Supplemental Figure Legends

Supplemental Figure 1. Germline and somatic mutations in the *NF1* gene. (a) Table showing mutation positions compared to the canonical NF1 gene transcript (ENST00000358273) and HS canonical protein product (ENSP00000351015). Reads showing reference sequence or variant alleles and the percent of variant reads were visually verified using IGV. **(b)** The location of predicted germline and somatic mutations in the canonical NF1 protein detected in PNF and DNF. Same-position mutations found in different cohorts are marked.

Supplemental Figure 2. QQ plots of burden test and lollipop plots of some candidate germline mutated genes.

(a) Burden testing was performed using TRAPD and results were then visually represented by quantile-quantile (QQ) plots. Because ExAC samples and the case samples are not jointly processed and variant-called, the level of technical artifacts potentially introduced in this analysis (artifact inflation/deflation) were measured using the λΔ95 metric implemented in TRAPD. We detected small levels of inflation (i.e. above the optimal λΔ95 > 1.00), likely due to the small case sample size and incomplete ancestry information. **(b)** Lollipop plots of COL14A1, CELSR2, CUBN and FCGBP. Each variant's functional impact was predicted by three bioinformatics tools and ClinVar dataset (see Methods and supplemental tables) but only effect predicted by SIFT was shown in the figure.

Supplemental Figure 3. Neurofibroma Schwann cell somatic variants. (a) Summary table of non-synonymous rare somatic variants in PNF and DNF datasets. (i.e. number and percent of patients harboring somatic variants from whole exome sequencing (WES) data from SC versus matched FB; percent of NF1 patients harboring variants in a WGS dataset from dermal neurofibroma versus matched blood, or in a WES dataset from PNF[1] (**Material and Methods**). Notably, compared with low mutation ratio of the designated genes in cancers in COSMIC, these genes show a high percentage of variants in neurofibroma SCs, and also showed variants in the DNF[2] and/or PNF cohorts. Color bars on the right indicate the variant type, based on predicted effects on mRNA. Numbers are the number of tumors harboring variants in each gene in the PNFSC cohort. **(b)** Heat map of mRNA expression of those genes with expression level changes, in either tumor SC vs. normal SC and/or tumor tissue vs. nerve databases. **(C)** Lollipop plots showing the predicted effects of variants by SIFT on proteins. The genomic position of each variant detected in PNF was converted into the amino acid position of the corresponding protein product.

Supplemental Figure 4. Variants in *OBSCN* and *PKHD1L1*. (a) Summary table, Non-synonymous rare variants and the percentages of tumor samples with variants in these genes in PNF and DNF datasets **(b)** Gene variants are shown in 3 groups (I, II, and III); those with a germline variant and a 2nd allele somatic variant (I); germline variant only (II) or somatic variant only (III). FB: fibroblast; SC: Schwann cell. Color bars indicate variant types, based on predicted effects on the mRNA. Numbers at right of color bars are the total number of PNF (of 9) harboring variants in each gene. Individual neurofibromas

harbor different variants in *OBSCN* and *PKHD1L1* (i.e. for *OBSCN*, 1 neurofibroma showed a germline plus a second allele somatic mutations*;* 5 tumors harbored only a germline variant*;* 1 tumor sample harbored only a somatic variant). **(c)** Lollipop plots showing the predicted effects of variants on *OBSCN* and *PKHD1L1* protein function by SIFT. The genomic position of each variant detected in PNF was converted into the amino acid position of the corresponding protein product.

Supplemental Figure 5. Oncoprint analysis of candidate genes.

Oncoprint analysis of top 47 candidate (somatic and germline) genes in 9 PNF tumors generated by cBioPortal.

Supplemental Figure 6. Detailed description of *ATM* variants. Y: yes; N: Not found yet. The number (#) of databases that predicted a variant to be damaging.

Supplemental Figure 7. *ATM* germline variant confirmation, and proliferation in neurofibromas with *ATM* WT or *ATM* MU. (a, b) Sanger sequencing results for 2 different *ATM* germline variants. **(c)** Percent of Ki67 positive cells in neurofibromas with *ATM* MU vs. *ATM* WT (un-paired Student's *t*-test).

Supplemental Figure 8. Guide RNA and donor DNA sequences for variant construction, and shATM effects on SC growth. (**a**) ATM gene and amino acid sequence are highly conserved between human and mouse around the G2023R and S707P positions. In addition to a point mutation change in the donor DNA sequences, to

avoid Cas9 re-cutting we wobbled DNA sequences (green text) that were not predicted to change the amino acid sequence. In the donor DNA, there are ~50nt homology arms flanking the cut site. **(b, d)** sh*ATM* reduced *Atm* mRNA (RT-PCR) and ATM protein (western blot) in immortalized human Schwann cells (iHSC) and primary mouse embryonic SC (eSC). **(c, e)** sh*ATM* does not affect iHSC or eSC cell viability (MTS assays).

Supplemental Figure 9. Histological analysis of neurofibromas in *Nf1 fl/fl; DhhCre* and *Atm+/-; Nf1 fl/fl; DhhCre* mice. Neurofibromas were embedded in paraffin (n=10 tumors from 3 mice per genotype). (**a**) H&E staining, S100 marks SC, Toluidine blue metachromasia (purple) staining marks mast cells. (**b**) Statistical analysis of Toluidine blue+ mast cells/HPF (high power field) indicates an increased number of mast cells in neurofibromas from *Atm+/-; Nf1 fl/fl; DhhCre* mice. (Un-paired t-test, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001)

SUPPLEMENTARY TABLES

Supplementary Table 1. Each sample's depth of coverage.

Supplementary Table 2. *NF1* gene Germline and somatic mutation status in NF1 patients

Supplementary Table 3: Top 22 genes showing germline variants

Supplementary Table 4: Top 25 genes showing somatic variants.

Supplementary Table 5: Germline variants in top 22 genes present in DNF[2] data

Supplementary Table 6: Somatic variants in top 25 genes present in DNF[2] data

Supplementary Table 7: Germline variants in top 22 genes present in PNF[1] data

Supplementary Table 8: Somatic variants in top 25 genes present in PNF[1] data

Supplementary Table 9: Datasets used for variant function prediction

Supplementary Methods

Germline Variant Calling

We considered the same rare and non-synonymous variant detected both in a SC sample and its matched FB sample as a probable germline mutation variant. Our "rare" variant filtering strategy is described below.

