Systematic evaluation of MANE Plus Clinical transcripts: implications for genomic diagnostics

SUPPLEMENTARY NOTES

Note S1: Examples of candidate transcripts assessed for potential inclusion in future MANE PC releases.

Three clustered (L)P variants cluster in intron 23 of *SCN1A* (NM_001165963.4) (see https://genome.ucsc.edu/s/ExeterGenetics/MANE_PC_candidate_review_SCN1A). However, these are known to upregulate a poison exon and have been shown to cause Dravet syndrome and related epilepsies[1]. A related situation is seen for Gitelman syndrome (MIM #263800), which is an autosomal recessive renal tubular salt-wasting disorder linked to variants in *SLC12A3*. Two P/LP variants NM_001126108.2:c.2419+13C>G (VCV003581057.1) and NM_001126108.2:c.2419+13C>A (VCV002698822.3) both predict premature terminations (c.2432C>A c.2432C>G, p.Ser811Ter) based on the alternative transcript (NM_000339.3) that is used in HGMD, where exon 20 is extended. Both isoforms show strong expression in kidney cortex and evolutionary conservation extends into the extended part of this exon (see https://genome.ucsc.edu/s/ExeterGenetics/MANE_PC_candidate_review_SLC12A3). However, as both variants are also predicted to decrease the efficiency of splice donor site for the MS transcript (DL=0.23 and 0.38 respectively, DP = +14), it is unclear whether these variants act through decreased efficiency of splicing for the MS transcript or whether they support clinical/functional relevance for the alternative isoform.

At first glance, the NM_000328.3 isoform of *RPGR* (Retinitis pigmentosa GTPase regulator) appears to show better conservation and higher expression in GTEx than the Mane Select isoform (NM_001034853.2) (see https://genome.ucsc.edu/s/ExeterGenetics/MANE_PC_candidate_review_RPGR). However, HGMD uses both NM_000328.3 and NM_001034853.2. It is the latter ("ORF15") isoform that is expressed in retina and that contains a high proportion (around 80%) of all pathogenic variants. This retinal isoform is not expressed in other tissues, and the Glu-Gly rich repetitive region may explain why it appears to be not conserved. Of the pathogenic variants found in this exon, most variants are frameshifts.[2] Nevertheless, the longer isoform could well retain other important functional role as those exons are conserved and expressed in other tissues. Although there are relatively few P/LP variants in the last 4 exons of NM_000328.3, we tentatively propose that NM_000328.3 should be the MS and NM_001034853.2 the MPC transcript.

Although *NF1* is newly included in MANE V1.5, the MPC transcript (NM_000267.3) simply skips an exon (exon 31) present in the MS transcript. Other *NF1* transcripts with conserved alternate exons span single ClinVar P/LP variants and so these should be further scrutinised as additional data becomes available (see https://genome.ucsc.edu/s/ExeterGenetics/MANE_PC_candidates_NF1). For *TH* (Tyrosine hydroxylase), several (L)P variants lie in an exon specific to the NM_199292.3 transcript (see https://genome.ucsc.edu/s/ExeterGenetics/MANE_PC_candidate_review_TH). However, this exon does not show evolutionary conservation, GTEx data shows the exon is not expressed in the adrenal gland and none of the ClinVar submissions are particularly convincing. *TSC2* contains P/LP in two alternate exons (see https://genome.ucsc.edu/s/ExeterGenetics/MANE_PC_candidate_review_TSC2) – however these are mostly annotated as UTR exons and with just a single P/LP variant in each region, further data would be required to confirm any clinical relevance. Similarly, *PTEN* contains transcripts with moderately conserved alternate exons that harbour P/LP variants (https://genome.ucsc.edu/s/ExeterGenetics/MANE_PC_candidate_review_PTEN). More data is needed to assess whether these transcripts are functional or just represent upregulation of poison/non-functional transcripts.

We identified several genes where the clustered ClinVar P/LP variants are linked to transcription regulation. For instance, for *HBB* we note that use of ENST00000647020.1 could help capture an abundance of promoter variants which could be otherwise missed (see

https://genome.ucsc.edu/s/ExeterGenetics/MANE_PC_candidates_HBB). Similarly, for *RMRP* there is also a large number of P/LP variants clustering just upstream of the MS transcript. Genome analysts unaware of this promoter hotspot region (see https://genome.ucsc.edu/s/ExeterGenetics/MANE_PC_candidates_RMRP), could potentially overlook variants that lie outside the transcribed sequence. However, this is a gene that encodes an RNA component of mitochondrial RNA processing endoribonuclease and longer alternative isoforms are not known. A cluster of 17 P/LP variants lying in intron 5 of *LMBR1* were also identified by our analysis (see https://genome.ucsc.edu/s/ExeterGenetics/MANE_PC_candidate_review_LMBR1). However, these are localised in a regulatory region named the "Zone of Polarizing Activity Regulatory Sequence" which activates *SHH* and was described over 20 years ago.[3]

We anticipate that, as the future releases of ClinVar and other updates to expression/functional annotation datasets are included in UCSC genome browser, the set of custom UCSC sessions listed above will automatically update and thus become increasingly informative.

Note S2: Outlier genes due to "fix patch" sequences in GRCh38.

SHANK3 is a gene linked to Phelan-McDermid syndrome (MIM 606232). All 191 (L)P variants we identified from ClinVar were apparently outside the MANE target regions (annotated exons ±8bp). This is because the MANE Select transcript for *SHANK3* lies on a "fix patch" sequence (chr22_KQ759762v2) in GRCh38/hg38. Patches in the refence genome are due to protein-coding genes which cannot be represented accurately on the GRCh38 primary reference genome due to assembly errors. In the case of *SHANK3*, the patch is due to a 2bp insertion in the 5' coding region which would result in a frameshift. Similarly, the MANE Select transcripts for *SLC37A4* (Glycogen storage disease; MIM 232220 and 232240), *POLR2A* (Neurodevelopmental disorder with hypotonia and variable intellectual and behavioural abnormalities; MIM 618603) and *ORAI1* (Immunodeficiency 9; MIM 612782) lie on chr11_KN196481v1_fix, chr17_KV575245v1_fix and chr12_KZ208917v1_fix, respectively. ClinVar variants are not mapped to alternate these chromosome scaffolds and so this explains why these genes each appeared to harbour multiple P/LP ClinVar variants apparently outside the MANE Select exon boundaries. In total, MANE Select transcripts of 64 genes are annotated on genomic patch sequences in MANE v1.5 (www.ncbi.nlm.nih.gov/refseq/MANE/#mane-select-on-genomic-patch-seq).

SUPPLEMENTARY TABLES

Tables S1, S2 and S3 are provided as separate .xlsx files.

Table S4: Details of positive vs negative strand start codon overlaps in the human genome. These overlaps are also shown in **Figure S6 and S7**.

