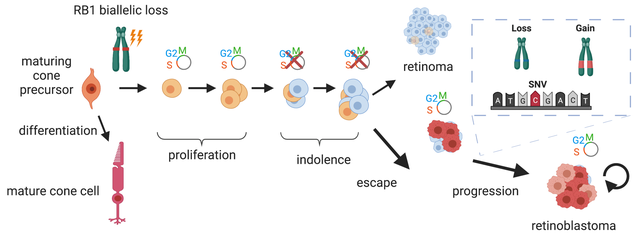
Non-Synonymous, Synonymous, and Non-Coding Nucleotide Variants Converge on Mitotic, RNA Processing, and PRC1 Dysregulation During Retinoblastoma Progression

Kevin Stachelek[1,2], Narine Harutyunyan[1], Susan Lee[1], Assaf Beck[1], Jonathan Kim[1,3], Liya Xu[1,3], Jesse L. Berry[1,3,6], Aaron Nagiel[1,3], C. Patrick Reynolds[4], A. Linn Murphree[1,3], Thomas C. Lee[1,3], Jennifer G. Aparicio[1], and David Cobrinik[1,3,5,6*]

[1]The Vision Center and Saban Research Institute, Children's Hospital Los Angeles, Los Angeles, CA ; [2]Cancer Biology and Genomics Program, Keck School of Medicine, University of Southern California, Los Angeles, CA; [3]Department of Ophthalmology and Roski Eye Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA; [4]Department of Pediatrics, Cancer Center, Texas Tech University Health Sciences Center, Lubbock, TX; [5]Department of Biochemistry & Molecular Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA; [6]Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA

Background

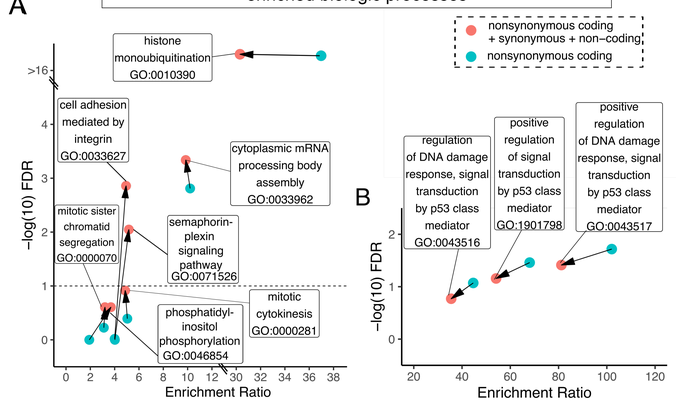
- Retinoblastomas initiate due to biallelic RB1 loss or MYCN amplification and progress through secondary mutations.
- These mutations include non-recurrent nucleotide variants whose significance has been unclear.
- Whole genome or exome sequencing (WGS/WES) of 156 treatment-naïve retinoblastomas has revealed no significantly recurrently altered genes with exome tier-1 mutations



Approach

- Call somatic variants using GATK MuTect2 and Strelka2 in retinoblastoma tumors and cell lines relative to matched normal DNAs
- Identify recurrently altered progression-related processes by integrating whole exome sequencing of 12 retinoblastomas and matched cell lines with 156 retinoblastomas from prior studies.
- Determine contributions of noncoding and synonymous variants in overrepresented processes
- Identify transcriptional effects of 3' UTR sequences of *PCGF3* and *CDC14B* by luciferase reporter assay.
- Test post-transcriptional effects of synonymous mutation in *DYNC1H1* using CRISPR base editing and immunoblotting

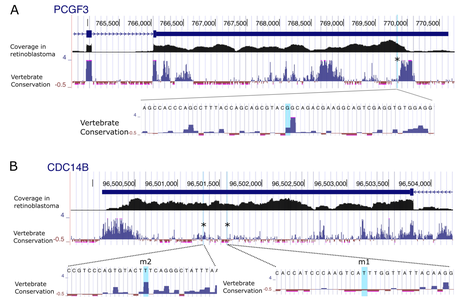
Non-coding and synonymous variants enhance detection of enriched biologic processes



