

Germline pathogenic variants in 786 neuroblastoma patients

Jung Kim, PhD^{1*}; Zalman Vaksman, PhD^{2,3*}; Laura E. Egolf, PhD^{3,4}; Rebecca Kaufman, MS³;
J. Perry Evans, PhD^{2,3}; Karina L. Conkrite, BS³; Arnavaz Danesh, MS⁵; Gonzalo Lopez, PhD³; Michael P.
Randall, MD³; Maiah H. Dent, MS³; Lance M. Farra, BS³; Neil Menghani, BA³; Malwina Dymek, BS³;
Heena Desai MS^{6,7}; Ryan Hausler MS^{6,7}; Penn Medicine BioBank^{8,⊗}; Regeneron Genetics Center^{9,⊗}; Cancer
Genomics Research Laboratory¹⁰; Jaime Guidry Auvil, PhD¹¹; Daniela S. Gerhard, MD¹¹; Hakon Hakonarson,
MD, PhD^{12,13}; Kara N. Maxwell, MD^{6,7}; Kristina A. Cole, MD, PhD^{3,13}; Trevor J. Pugh, PhD^{5,14}; Kristopher R.
Bosse, MD^{3,7,13}; Javed Khan, MD¹⁵; Jun S. Wei, PhD¹⁵; John M. Maris, MD^{3,7,13}; Douglas R. Stewart, MD^{1**};
Sharon J. Diskin, PhD^{2,3,7,13,**,Λ}

¹ Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Rockville, MD, USA

² Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, PA, USA

³ Division of Oncology and Center for Childhood Cancer Research, Children's Hospital of Philadelphia, Philadelphia, PA, USA

⁴ Cell and Molecular Biology Graduate Group, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

⁵ Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, ON, M5S Canada

⁶ Division of Hematology/Oncology, Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

⁷ Abramson Family Cancer Research Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

⁸ Penn Medicine BioBank, Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

⁹ Regeneron Genetics Center, Regeneron Pharmaceuticals Inc., Tarrytown, NY, USA

¹⁰ Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Rockville, MD, USA

¹¹ Office of Cancer Genomics, National Cancer Institute, Bethesda, MD, USA

¹² Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

¹³ Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

¹⁴ Department of Medical Biophysics, University of Toronto, Toronto, Ontario, ON, M5S Canada

¹⁵ Oncogenomics Section, Genetics Branch, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD, USA

* First author equal contribution

** Senior author equal contribution

⊗ Full author list in supplement

^ Correspondence: Sharon J. Diskin, PhD (diskin@chop.edu)

Short title: Rare germline pathogenic variation in neuroblastoma

Key words: Neuroblastoma, genetic predisposition, germline, survival, DNA repair

Key Points

Question: What is the prevalence and clinical significance of germline pathogenic variants in cancer predisposition genes (CPGs) in neuroblastoma patients?

Findings: Among 786 neuroblastoma patients with germline DNA sequencing, 13.9% harbored a pathogenic (P) or likely pathogenic (LP) variant in a CPG. The number of patients with germline P-LP variants in *BARD1* and other DNA repair genes was significantly greater than observed in two cancer-free control cohorts. The presence of a germline P-LP variant was independently predictive of worse overall survival.

Meaning: Germline sequencing should be performed for all neuroblastoma patients at diagnosis to inform genetic counseling and frequency of clinical follow-up. Centralization of these data will facilitate longitudinal and mechanistic studies needed to identify specific actionable events and improve patient outcomes.

Abstract

Importance: Neuroblastoma accounts for 12% of childhood cancer deaths. The genetic contribution of rare pathogenic germline variation in patients without a family history remains unclear.

Objective: To define the prevalence, spectrum, and clinical significance of pathogenic germline variation in cancer predisposition genes (CPGs) in neuroblastoma patients.

Design, Setting and Participants: Germline DNA sequencing was performed on the peripheral blood from 786 neuroblastoma patients unselected for family history. Rare variants mapping to CPGs were evaluated for pathogenicity and the percentage of cases harboring pathogenic (P) or likely pathogenic (LP) variants was quantified. The frequency of CPG P-LP variants in neuroblastoma cases was compared to two distinct cancer-free control cohorts to assess enrichment. Matched tumor DNA sequencing was evaluated for “second hits” at CPGs and germline DNA array data from 5,585 neuroblastoma cases and 23,505 cancer-free control children was analyzed to identify rare germline copy number variants (CNVs) affecting genes with an excess burden of P-LP variants in neuroblastoma. Neuroblastoma patients with germline P-LP variants were compared to those without P-LP variants to test for association with clinical characteristics, tumor features, and patient survival.

Main Outcomes and Measures: Rare variant prevalence, pathogenicity, enrichment, and association with clinical characteristics, tumor features, and patient survival.

Results: We observed 116 P-LP variants in CPGs involving 13.9% (109/786) of patients, representing a significant excess burden of P-LP variants compared to controls (9.1%; $P = 5.14 \times 10^{-5}$, Odds Ratio: 1.60, 95% confidence interval: 1.27-2.00). *BARD1* harbored the most significant burden of P-LP variants compared to controls (1.0% vs. 0.03%; $P = 8.18 \times 10^{-7}$; Odds Ratio: 32.30, 95% confidence interval: 6.44-310.35). Rare germline CNVs disrupting *BARD1* were also identified in neuroblastoma patients (0.05%) but absent in controls ($P = 7.08 \times 10^{-3}$; Odds Ratio: 29.47, 95% confidence interval: 1.52 – 570.70). Overall, P-LP variants in DNA repair genes in this study were enriched in cases compared to controls (8.1% vs. 5.7%; $P = 0.01$; Odds Ratio: 1.45, 95% confidence interval: 1.08-1.92). Neuroblastoma patients harboring a germline P-LP variant had a worse overall survival when compared to patients without P-LP variants ($P = 8.6 \times 10^{-3}$), and this remained significant in a multivariate Cox proportional-hazards model ($P = 0.01$).

Conclusions and Relevance: Neuroblastoma patients harboring germline P-LP variants in CPGs have worse overall survival and *BARD1* is an important predisposition gene affected by both common and rare pathogenic variation. Germline sequencing should be performed for all neuroblastoma patients at diagnosis to inform genetic counseling and support future longitudinal and mechanistic studies. Patients with a germline P-LP variant should be closely monitored, regardless of risk group assignment.

Introduction

Neuroblastoma is an embryonal malignancy of early childhood that arises from developing postganglionic sympathetic neurons and accounts for 12% of all childhood cancer-related deaths¹. Patients are classified into low-, intermediate- and high-risk based on a series of clinical and tumor biological features and this risk group is used for treatment stratification purposes¹. Despite aggressive multi-modal therapy, nearly 50% of high-risk neuroblastoma patients diagnosed over 18 months of age eventually succumb to their disease. A subset of these tumors harbor somatic *MYCN* amplification and/or an activating somatic *ALK* mutation or gene amplification¹. However, sequencing studies of neuroblastoma tumors have revealed a low overall somatic mutation rate and few recurrently mutated genes^{2,3}. The young median age at diagnosis and standardized incidence ratio of siblings of children with neuroblastoma of ~9.7⁴ are consistent with an underlying genetic etiology.