Somatic Variant Calling

Variant callers use different approaches to call variants, so we integrated results from methods described in Xu *et al*[3]. and Cai *et al*[4]. We used GATK-naiveSubtract, MuTect1, Strelka, SomaticSniper and VarScan2 to predict somatic point mutations (SNPs). INDELs were predicted using GATK-naiveSubtract, Strelka and VarScan2.

GATK-naiveSubtract The GATK-naiveSubtract method[3] was used to call SNV and INDEL variants present in a SC sample VCF file but not in its matched control (FB) sample VCF file. Variants set for case and control samples were called independently using

GATK-HaplotypeCaller. For stringent filtering, we considered PASS-tagged variants passing GQ > 20 and DP > 20 cutoffs.

MuTect1 (v1.1.7) MuTect1[5] detects SNPs only and is bundled into the GATK pipeline. To run this version of MuTect1, we used GATK v2.8 with java v1.7.0u40. We also used human genome (v37) bundled with GATK bundle 2.8. COSMIC v54 and dbSNP 132 (b37).

VarScan2 (v2.3.9) VarScan2[6] detects SNPs, INDELs, and loss of heterozygosity events (LOH) in NGS data. SAMTOOLS v1.3 was used to make mpileup (-B q 1 –f). To perform somatic filtering, we used bam-readcount (-q 1 -b 20). After variant calling, we applied recommended somatic filters together with bam-readcount (-q 1 –b 20 –f). Java 1.8.0_40 was used to run this version of VarScan2.

Strelka (v1.0.15): Strelka[7] detects SNPs and INDELs. SAMTOOLS v1.3, java v1.7.0u40 were used to run Strelka and a recommended built-in post-calling filter was used to remove possible false positives.

SomaticSniper (v1.0.5): SomaticSniper[8] detects SNPs. SAMTOOLS v1.3 and bcftools v1.3 were used for the SomaticSniper pipeline. We set the somatic quality threshold of SomaticSniper to 15 (the author recommended 15-40). Raw call sets generated by SomaticSniper were filtered by pipelines proposed by the developers.

Variant Filtering Strategy We defined common variants (minor alternative reads ratio [MAF], >1%) in a public data set as lower priority. We adopted dbSNP's GMAF < 1% and ExAC MAF < 1% as MAF-filtering criteria. If both MAF scores were <1%, then we chose ExAC MAF. We also included possible novel variants that do not have defined MAF scores. We focused on non-synonymous variants found in protein coding canonical gene

transcript exon and splicing regions predicted to have HIGH / MODERATE / MODIFIER-impact functional effects. “HIGH” includes variants predicted to have disruptive impact on a protein, e.g. stop gained or frameshift variants causing protein truncation, loss of function or triggering nonsense mediated decay. “MODERATE” is a predicted non-disruptive variant that might change protein structure or function (e.g. missense variant, inframe deletion). “MODIFIER” includes predicted non-coding variants or variants affecting non-coding genes, where prediction is difficult or there is no evidence of impact; these were ignored in our analysis. We also ignored variants labeled “LOW”, alterations assumed to be harmless or unlikely to change protein behavior.

For somatic mutations, we considered HIGH, MODERATE, and MODIFIER-impact variants with min [GMAF < 1%, ExAC MAF < 1%], or novel. For germline mutations, we considered HIGH/MODERATE/MODIFIER-impact variants whose MAF < 1%. For analysis of the *NF1* gene, we considered HIGH- and MODERATE-impact variants satisfying these MAF-filtering criteria. The numbers of rare and non-synonymous HIGH, MODERATE, MODIFIER, and LOW-impact variants were plotted. Final variants were visually examined using an Integrative Genomics Viewer (IGV, https://www.software.broadinstitute.org/software/igv).

Low-Complexity Regions. Based on 1,462,754 low-complexity region data (human v37 whole genome) predicted by mdust[9], we checked +/- 20 nucleotide-long subsequences from called variant loci to determine if they fall into predicted low complexity regions. This does not mean the sequencing qualities of these variants are low. We used this information to prioritize candidates.

Variants effect prediction: Variant Annotation Variants from these 5 methods were annotated using Variant Effect Predictor (VEP) v83[10] using Ensemble human genome (v37) and add-in databases and online algorithms including dbNSFP (v2.9)[11] and ClinVar (release: 20170104, https://www.ncbi.nlm.nih.gov/clinvar/)[12], dbSNP (v144, https://www.ncbi.nlm.nih.gov/projects/SNP/), Exome Aggregation Consortium (ExAC v0.3.1, http://exac.broadinstitute.org/), SIFT (v5.2.2), and HGMG-PUBLIC (v20152, http://www.hgmd.cf.ac.uk/ac/index.php). Annotated VCF files were converted to tab-delimited tables for data exploration using in-house PERL and BASH scripts. Mutation impact was annotated using ClinVar[12], SIFT[13], MetaLR[14] and MetaSVM[14], included in the dbNSFP database (**Suppl. Table 9**).

MTS assay iHSC cells cultured in basic medium (10% Gem-Cell Fetal Bovine Serum in Gibco DMEM, 1% HyClone Penicillin-streptomycin solution) and primary mouse embryonic Schwann cells cultured in basic medium with 2uM forskolin (Calbiochem) and 10ng/ml β-heregulin (R&D system) were plated at 500 cells/well in 96 well plates (Corning, Black). After 3 days of culture, cell titer aqueous one solution for cell proliferation assay (Promega) was added to the wells according to the manufacturer's protocol. The plates then were read at 490nm absorbance 4 hours later. Three independent experiments with 5 replicates/sample in each experiment were performed and used for statistical analysis.

Un-paired t-test. To test if germline variants in the *ATM* gene cause differential effects to the total variants number, we calculated the Δ in number of total variants in SC vs. FB from each individual neurofibroma. We applied un-paired t-tests between the groups of

neurofibromas with and without *ATM* variants. The p-value is derived from calculated t-statistics. We used the t. test function, implemented in Prism.