geneSet	Ontology Description	FDR	Variant Genes in Retinoblastoma Tumors
GO:0010390	Histone monoubiquitination	< 1E-16	***BCOR* (18)**; *DDB1; RNF20; PCGF3*
GO:0016569	Covalent chromatin modification	2.25E-08	***BCOR* (18); *CREBBP* (2)**; *BRCA2; BRMS1; CHD1; DDB1; EHMT1; EYA1; HDAC10; HIST1H1E; KAT6A; KDM8; NSD1; PADI4; PRKCA; RNF20; SUPT6H; TAF1; TAF1L; TAF9; PCGF3, HDAC9*
GO:0033962	Cytoplasmic mRNA processing (P) body assembly	4.60E-04	***PAN2* (2)**; *CNOT1; CNOT2; DYNC1H1*
GO:0022618	Ribonucleoprotein complex assembly	5.36E-04	***G3BP1* (2); *NAF1* (2); *PAN2* (2)**; *CELF3; CNOT1; CNOT2; EIF3E; EIF4B; GEMIN4; NLE1; TAF9; DYNC1H1; GEMIN8; RPF2; USP4*
GO:0033627	Cell adhesion mediated by integrin	1.38E-03	***HRG* *HRG***; *CRK; ITGB4; ITGB6; ITGB1; ACER2; ITGB2*
GO:0071526	semaphorin-plexin signaling pathway	9.00E-03	***NCAM1* (2)**; *SEMA3G; SEMA4D; MET*
GO:0000281	Mitotic cytokinesis	0.121	*BRCA2; ANK3; INCENP; NUP62; SEPT10; SPTBN1; STAMBP; UNC119; MYH10; ZFYVE26*
GO:0046854	phosphatidylinositol phosphorylation	0.248	*CISH; FGF18; PI4KA; PIKFYVE; ICOS; MET; PIK3CD; PIK3R6; PIP5K1C; SRC*
GO:0000070	Mitotic sister chromatid segregation	0.248	*BOD1; BUB1B; DYNC1LI1; HIRA; INCENP; KIF14; NUP62; PIBF1; PRC1; CDC14B (2); RAB11A; TTN*

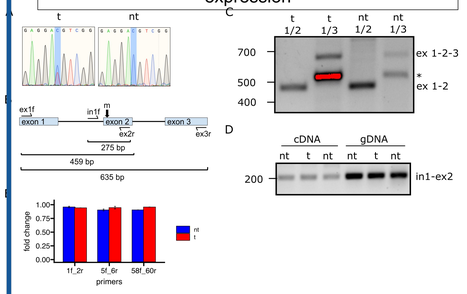
Biological process gene ontology terms over-represented in exome variants in retinoblastoma whole exome or whole genome sequencing. performed separately for non-synonymous variants affecting protein amino acid sequence (blue) and for all exomic variants including synonymous coding mutations and non-coding 5' or 3' UTR mutations (red)

Noncoding variants of interest in 3' UTR



Expression and conservation of PCGF3 and CDC14B variant sequences. UCSC Browser views of PCGF3 (A) and CDC14B (B) 3' UTRs, displaying RNAseq expression in a representative retinoblastoma tumor (GSE125903) and vertebrate conservation at low and high resolution. Note genomic CDC14B sequences represent the cDNA minus strand. Asterisks and blue shading indicate variant positions

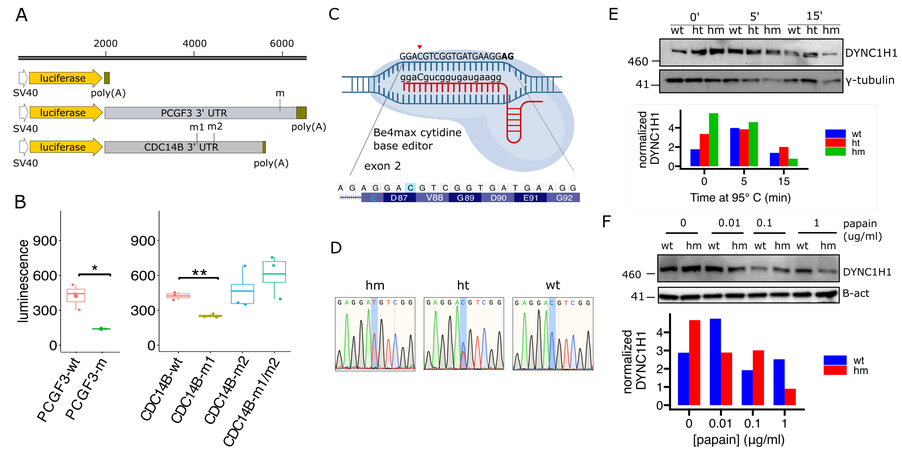
DYNC1H1 variant effects do not impact splicing or expression



DYNC1H1 splicing and RNA expression not impacted by C.261C>T. A. Sanger sequence traces of base-edited DYNC1H1 region B. DYNC1H1 exons 1,2,3 with PCR primer positions and predicted product sizes. C. RT-PCR analysis of RNA from base-edited (t) or non-targeted (nt) bulk RB31 cultures. D. Amplification of cDNA (generated with DNase-treated RNA) and gDNA products

