A genotype-first approach identifies high incidence of *NF1* pathogenic variants with distinct disease associations

A Safonov,[1] TT Nomakuchi,[2] Elizabeth Chao,[3] Carrie Horton,[3] Jill S Dolinsky,[3] Amal Yussuf,[3] Marcy Richardson,[3] Virginia Speare,[3] Zoe Powis,[3] Melissa Pronold,[3] Chia-Ling Gau,[3] Holly Laduca,[3] Shuwei Li,[3] Kyle Allen,[3] Penn Medicine BioBank, Zoe C Bogus,[4] Maria Bonanni,[4] Anna Raper,[4] Staci Kallish,[2,4] Marylyn D Ritchie,[5] Regeneron Genetics Center, Katherine L Nathanson,[4,5,6] Theodore G Drivas[4]

[1] Department of Medicine, University of Pennsylvania, Philadelphia, PA, [2] Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, [3] Department of Clinical Diagnostics, Ambry Genetics, Aliso Viejo, CA, [4] Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, [5] Basser Center for BRCA and Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA

Abstract

Loss of function variants in the NF1 gene cause neurofibromatosis type 1 (NF1), an autosomal dominant genetic disorder classically characterized by complete penetrance, a prevalence of 1 in 3,000, characteristic physical exam findings, and a substantially increased risk for malignancy. However, our understanding of the disorder is entirely based on patients ascertained through phenotype-first approaches. We have recently been referred three patients with incidentally discovered pathogenic NF1 variants, but with no features of the syndrome on exam or history. We hypothesized that the true population-level incidence of NF1 pathogenic variants might be higher than reported, with reduced penetrance or a higher incidence of somatic mosaicism than is currently known.

To investigate this hypothesis, we evaluated two unique large patient cohorts from independent datasets that had undergone comprehensive sequencing of the NF1 gene: the population-level Penn Medicine Biobank (PMBB, n = 43,731) and a database of patients clinically sequenced for cancer risk evaluation by Ambry Genetics (n = 118,768). We identified an unexpectedly high prevalence (1 in 450-750) of pathogenic variants in NF1, more than four times the rate expected given the reported prevalence of NF1. Half of these individuals lacked any evidence of syndromic NF1, and 15-30% of these individuals appeared to be post-zygotic mosaic for the NF1 variant identified. The discovery of an incidental NF1 pathogenic variant did not correlate with the presence of classic symptoms of NF1 but was associated with a significantly greater incidence of certain malignancies compared to a matched control population, including ovarian cancer (p=0.01), sarcoma (p=0.04), adrenal cancers (p=1.5e-11), CNS cancers (p=0.04), and hematologic malignancies (p=3.8e-04). Our findings suggest that NF1 pathogenic variants are substantially more common than previously thought, often characterized by somatic mosaicism and reduced penetrance, and are important contributors to cancer risk in the general population.

Our experience with NF1 led us to examine the incidence of somatic mosaicism on a larger, population-level scale. Within PMBB, there is clear evidence that nearly all individuals harbor multiple somatic-mosaic variants in various genes, with certain genes being significantly enriched for somatic mosaic variants, at least in peripheral blood. This identification of widespread mosaicism has major implications for future genetic testing and biobanking efforts, and the further investigation of this finding will be critical for accurate counseling of patients and families.

Patients with *NF1* Pathogenic Variants Lacking NF1 Features

We have evaluated four patients in our adult medical genetics practice with *NF1* pathogenic/likely pathogenic variants identified on genetic testing, but no or minimal features of NF1 on comprehensive physical exam. None met diagnostic criteria for NF1. These cases are summarized in the table on the right.

Patient	Case 1	Case 2	Case 3	Case 4
Age (years)	65	71	53	54
Sex	Male	Male	Female	Female
NF1 Variant Identified	c.4158del (p.E1387Kfs*19)	c.1466A>G (p.Tyr489Cys)	c.889-1G>T	c.2410-16A>G
NF1 Variant Classification	Pathogenic	Pathogenic	Likely Pathogenic	Pathogenic
NF1 Variant VAF	48-49%	20-30%	Unknown	31%
Indication for NF1 testing	CLL	Hepatocellular carcinoma	Breast cancer	Subcutaneous nodules
Significant PMH	CLL s/p chemotherapy	HCV-associated cirrhosis & HCC, s/p liver transplant	Bilateral DCIS, left LCIS	Steatocystoma multiplex, cataracts
Cafe-au-lait macules	None	None	1 classical, 3 irregular	None
Freckling	None	1 axillary freckle	2 submammary	None
Lisch nodules	None	None	Unknown	None
Neurofibromas	None	None	3 subdermal nodules	None
Offspring	None	None	None	3 unaffected