Sanger sequencing primers for ATM germline variants confirmation:

Chr11: 108198394 C>A

Forward: TGGCAAAAGCAGATGAGGAAAAAC

Reverse: TCACTCCACCCTAGAGACTATACA

Chr11: 108098576 C>G

Forward: CTGCTGCCGTCAACTAGAACAT

Reverse: AAATGCCAAATTCATATGCAAGGC

Chr11: 108123621 T>G

Forward: ATGGTTGTCCTCCTTAAATTGTCCT

Reverse: AAACAACCTCTTCCCTGGCTAA

Full sequence of Donor DNA for G2023R:

AATTACTATCTAGAAAGTGCAGTTTACCTAGTAAGGGGTTGTAACATTTTCCCTCTC

CCGCAACCATACAGGCTGTCCGGCTCTCCTATACTTCTGTAGATCTCTAAGAGAAG

ATCCTGCAAACAGA

Full sequence of Donor DNA for S707P:

AACCCATAAATGCTCAAGAATACACCTGAATTCGTTTTAAGAAATCTCACCTCAGGA

GAATAGTTGCTTAAGAGCTGTTCTGATAATCCCAGAAGATAGTGATCCAATGATTCC

TTGAGATTTTGCT

Sanger sequencing primer for G2023R colony confirmation:

Forward: CTGCGTAAACTTCACTGATACAC

Reverse: TTCAAGTCTCTGCCTTTGGTATT

Sanger sequencing primer for S707P colony confirmation:

Forward: GCACTCCTTCCCACTAACCTA

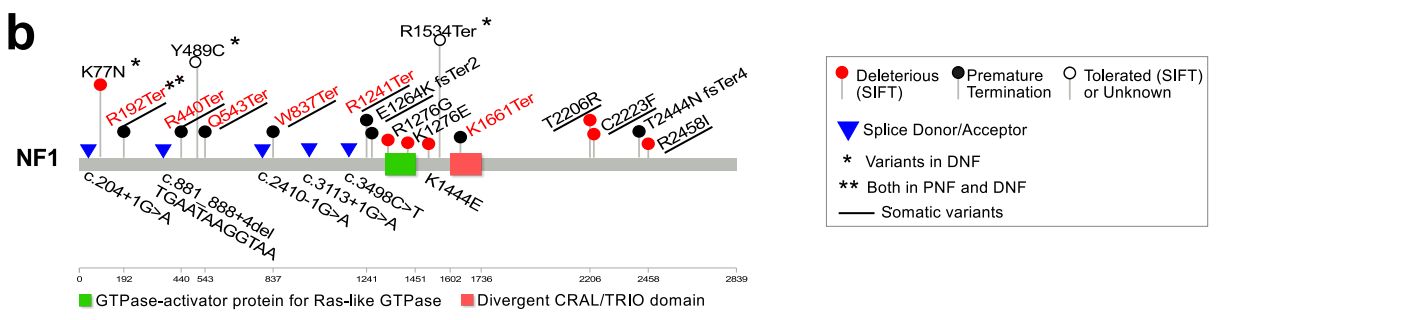
Reverse: AAGTGAACAACACTGCGAAGATAA

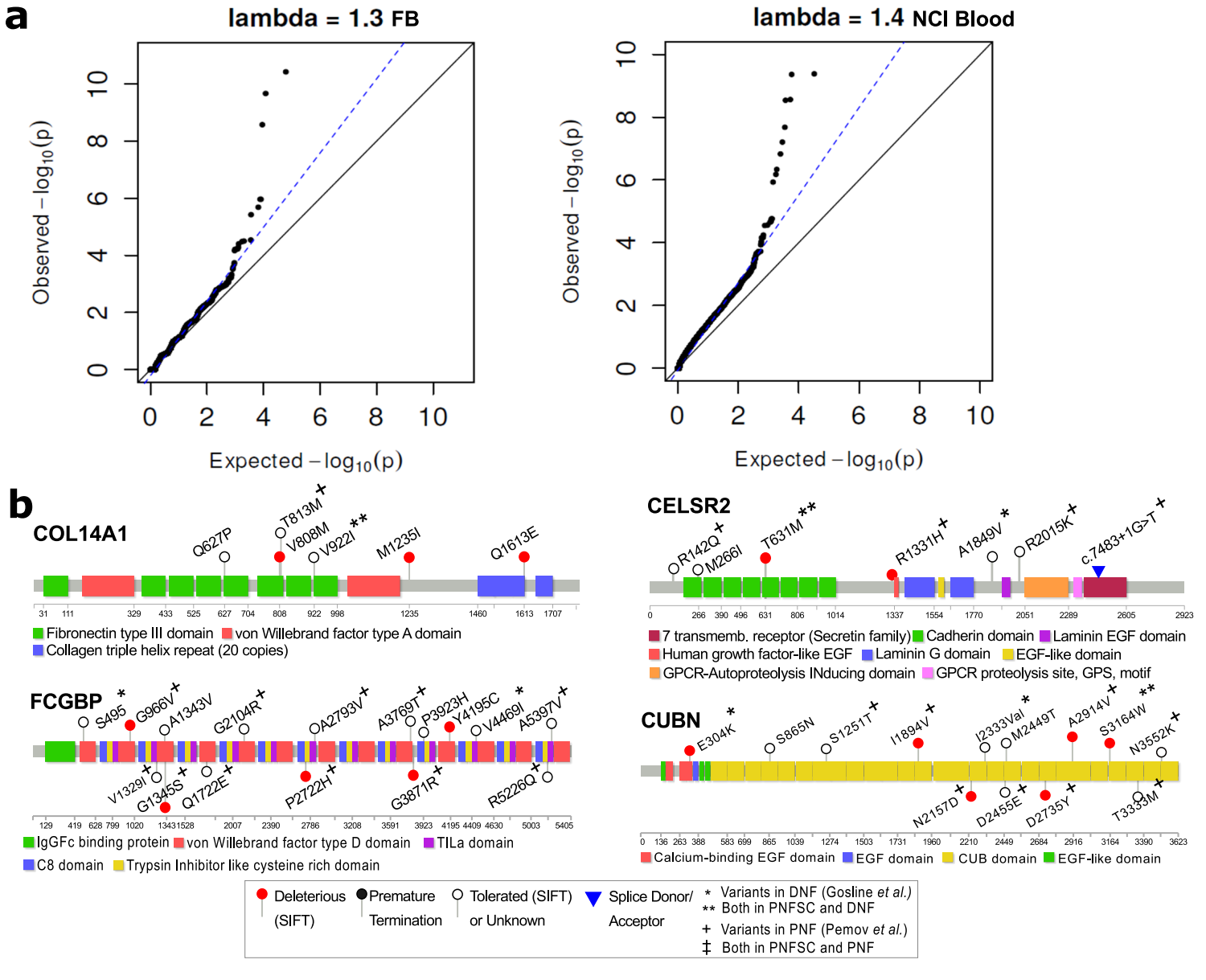
Reference

1. Pemov, A. *et al.* The primacy of NF1 loss as the driver of tumorigenesis in neurofibromatosis type 1-associated plexiform neurofibromas. *Oncogene* **36**, 3168-3177 (2017).
2. Gosline, S.J. *et al.* A high-throughput molecular data resource for cutaneous neurofibromas. *Sci Data* **4**, 170045 (2017).
3. Xu, H., DiCarlo, J., Satya, R.V., Peng, Q. & Wang, Y. Comparison of somatic mutation calling methods in amplicon and whole exome sequence data. *BMC Genomics* **15**, 244 (2014).
4. Cai, L., Yuan, W., Zhang, Z., He, L. & Chou, K.C. In-depth comparison of somatic point mutation callers based on different tumor next-generation sequencing depth data. *Sci Rep* **6**, 36540 (2016).
5. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* **31**, 213-9 (2013).
6. Koboldt, D.C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* **22**, 568-76 (2012).
7. Saunders, C.T. *et al.* Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* **28**, 1811-7 (2012).
8. Larson, D.E. *et al.* SomaticSniper: identification of somatic point mutations in whole genome sequencing data. *Bioinformatics* **28**, 311-7 (2012).
9. Li, H. Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics* **30**, 2843-51 (2014).
10. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122 (2016).
11. Liu, X., Jian, X. & Boerwinkle, E. dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. *Hum Mutat* **34**, E2393-402 (2013).
12. Landrum, M.J. *et al.* ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* **42**, D980-5 (2014).
13. Ng, P.C. & Henikoff, S. Predicting deleterious amino acid substitutions. *Genome Res* **11**, 863-74 (2001).
14. Dong, C. *et al.* Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum Mol Genet* **24**, 2125-37 (2015).

a *NF1* germline and somatic mutation status in different NF1 patients

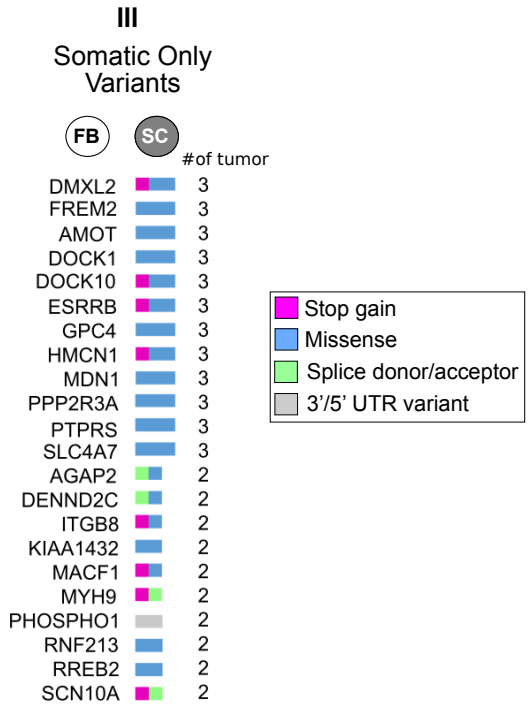
Sample	Position	REF	ALT	Effect	HGMD ID	ClinVar	Cell type	Status	Reads with reference allele(FB/SC)	Reads with mutation allele(FB/SC)	% of mutation reads
T1	17:29652983	A	T	stop gained			FB/SC	germline	42/53	48/52	53/50
	17:29664575	C	G	missense			SC	somatic	79	37	47
	17:29664862	G	T	missense			SC	somatic	41	4	9
T2	17:29562746	C	G	missense	CM040785	likely pathogenic	FB/SC	germline	35/33	36/34	51/51
	17:29509674	CATGAAT AAGGTA	C	splicing			SC	somatic	35	7	16
	17:29677252	G	T	missense			SC	somatic	117	6	5
T3	17:29497003	C	T	stop gained	CM000774	pathogenic	SC	somatic	52	13	25
T4	17:29585518	A	G	missense		pathogenic	FB/SC	germline	79/44	61/39	44/47
	17:29556042	G	A	splicing		pathogenic	SC	somatic	35	6	15
T5a	17:29483145	G	A	splicing		pathogenic	FB/SC	germline	109/136	116/109	51/44
	17:29533315	C	T	stop gained	CM000780	pathogenic	SC	somatic	192	85	44
T5c	17:29483145	G	A	splicing		pathogenic	FB/SC	germline	118/210	130/118	52/36
	17:29562641	C	T	stop gained	CM000799	pathogenic	SC	somatic	191	32	14
T6	17:29560021	C	T	splicing/synonymous			FB/SC	germline	41/12	29/46	41/79
	17:29546122	C	T	stop gained	CM020463	pathogenic	SC	somatic	74	20	21
	17:29556143	G	A	stop gained	CM076345	pathogenic	SC	somatic	41	8	16
T7	17:29557401	G	A	splicing		pathogenic	FB/SC	germline	42/0	34/67	45/100
	17:29557401	G	A	splicing		pathogenic	SC	Somatic (LOH)	0	67	100
T8	17:29677208	T	TA	frame shift		pathogenic	FB/SC	germline	272/276	104/116	28/30
	17:29562709	AG	A	frame shift			SC	somatic	147	73	50