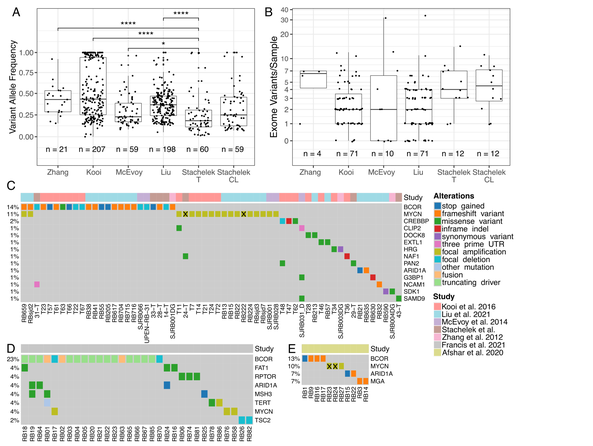
Noncoding variants in PCGF3 and CDC14B impact transcription

Synonymous variant in DYNC1H1 impacts post-transcriptional stability



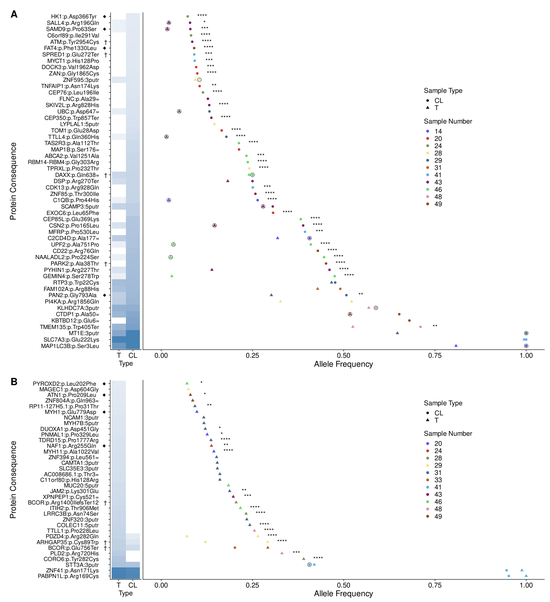
Effects of non-coding and synonymous mutations in *PCGF3*, *CDC14B* and *DYNC1H1*. A. pGL3-SV40 luciferase reporter plasmids showing *PCGF3* and *CDC14B* 3' UTRs and position of mutations (m, m1, m2). B. Luciferase assays in retinoblastoma cells transfected with the indicated reporter constructs. C. Base-editing strategy for *DYNC1H1* c.261C>T using BE4max cytosine base editor. D. Sanger sequence traces of edited homozygous (hm), heterozygous (ht) and wild-type (wt) RB31 clones. E. DYNC1H1 immunoblot of whole cell lysates of homozygous and wild-type CHLA-VC-RB31 clones digested with papain at increasing concentrations. b-actin is a loading control. Graphs indicate DYNC1H1/ b-actin ratios for each condition.

Variant allele frequencies, variant numbers and mutated genes in retinoblastoma WES/WGS studies



Somatic variant frequencies and repeatedly mutated genes beyond *RB1* in 168 fully reported retinoblastoma tumors. A. Variant allele frequencies (VAF) of all exomic non-*RB1* single nucleotide and indel variants in this and four prior studies. B. Number of variants detected per sample in each study displayed with pseudo log scale. C-E. Repeatedly mutated genes beyond *RB1* identified in retinoblastoma whole genome or whole exome sequencing (C) or in targeted sequencing (D, E). Tumors with focal *MYCN* amplification and no annotated RB1 lesion are indicated with x.

Enrichment and depletion of somatic variants during retinoblastoma cell line establishment



Curated somatic variants in CHLA-VC-RB tumor and cell lines. Variants in tumors (triangles) and cell lines (circles) are displayed in ascending VAF order for samples with increased variant frequency in tumor-derived cell lines. Circled icons indicate variants initially detected in a cell line and subsequently detected exclusively in the matched tumor. ◆, alleles confirmed by targeted sequencing. †, genes listed in targeted sequencing panels (UCSF500 or MSK-IMPACT). Asterisks, significantly altered allele frequencies in tumor-cell line pairs (*, p<0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001; Fisher exact test).

Conclusions

- Identified genomic abnormalities and altered cellular processes that contribute to retinoblastoma progression.
- Over-represented ontologies were more significantly over-represented when synonymous coding and non-coding UTR mutations were included
- Our analyses reveal that retinoblastoma progression is associated with accumulation of subclonal gene variants that affect cell signaling pathways different from those queried in current cancer gene sequencing panels.
- Most of the affected biological process ontologies were more significantly over-represented when including non-coding UTR variants that alter the levels and thus the stoichiometry of multi-protein complex components
- Noncoding and synonymous variants of unknown significance are demonstrated to impact transcription and post-transcriptional processes