Positions of ATG overlap†	Gene symbols (-ve/+ve strand)	OMIM disease-gene associations‡
1:45340253-45340254	*MUTYH*(ENST00000710952.2)-*TOE1*	*MUTYH:* Adenomas, multiple colorectal and *TOE1:* Pontocerebellar hypoplasia, type 7
2:26346233-26346234	*ADGRF3-SELENOI*	*SELENOI:* Spastic paraplegia 81 autosomal recessive
3:158732414-158732415	*RARRES1/MFSD1*(ENST00000486568.5)	-
7:100119745-100119746	*TAF6*(ENST00000437822.6)/*CNPY4*	*TAF6:* Alazami-Yuan syndrome
9:32552435-32552436	*TOPORS/SMIM27*	*TOPORS:* Retinitis pigmentosa 31
10:116849567-116849568	*HSPA12*(ENST00000635765.1)/*ENO4*	
11:62671746-62671747	*LBHD1*(ENST00000415855.6)/*UQCC3*	*UQCC3:* ?Mitochondrial complex III deficiency, nuclear type 9
17:7217146-7217147	*DLG4/ACADVL*(ENST00000543245.6)	*DLG4:* Intellectual developmental disorder, autosomal dominant 62 and *ACADVL:* VLCAD deficiency

†Genomic positions are reported on GRCh38. ‡OMIM disease-gene associations are as available 10th December 2025. Transcripts are from the MANE select collection, unless otherwise stated.

Table S5: Details of 15 predicted loss-of-function variants in *PCDH15* from ClinVar that support the clinical relevance of the CD2 isoform. Although not included in the MANE Plus Clinical set of transcripts (both v1.4 and v1.5), we note that this isoform is included in the Human Gene Mutation Database, designated as "*PCDH15* transcript variant I" (pdch15tvi).

cDNA change based on CD2 isoform (NM_001142769.3)	Protein change (NP_001136241.1)	ClinVar accession	Assessment	Submitter	Publication
c.5073del	p.(Glu1692AsnfsTer128)	VCV003068472.2	LP	University of Leipzig Medical Centre	-
c.5051_5052dup	p.(Glu1685ArgfsTer136)	VCV003601629.1	P	Sichuan University	-
c.5050G>T	p.Glu1684Ter	VCV003381831.1	P	Columbia University	-
c.5047A>T	p.Lys1683Ter	VCV004685461.1	LP	Tel Aviv University	-
c.4972_4975dup	p.(Lys1659ArgfsTer17)	VCV003601628.1	P	Sichuan University	-
c.4961_4962del	p.(Glu1654ValfsTer20)	VCV003601627.1	P	Sichuan University	-
c.4938_4941dup	p.(Glu1648IlefsTer28)	VCV003601626.2	P/LP	Sichuan University, Zhengzhou University	Kang *et al* [4]
c.4895_4896dup	p.(Glu1634GlnfsTer187)	VCV003601624.1	P	Sichuan University	-
c.4893dup	p.Glu1632Ter	VCV003601623.1	LP	Sichuan University	-
c.4858_4862dup	p.(Phe1621LeufsTer201)	VCV002445660.1	P	University of Washington	Carlson *et al* [5]
c.4816G>T	p.Glu1606Ter	VCV003601622.1	P	Sichuan University	-
c.4727dup	p.(Arg1577GlufsTer4)	VCV003601621.1	LP	Sichuan University	-
c.4726C>T	p.(Gln1576Ter)	VCV000402273.4	P	University of Washington, University of Social Welfare and Rehabilitation Sciences (Iran), Bethlehem University	Sloan-Heggen *et al* [6], Abu Rayyan *et al* [7]†
c.4673_4676dup	p.(Arg1559SerfsTer2)	VCV003601620.1	P	Sichuan University	-
c.4542dup	p.(Pro1515ThrfsTer4)	VCV000505342.5	VUS but "favouring pathogenic"	Mass General Brigham Personalized Medicine	Pepermans *et al* [8] ‡

†In the study by Sloan-Heggen *et al*, the most frequently identified variant, p.Gln1576Ter, was identified in four families, three of whom were of Fars ethnicity. Although the phenotype was described as "profound prelingual non-syndromic hearing loss", it was noted that families with this variant exhibited varying phenotypes. All homozygous individuals segregated the same haplotype. Abu Rayyan *et al* also described 14 individuals from 6 Palestinian families with severe-to-profound nonsyndromic hearing loss due to the same *PCDH15* variant. ‡, the patient from family CPID4744 carried this variant in the homozygous state, and the two patients from family CPIDS6-10 carried it in the heterozygous state, *in trans* with c.400C>T (p.R134*), located in *PCDH15* exon 6, which is common to the three Pcdh15 isoforms. Five siblings of the two families (individual IV.1 from family CPID4744 and individuals IV.1, IV.2, IV.3 and IV.6 from family CPIDS6-10) with normal hearing did not carry the c.4542dup variant in the homozygous state (family CPID4744) or in the compound heterozygous state (family CPIDS6-10). Homozygous occurrence of this variant was also detected in a 100kGP participant with congenital hearing loss (**Figure S9**).

Table S6: Apparent clustering of pathogenic deletions outside MANE Select/Plus Clinical exonic regions due to LiftOver errors. All deletions were based on those reported in Table 3 of the study by Flynn *et al*,[9] and were assessed with respect to Fanconi anaemia complementation group A (MONDO:0009215). Variants were submitted to ClinVar via the Leiden Open Variation Database (organization 506264). The initial LiftOver was done from hg19 to GRCh38 instead of from hg18 to GRCh38. This set of variants (corrected and uncorrected) is shown in **Figure S10**. All have been flagged for correction in future ClinVar updates.

