

Project 1: Analysis of somatic variants from intra-patient tumors to understand progression of benign tumors to malignancy

true

27 July, 2023

Contents

0.1	Introduction	1
0.2	Methods:	1
0.3	Results:	2

0.1 Introduction

Approximately 50% of people with NF1 develop plexiform neurofibromas (PN), and approximately 10% of patients with PN develop malignant peripheral nerve sheath tumors (MPNSTs) with a low 5-year survival rate of 30-50%. Recent studies suggest that many PNs transform into atypical neurofibromas, which in turn can develop into MPNST. The mechanisms underlying such tumor progression are poorly understood, limiting therapeutic options for both PNs and MPNSTs. While selumetinib (FDA approved) and cabozantinib (in clinical trials) are possible treatment options for a subset of PNs, once progressed to MPNSTs, there are no pharmacologic therapeutic options available to the patients beyond traditional chemotherapies. Identifying genetic indicators of PN to MPNST progression may help improve tumor treatment and identify new therapeutic avenues.

Hypothesis:

The JHU Biobank has generated samples and standardized sequencing data from a unique set of patients with both benign and malignant forms of NF1 tumors. The data from these patient triads include samples from normal blood, benign tumor, and malignant tumor from the same patient (n=8 patients, with more expected in next batches). **We hypothesize that genetic changes identified in these patient triads can inform our understanding of the cellular pathways involved in NF1 tumor progression and help identify candidate drugs that specifically target progressing tumors.**

Specific Aim 1: Identify intra-patient and inter-patient genetic changes that correlate with progression from benign to malignant NF1 tumors

Experiment 1.1: Analysis of somatic single nucleotide variants (SNVs) in whole exome sequencing data from all NF1 patients to understand inter-patient landscape of genomic variants

Experiment 1.2: Analysis of somatic single nucleotide variants (SNVs) in whole exome sequencing data from patients with known benign to malignant tumor progression

0.2 Methods:

0.2.1 Somatic Variant Calling

Raw fastq data files were quality checked using FastQC v0.11.9 and a report was generated using MultiQC v1.8. Fastq files were aligned to GRCh38 using BWA 0.7.17-r1188. Duplicates were marked using GATK MarkDuplicates, and bases recalibrated using GATK BaseRecalibrator and GATK ApplyBQSR (GATK

v4.3.0.0). Somatic single nucleotide variants were then called using Strelka2 software (Strelka v2.9.10). Strelka 2 was shown to have high precision and recall for single nucleotide variants compared to others as tested in the precisionFDA challenge (Ref). The variants were annotated using Variant Effect Predictor (VEP v99.2) and converted to MAF files using vcf2maf (vcf2maf v1.6.21). All of these steps were completed on Nextflow Tower running the standardized nf-core pipeline sarek v3.1.2.

0.3 Results:

0.3.1 Sample Summary

The table below shows the count of all different tumor samples present in batches 1, 2, and 3 of the JHU NF1 Biospecimen Repository.

as.factor(tumorType)	n
normal	72
Plexiform Neurofibroma	63
Malignant Peripheral Nerve Sheath Tumor	40
Cutaneous Neurofibroma	12
Diffuse Infiltrating Neurofibroma	9
Atypical Neurofibroma	4
Nodular Neurofibroma	1

0.3.2 Elucidating the somatic variant landscape of the different tumor types among the biobank samples:

First we examine SNVs detected in all samples available from the repository. Figure 1 shows that all variants that were called with high confidence and passed the common variant filter of Strelka2 were found to be single nucleotide variants. They include missense, nonsense, splice-site, nonstop, and translation start-site variants. The NF1 gene features as one of the top mutated genes in the samples provided.