NF1 Pathogenic Variants are found in 1 in 750 individuals in PMBB and in 1 in 450 individuals in the Ambry Data Set
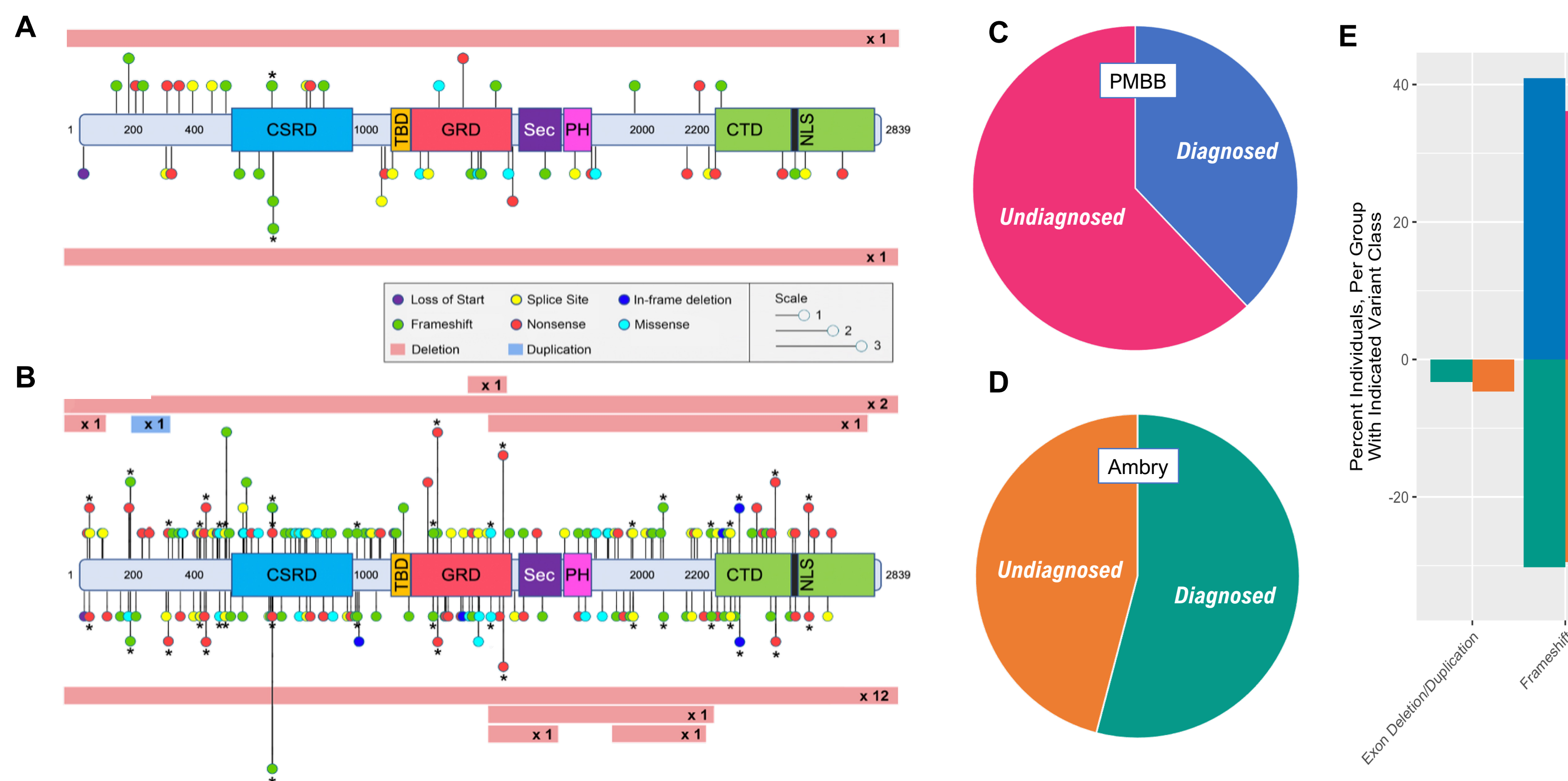