HGVS (GRCh38, incorrect -> correct)	Size (bp)	ClinVar accession	Clinical significance	Review status	Gene	Origin	Cytoband	Submission accession
NC_000016.10:g.88310445_88314136del -> NC_000016.10:g.89750142_89753833del	3692	VCV000974152.2	Pathogenic	No assertion criteria provided	ZNF469	germline	16q24.2	SCV001425839.1
NC_000016.10:g.88314822_88320369del -> NC_000016.10:g.89754519_89760066del	5548	VCV000974045.2	Pathogenic	No assertion criteria provided	LOC112486220, ZNF469	germline	16q24.2	SCV001425699.1
NC_000016.10:g.88317551_88325754del -> NC_000016.10:g.89757248_89765451del	8204	VCV000974146.2	Pathogenic	No assertion criteria provided	ZNF469	germline	16q24.2	SCV001425830.2
NC_000016.10:g.88324127_88331653del -> NC_000016.10:g.89763824_89771350del	7527	VCV000974324.2	Pathogenic	No assertion criteria provided	ZNF469	germline	16q24.2	SCV001426072.1
NC_000016.10:g.88334974_88341919del -> NC_000016.10:g.89774671_89781616del	6946	VCV000974319.2	Pathogenic	No assertion criteria provided	ZNF469	germline	16q24.2	SCV001426066.2
NC_000016.10:g.88335235_88340023del -> NC_000016.10:g.89774932_89779720del	4789	VCV000974021.2	Pathogenic	No assertion criteria provided	ZNF469	germline	16q24.2	SCV001425666.2
NC_000016.10:g.88341728_88343606del -> NC_000016.10:g.89781425_89783303del	1879	VCV000974219.2	Pathogenic	No assertion criteria provided	ZNF469	germline	16q24.2	SCV001425930.2
NC_000016.10:g.88342444_88343618del -> NC_000016.10:g.89782141_89783315del	1175	VCV000974220.2	Pathogenic	No assertion criteria provided	ZNF469	germline	16q24.2	SCV001425931.1
NC_000016.10:g.88342520_88343694del -> NC_000016.10:g.89782217_89783391del	1175	VCV000974218.2	Pathogenic	No assertion criteria provided	ZNF469	germline	16q24.2	SCV001425929.1
NC_000016.10:g.88342678_88348626del -> NC_000016.10:g.89782375_89788323del	5949	VCV000974113.2	Pathogenic	No assertion criteria provided	ZNF469	germline	16q24.2	SCV001425789.2

SUPPLEMENTARY FIGURES

Figure S1: UCSC browser session showing that the internal exon specific to the PC transcript of *SYNE1*. *SYNE1* encodes nesprin-1, a structural protein that links the actin cytoskeleton to the nuclear plasma membrane.[10] Variants in this gene can cause arthrogryposis multiplex congenita 3, myogenic type (MIM 618484), Emery-Dreifuss muscular dystrophy 4, autosomal dominant (MIM 612998) and Spinocerebellar ataxia, autosomal recessive 8 (MIM 610743). The PC specific exon contains two ClinVar variants that are P/LP. For one of these (NM_001347702.2:c.1473T>G; p.Tyr491Ter, VCV002132078.3), the prediction from SpliceAI-visual[11] is that there is exon skipping, or in other words, decreased inclusion of the exon specific to the PC isoform (acceptor loss=0.37, donor loss = 0.33). The residual scores are very low, highlighting the benefits of evaluating absolute as well as the delta SpliceAI scores. The other variant is NM_001347702.2:c.1455del, p.(Glu486LysfsTer7) (VCV000524033.2). An interactive version of this figure is available at https://genome.ucsc.edu/s/AlistairP/SYNE1_splice_example_V2.

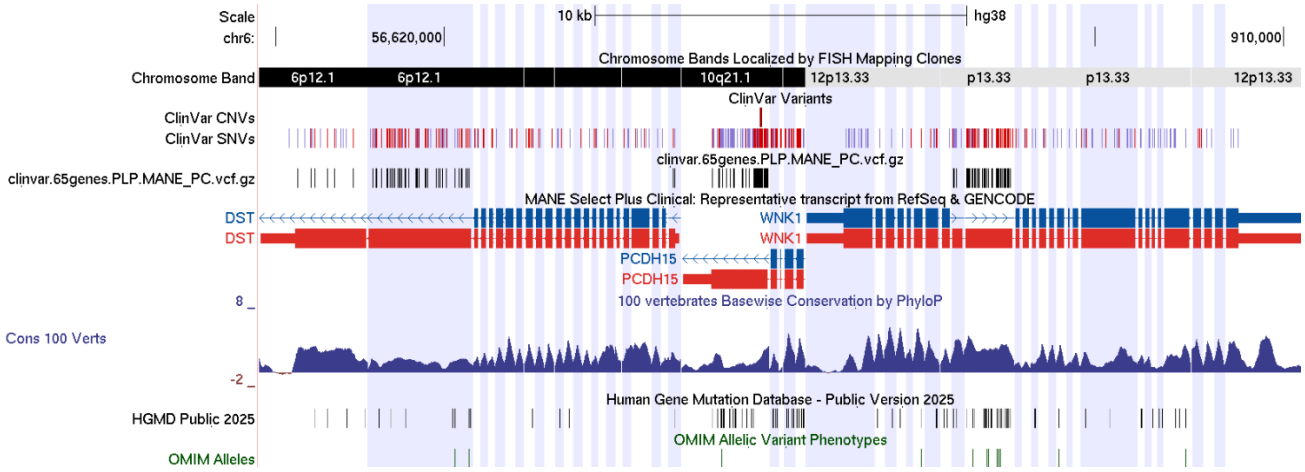

Figure S2: Outlier genes where the highest number of ClinVar (likely) pathogenic variants lie in the regions that are exclusive to the PC transcript. Figure available as an interactive UCSC Genome Browser session https://genome.ucsc.edu/s/ExeterGenetics/ClinVar_intersection. The filtered vcf downloaded from the ClinVar FTP site and uploaded as a custom track is more up-to-date than the public ClinVar track used in browser, e.g. VCV004532280.1 from Nov 2025 is currently missing from the ClinVar track used in UCSC. For these three outlier genes, the high number of PC-specific ClinVar (L)P variants is due to the relatively large size of the PC-specific exons.