The genetic basis of neuroblastoma predisposition has come into focus over the last decade. Familial neuroblastoma, which accounts for 1-2% of cases, arises primarily from pathogenic germline variants in *ALK*⁵, with rarer neurocristopathy syndrome cases explained by germline pathogenic variants in *PHOX2B*^{6,7}. However, the vast majority of neuroblastomas appear to arise sporadically, without a family history. Genome-wide association studies (GWAS) have identified common variants associated with sporadic neuroblastoma at over a dozen loci. These genetic associations have implicated multiple candidate genes including *CASC15*, *NBAT1*, *BARD1*, *LMO1*, *DUSP12*, *DDX4*, *IL31RA*, *HSD17B12*, *HACE1*, *LIN28B*, *TP53*, *RSRC1*, *MLF1*, *CPZ*, *MMP20*, *KIF15*, and *NBPF23*⁸⁻¹⁶. Several susceptibility genes identified by GWAS not only influence disease initiation, but also drive tumor aggressiveness and/or maintenance of the malignant phenotype^{10,12,14,17-20}. A rare 16p11.2 microdeletion syndrome has also been associated with neuroblastoma²¹. Finally, recent sequencing efforts have reported rare pathogenic germline variants in multiple cancer predisposition genes^{2,22-32}; however, the prevalence and clinical significance of these and other rare variants in neuroblastoma remains unclear and requires evaluation in larger patient cohorts with detailed phenotypic data.

Here, we analyzed germline whole genome sequencing (WGS), whole exome sequencing (WES) and targeted capture sequencing (CAP) data from 786 children diagnosed with neuroblastoma and profiled through the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) initiative. Our aims were to 1) determine the prevalence, spectrum, and pathogenicity of rare germline variants in known cancer predisposition genes (CPGs), 2) test for enrichment of rare variants in children with neuroblastoma compared to cancer-free control populations, and 3) evaluate clinical features and outcomes in neuroblastoma patients with and without germline pathogenic (P) or likely pathogenic (LP) variants in CPGs to identify translational opportunities.

Materials and Methods

Detailed methods are provided in **Supplemental Methods**. Briefly, the study cohort consisted of 786 neuroblastoma patients accrued through the Children's Oncology Group (COG) ANBL00B1 biology study, unselected for family history (**Table 1**, **eTable 1**). Patients are assigned a Universal Subject Identifier (USI) by the COG and are not identifiable. Germline DNA and matched diagnostic tumor DNA and RNA were sequenced through the TARGET initiative. The original set of tumor-normal pairs were sequenced with Complete Genomics WGS (n=134) and/or Illumina WES (n=222), as previously described^{2,33}. A total of 59 samples were sequenced by WGS and WES provided internal validation. We have previously reported a small number of germline variants based on the WES cohort²; however, an in-depth study of pathogenic germline variation in these children was not performed at that time. Germline DNA from an independent neuroblastoma cohort (n=489) was sequenced using Illumina custom capture (CAP) panels, including a germline panel newly designed for this study (n=166 genes; **eTable 2**). Cancer-free control data were obtained from the Penn Medicine BioBank (PMBB; n=6,295)^{34,35} and gnomAD v2.1 without cancer (n=15,708). Ancestry for neuroblastoma and PMBB participants was inferred through principal component analysis using matched germline DNA array data. Germline variants were called using GATK best practices (WES and CAP data) or the Complete Genomics pipeline (v2) with custom filtering³⁶ (WGS data), then annotated with SnpEff³⁷ (v4.3t) and ANNOVAR³⁸. For all cohorts (neuroblastoma, PMBB and gnomAD), rare germline variants (<0.1% across

each population in public control databases) in 166 cancer predisposition genes were then assessed for pathogenicity with a clinically-focused pipeline incorporating ClinVar evidence and a modified implementation of InterVar (**eFigure 1**). A subset of germline variants were validated through Sanger sequencing. Cases harboring a germline P-LP variant in a cancer predisposition gene were further assessed using matched tumor sequencing data when available. Fisher's exact test was used to compare the enrichment of P-LP variants in neuroblastoma cases to controls and across clinical and biological subsets of neuroblastoma. P-LP variant enrichment was also compared at gene and pathway levels, using a Bonferroni correction for multiple testing. Kaplan-Meier analyses of event-free and overall survival were performed to compare outcomes of patients with and without germline P-LP variants. A multivariate Cox proportional-hazards regression model was used to assess if the presence of a CPG P-LP variant was independently predictive of survival.

Results

Neuroblastoma patient characteristics

A total of 786 neuroblastoma patients were included in the study (**Table 1, eTable 2**). Overall, 564 cases (71.8%) were classified as high-risk based on the COG risk stratification system, and 546 of these patients had stage 4 disease according to the International Neuroblastoma Staging System (INSS) criteria. Cases profiled by WGS or WES were intentionally enriched for high-risk disease, consistent with the overall goals of the TARGET initiative. In contrast, cases that underwent targeted capture (CAP) sequencing were representative of the general neuroblastoma risk group profile. A total of 769 patients had available matched germline DNA array data and were evaluated for ancestry by principal component analysis (PCA). As expected, the majority of cases (66.8%) were inferred to be of European ancestry (**eFigure 2**).

Frequency of P-LP variants in known cancer predisposition genes

We observed 116 P-LP variants involving 54 of the 166 cancer predisposition genes studied (**Figure 1, eTable 3**). Of these variants, 73 were classified as P-LP based on ClinVar evidence and 43 variants were assigned P-

LP based on our revised InterVar assessment. Overall, P-LP variants were detected in 109/786 (13.9%) of cases. Classic familial neuroblastoma germline variants were observed in 0.4% (3/786) of cases. These included two patients with *ALK* (p.R1275Q) activating variants and a single patient with a *PHOX2B* splice variant (NM_003924:exon3:c.430-2A>G). An additional *ALK* variant (p.I1250T) was predicted to be LP; however, this variant was not found to be activating in a previous study³⁹. Six cases harbored more than one P-LP variant in a known cancer predisposition gene (**eTable 4**). A total of 27 genes harbored P-LP variants in two or more cases, with variants in the *BARD1* gene being most frequent (8/786 cases or 1.0% overall; **eFigure 3, eTable 3**). Select variants were validated by Sanger sequencing (**eTable 5**). Notably, of the nine neuroblastoma genes identified by GWAS and included in this study, only *BARD1* and *TP53* harbored rare P-LP variants.

Matched tumor DNA sequencing reveals P-LP variants are retained and second hits are rare

We examined somatic alterations affecting genes with germline P-LP variants using matched published neuroblastoma tumor WGS and WES variant calls³³ and our analysis of available CAP sequencing from TARGET. Nearly all germline P-LP variants were detected in matched tumor DNA, when available (95%, 73/77 variants; **eTable 3**). For variants not detected in the matched tumor (n=4), germline variant allele fraction (VAF) ranged from 0.25 to 0.31. No second hit SNVs or indels were observed in the tumor DNA of patients harboring a germline P-LP variant in a CPG. We detected one focal somatic deletion within *EZH2* observed in the tumor from PASEGA, a patient who also harbored a pathogenic germline *EZH2* variant (p.T536fs). Phase could not be determined, but this somatic event was confirmed by Sanger sequencing of the tumor DNA (**eFigure 4**).

