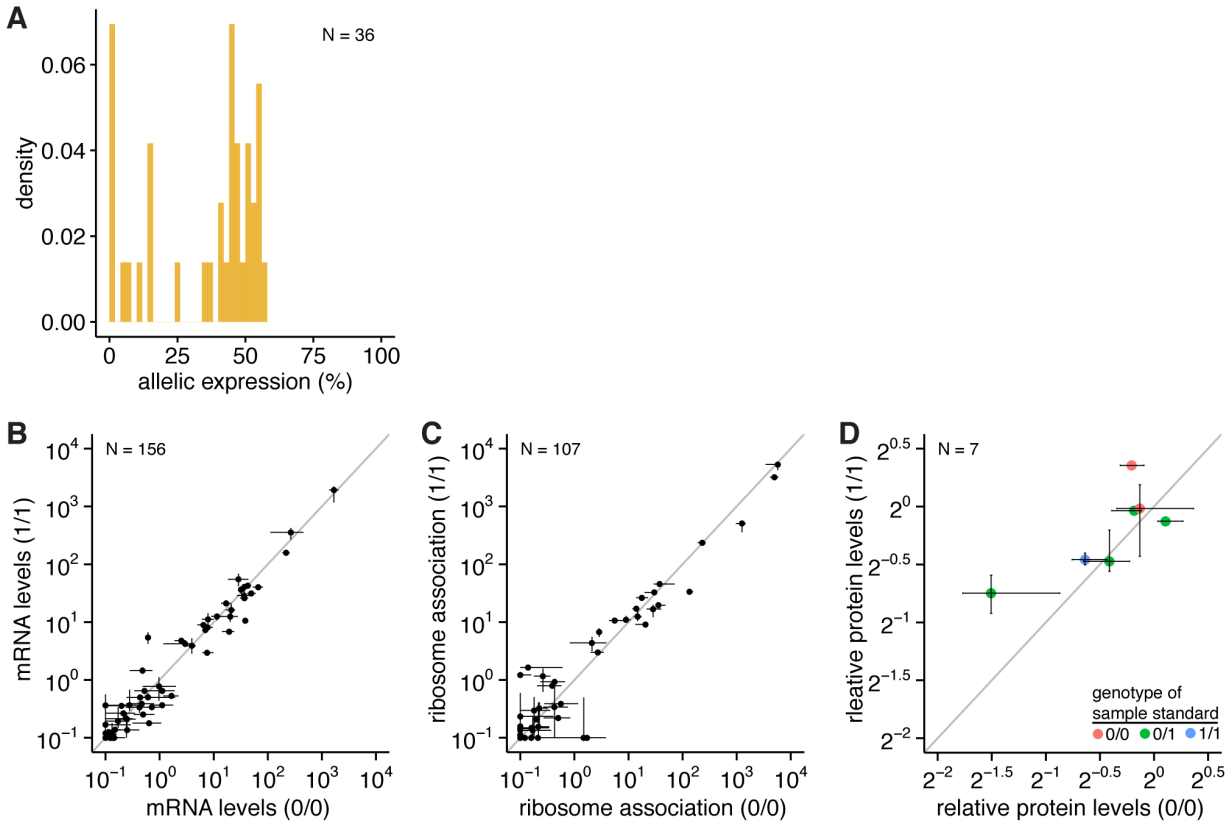

Figure S1. Genotyping errors do not explain robust protein production from putatively disabled genes.

(**A**) Histogram of allele-specific expression (ASE) of the alternate allele in heterozygous individuals (genotype 0/1). Plot restricted to nonsense variants analyzed in Figure 1C for which at least one heterozygous sample had at least 20 RNA-seq reads covering each variant. Plot illustrates the median ASE over all such heterozygous samples.

(**B-D**) As Figure 1C-E, but restricted to variants and samples for which we could confirm genotypes called by genome sequencing using RNA-seq data.