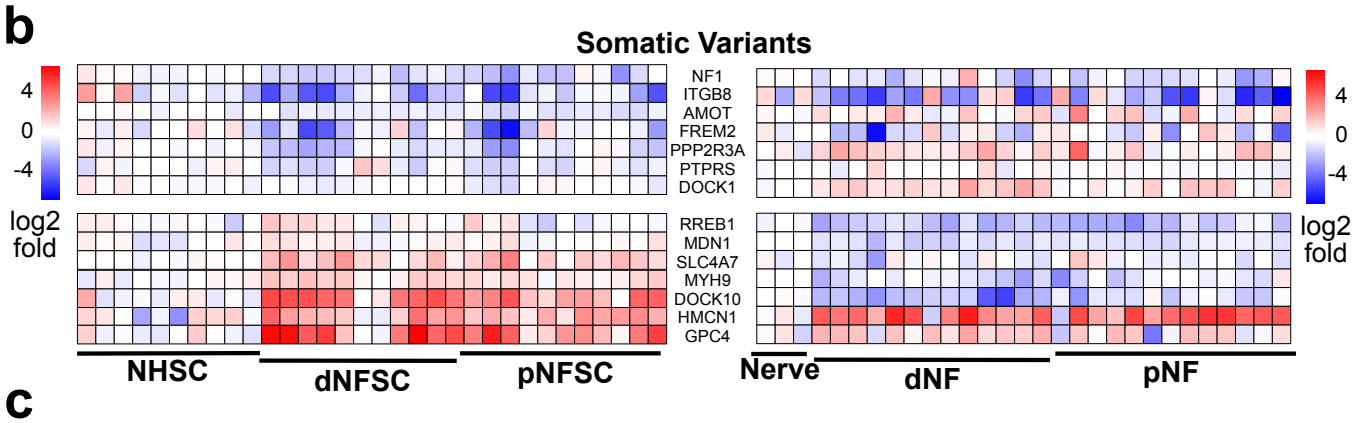

a
lambda = 1.3 FB
lambda = 1.4 NCI Blood
Observed −log10(p)
Expected −log10(p)
b
COL14A1
Q627P
T813M+
V808M
V922I**
M1235I
Q1613E
Fibronectin type III domain
von Willebrand factor type A domain
Collagen triple helix repeat (20 copies)
FCGBP
S495*
G966V+
A1343V
G2104R+
A2793V+
A3769T+
P3923H
V4195C
V4469I*
A5397V+
V1329I+
G1345S+
Q1722E+
P2722H+
G3871R+
R5226Q+
IgGFc binding protein
von Willebrand factor type D domain
TILa domain
C8 domain
Trypsin Inhibitor like cysteine rich domain
CELSR2
R142Q+
M266I
T631M**
R1331H+
A1849V*
R2015K+
c.7483+1G>T+
7 transmemb. receptor (Secretin family)
Cadherin domain
Laminin EGF domain
Human growth factor-like EGF
Laminin G domain
EGF-like domain
GPCR-Autoproteolysis INducing domain
GPCR proteolysis site, GPS, motif
CUBN
E304K*
S865N
S1251T+
I1894V+
I2333Val*
M2449T
A2914V+
S3164W**
N3552K+
N2157D+
D2455E+
D2735Y+
T3333M+
Calcium-binding EGF domain
EGF domain
CUB domain
EGF-like domain
Deleterious (SIFT)
Premature Termination
Tolerated (SIFT) or Unknown
Splice Donor/ Acceptor
* Variants in DNF (Gosline et al.)
** Both in PNFSC and DNF
+ Variants in PNF (Pemov et al.)
‡ Both in PNFSC and PNF

a

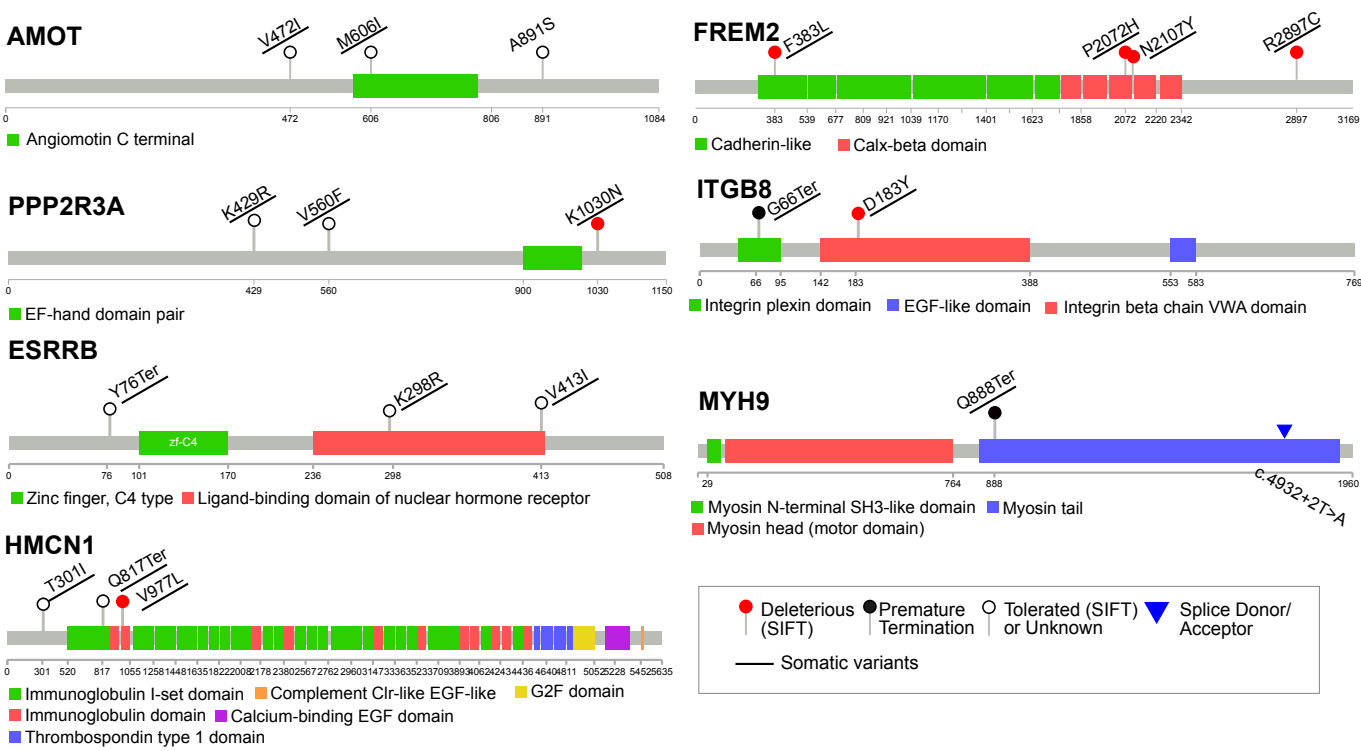
Gene Symbol	# of PNFSC tumors with I	II	III	PNFSC % NF1 patients with variants	DNF Gosline *et.al* % NF1 patients with variants	PNF Pemov *et.al* % NF1 patients with variants	COSMIC % with mutations
DMXL2			3	33	8		3.2
FREM2			3	33			5.0
AMOT			3	33			0.7
DOCK1			3	33			4.9
DOCK10			3	33		5	4.4
ESRRB			3	33			0.7
GPC4			3	33			1.8
HMCN1			3	33			9.5
MDN1			3	33		5	4.3
PPP2R3A			3	33			2.1
PTPRS			3	33			2.1
SLC4A7			3	33			1.5
AGAP2			2	22			0.8
DENND2C			2	22			1.2
ITGB8			2	22			2.0
KIAA1432			2	22			1.7
MACF1			2	22	15		3.9
MYH9			2	22	8		2.5
PHOSPHO1			2	22			0.1
RNF213			2	22			2.4
RREB1			2	22		5	1.3
SCN10A			2	22			3.9