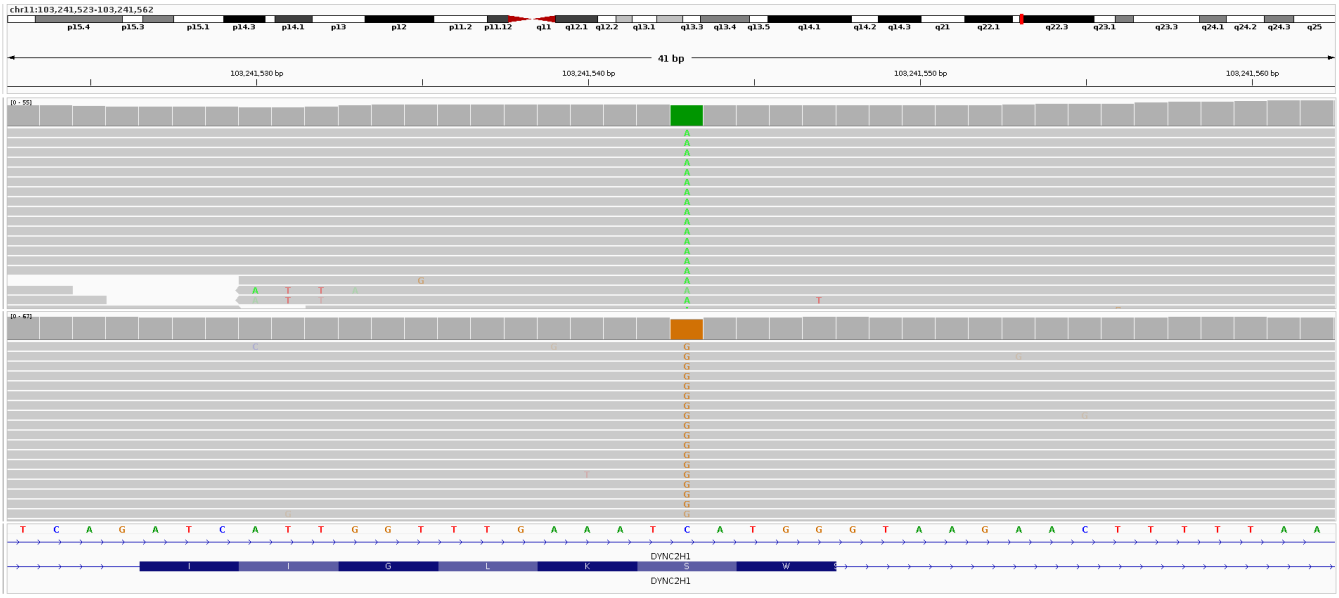

Figure S3: Read alignments supporting two different stop-gain variants at the same position in the PC-specific exon of *DYNC2H1*. NM_001080463.2:c.9836C>G (VCV000978199.2), and NM_001080463.2:c.9836C>A. Both variants are observed in the homozygous state and predict p.Ser3279Ter.

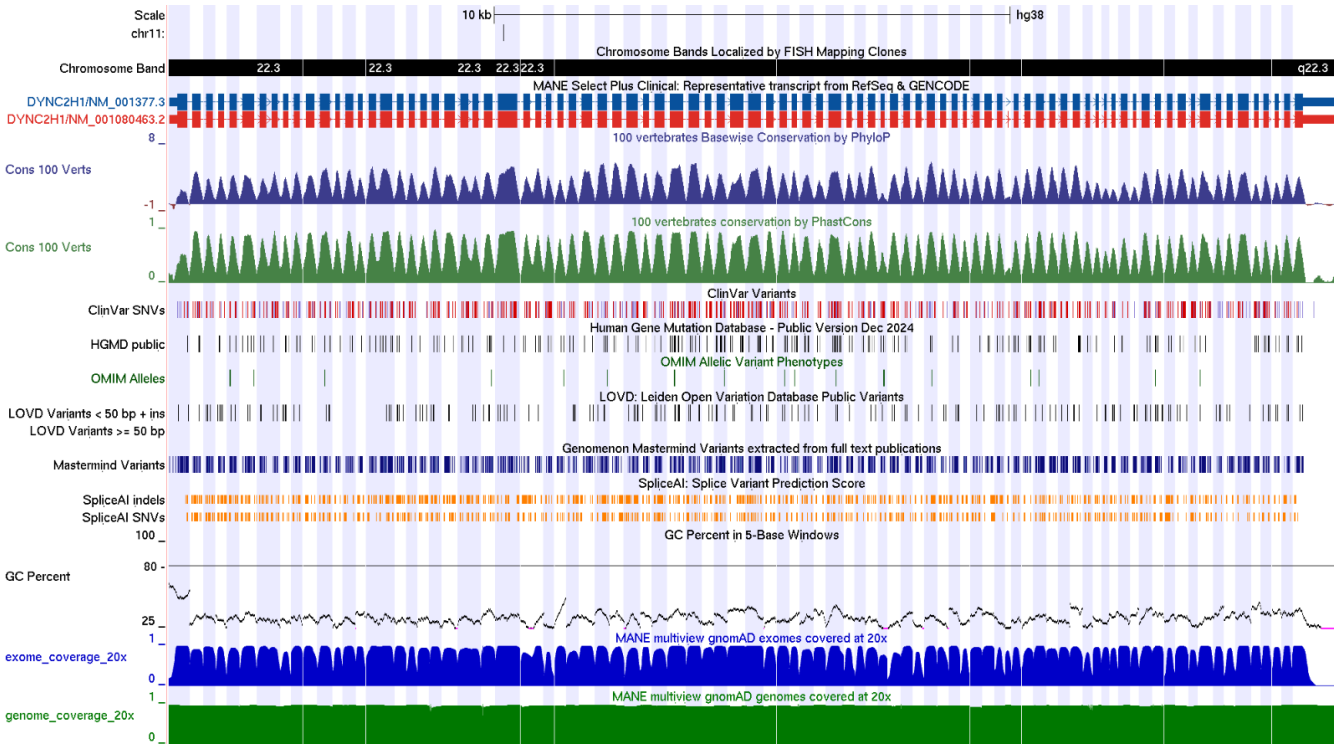

Figure S4: Example of *DYNC2H1* where the MANE Select (NM_001377.3) and PC (NM_001080463.2) transcripts differ by the inclusion of a single 21 bp exon. This small exon lies within intron 63/88 of the Select transcript. This change would be expected to increase the protein length from 4307aa to 4314aa, representing a 0.16% increase in size. An interactive version of this figure is available at https://genome.ucsc.edu/s/ExeterGenetics/MANE_DYNC2H1. The two pathogenic variants in ClinVar linked in this PC-specific exon are NM_001080463.2:c.9836C>G (p.Ser3279Ter, VCV000978199.2) and NM_001080463.2:c.9839G>A (p.Trp3280Ter, VCV000567216.5).