Given our experience in the above cases, we took a genotype-first approach to identify all individuals with an *NF1* pathogenic variant (PV) in the population-level Penn Medicine Biobank (PMBB) of 43,731 patients, and in the dataset of 118,769 patients who had undergone genetic sequencing with a gene panel containing the NF1 gene at Ambry genetics. 1 in 750 individuals in PMBB and 1 in 450 individuals in the Ambry dataset were found to have an NF1 pathogenic variant, much higher than the 1 in 3,000 incidence of the NF1 syndrome. The *NF1* variants identified in (A) the PMBB dataset and (B) the Ambry dataset are displayed along a schematic of the NF1 protein.

On medical record review, it became clear that half of all patients with an NF1 pathogenic variant identified on genetic testing lacked any evidence of an NF1 diagnosis. In PMBB (C) only 39.7% of NF1 PV carriers had a diagnosis of NF1, and in Ambry (D) only 54% had an NF1 diagnosis. We divided our cohorts into the groups with a known NF1 diagnosis (the Clinical-NF1 group) and into the group without a known NF1 diagnosis (the OV-Only group).

Most *NF1* PVs identified were frameshift, nonsense, or splice-site variants (E), consistent with what is known about the gene. Comparison of different predicted protein effects for the different NF1 PVs identified in both Ambry and PMBB between the Clinical-NF1 and PV-Only groups found no differences, with the exception that the PV-Only group was significantly enriched for whole gene deletions in the Ambry cohort. This finding did not replicate in PMBB.

## Evidence for Somatic Mosaicism in Many *NF1* Variant Carriers
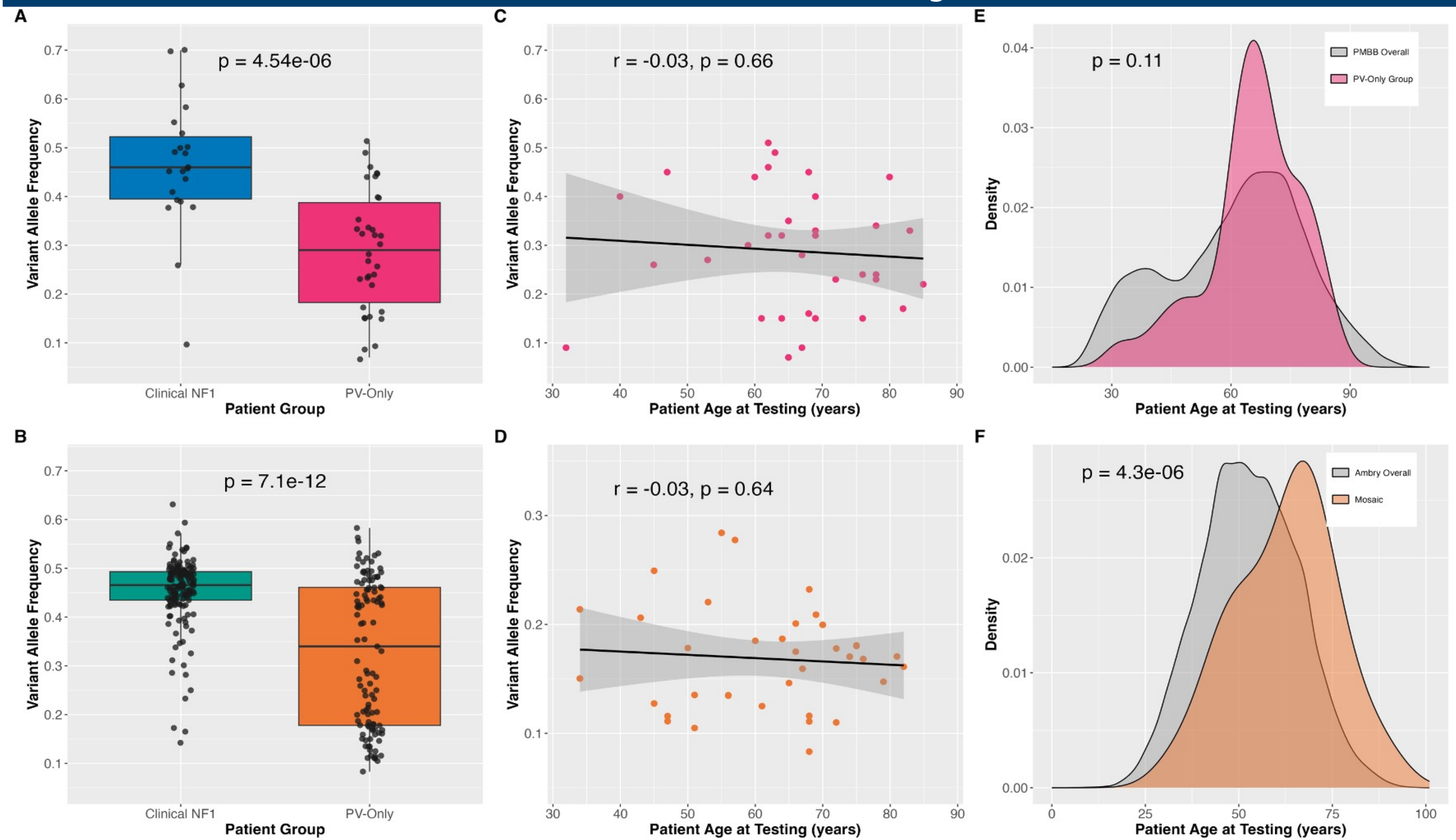
We next asked if there might be evidence of somatic mosaicism of the variant in the PV-Only group. The variant allele fraction (VAF) for each NF1 PV identified in individuals in the Clinical-NF1 group and PV-Only group are displayed for PMBB (A) and Ambry (B). The PV-Only group had a significantly lower mean VAF than the Clinical-NF1 group, suggesting somatic mosaicism for the *NF1* PV identified. We found no correlation between patient age and *NF1* PV VAF in either group (C-D), and PV-Only patients in PMBB were not significantly older than the overall PMBB patient population (E), although in Ambry the mosaic *NF1* PV-Only individuals were observed to be significantly older than the overall testing cohort (F). Together this argues against an age-related phenomenon as clonal hematopoesis (CH) as the only driving force the the somatic mosaicism identified.