b

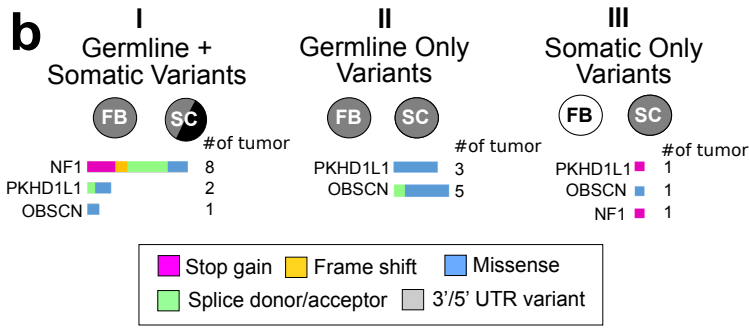


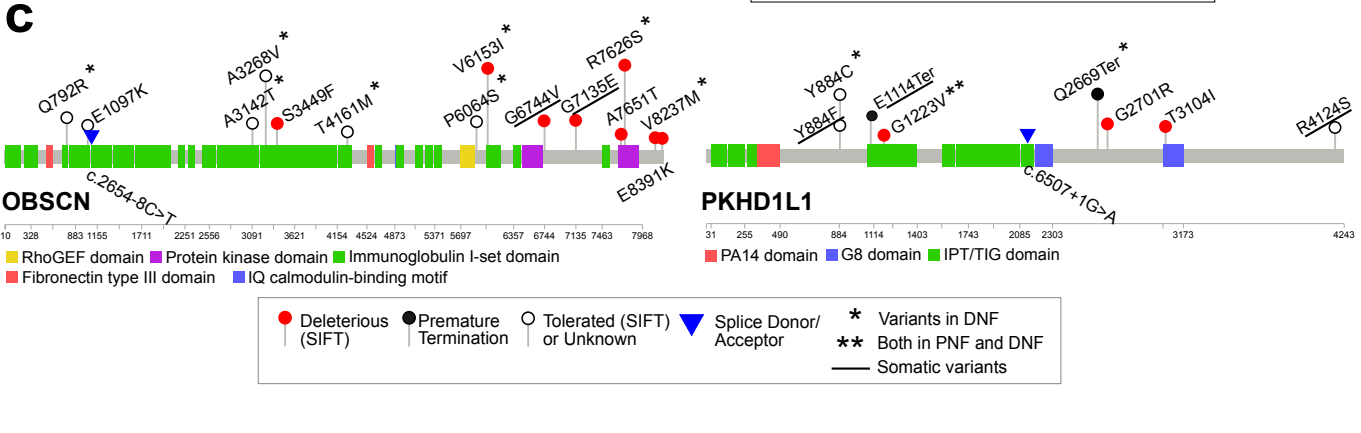
c

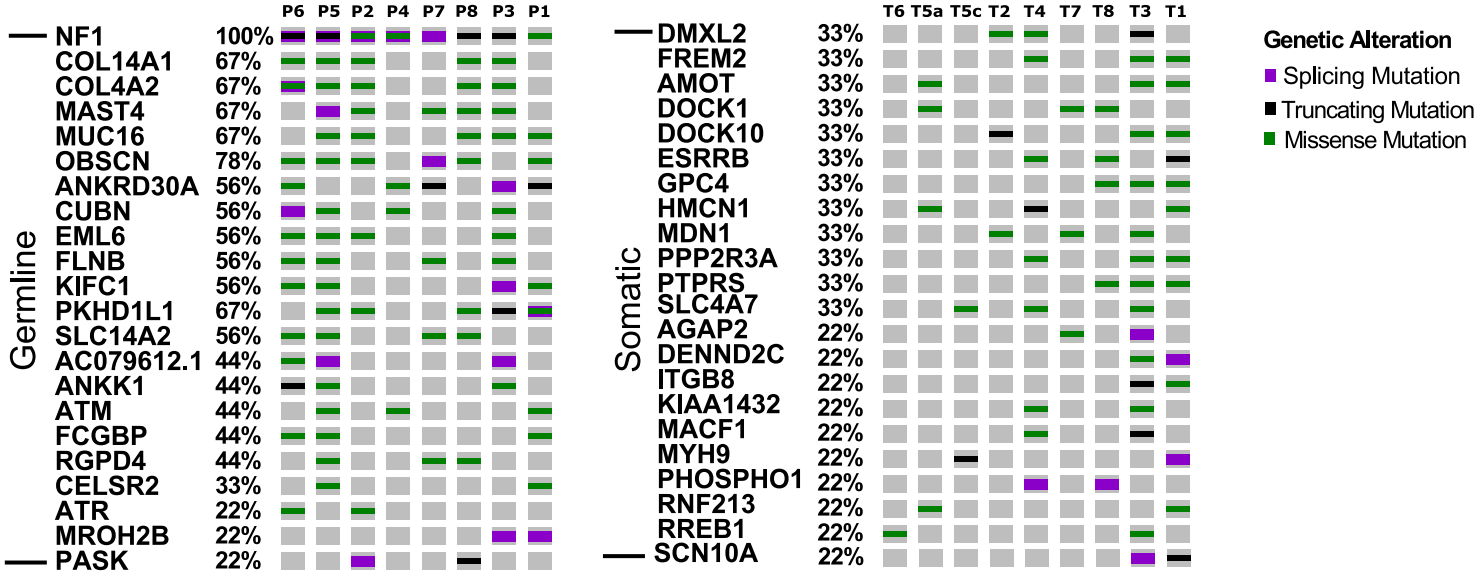


a

Gene Symbol	# of PNF tumors with			% NF1 patients with variants in	% NF1 patients with variants in syn4984604
	I	II	III	PNF	DNF
NF1	8	0	1	100	31
OBSCN	1	5	1	78	39
PKHD1L1	2	3	1	67	39