Germline P-LP variants in cancer predisposition genes are enriched in neuroblastoma

To assess whether children diagnosed with neuroblastoma harbor an excess of rare pathogenic germline variation in the 166 CPGs considered in this study, we compared the P-LP burden in neuroblastoma cases to two independent control cohorts (**Figure 2A, top panel**). First, we applied our full analytic pipeline (alignment, variant calling, quality control, and pathogenicity assessment) to subjects sequenced through the Penn

Medicine Bio Bank (PMBB) and without a history of cancer or benign tumors (n=6,295, **eFigure 5**). Germline P-LP variants were significantly enriched in neuroblastoma patients compared to PMBB ($P_{\text{PMBB}} = 5.14 \times 10^{-5}$; Odds Ratio: 1.60; 95% confidence interval: 1.27-2.00). To assess reproducibility of this result, identical filtering and pathogenicity assessment was applied to rare variants in gnomAD excluding cancer samples (n=15,708 individuals). This confirmed the excess burden of P-LP variants in neuroblastoma ($P_{\text{gnomAD}} = 1.82 \times 10^{-3}$; Odds Ratio: 1.41; 95% confidence interval: 1.34-1.74).

Germline P-LP variants in BARD1 are enriched in neuroblastoma patients

Next, we performed gene-based rare variant burden testing comparing neuroblastoma cases to PMBB and gnomAD control cohorts (**eTable 6, eTable 7**). Five genes (*BARD1*, *EZH2*, *ALK*, *PTCH1* and *MSH3*) exhibited significant enrichment ($P < 0.05$) and odds ratio lower 95% confidence interval > 1.0 in both control cohort comparisons (**Figure 2B**). Pathogenic variants in *BARD1* and *EZH2* were validated by Sanger sequencing in neuroblastoma cases when DNA was available (**eFigure 3** and **eFigure 4**). *ALK* and *BARD1* remained significant after Bonferroni adjustment for multiple testing in at least one control comparison. *ALK* is the main major familial neuroblastoma predisposition gene⁵. Only one LP variant in *ALK* was detected in PMBB ($P_{\text{PMBB}} = 5.00 \times 10^{-3}$; Odds Ratio: 24.09, 95% confidence interval: 1.93-1255.78). This variant (p.A1168T) was classified LP based on InterVar and was not reported in ClinVar. No P-LP *ALK* variants were observed in gnomAD v2.1 whole genome controls ($P_{\text{gnomAD}} = 1.08 \times 10^{-4}$; Odds Ratio: 140.3, 95% confidence interval: 7.24-2719.0).

BARD1 is the only gene that passed a Bonferroni adjustment in both control comparisons (**Figure 2B**). Common variation at the *BARD1* locus is known to be associated with high-risk neuroblastoma from our prior GWAS⁹. We and others have also reported germline *BARD1* rare variants in neuroblastoma patients^{2,40,41}. However, to date, the number of cases analyzed has been limited. Here, we observed rare P-LP variants in *BARD1* in 8/786 (1.0%) of neuroblastoma cases, all predicted to be loss-of-function (**Figure 2C** and **eTable 8**). Moreover, all but one variant was observed in the high-risk subset (**Figure 2D**). Only 2/6,295 (0.03%) of control subjects in PMBB harbored a P-LP germline variant in *BARD1* ($P_{\text{PMBB}} = 8.18 \times 10^{-7}$; Odds Ratio: 32.30, 95%

confidence interval: 6.44-310.35). Similarly, only 15/15,708 (0.09%) control subjects in gnomAD harbored a P-LP germline variant in *BARD1* ($P_{\text{gnomAD}} = 6.64 \times 10^{-6}$; Odds Ratio: 10.75, 95% confidence interval: 3.93-27.13). Enrichment of P-LP variants in *BARD1* remained significant when we restricted the analysis to cases and controls of European ancestry, considering both PMBB ($P_{\text{PMBB}} = 2.70 \times 10^{-5}$; Odds Ratio: 21.36, 95% confidence interval: 4.55-131.80) and gnomAD ($P_{\text{gnomAD}} = 1.55 \times 10^{-4}$; Odds Ratio: 10.13, 95% confidence interval: 3.75-28.74.86), suggesting this result is not likely due to population stratification.

Germline P-LP variants in DNA repair genes are enriched in neuroblastoma patients

BARD1 is known to bind *BRCA1* and influence DNA repair⁴². We hypothesized that neuroblastoma patients harbor an excess burden of P-LP variants in DNA repair pathway genes overall. To explore this hypothesis, we interrogated genes in our 166-gene panel that intersected published DNA repair genes.⁴³ A total of 48 DNA repair genes were assayed in the full study cohort and included in the analysis (**eTable 2**). We observed 68 P-LP variants in 27 distinct DNA repair genes, affecting 64/786 (8.1%) of neuroblastoma patients (**Figure 2A, bottom panel**). In contrast, only 362/6,295 (5.8%) of PMBB controls harbored a P-LP variant ($P_{\text{PMBB}} = 0.011$; Odds Ratio: 1.45, 95% confidence interval: 1.08-1.92). Similarly, only 959/15,708 (6.1%) of gnomAD controls harbored a P-LP variant in a DNA repair gene ($P_{\text{gnomAD}} = 0.028$; Odds Ratio: 1.36, 95% confidence interval: 1.03-1.77).

Germline copy number variants (CNVs) disrupting BARD1 are enriched in neuroblastoma

Given the enrichment of *BARD1* rare P-LP SNVs and indels in this study and the previous association of *BARD1* common variants with neuroblastoma through GWAS, we next sought to determine if rare germline CNVs at *BARD1* also associate with neuroblastoma. We analyzed CNVs in germline DNA array data from 5,585 neuroblastoma cases and 23,505 cancer-free control children genotyped in our neuroblastoma GWAS efforts²¹. We detected three focal germline deletions fully or partially encompassing *BARD1* in neuroblastoma (0.05%; **Figure 3A-E, eTable 9**). No CNVs affecting *BARD1* were observed in 23,505 chip-matched GWAS controls, and no protein coding deletions affecting *BARD1* were observed in 10,847 individuals in the gnomAD

v2.1 structural variant dataset⁴⁴ (**Figure 3A**). The rare germline deletions at *BARD1* were significantly associated with neuroblastoma ($P = 7.08 \times 10^{-3}$; Odds Ratio: 29.47, 95% confidence interval: 1.52 – 570.70; **Figure 3F**) and were detected only in high-risk neuroblastoma patients (**Figure 3G**). Collectively, these data suggest that *BARD1* alterations are an important genetic determinant of neuroblastoma, including common germline variants and rare germline SNVs, indels, and structural variants.

Neuroblastoma patients harboring germline P-LP variants in CPGs have worse overall survival

Finally, we investigated whether rare P-LP variants (SNVs and indels) in CPGs were associated with specific clinical and tumor biological characteristics and patient survival. A nominally significant enrichment of P-LP variants was observed in patients with tumors harboring loss of heterozygosity (LOH) of chromosome 11q ($P = 0.012$); however no association with age at diagnosis, stage, *MYCN* amplification status, COG risk group, or other characteristics was detected (**eTable 10**). We repeated this analysis using the CAP data only; however, the results were similar (**eTable 11**). We next evaluated overall survival probability based on the presence or absence of germline P-LP variants in CPGs. We observed that subjects with a germline P-LP variant have worse overall survival compared to subjects without a germline P-LP variant in a CPG (log-rank test $P = 8.6 \times 10^{-3}$; **Figure 4A**). Furthermore, if restricted to only low- and intermediate-risk subjects, overall survival remains worse for patients with a germline P-LP variant (log-rank test: $P = 1.3 \times 10^{-4}$; **Figure 4B**). A similar trend was observed when restricted to high-risk only, though this did not reach statistical significance (log-rank test $P = 0.1049$; **Figure 4C**). Finally, a multivariate Cox proportional-hazards regression model revealed the presence of a germline P-LP variant was independently predictive of overall survival when considering age and diagnosis, INSS stage, *MYCN* amplification status, and COG risk group (**Figure 4D**; $p=0.017$; hazard ratio= 1.44, 95% confidence interval: 1.07-1.96; **eTable 12**). Taken together, these data demonstrate that the presence of a germline P-LP variant in a CPG is associated with worse survival, independent of risk group.