NF1 PheWAS Identifies Associations Only in Clinical-NF1 Group

Leveraging the deep phenotypic data available for PMBB participants, we completed a Phenome-Wide Association Study (PheWAS) across 9,030 ICD-10 code-based phenotypes to discover, in an unbiased way, patient phenotypes significantly associated with the presence of an NF1 PV and identified 53 significant associations (A). The most statistically significant associations were for Neurofibromatosis, unspecified and Neurofibromatosis, type 1. The remaining 51 significant associations, all known features of syndromic NF1, were for benign/malignant neoplasms, leukemia, stress fracture, and scoliosis.
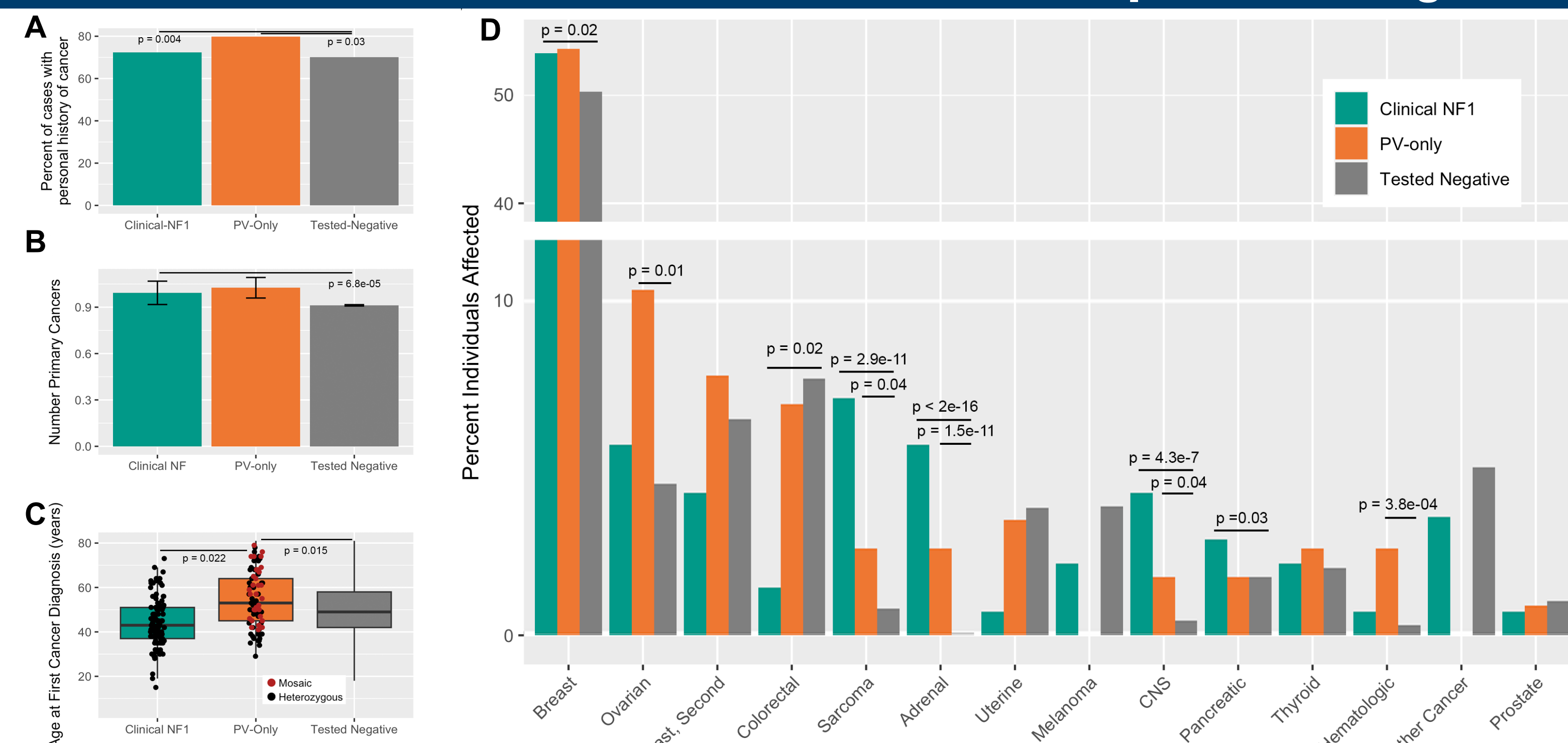