Through various experimental approaches various genes have been identified to be of interest to the NF1 research community. We consulted six seminal articles (Cortes-Ciriano et al, 2023, Pollard et al, 2020, Zhang et al, 2014, Lee et al, 2014, De Raedt et al, 2014, Sohier et al, 2017) to generate a list of genes that are deemed to be of interest to the community. Figure 2 below shows an oncplot of all somatic variants identified in this list of genes of interest.

Once again, the NF1 gene floats to the top as having variants in most number of samples. Given that most of the samples belong to Plexiform Neurofibroma (PNF) and MPNST tumor types, and the plexiform neurofibromas are often likely to undergo malignant transformation to MPNST, we dive deeper into these two tumor types in the next few sections.

0.3.3 Somatic variants in PNF and MPNST samples:

Figure 3 shows a side-by-side comparison of the variants in genes of interest in these two tumor types.

We note that 20% of PNF samples and 12% of MPNST samples show the presence of somatic SNVs in the NF1 gene. There may be two reasons for this: a) Samples may contain microdeletions or copy number variations in NF1 gene which would not be detected in this analysis, b) Samples may have lower tumor purity or depth of sequencing resulting in low detection range for NF1 variants.

0.3.3.1 Variants in NF1 gene: We further dive in to visualize the main SNVs and their protein consequences in the PNF and MPNST samples. Figure 4 shows a lollipop plot that visualizes the main SNVs

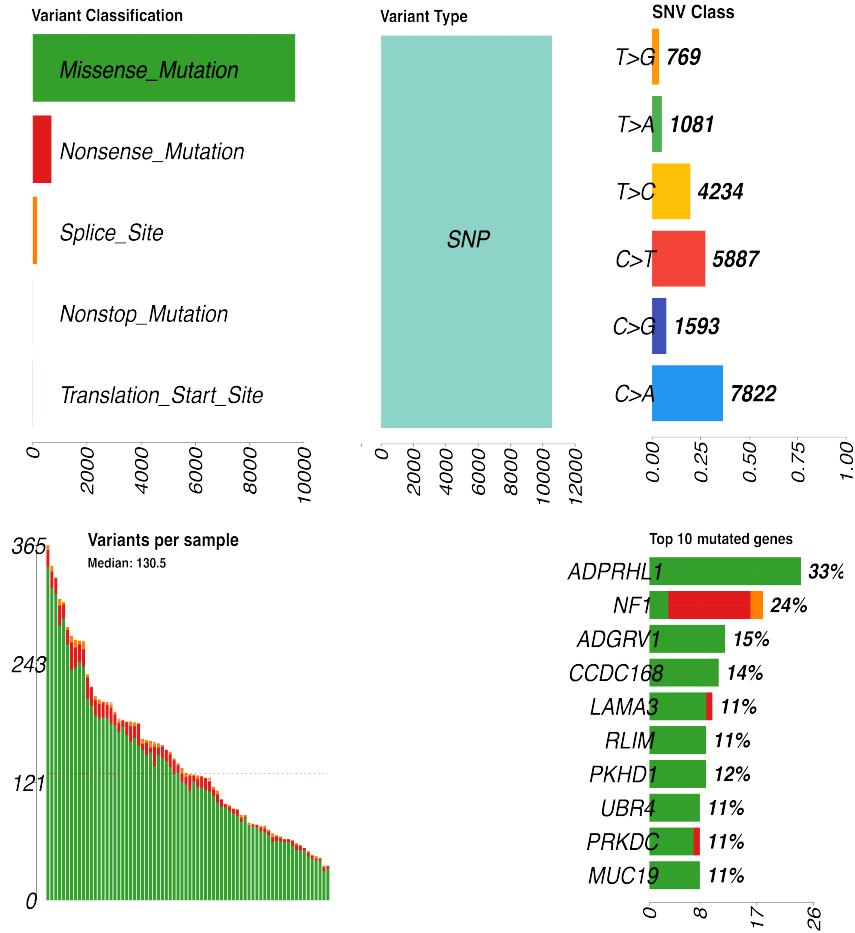


Figure 1: Summary of Single Nucleotide Variants (SNVs) of all samples present in biobank. The plots highlight Variant classification, variant type, SNV classes, Variants per sample, and the top ten mutated genes in the samples provided.