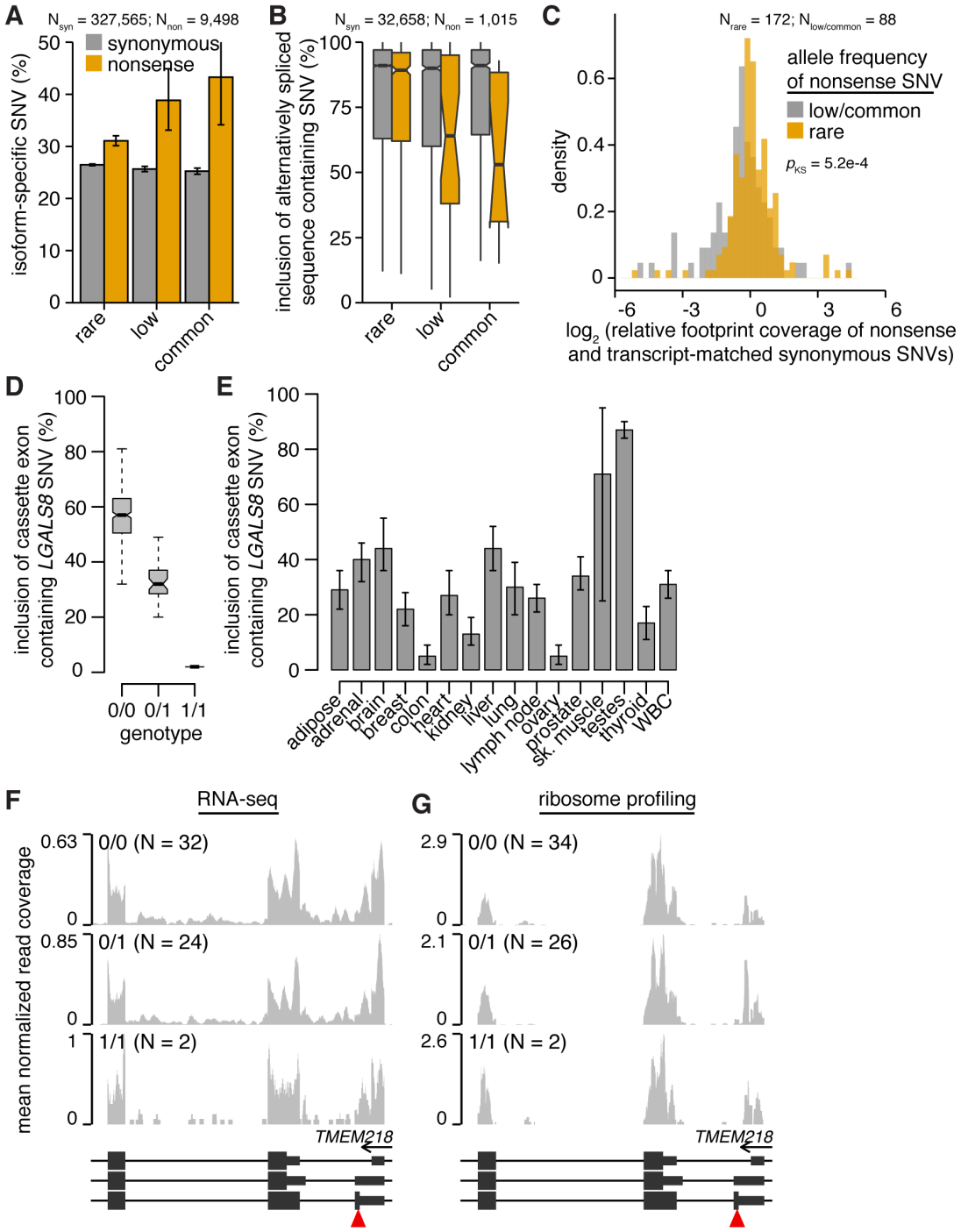

Figure S2. Alternative RNA processing of nonsense variants.
(**A-B**) As Figure 2A-B, but with variants within any genes—not just those containing nonsense variants—illustrated. Figure 2A-B and these analyses together control for the fact that nonsense variants preferentially occur in specific gene classes (MacArthur et al. 2012; Sulem et al. 2015).
(**C**) Histogram of coverage of nonsense versus transcript-matched synonymous variants, stratified by allele frequency of the nonsense variant. Coverage of nonsense variants was computed as the number of footprints overlapping each variant; coverage of transcript-matched synonymous variants was estimated as the median number of footprints overlapping each synonymous variant within the parent transcript

containing the nonsense variant (median taken over all synonymous variants within the transcript). All footprint coverage calculations were restricted to 0/0 samples to avoid potentially confounding effects of nonsense variants, and the plotted values indicate medians over those samples. Low-frequency and common variants were grouped together to increase statistical power. p_{KS}, p-value computed with the two-sided Kolmogorov–Smirnov test for a difference in distribution.