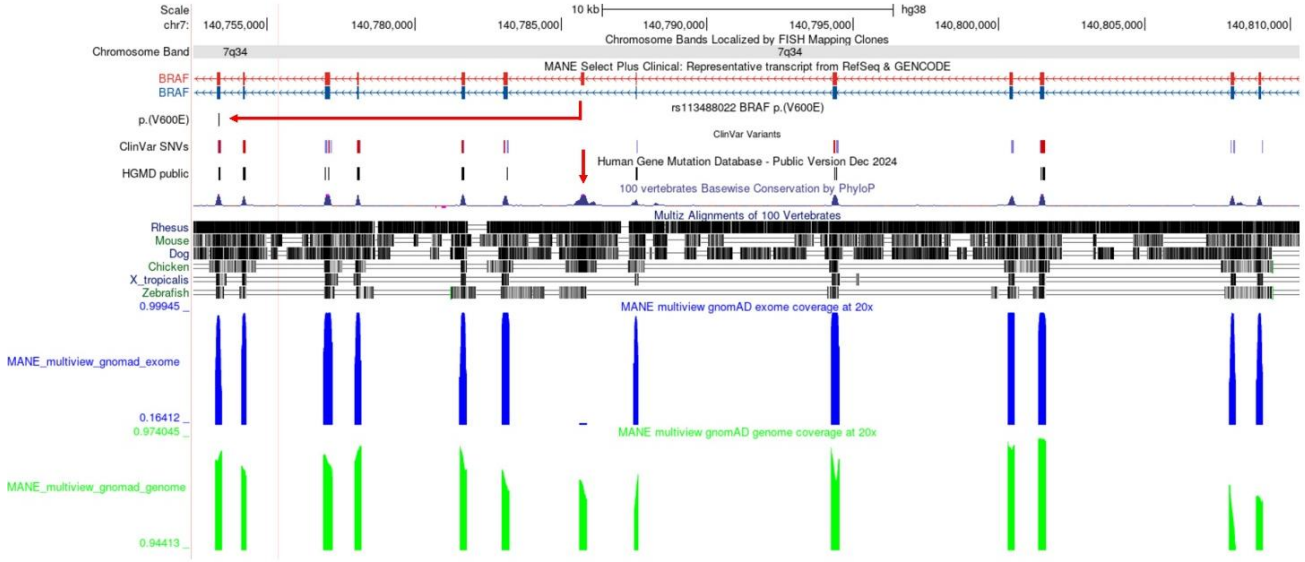

Figure S5: Conservation of an alternatively spliced internal exon in *BRAF,* where inclusion alters the codon numbering for a well-established hotspot mutation. Arrow highlights that inclusion of the 120 bp exon specific to the MANE PC transcript (red) would change annotation of a well-known pathogenic variant that lies in a downstream exon, from p.V600E to p.V640E. Another red arrow highlights the evolutionary sequence conservation seen for the alternatively spliced exon. An interactive session is also available at https://genome-euro.ucsc.edu/s/AlistairP/BRAF_example. For visualisation purposes, the large intron 13/14 is truncated. Switching of the PC and Select transcripts was actioned between versions 36 and 37 of the ClinVar record for the recurrent variant; in accession VCV000013961.36 and in earlier versions it is described as NM_001374258.1:c.1919T>A, p.(Val640Glu), whilst in VCV000013961.37, up to the current version (VCV000013961.140), it is NM_004333.6:c.1799T>A, p.(Val600Glu).

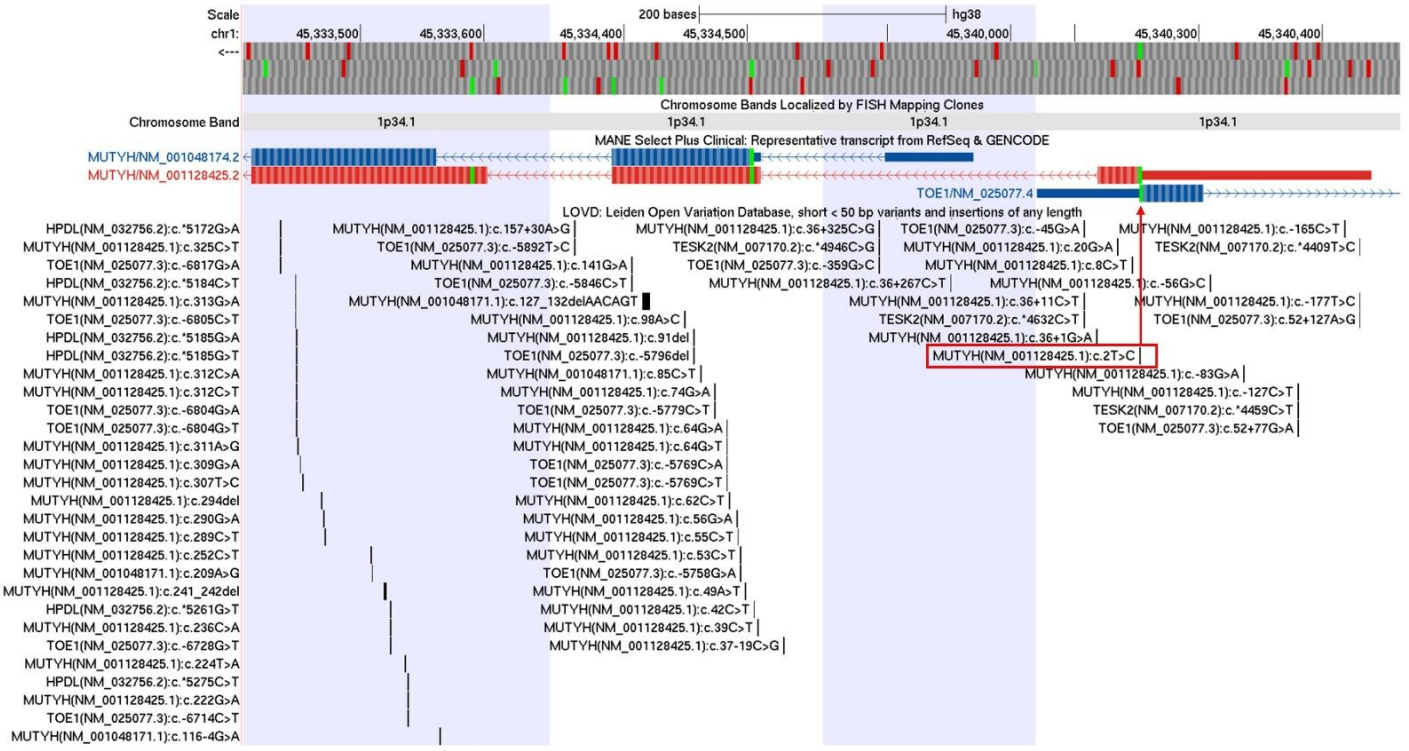

Figure S6: Alternative 5'-UTR exons for *MUTYH* and an overlapping start codon. Comparison of the MANE Select (blue) and MANE Plus Clinical (MPC; red) transcripts shows the latter to have a longer version of exon 3. There is also a different exon 1 that encodes an additional 14 amino-acids and overlaps the start codon of *TOE1*. The NM_001128425.2:c.2T>C

variant from the LOVD database (highlighted with a red rectangle) is a start-loss for both *MUTYH* and for *TOE1*. This NC_000001.11:g.45340253A>G variant can therefore also be annotated as NM_025077.4:c.1A>G. The variant was identified in an individual with colorectal cancer, clinically-confirmed polyposis, with all 4 mismatch-repair genes expressed and showing no evidence for microsatellite instability; and so classified as likely pathogenic (#0000439057) in LOVD. However, the variant is listed as a VUS in ClinVar (VCV000230848.44). An interactive version of this figure is available at https://genome.ucsc.edu/s/ExeterGenetics/MUTYH. Analysis using SpliceAI-visual[11] shows that two of the HGMD variants at the start of exon 3 (NC_000001.11(NM_001128425.2):c.158-1G>T and NC_000001.11(NM_001048171.2):c.116-4G>A) are predicted to result in increased splicing efficiency of exon 3 acceptor sites that are more 3' when compared to the site used in the MPC transcript, see https://genome.ucsc.edu/s/ExeterGenetics/MUTYH_exon3_splicing.