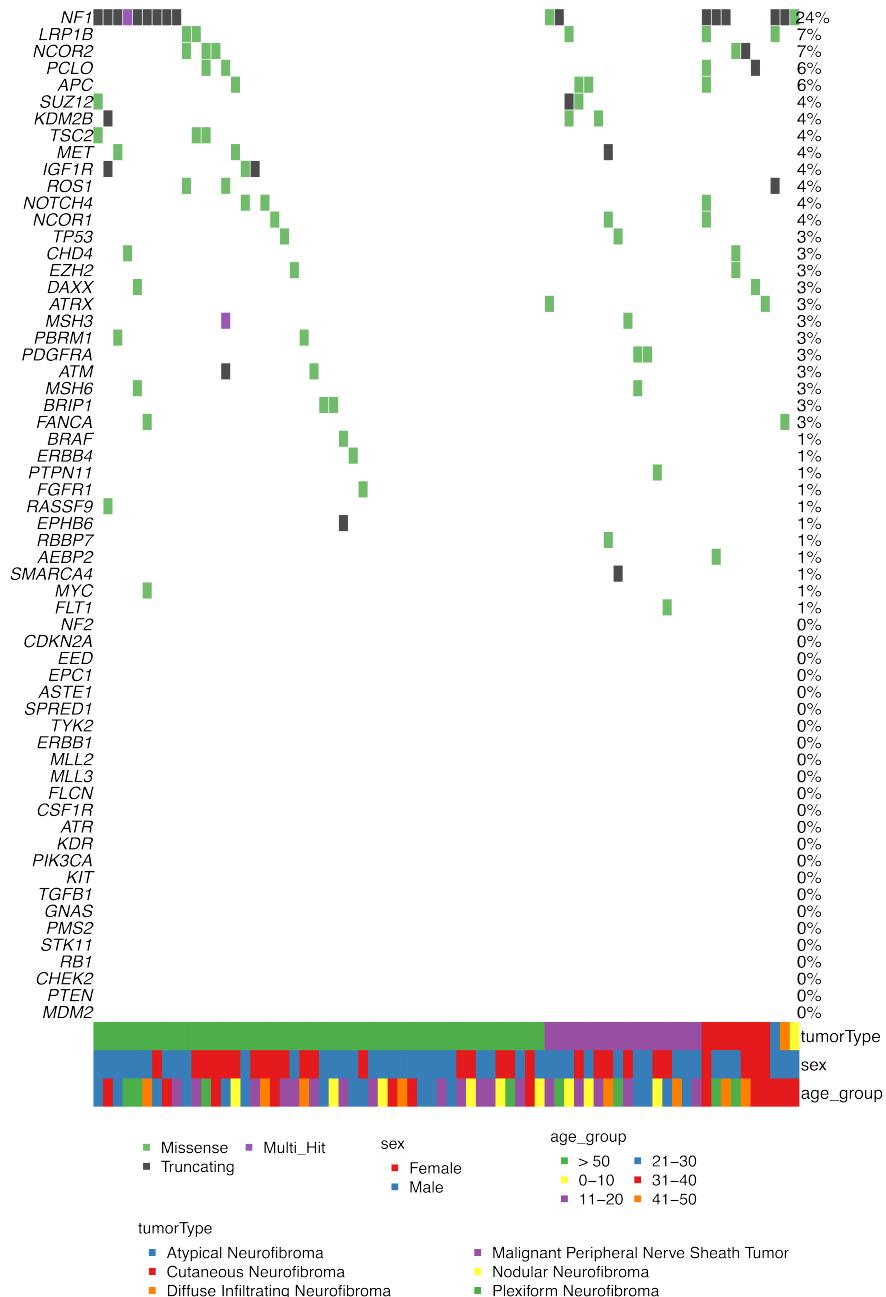


Figure 2: Somatic SNVs in genes of interest for all samples.