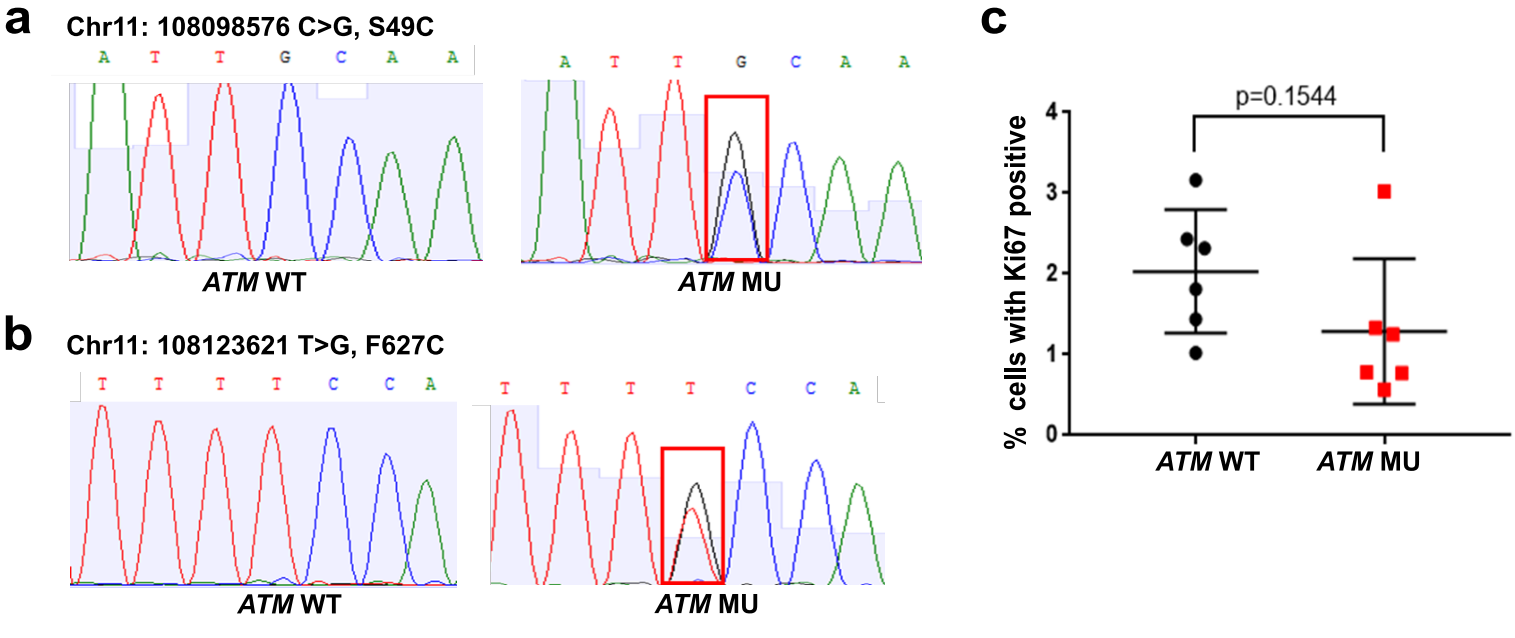


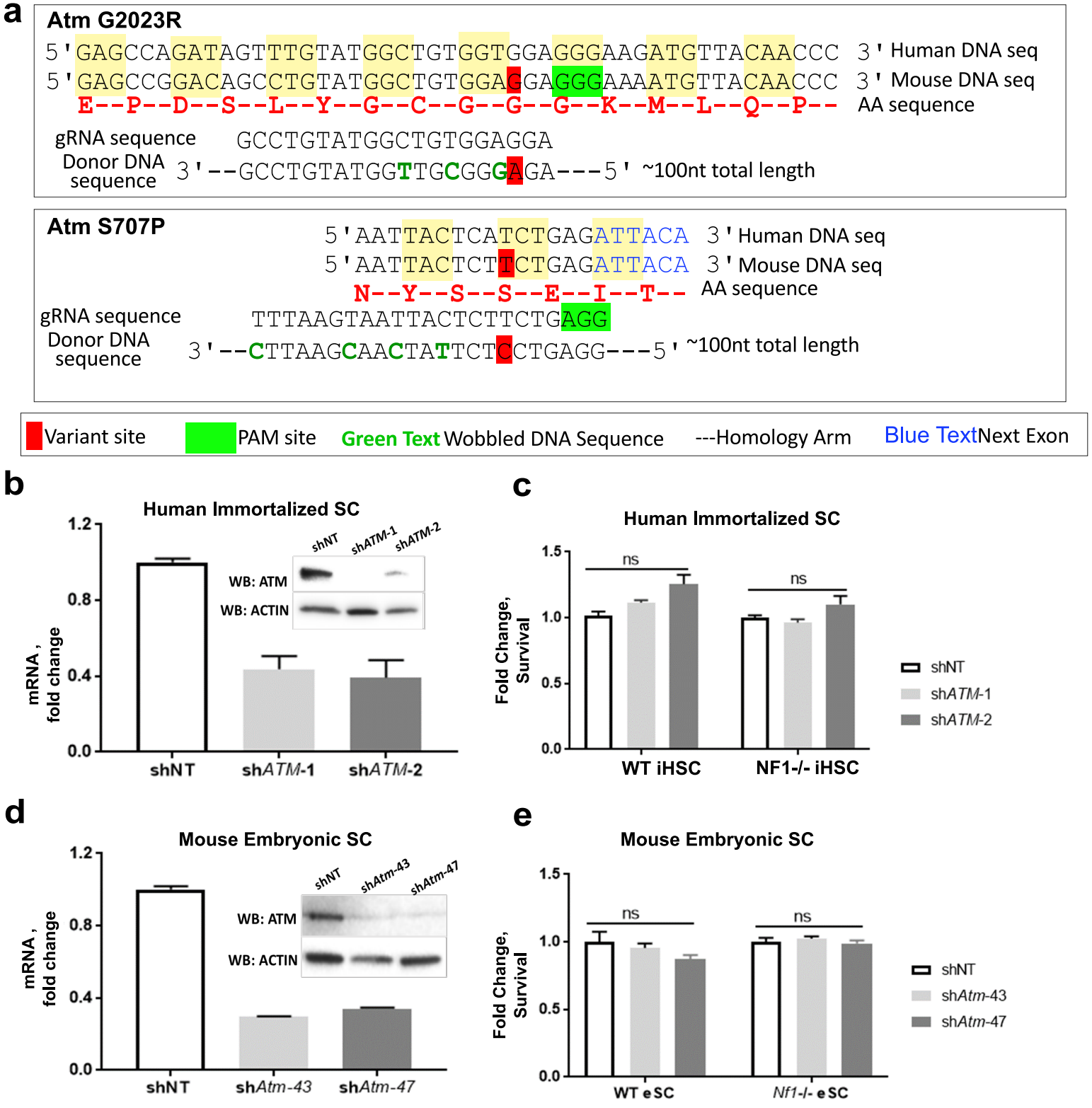


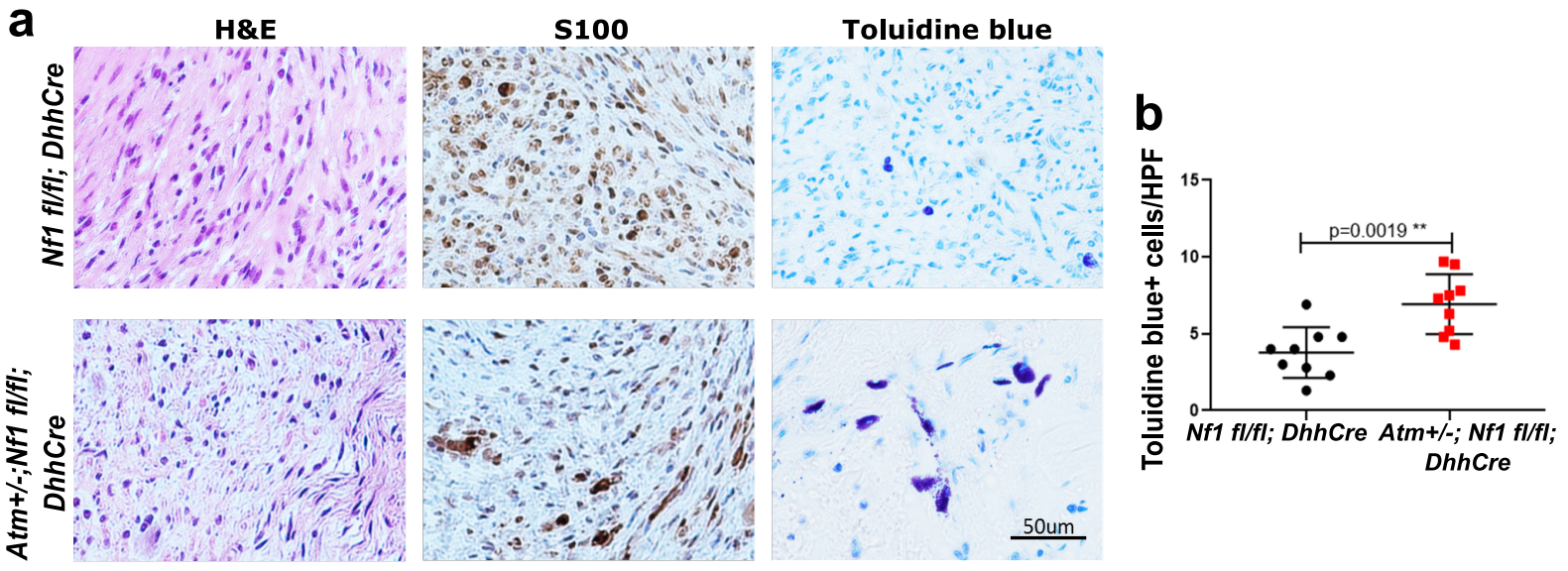

P6 P5 P2 P4 P7 P8 P3 P1
Germline
NF1 100%
COL14A1 67%
COL4A2 67%
MAST4 67%
MUC16 67%
OBSCN 78%
ANKRD30A 56%
CUBN 56%
EML6 56%
FLNB 56%
KIFC1 56%
PKHD1L1 67%
SLC14A2 56%
AC079612.1 44%
ANKK1 44%
ATM 44%
FCGBP 44%
RGPD4 44%
CELSR2 33%
ATR 22%
MROH2B 22%
PASK 22%
T6 T5a T5c T2 T4 T7 T8 T3 T1
Somatic
DMXL2 33%
FREM2 33%
AMOT 33%
DOCK1 33%
DOCK10 33%
ESRRB 33%
GPC4 33%
HMCN1 33%
MDN1 33%
PPP2R3A 33%
PTPRS 33%
SLC4A7 33%
AGAP2 22%
DENND2C 22%
ITGB8 22%
KIAA1432 22%
MACF1 22%
MYH9 22%
PHOSPHO1 22%
RNF213 22%
RREB1 22%
SCN10A 22%
Genetic Alteration
Splicing Mutation
Truncating Mutation
Missense Mutation

Detailed Descriiption of ATM genomic Variants in neurofibromas

Gene Symbol	Variants HGVSp	Variant sample	Found in A-T Patients	Found in Breast Cancer	Found in Other Cancer type(s)	# of Datasets Predicted as Damaging	Known effects on Protein function or Tumorigenesis	Reference (PMID)
ATM	ENSP00000278616.4 :p.Ser49Cys	PNF(This study & Pemov)	Y	Y	Y(Melonoma, Prostate cancer and oropharyngeal cancer)	3	May be breast cancer susceptibility allele; associated with other types of cancers; undertermined for protein function.	16652348; 18565893; 8665503; 10534763
