIII_2N | NNCATGTCAGAGTAGTGCGACGA |
| | ArrayMIP_11_REV | CAAGCATTCAGCTCTACCTGG | Guide_11_NlaIII_GC | CGCATGTCAGAGTAGTGCGACGA |
| | | | Guide_11_NlaIII_GD | DGCATGTCAGAGTAGTGCGACGA |
| Sequencing Primers | MIPBC_SEQ_FOR | CATACGAGATCCGTAATCGGGAAGCTGAAG | | |
| | MIPBC_SEQ_REV | ACACGCACGATCCGACGGTAGTGT | | |
| | MIPBC_SEQ_INDX 1 | ACACTACCGTCGGATCGTGCGTGT | | |
| | MIPBC_SEQ_INDX 2 | CTTCAGCTTCCCGATTACGGATCTCGTATG | | |

**Supplemental References**


1. Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B.P., Stessman, H.A., He, Z.-X., et al. (2015). Excess of rare, inherited truncating mutations in autism. Nat Genet 47, 582-588.
2. Stead, L.F., Sutton, K.M., Taylor, G.R., Quirke, P., and Rabbitts, P. (2013). Accurately identifying low-allelic fraction variants in single samples with next-generation sequencing: applications in tumor subclone resolution. Hum Mutat 34, 1432-1438.
3. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754-1760.
4. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.
5. Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22, 568-576.
6. Wilm, A., Aw, P.P., Bertrand, D., Yeo, G.H., Ong, S.H., Wong, C.H., Khor, C.C., Petric, R., Hibberd, M.L., and Nagarajan, N. (2012). LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res 40, 11189-11201.
7. Challis, D., Yu, J., Evani, U.S., Jackson, A.R., Paithankar, S., Coarfa, C., Milosavljevic, A., Gibbs, R.A., and Yu, F. (2012). An integrative variant analysis suite for whole exome next-generation sequencing data. BMC Bioinformatics 13, 8.
8. Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature 515, 216-221.
9. O'Roak, B.J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B.P., Levy, R., Ko, A., Lee, C., Smith, J.D., et al. (2012). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature 485, 246-250.
10. Boyle, E.A., O'Roak, B.J., Martin, B.K., Kumar, A., and Shendure, J. (2014). MIPgen: optimized modeling and design of molecular inversion probes for targeted resequencing. Bioinformatics 30, 2670-2672.
11. O'Roak, B.J., Stessman, H.A., Boyle, E.A., Witherspoon, K.T., Martin, B., Lee, C., Vives, L., Baker, C., Hiatt, J.B., Nickerson, D.A., et al. (2014). Recurrent de novo mutations implicate novel genes underlying simplex autism risk. Nat Commun 5, 5595.
12. Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics 30, 614-620.
13. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol 29, 24-26.
14. Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. Ann Statist, 1165-1188.
15. Roche, A.F., Mukherjee, D., Guo, S.M., and Moore, W.M. (1987). Head circumference reference data: birth to 18 years. Pediatrics 79, 706-712.
16. Sanders, S.J., He, X., Willsey, A.J., Ercan-Sencicek, A.G., Samocha, K.E., Cicek, A.E., Murtha, M.T., Bal, V.H., Bishop, S.L., Dong, S., et al. (2015). Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. Neuron 87, 1215-1233.
17. Blokzijl, F., Janssen, R., Van Boxtel, R., and Cuppen, E. (2016). MutationalPatterns: an integrative R package for studying patterns in base substitution catalogues. bioRxiv.
18. Alexandrov, L.B., Ju, Y.S., Haase, K., Van Loo, P., Martincorena, I., Nik-Zainal, S., Totoki, Y., Fujimoto, A., Nakagawa, H., Shibata, T., et al. (2016). Mutational signatures associated with tobacco smoking in human cancer. Science 354, 618-622.
19. Krishnan, A., Zhang, R., Yao, V., Theesfeld, C.L., Wong, A.K., Tadych, A., Volfovsky, N., Packer, A., Lash, A., and Troyanskaya, O.G. (2016). Genome-wide prediction and functional characterization of the genetic basis of autism spectrum disorder. Nature Neurosci 19, 1454-1462.
20. Iossifov, I., Levy, D., Allen, J., Ye, K., Ronemus, M., Lee, Y.H., Yamrom, B., and Wigler, M. (2015). Low load for disruptive mutations in autism genes and their biased transmission. Proc Natl Acad Sci U S A 112, E5600-5607.