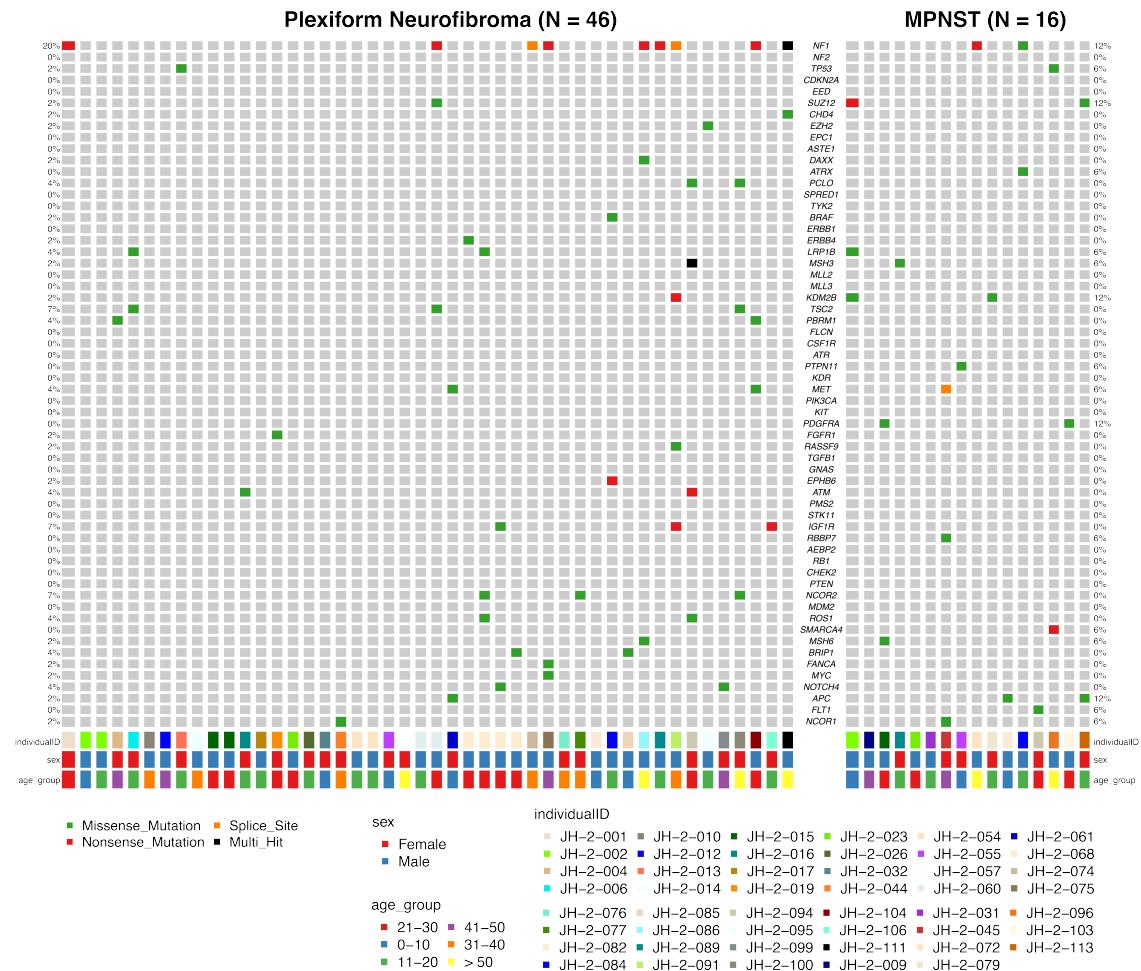


Figure 3: Somatic SNVs in genes of interest specifically in PNF and MPNST samples