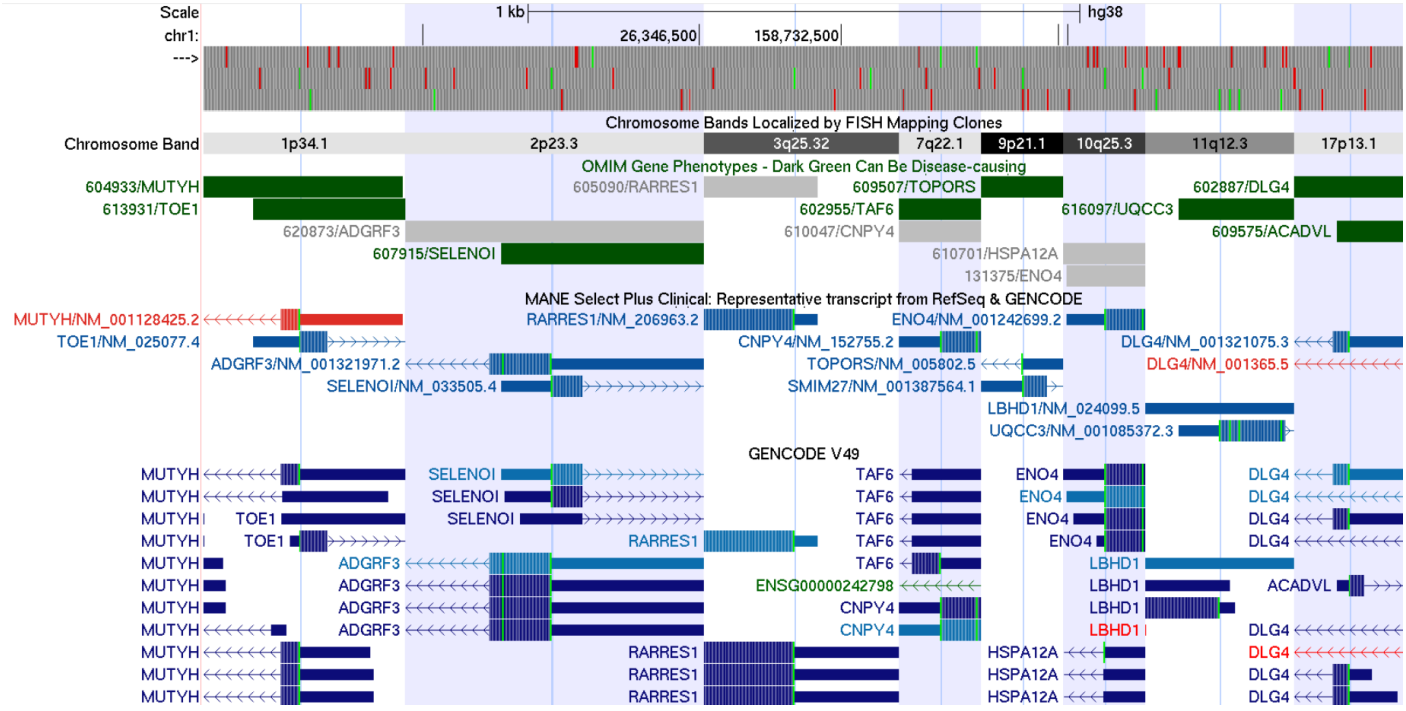

Figure S7: Start codon overlaps in the human genome. Multi-window UCSC session showing 8 gene-pairs where the AT of the ATG start codon are overlapping on opposing strands. An interactive version of this figure is available at https://genome.ucsc.edu/s/ExeterGenetics/Start_overlaps.