Discussion

Neuroblastoma is a cancer of the developing sympathetic nervous system with an established genetic basis. Patients who present with a family history of the disease most commonly harbor rare pathogenic variants in *ALK*⁵ or *PHOX2B*^{6,7}. In contrast, GWAS studies have identified common variation associated with sporadic neuroblastoma implicating over a dozen susceptibility genes⁴⁵, including *BARD1*⁹. Germline sequencing studies have reported rare pathogenic variation in CPGs, including *APC*, *AXIN2*, *BARD1*, *BRCA1*, *BRCA2*, *CHEK2*, *LZTR1*, *PALB2*, *PINK1*, *SDHB*, *SMARCA4*, and *TP53*⁴⁶. Neuroblastoma has also been reported in several childhood-onset tumor-predisposition syndromes⁴⁶. However, a study of germline pathogenic variants in a clinically annotated cohort of patients large enough to investigate excess burden of pathogenic variation and clinical features associated with such variation has not been previously reported. Thus, in this study, we sought to define the prevalence, spectrum, and clinical significance of rare pathogenic germline variants in CPGs in neuroblastoma.

To accomplish these goals, we analyzed germline DNA sequencing from 786 neuroblastoma patients with detailed clinical covariate and outcomes data. Using a conservative, clinically-focused pipeline to classify pathogenicity of rare variants in cancer predisposition genes, we observed P-LP germline variants in a substantial portion of cases studied (13.9%). This percentage is slightly higher, but overall in-line with recent pan-childhood cancer germline studies^{22,23,28,31,32,40}. Two genes (*ALK* and *BARD1*) showed enrichment of P-LP variants in neuroblastoma cases compared to independent cancer-free control cohorts after adjusting for multiple testing. Neuroblastoma patients carrying a germline P-LP variant had worse overall survival compared to those without P-LP variants, independent of age at diagnosis, INSS stage, *MYCN* amplification, and COG risk group.

The greatest number of pathogenic germline variants were observed in *BARD1*, *BRCA2*, *ERCC2*, *CHEK2*, and *MSH3*. Notably, *BARD1*, *BRCA2*, *CHEK2* and *MSH3* variants were primarily classified as P-LP based on ClinVar annotation, suggesting that these variants have previously been observed in patients in a clinical lab. All five genes are involved in DNA repair, and indeed we observed an overall enrichment of P-LP variants in DNA repair genes considered here. While this finding requires validation using a full repertoire of DNA repair

genes, the result suggests that neuroblastoma is another cancer initiated by germline defects in DNA repair. The current study was large enough to demonstrate a significant enrichment of rare *BARD1* P-LP variants (SNVs and CNVs) in neuroblastoma, adding to the common variants in *BARD1* identified by GWAS and previously implicated in disease pathogenesis. In addition to *BARD1*, we observed 1.7% (13/786) of children in our cohort with a P-LP variant in *BRCA1*, *BRCA2* or a mismatch repair gene, which is similar to that observed (~1.2%) in a large (n=3,975) meta-analysis of childhood cancer studies⁴⁷. Notably, two of these subjects harbored multiple P-LP variants in these genes, including one patient with a P-LP variant in both *BRCA1* and *BRCA2* another patient with P-LP variants in *BRCA2* and *MSH3*.

In a companion study (co-submitted), multiple lines of evidence demonstrated the impact of *BARD1* germline P-LP variants identified here on DNA repair processes in neuroblastoma⁴⁸. Briefly, a subset of the rare *BARD1* variants identified in the current study were introduced as monoallelic “knock-ins” in neuroblastoma cell models via CRISPR/Cas9 genome editing. These heterozygous variants induced *BARD1* haploinsufficiency leading to DNA repair deficiency and increased DNA damage. Taken together, these data further implicate *BARD1* and DNA repair as important driving factors in neuroblastoma tumorigenesis.

Evidence for bi-allelic inactivation and/or loss of heterozygosity (LOH) in patients with P-LP germline variants was observed in only one tumor and involved *EZH2*. There are several possible explanations for the low rate of bi-allelic inactivation in neuroblastoma. First, haploinsufficiency may be sufficient to tumorigenesis, as seen in our companion *BARD1* functional studies (co-submitted)⁴⁸. Alternatively, other inactivation mechanisms may be present but not detected by our approach (e.g., epigenetic and non-coding alterations). Finally, there have been limited number of subjects with germline P-LP variants and matched tumor data evaluated to date. Functional studies, such as those presented for *BARD1*⁴⁸, and large cohort analyses incorporating the full spectrum of potential inactivation mechanisms are needed to resolve these important questions.

Given the prevalence and relevance of germline pathogenic variants to survival in neuroblastoma, we propose that all newly diagnosed children should undergo paired germline-tumor sequencing and efforts should be

made to centralize these data. This is important for multiple reasons: 1) from a prognostic standpoint, identification of a P-LP variant predicts worse overall survival, independent of clinical risk stratification, 2) identification of P-LP variants in some genes (e.g. *BARD1*) may suggest eligibility for specific therapies, especially at time of relapse, 3) cascade testing of adult family members may guide gene-specific surveillance and therapies (e.g., in *BARD1*, *BRCA1*, *BRCA2*, *CHEK2*, Lynch, *PALB2* and *TP53*), 4) cascade testing of children for select genes (e.g., *TP53*, *PTPN11*, *DICER1*) may guide surveillance and therapy and, 5) identification of pathogenic variation in genes associated with specific syndromes (e.g., *PTCH1*/Gorlin, *EZH2*/Weaver) may guide prognosis and clinical management. While the field is adopting paired germline-tumor DNA sequencing at the time of diagnosis for neuroblastoma and other pediatric cancer patients, there remains confusion on how to act on findings. Centralization of these data, through the Childhood Cancer Data Initiative (CCDI) or similar efforts, will facilitate longitudinal studies of patient survival and provide a resource for prioritizing functional and mechanistic studies to identify specific actionable insights to improve outcomes.

There are some limitations to this study. Patients studied were enrolled in the North American neuroblastoma biology study, ANBL00B1. These cases are predominantly of European ancestry, and may not be representative of other geographic locations and ancestries. Second, while we attempted to control for this in our burden testing, the use of different sequencing methods (WGS, WES, CAP) may affect capture and variant detection at some genetic loci. Third, non-coding variants, epigenetic alterations and an exhaustive set of structural variants were not analyzed here due to the different sequencing methods used. Finally, due to lack of parental DNA, we cannot say whether the P-LP identified in this study are inherited or acquired *de novo*. Additional large studies from diverse populations, and including parental DNA, are needed to replicate and extend our results.

In conclusion, this study of 786 neuroblastoma patients found that 13.9% harbor rare germline pathogenic variants in one or more CPGs. Rare pathogenic variants (SNVs and CNVs) in *BARD1* and other DNA repair

genes were significantly enriched in neuroblastoma compared to cancer-free controls. The presence of one or more germline P-LP variants in a CPG was independently associated with worse overall survival. These data may be used to inform decision-making regarding genetic testing and potential therapeutic options for children diagnosed with neuroblastoma.