and the protein effect of the variant in NF1 gene. The top half of the plot highlights position of variants detected in different plexiform neurofibroma samples. The bottom half highlights variants detected in MPNST samples.

```
##   HGNC    refseq.ID  protein.ID aa.length
## 1:  NF1     NM_000267  NP_000258      2818
## 2:  NF1 NM_001042492 NP_001035957     2839
##   HGNC    refseq.ID  protein.ID aa.length
## 1:  NF1     NM_000267  NP_000258      2818
## 2:  NF1 NM_001042492 NP_001035957     2839
```

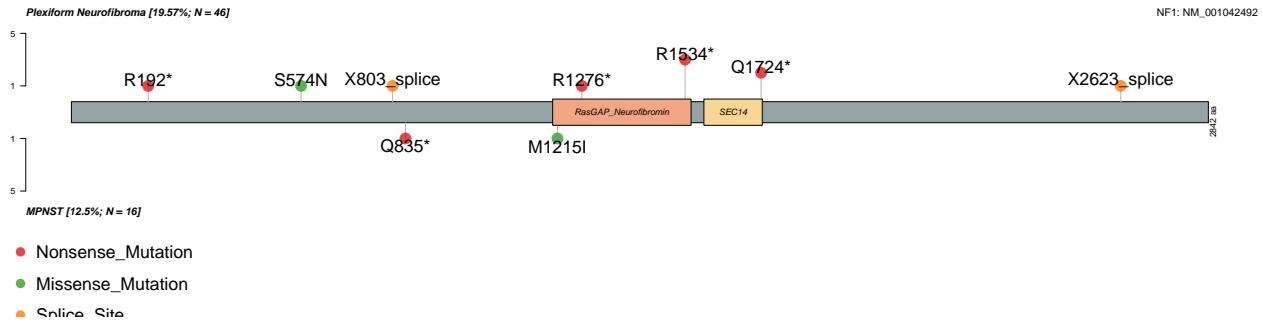


Figure 4: Lollipop plot of NF1 gene

0.3.3.2 Oncogenic pathways in PNF and MPNST samples: The genetic variants in the PNF and MPNST samples may cause disruption of various oncogenic pathways. Figure 5A-B below show that RAS and NOTCH are mainly affected by the variants present in PNF samples. However, in MPNST samples, the genetic variants mostly affect the Notch and Hippo pathways (Figure 5 C-D).

0.3.4 Somatic variants in intra-patient tumors:

The JHU NF1 Biospecimen repository with great effort have collected multiple samples of progressing tumor types from 8 individual patients. These “triad” samples include a trio of normal, benign, and malignant tissue from the **same patient**. The patients with triad samples are: “JH-2-002”, “JH-2-015”, “JH-2-016”, “JH-2-023”, “JH-2-031”, “JH-2-045”, “JH-2-055”, “JH-2-084”.

However, after updates to the sample information all tumor samples related to JH-2-031 were found to be MPNSTs. So JH-2-031 was excluded from the analyses below.

Figure 6 below shows somatic variants in our list of genes of interest detected in the triad samples. Like before we note that many of the samples do not show a pre-existant somatic variant in NF1 gene. The Biobank is currently looking at tumor purity information for these samples to rule out any purity related issues.

0.3.4.1 Oncogenic pathways in benign and malignant TRIAD samples: Figure 7 below shows the effect of the SNVs in benign and malignant tumors. We note here that while the genomic variants mostly affect the Wnt pathway in the benign tumors, the malignant tumors have most samples with variants affecting the Hippo Pathway. Thus the Hippo pathway may be a candidate that plays an important role in the transition from a benign to malignant tumor type in a plexiform neurofibroma patient. As noted in many other studies(Feltri et al, 2018, Brosseau et al, 2020 and others), the Hippo pathway plays a significant role