(**D**) Inclusion of the *LGALS8* cassette exon illustrated in Figure 2D, stratified by genotype. Plot is a quantification of Figure 2D.
(**E**) Inclusion of the *LGALS8* cassette exon across human tissues. Tissue RNA-seq data is from the Illumina Body Map 2.0 dataset, restricted to the samples that were sequenced with paired-end reads. Tissue samples were obtained from different individuals. All samples were 0/0 with the exception of breast and ovary, which were 0/1. Individuals were manually genotyped based on allelic expression at the *LGALS8* variant. WBC, white blood cells; sk. muscle, skeletal muscle.
(**F**) RNA-seq and (**G**) ribosome profiling read coverage of a nonsense variant lying within alternatively spliced sequence of *TMEM218* that is either coding or non-coding, depending upon which start codon is used.

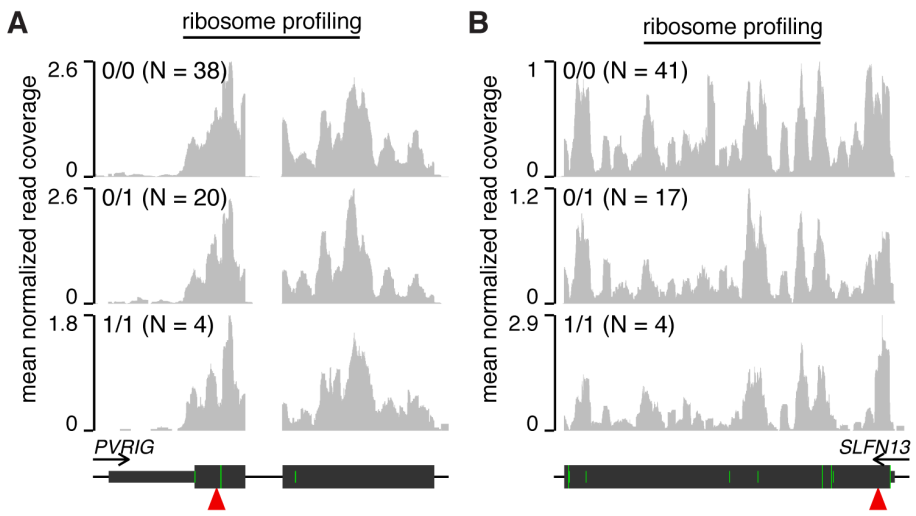

Figure S3. Stop codon readthrough.
As Figure 3, but illustrating zoomed-in plots for all genotypes.