Acknowledgments

This work was supported by a supplement to the Children's Oncology Group Chair's grant CA098543 and with federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E to S.J.D and Complete Genomics. This work was also supported by National Institutes of Health grants R01CA204974 (S.J.D.), R01CA237562 (S.J.D.), R03CA230366 (S.J.D.), R35CA220500 (J.M.M), K08CA230223 (K.R.B.), a Howard Hughes Medical Institute Medical Fellows grant (M.P.R.), Alex's Lemonade Stand Foundation (K.R.B.), the EVAN Foundation (K.R.B.), and the Intramural Research Program of the Division of Cancer Epidemiology and Genetics and Center for Cancer Research of the National Cancer Institute, Bethesda, MD. K.R.B. is a Damon Runyon Physician-Scientist supported (in part) by the Damon Runyon Cancer Research Foundation (PST-07-16). This work used resources of the NIH High Performance Computing Biowulf cluster and the Children's Hospital of Philadelphia High Performance Compute cluster.

References

1. Matthay KK, Maris JM, Schleiermacher G, et al. Neuroblastoma. *Nature Reviews Disease Primers*. 2016;2(1):16078.
2. Pugh TJ, Morozova O, Attiyeh EF, et al. The genetic landscape of high-risk neuroblastoma. *Nature genetics*. 2013;45(3):279-284.
3. Sausen M, Leary RJ, Jones S, et al. Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. *Nature genetics*. 2013;45(1):12-17.
4. Friedman DL, Kadan-Lottick NS, Whitton J, et al. Increased risk of cancer among siblings of long-term childhood cancer survivors: a report from the childhood cancer survivor study. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2005;14(8):1922-1927.
5. Mosse YP, Laudenslager M, Longo L, et al. Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature*. 2008;455(7215):930-935.
6. Mosse YP, Laudenslager M, Khazi D, et al. Germline PHOX2B mutation in hereditary neuroblastoma. *American journal of human genetics*. 2004;75(4):727-730.
7. Trochet D, Bourdeaut F, Janoueix-Lerosey I, et al. Germline mutations of the paired-like homeobox 2B (PHOX2B) gene in neuroblastoma. *American journal of human genetics*. 2004;74(4):761-764.
8. Maris JM, Mosse YP, Bradfield JP, et al. Chromosome 6p22 locus associated with clinically aggressive neuroblastoma. *The New England journal of medicine*. 2008;358(24):2585-2593.
9. Capasso M, Devoto M, Hou C, et al. Common variations in BARD1 influence susceptibility to high-risk neuroblastoma. *Nature genetics*. 2009;41(6):718-723.
10. Wang K, Diskin SJ, Zhang H, et al. Integrative genomics identifies LMO1 as a neuroblastoma oncogene. *Nature*. 2011;469(7329):216-220.
11. Nguyen le B, Diskin SJ, Capasso M, et al. Phenotype restricted genome-wide association study using a gene-centric approach identifies three low-risk neuroblastoma susceptibility Loci. *PLoS Genet*. 2011;7(3):e1002026.
12. Diskin SJ, Capasso M, Schnepf RW, et al. Common variation at 6q16 within HACE1 and LIN28B influences susceptibility to neuroblastoma. *Nature genetics*. 2012;44(10):1126-1130.
13. Diskin SJ, Capasso M, Diamond M, et al. Rare variants in TP53 and susceptibility to neuroblastoma. *J Natl Cancer Inst*. 2014;106(4):dju047.
14. McDaniel LD, Conkrite KL, Chang X, et al. Common variants upstream of MLF1 at 3q25 and within CPZ at 4p16 associated with neuroblastoma. *PLoS Genet*. 2017;13(5):e1006787.
15. Hungate EA, Applebaum MA, Skol AD, et al. Evaluation of Genetic Predisposition for MYCN-Amplified Neuroblastoma. *J Natl Cancer Inst*. 2017;109(10).
16. Diskin SJ, Hou C, Glessner JT, et al. Copy number variation at 1q21.1 associated with neuroblastoma. *Nature*. 2009;459(7249):987-991.
17. Bosse KR, Diskin SJ, Cole KA, et al. Common variation at BARD1 results in the expression of an oncogenic isoform that influences neuroblastoma susceptibility and oncogenicity. *Cancer Res*. 2012;72(8):2068-2078.
18. Oldridge DA, Wood AC, Weichert-Leahey N, et al. Genetic predisposition to neuroblastoma mediated by a LMO1 super-enhancer polymorphism. *Nature*. 2015;528(7582):418-421.
19. Schnepf RW, Khurana P, Attiyeh EF, et al. A LIN28B-RAN-AURKA Signaling Network Promotes Neuroblastoma Tumorigenesis. *Cancer Cell*. 2015;28(5):599-609.
20. Russell MR, Penikis A, Oldridge DA, et al. CASC15-S Is a Tumor Suppressor lncRNA at the 6p22 Neuroblastoma Susceptibility Locus. *Cancer Res*. 2015;75(15):3155-3166.
21. Egolf LE, Vaksman Z, Lopez G, et al. Germline 16p11.2 Microdeletion Predisposes to Neuroblastoma. *American journal of human genetics*. 2019;105(3):658-668.
22. Zhang J, Walsh MF, Wu G, et al. Germline Mutations in Predisposition Genes in Pediatric Cancer. *The New England journal of medicine*. 2015;373(24):2336-2346.