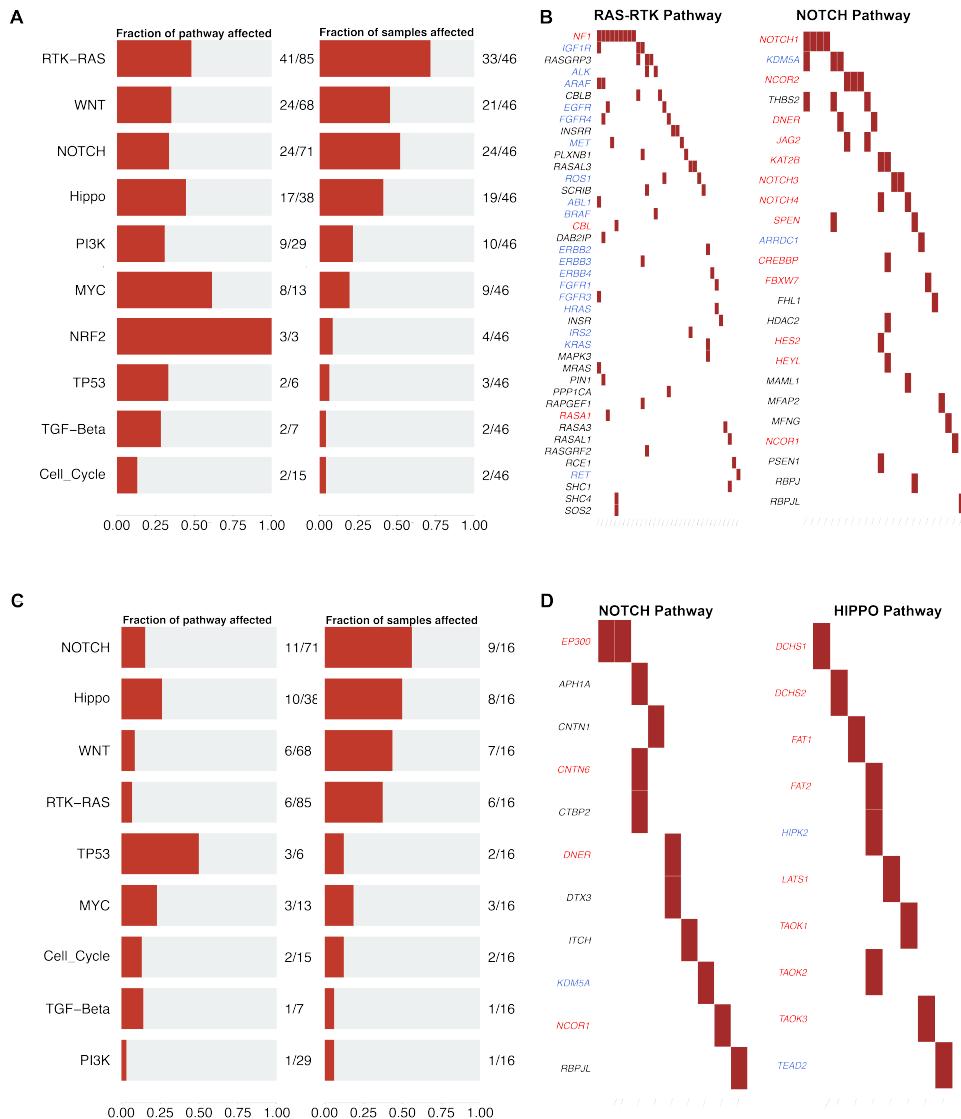


Figure 5: Oncogenic pathways affected in PNF and MPNST samples. Tumor suppressor genes are in red and oncogenes are in blue font.