We repeated the PheWAS but excluded either the 35 PV-Only individuals (B, top panel) or the 23 Clinical-NF1 individuals from analysis (B, bottom panel). PheWAS results considering only the Clinical-NF1 individuals identified 43 statistically significant phenotypic associations, of which, 39 (89%) had also been identified in our initial analysis of all 58 NF1 PV carriers. On the other hand, PheWAS results considering only the PV-Only group, identified no significant disease associations. With the caveat that this sub-analysis is relatively underpowered, these results suggests that the presence of an incidentally discovered NF1 PV in blood confers little risk for phenotypes classically associated with syndromic NF1.
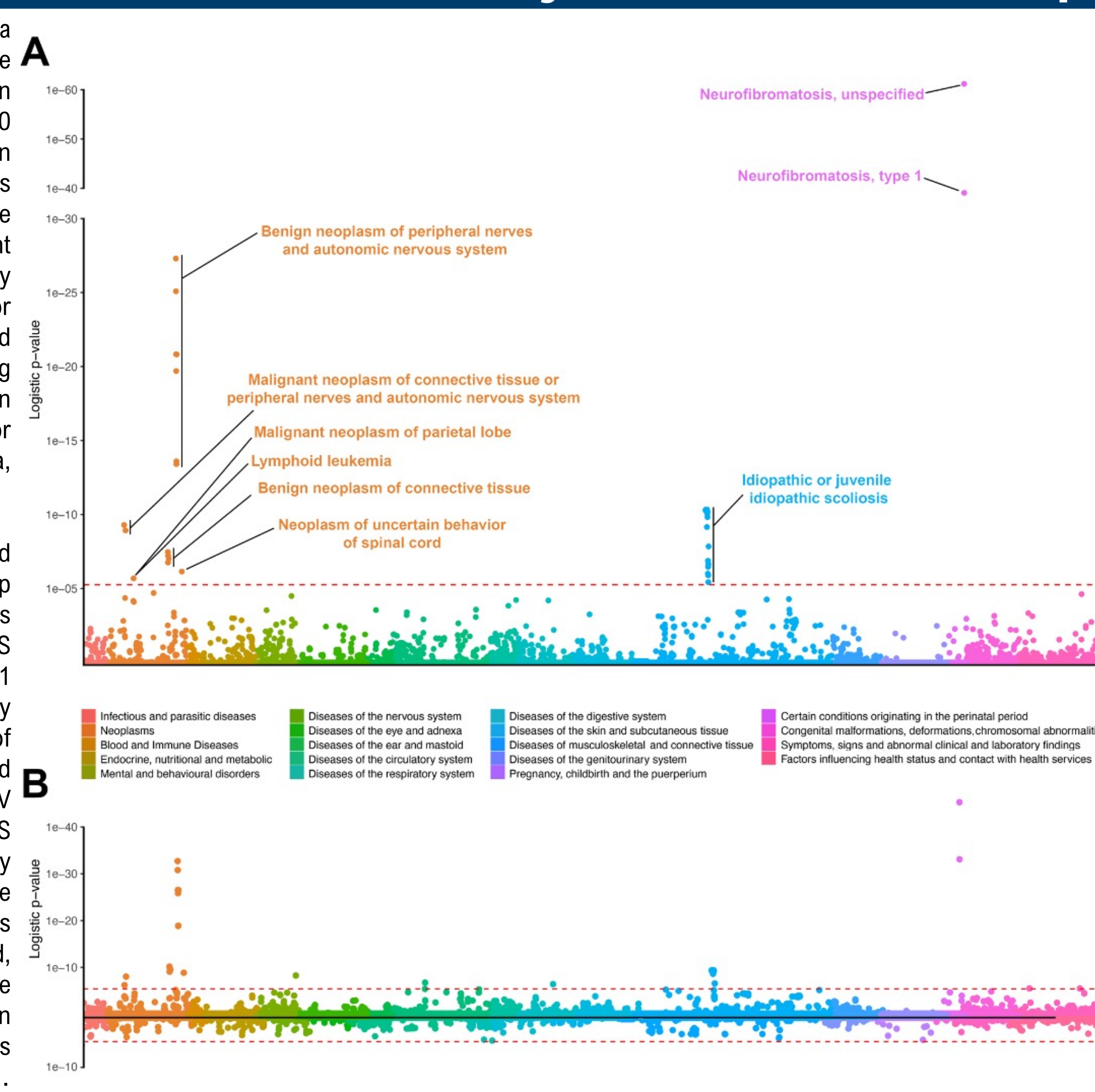
Incidental *NF1* Pathogenic Variants Are Associated with Increased Incidence of Specific Malignancies