23. Parsons DW, Roy A, Yang Y, et al. Diagnostic Yield of Clinical Tumor and Germline Whole-Exome Sequencing for Children With Solid Tumors. *JAMA Oncol.* 2016.
24. Grobner SN, Worst BC, Weischenfeldt J, et al. The landscape of genomic alterations across childhood cancers. *Nature.* 2018;555(7696):321-327.
25. Wang Z, Wilson CL, Easton J, et al. Genetic Risk for Subsequent Neoplasms Among Long-Term Survivors of Childhood Cancer. *J Clin Oncol.* 2018;36(20):2078-2087.
26. Fransson S, Martinez-Monleon A, Johansson M, et al. Whole-genome sequencing of recurrent neuroblastoma reveals somatic mutations that affect key players in cancer progression and telomere maintenance. *Sci Rep.* 2020;10(1):22432.
27. Akhavanfard S, Padmanabhan R, Yehia L, Cheng F, Eng C. Comprehensive germline genomic profiles of children, adolescents and young adults with solid tumors. *Nat Commun.* 2020;11(1):2206.
28. Fiala EM, Jayakumaran G, Mauguen A, et al. Prospective pan-cancer germline testing using MSK-IMPACT informs clinical translation in 751 patients with pediatric solid tumors. *Nat Cancer.* 2021;2:357-365.
29. Kim J, Gianferante M, Karyadi DM, et al. Frequency of Pathogenic Germline Variants in Cancer-Susceptibility Genes in the Childhood Cancer Survivor Study. *JNCI Cancer Spectrum.* 2021;5(2).
30. Wagener R, Taebner J, Walter C, et al. Comprehensive germline-genomic and clinical profiling in 160 unselected children and adolescents with cancer. *Eur J Hum Genet.* 2021;29(8):1301-1311.
31. Wong M, Mayoh C, Lau LMS, et al. Whole genome, transcriptome and methylome profiling enhances actionable target discovery in high-risk pediatric cancer. *Nat Med.* 2020;26(11):1742-1753.
32. Oberg JA, Glade Bender JL, Sulis ML, et al. Implementation of next generation sequencing into pediatric hematology-oncology practice: moving beyond actionable alterations. *Genome Med.* 2016;8(1):133.
33. Brady SW, Liu Y, Ma X, et al. Pan-neuroblastoma analysis reveals age- and signature-associated driver alterations. *Nat Commun.* 2020;11(1):5183.
34. Park J, Levin MG, Haggerty CM, et al. A genome-first approach to aggregating rare genetic variants in LMNA for association with electronic health record phenotypes. *Genet Med.* 2020;22(1):102-111.
35. Wang L, Desai H, Verma SS, et al. Performance of polygenic risk scores for cancer prediction in a racially diverse academic biobank. *Genet Med.* 2022;24(3):601-609.
36. Brohl AS, Patidar R, Turner CE, et al. Frequent inactivating germline mutations in DNA repair genes in patients with Ewing sarcoma. *Genet Med.* 2017;19(8):955-958.
37. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin).* 2012;6(2):80-92.
38. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164.
39. Bresler SC, Weiser DA, Huwe PJ, et al. ALK mutations confer differential oncogenic activation and sensitivity to ALK inhibition therapy in neuroblastoma. *Cancer Cell.* 2014;26(5):682-694.
40. Mody RJ, Wu YM, Lonigro RJ, et al. Integrative Clinical Sequencing in the Management of Refractory or Relapsed Cancer in Youth. *JAMA.* 2015;314(9):913-925.
41. Lasorsa VA, Formicola D, Pignataro P, et al. Exome and deep sequencing of clinically aggressive neuroblastoma reveal somatic mutations that affect key pathways involved in cancer progression. *Oncotarget.* 2016;7(16):21840-21852.
42. Tarsounas M, Sung P. The antitumorigenic roles of BRCA1-BARD1 in DNA repair and replication. *Nat Rev Mol Cell Biol.* 2020;21(5):284-299.
43. Knijnenburg TA, Wang L, Zimmermann MT, et al. Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. *Cell Rep.* 2018;23(1):239-254 e236.
44. Collins RL, Brand H, Karczewski KJ, et al. A structural variation reference for medical and population genetics. *Nature.* 2020;581(7809):444-451.
45. Ritenour LE, Randall MP, Bosse KR, Diskin SJ. Genetic susceptibility to neuroblastoma: current knowledge and future directions. *Cell Tissue Res.* 2018;372(2):287-307.
46. Capasso M, Montella A, Tirelli M, Maiorino T, Cantalupo S, Iolascon A. Genetic Predisposition to Solid Pediatric Cancers. *Front Oncol.* 2020;10:590033.

47. Kratz CP, Smirnov D, Autry R, et al. Heterozygous BRCA1/2 and Mismatch Repair Gene Pathogenic Variants in Children and Adolescents with Cancer. *J Natl Cancer Inst.* 2022.
48. Randall MP, Egolf LE, Vaksman Z, et al. BARD1 germline variants induce DNA repair defects in neuroblastoma. *Submitted (Companion Article)*.

Tables

Table 1. Clinical and tumor biological characteristics for 786 neuroblastoma cases.

Characteristic	Neuroblastoma sequencing cohort n (%)	
Age		
< 18 months	242	30.8%
≥ 18 months	544	69.2%
Sex		
Female	339	43.1%
Male	447	56.9%
COG Risk		
Low	103	13.2%
Intermediate	119	15.1%
High	564	71.8%
INSS Stage		
Stage 1	38	4.8%
Stage 2	60	7.7%
Stage 3	93	11.8%
Stage 4	546	69.5%
Stage 4S	49	6.2%
MYCN Status		
Not amplified	552	71.1%
Amplified	224	28.9%
Not available	10	
Histology		
Favorable	211	28.8%
Unfavorable	522	71.2%
Not available	53	
Ploidy		
Hyperdiploid	473	61.7%
Diploid	293	38.3%
Not available	20	

COG: Children's Oncology Group

INSS: International Neuroblastoma Staging System

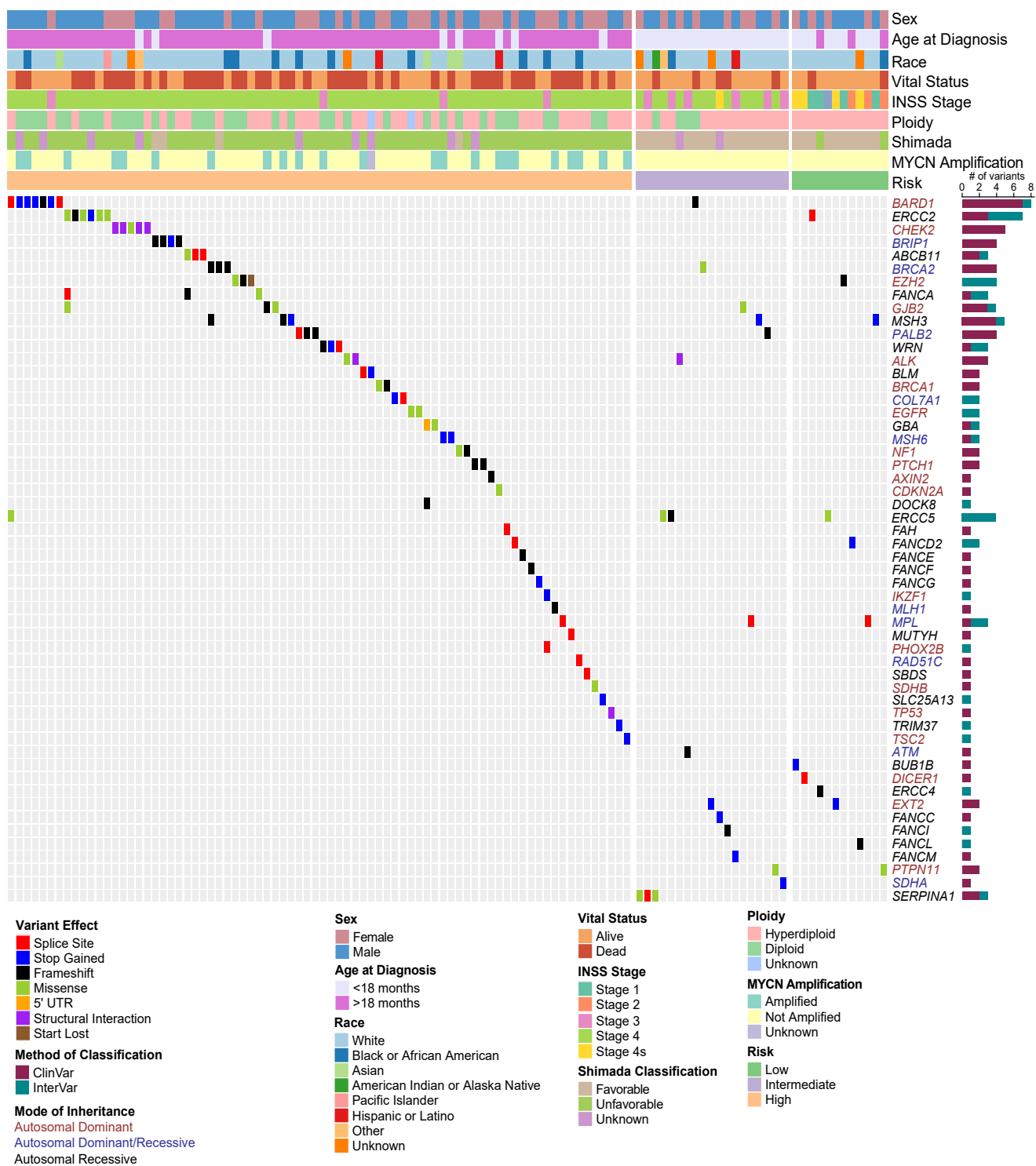


Figure 1. Germline P-LP variants in cancer predisposition genes across 786 neuroblastoma cases.