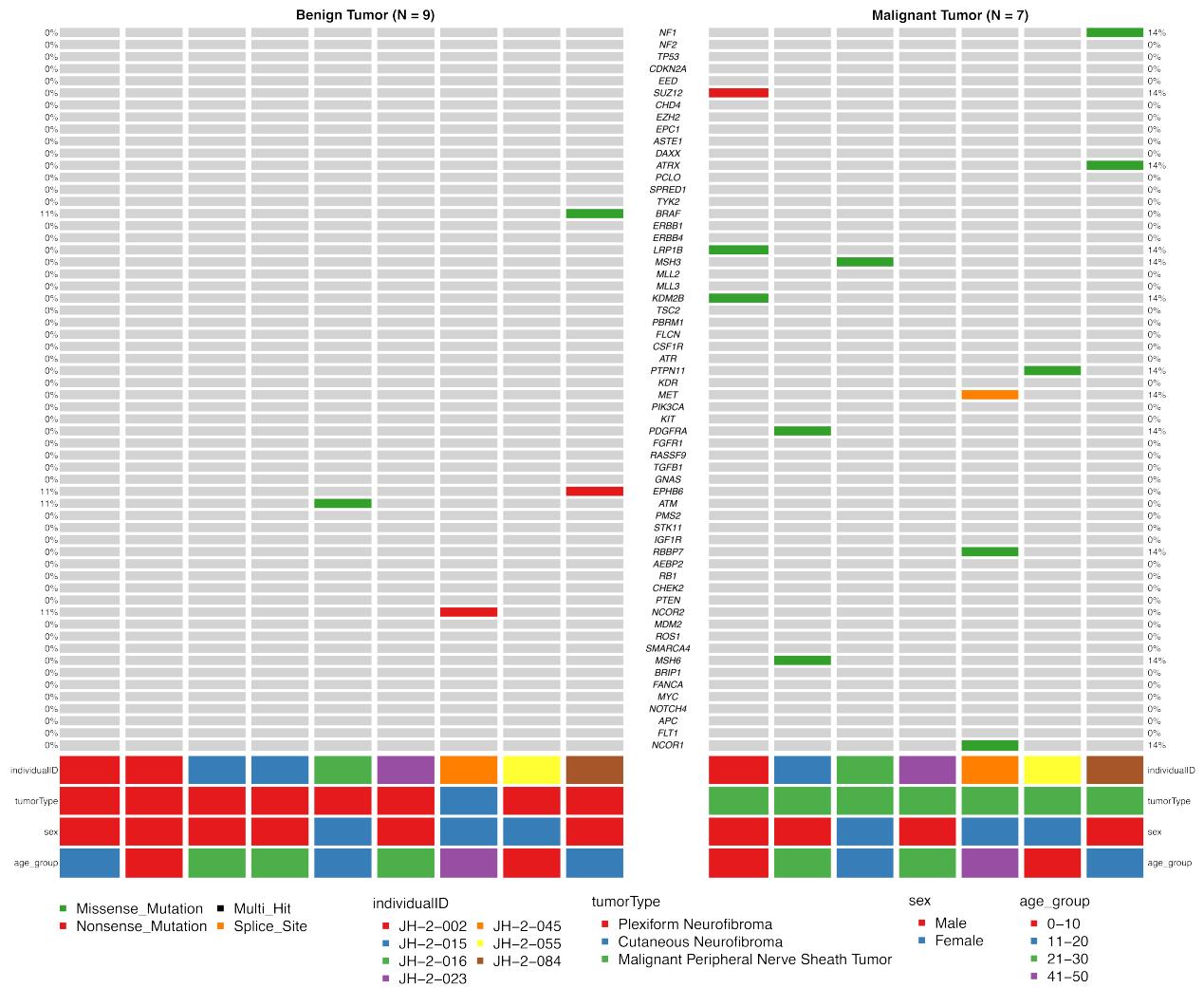


Figure 6: Somatic variants in triad samples.

in sporadic and NF1 associated MPNST tumorigenesis. The study of triad samples suggest that it may play a significant role in malignant transition from plexiform neurofibromas to MPNSTs. Further experiments are required to validate the role of Hippo pathway in the malignant transition from PNF to MPNST.