We defined a control group, the Tested-Negative group, to include all 31,598 patients who had completed genetic testing at Ambry with gene panels containing the NF1 gene, but whose genetic testing revealed no pathogenic or likely pathogenic variants in any cancer predisposition gene. 110 individuals (72.4%) in the Clinical-NF1 group, 103 (79.8%) in the PV-Only group, and 21,659 (70.2%) in the Tested-Negative group had a personal history of cancer (A). Adjusting for patient age, both the Clinical-NF1 and PV-Only groups were significantly more likely to have a personal history of cancer than the Tested-Negative group. Individuals in the Clinical-NF1 group also were found to have a significantly greater number of primary cancers compared to the Tested-Negative group (B) whereas no difference was seen in number of primary malignancies between the Tested Negative and PV-Only groups. Additionally, Individuals in the PV-Only group were significantly older (mean 54.2 years) than both the Clinical-NF1 (mean 44.0 years) and Tested-Negative groups (mean 49.8 years) at the time of first cancer diagnosis (C).

Dividing cancer diagnoses by type and adjusting for patient age (D), significant differences were seen between the three groups. Compared to the Tested-Negative group, the Clinical-NF1 group was significantly more likely to be affected by breast cancer, sarcoma, adrenal cancer, central nervous system (CNS) cancers, and pancreatic cancer. Patients in the PV-Only group were significantly more likely to be affected by ovarian cancer, sarcoma, adrenal cancers, CNS cancers, and hematologic malignancies compared to the Tested-Negative group. The increased risk for ovarian and hematologic malignancies, the rates of which were more than double what was observed in the Tested-Negative group, was unique to the PV-Only group, and was not seen in the Clinical-NF1 group. No significant differences were seen in rates of specific malignancies between the Clinical-NF1 and PV-Only groups.

Find our preprint here!