Oncoprint of known cancer predisposition genes (CPGs) harboring rare germline variants classified as pathogenic (P) or likely pathogenic (LP). Samples without P-LP variants in these genes are not shown. Samples are ordered by COG risk group and annotated with patient clinical and tumor biologic features. Genes are color-coded according to mode of inheritance, when known. Bar chart to the right indicates the number of variants detected for each gene and whether pathogenicity was determined based on ClinVar or our modified InterVar automated assessment. All variants were manually reviewed for quality and evidence of pathogenicity.

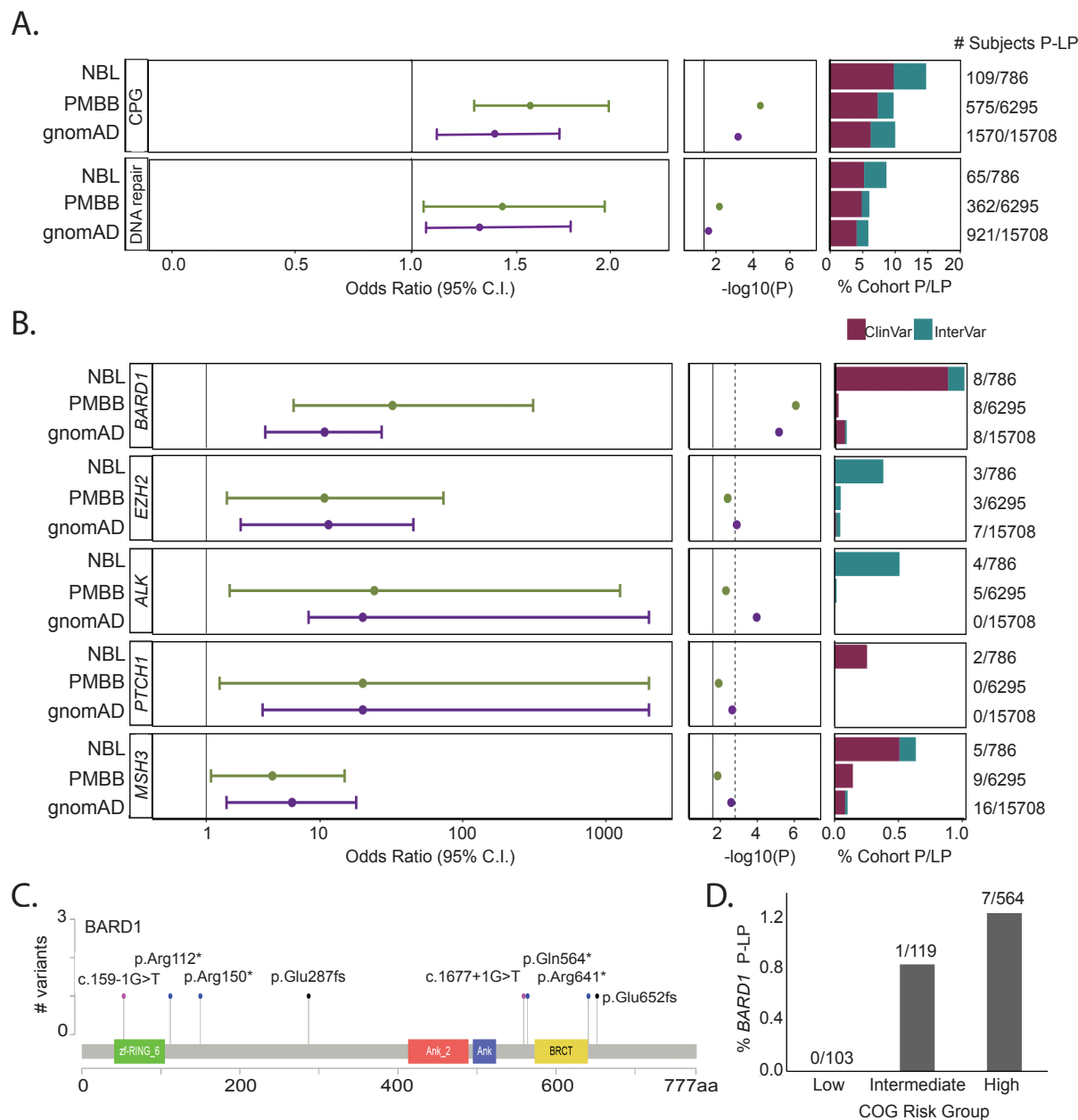


Figure 2. Neuroblastoma cases harbor an excess burden of rare P-LP variants in cancer predisposition genes. **A.** Overall excess burden of P-LP variants (SNVs and indels) in neuroblastoma cases vs. PMBB and gnomAD v2.1 controls is shown for cancer predisposition genes and the subset of genes studied involved in DNA repair. **B.** Gene-based rare variant burden test results comparing the number of neuroblastoma subjects with P-LP variants to those detected in gnomAD v2.1 and PMBB. **C.** Lollipop figure depicting eight germline P-LP variants in BARD1. **D.** Rare P-LP variants in BARD1 are observed predominantly in patients diagnosed with high-risk neuroblastoma.