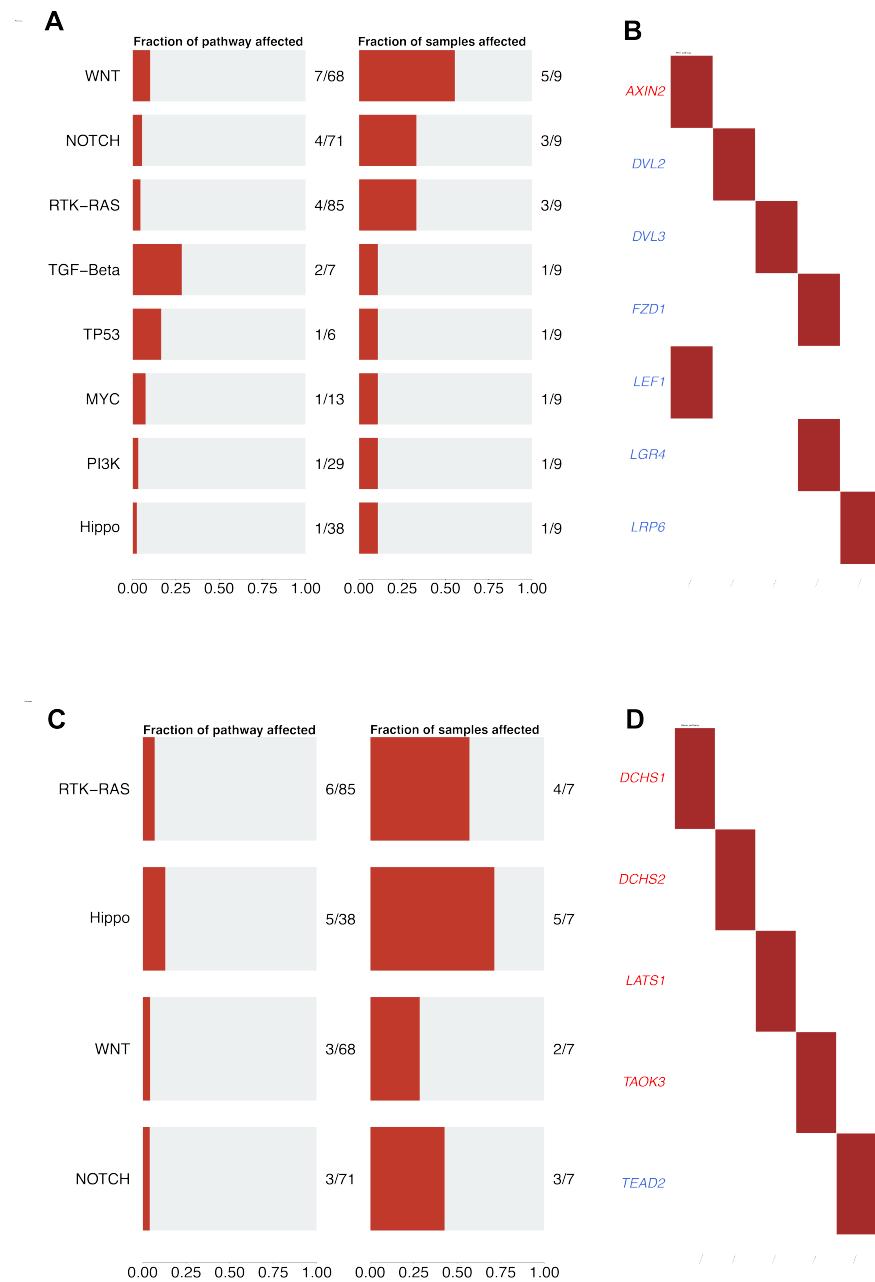


Figure 7: Oncogenic pathways affected in benign and malignant tumors