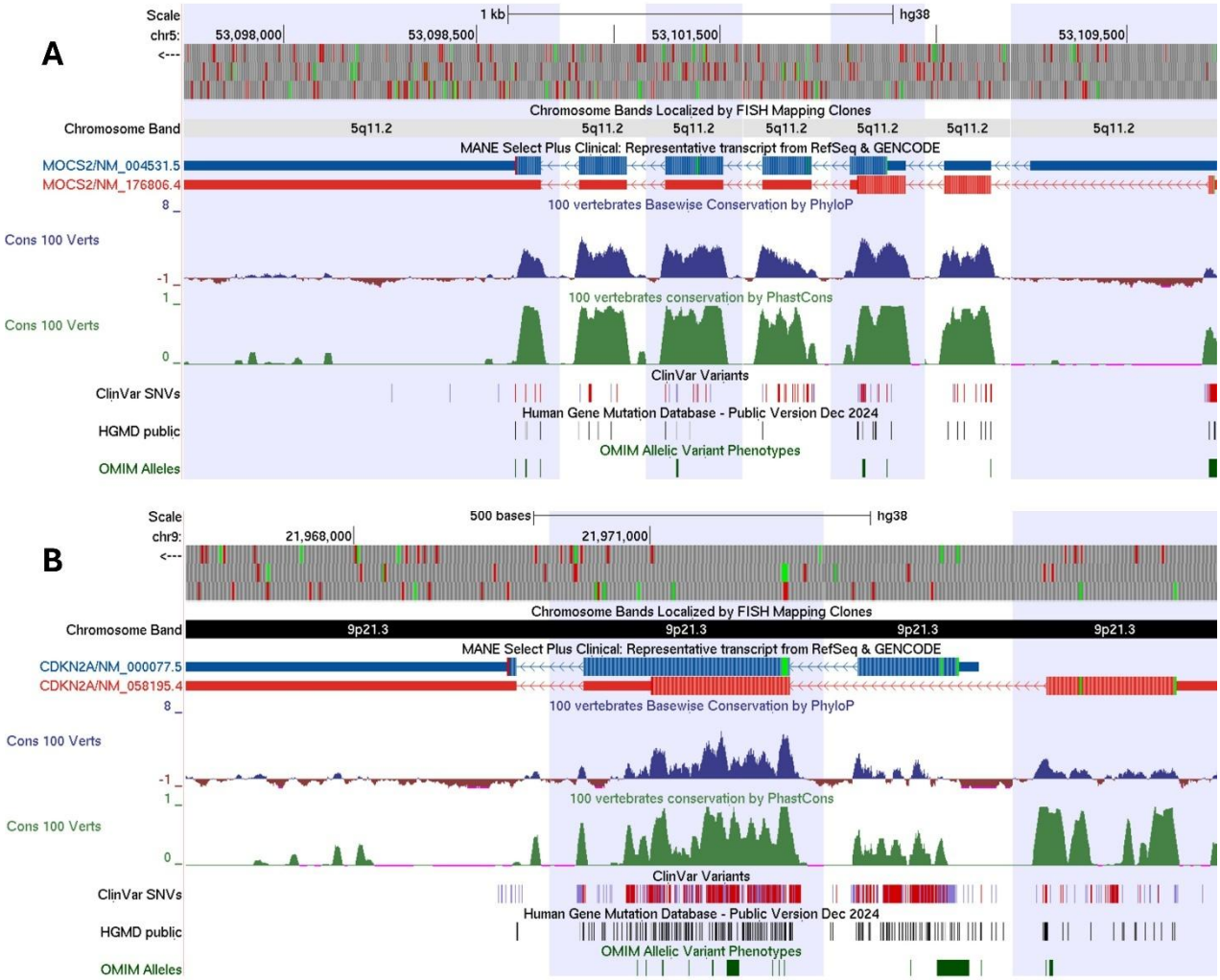

Figure S8: Comparison of the MANE Select and PC transcripts for *MOCS2* and *CDKN2A*. A) Genome browser graphic shows that although the MANE Select and PC transcripts for the *MOCS2* share overlapping exons, the different reading frames in exon 3 means that the two encoded proteins share no amino-acid similarity. B) There is a similar pattern seen for the MANE Select and PC transcripts of *CDKN2A* for which the coding sequence in exon 2 overlaps but is in a different reading frame. The ClinVar, HGMD and LOVD tracks indicates that for both these genes, variants considered to be disease causing are spread across both isoforms. Interactive sessions that contains that contains the information shown above, together with additional UCSC tracks are available at http://genome.ucsc.edu/s/ExeterGenetics/MANE_MOCS2 and http://genome.ucsc.edu/s/ExeterGenetics/MANE_CDKN2A.

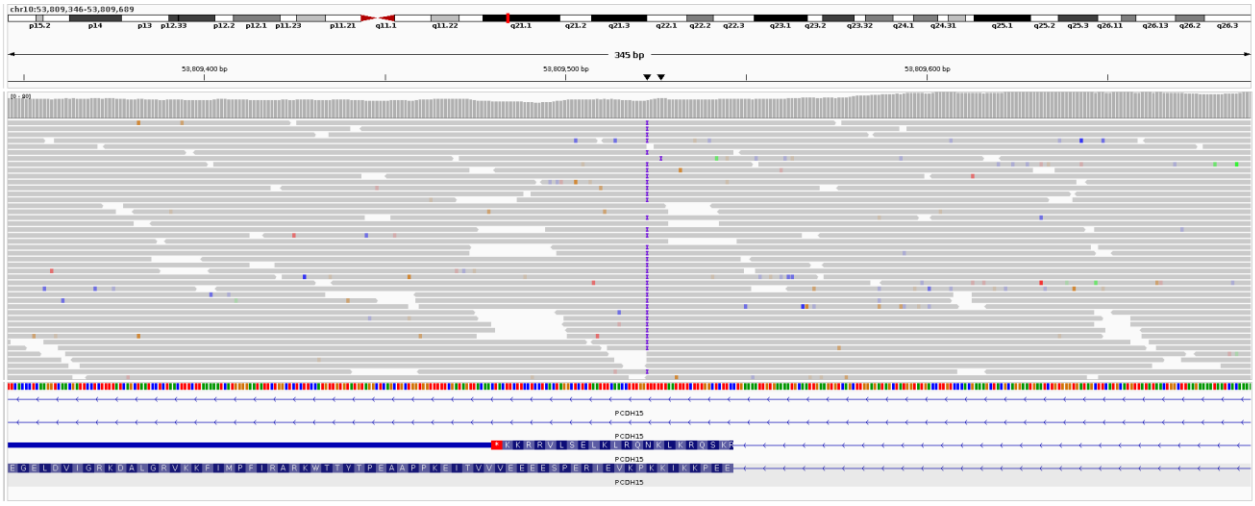

Figure S9: Read alignments supporting a genetic diagnosis of hearing loss due to a homozygous frameshift in the CD2 isoform of *PCDH15*. This 100k Genomes Project participant is homozygous for NM_001142769.3:c.4542dup[8] which predicts a frameshift (p.Pro1515ThrfsTer4) based on the "CD2" isoform, but is intronic based on the MANE Select isoform (NM_001384140.1:c.4671+1033dup) and lies >10kb downstream of the Plus Clinical isoform.

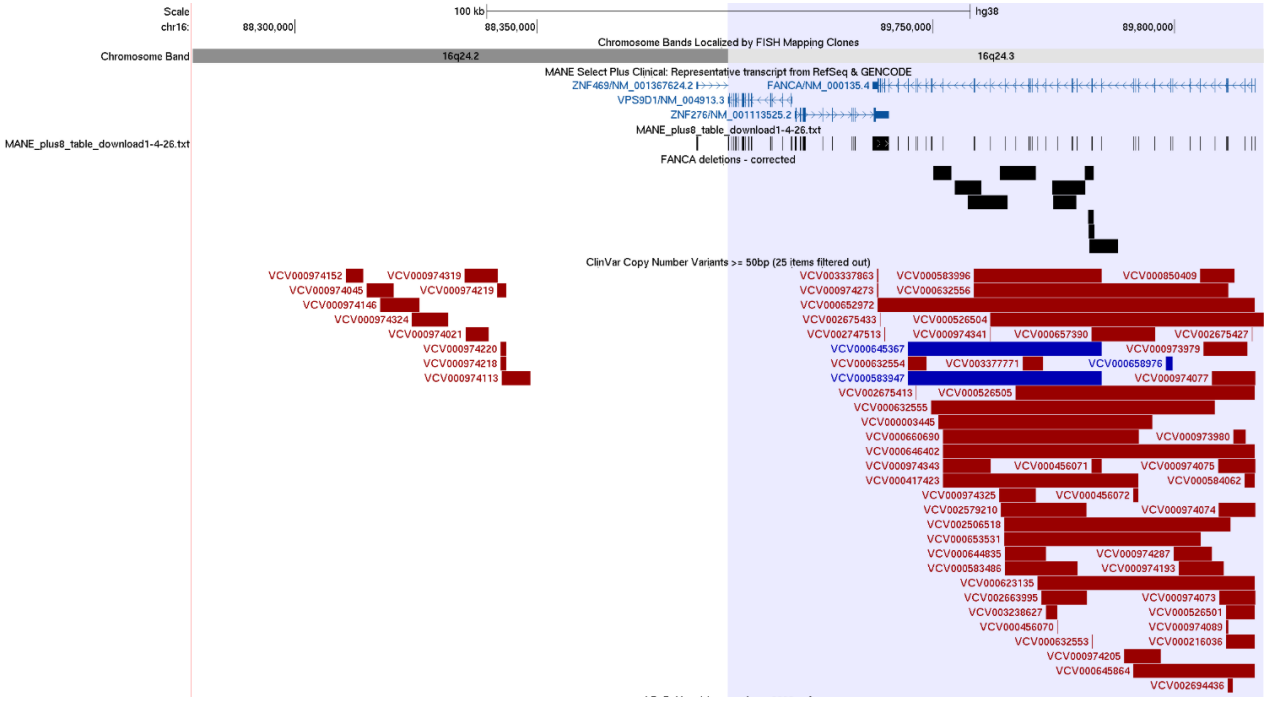

Figure S10: Image from multi-region UCSC genome browser session showing apparent clustering of (likely) pathogenic deletion upstream of *ZNF469*. These set of deletions were published by Flynn *et al* in 2014[9] with genomic coordinates based on build hg18, but these were incorrectly lifted over to GRCh38. Corrected coordinates are shown in the custom track in the RHS panel. For the interactive session (https://genome.ucsc.edu/s/ExeterGenetics/FANCA_liftover_errors), we note that the incorrectly mapped deletions on the LHS will likely disappear in future due to ClinVar/UCSC updates.