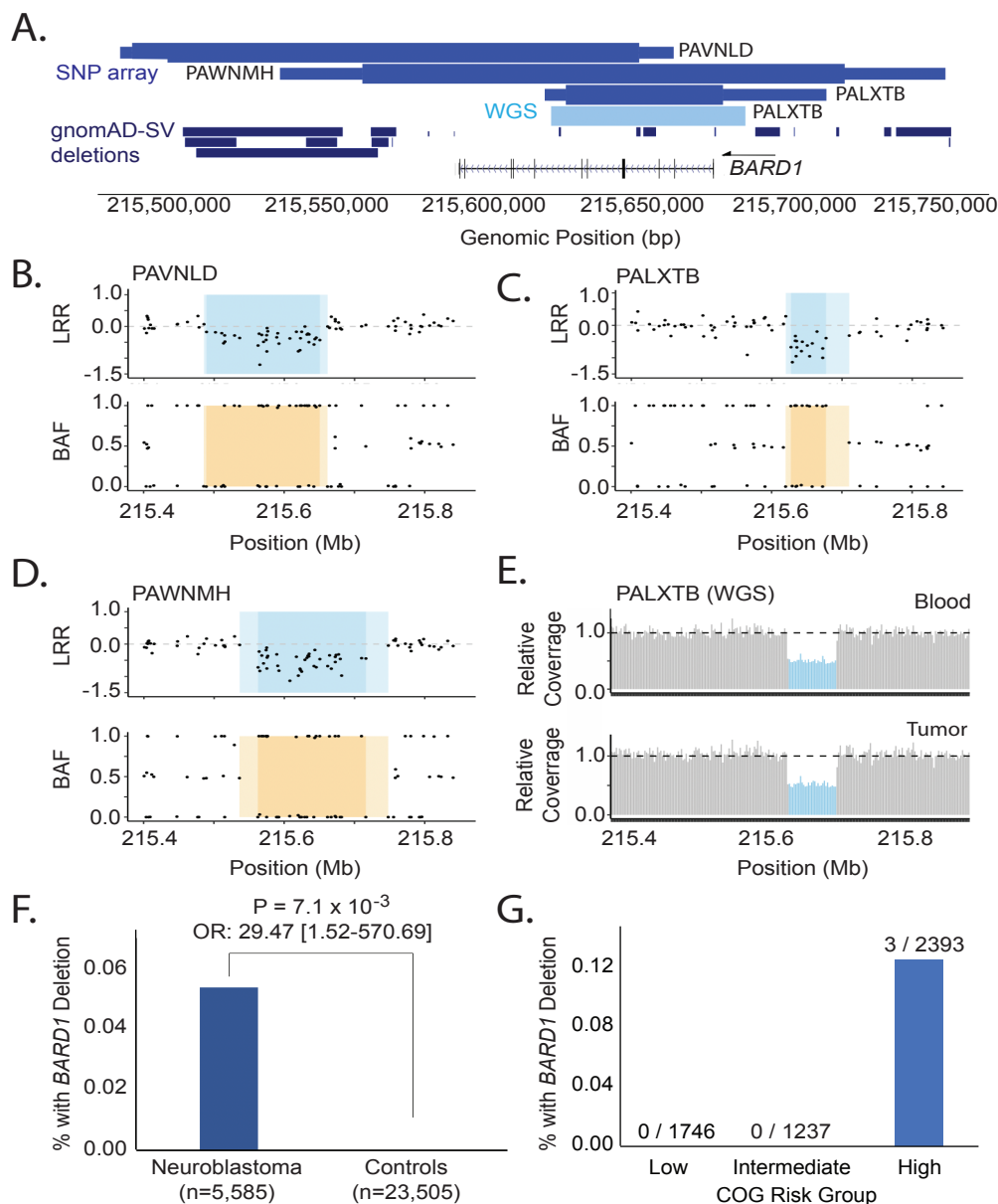


Figure 3. Rare germline CNVs disrupting *BARD1* in neuroblastoma. **A.** *BARD1* deletions were identified in 3 out of 5,585 neuroblastoma patients through copy number analysis of a large germline DNA array dataset (medium blue, top track). No deletions were observed in 23,505 chip-matched controls. The thick and thin bars represent minimum and maximum deletion coordinates, respectively. One deletion was validated and fine-mapped by WGS (light blue, middle track). No *BARD1* protein coding deletions were observed in 10,847 individuals in the gnomAD v2.1 structural variant dataset (dark blue, bottom track). **B-D.** The three array-based CNV calls are shown in log R ratio (LRR) and B allele frequency (BAF) plots. Darker shading indicates the minimum deleted region, whereas lighter shading indicates the maximum region. **E.** WGS validation for patient PALXTB is shown as relative sequencing coverage for matched blood and tumor samples. **F.** Rare *BARD1* deletion CNVs are enriched in neuroblastoma compared to cancer-free controls. **G.** Deletions disrupting *BARD1* were observed exclusively in patients diagnosed with high-risk subset of neuroblastoma.

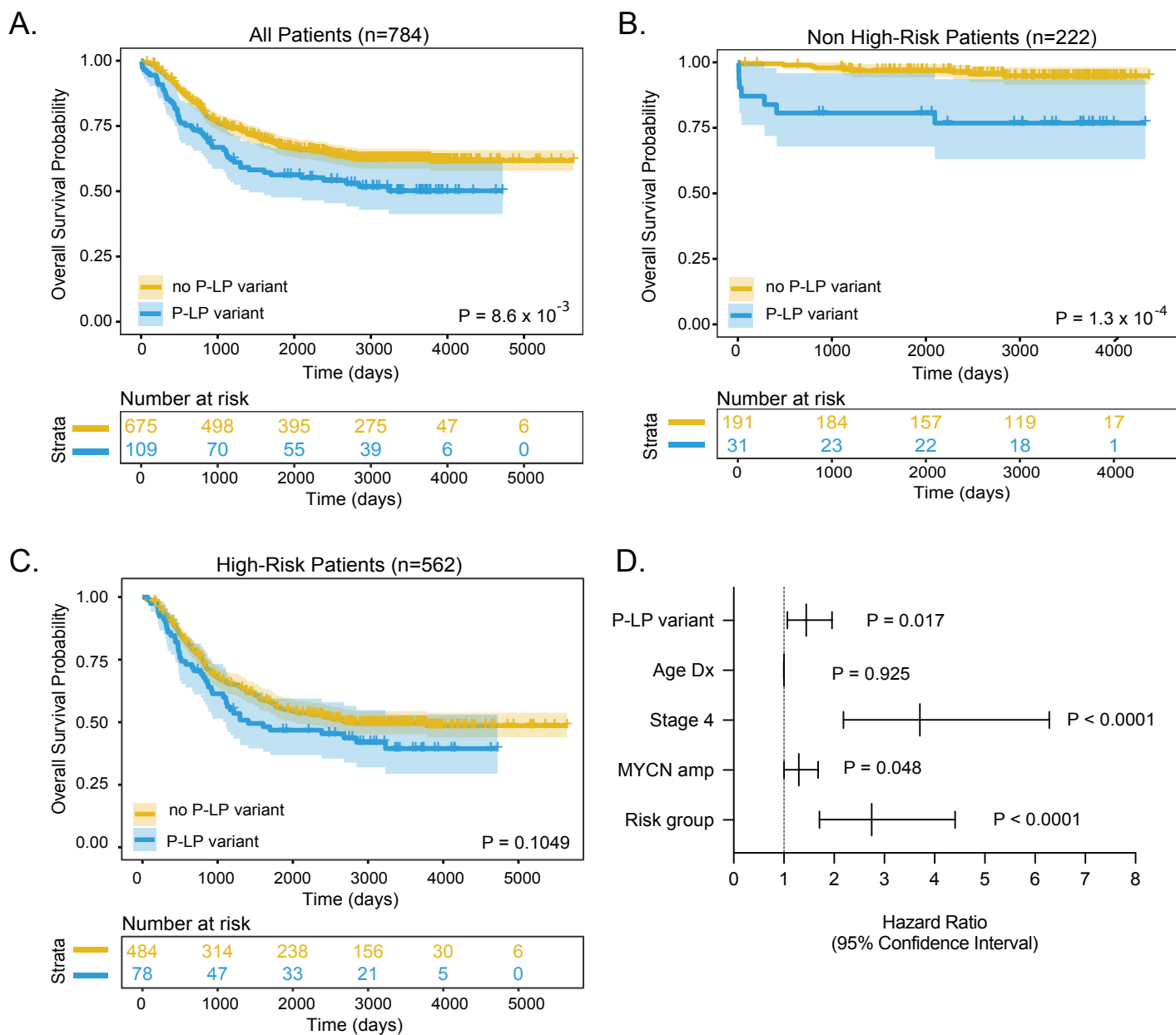


Figure 4. Patient carriers of germline P-LP variants in CPGs have worse outcome. Kaplan-Meier plots of overall survival probability in neuroblastoma patients with and without P/LP variants cancer predisposition genes. **A.** All patients **B.** Restricted to low- and intermediate risk (non-high-risk) groups. **C.** Restricted to high-risk group. Statistical significance in A-C assessed by log-rank test ($P < 0.05$). **D.** Forest plot of hazard ratios from Cox proportional-hazards model.