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

Anton Safonov,[1,*]
Tomoki T 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,#]

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 four 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.