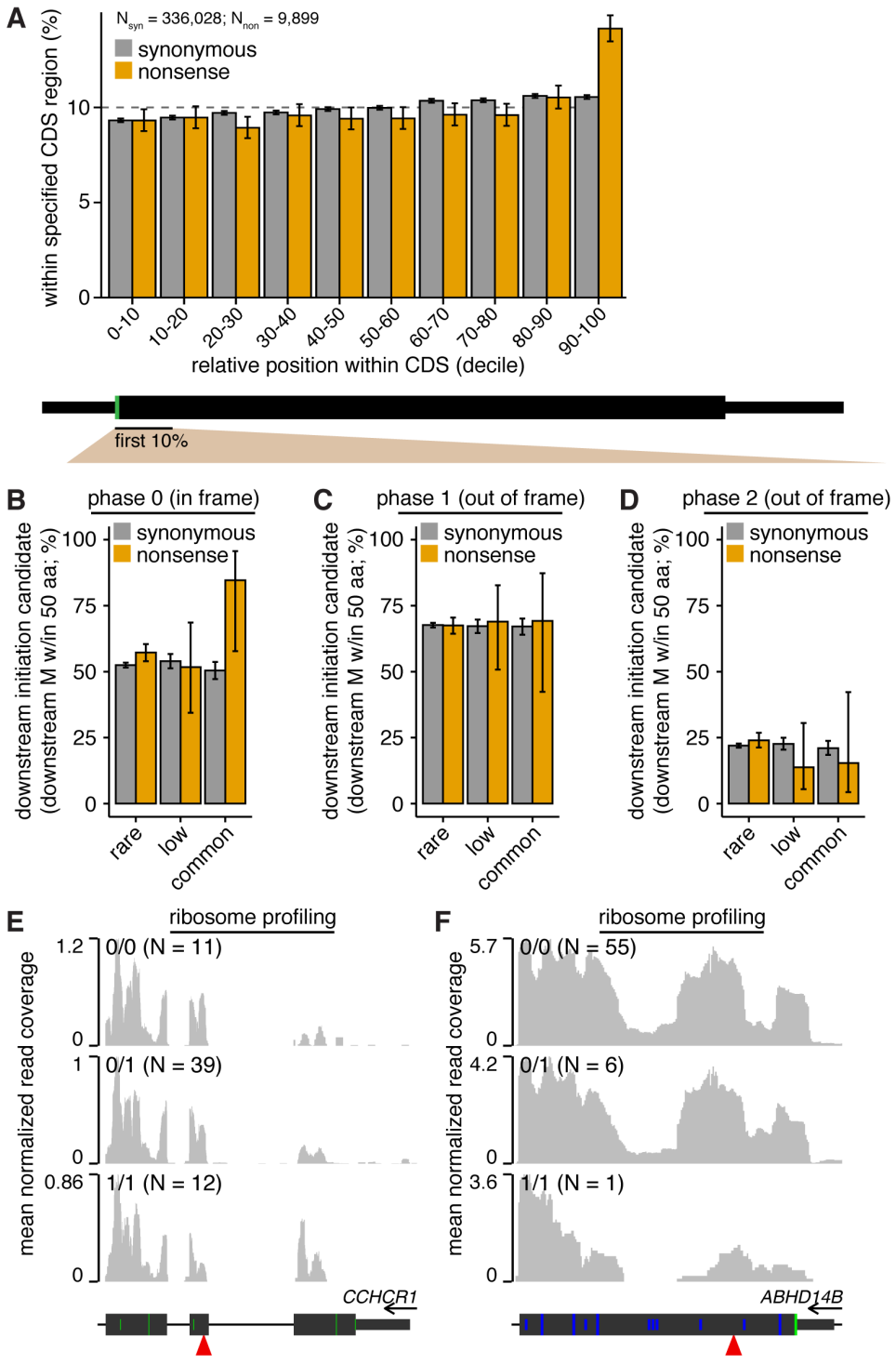

Figure S4. Protein truncation induced by nonsense variants.
(**A**) As Figure 4A, but with variants within any genes—not just those with nonsense variants—illustrated.
(**B-D**) As Figure 4B, but with downstream methionines that lie within each of the three possible reading frames of CDSs considered separately. Phases 0, 1, and 2 indicate the offset in nucleotides from the start of the true CDS used to define the reading frame in each panel, such that phase 0 corresponds to the true reading frame and phases 1 and 2 corresponds to frameshifts.
(**E-F**) As Figure 4E-F, but illustrating zoomed-in plots for all genotypes.