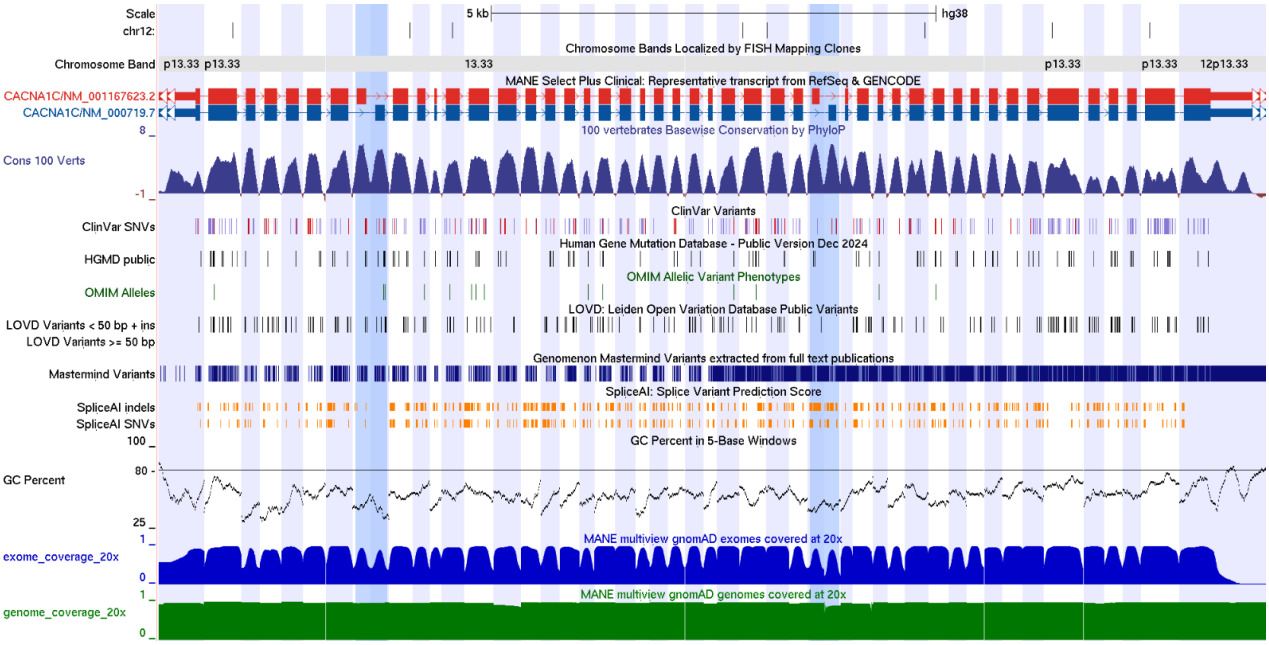

Figure S11: Image from multi-region UCSC genome browser session showing example of gene with two pairs of mutually exclusive exons. *CACNA1C* (MIM 114205) is a gene linked to both cardiac and neurodevelopmental diseases. For exons 8 and 30 (light blue highlighting), *CACNA1C* has different exons in the MANE Select (dark blue) and MANE PC (red) isoforms. All 4 variable exons contain ClinVar (L)P variants and are had good coverage in genome and exome sequencing. An interactive session is available at http://genome.ucsc.edu/s/ExeterGenetics/CACNA1C_example. This was the only MANE PC gene where there were two pairs of mutually exclusive exons. Long-read RNAseq data would be useful to assess whether all 4 combinations (i.e. 8A-30A, 8B-30B, 8A-30B, 8B-30A) are present in different tissues, or whether the inclusion of the A and B exons is co-regulated.

Supplementary references

1. Carvill, G.L. *et al.* Aberrant Inclusion of a Poison Exon Causes Dravet Syndrome and Related SCN1A-Associated Genetic Epilepsies. *Am J Hum Genet* **103**, 1022-1029 (2018).
2. Tran, M. *et al.* Clinical characteristics of high myopia in female carriers of pathogenic RPGR mutations: a case series and review of the literature. *Ophthalmic Genet* **44**, 295-303 (2023).
3. Lettice, L.A. *et al.* A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet* **12**, 1725-35 (2003).
4. Kang, H. *et al.* Clinical application of whole exome sequencing (WES) in the genetic diagnosis of 768 Chinese patients with bilateral hearing loss. *Eur J Hum Genet* **33**, 1484-1492 (2025).
5. Carlson, R.J. *et al.* Association of Genetic Diagnoses for Childhood-Onset Hearing Loss With Cochlear Implant Outcomes. *JAMA Otolaryngol Head Neck Surg* **149**, 212-222 (2023).
6. Sloan-Heggen, C.M. *et al.* Characterising the spectrum of autosomal recessive hereditary hearing loss in Iran. *J Med Genet* **52**, 823-9 (2015).
7. Abu Rayyan, A. *et al.* Genomic analysis of inherited hearing loss in the Palestinian population. *Proc Natl Acad Sci U S A* **117**, 20070-20076 (2020).
8. Pepermans, E. *et al.* The CD2 isoform of protocadherin-15 is an essential component of the tip-link complex in mature auditory hair cells. *EMBO Mol Med* **6**, 984-92 (2014).
9. Flynn, E.K. *et al.* Comprehensive analysis of pathogenic deletion variants in Fanconi anemia genes. *Hum Mutat* **35**, 1342-53 (2014).
10. Schuurs-Hoeijmakers, J.H. *et al.* Identification of pathogenic gene variants in small families with intellectually disabled siblings by exome sequencing. *J Med Genet* **50**, 802-11 (2013).
11. de Sainte Agathe, J.M. *et al.* SpliceAI-visual: a free online tool to improve SpliceAI splicing variant interpretation. *Hum Genomics* **17**, 7 (2023).