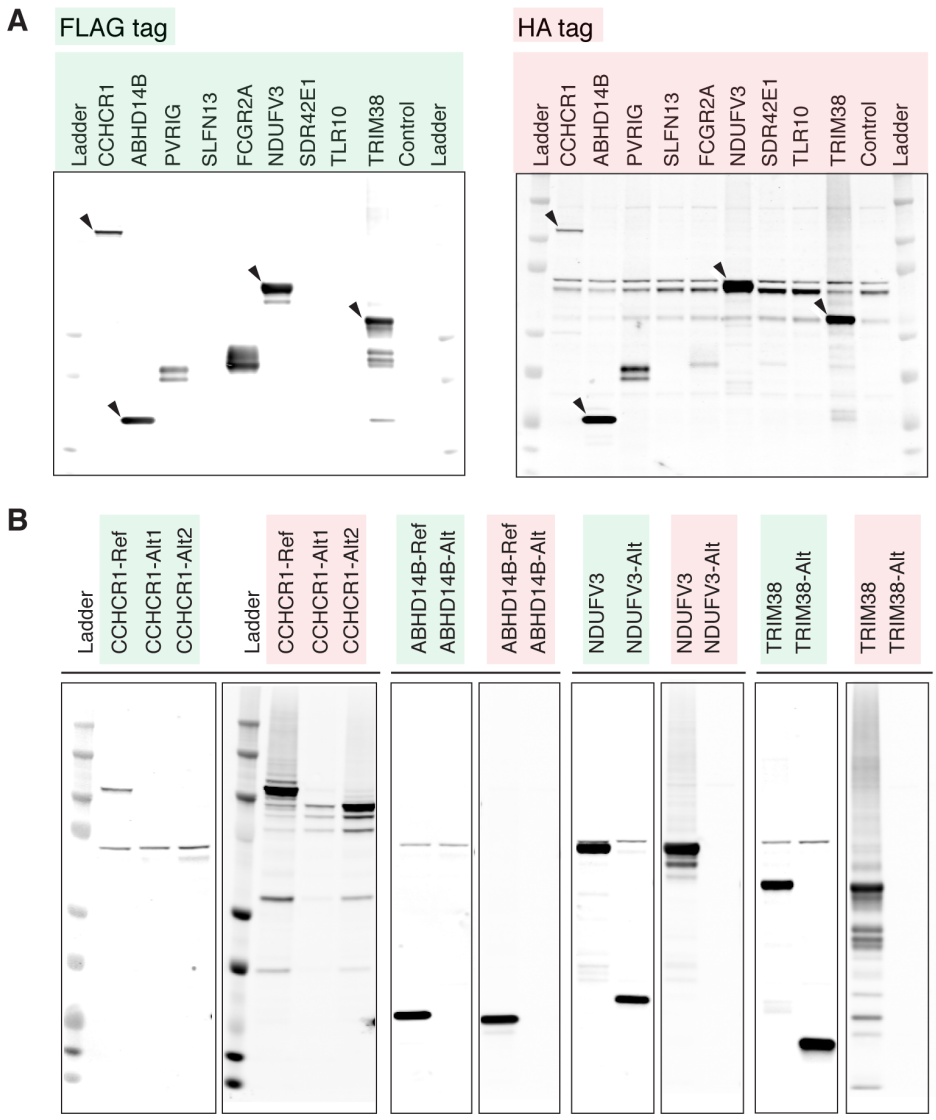

Figure S5. Heterologous reporter experiments.
(**A**-**B**) As Figure 5C-D, but with the Western blots for the FLAG and HA tags shown separately.