	ENSP00000278616.4 :p.Phe858Leu	PNF(This study)	Y	Y	Y(Lymphoma)	2	Increased risk for breast cancer development; no known risk for other cancers. May associate with LOH in breast cancer.	8797579; 10534763; 20826828; 12149228
	ENSP00000278616.4 :p.Gly2023Arg	PNF(This study)	Y	Y	Y(Lymphoma)	6	Loss of protein expression	27664052; 12149228; 23091097; 9005288
	ENSP00000278616.4 :p.Met1210Val	DNF(Gosline)	N	N	Y(Liver Cancer)	3	Unknown significance	COSMIC
	ENSP00000278616.4 :p.Val2079Ile	DNF(Gosline)	N	Y	N	0	Associated with loss of rare allele	19781682; 12917204; 8665503
	ENSP00000278616.4 :p.Leu2332Pro	DNF(Gosline)	N	N	N	2	Mutated in FAT domain	COSMIC; 28492532; 17344846
	ENSP00000278616.4 :p.Phe627Cys	PNF(Pemov)	N	Y	N	3	Unknown significance	COSMIC; 28492532; 19781682
	ENSP00000278616.4 :p.Ala740Thr	PNF(Pemov)	N	N	N	7	Unknown significance	N/A
	ENSP00000278616.4 :p.Thr2333Lys	PNF(Pemov)	N	N	N	1	Mutated in FAT domain	COSMIC; ClinVar; 28492532
	ENSP00000278616.4 :p.Cys2464Arg	PNF(Pemov)	N	Y	Y(Chronic Lymphocytic Leukemia, colorectal carcinomas)	3	Doesn't interfere with ATM kinase function	19781682; 17344846; 11805335; 27756406
	ENSP00000278616.4 :p.Ser707Pro	PNF(Pemov)	N	Y	Y(Lymphoma, Thyroid or endocrine cancers, Lung cancer)	0	Associated with increased risk of thyroid or endocrine cancers.	27413114; 18164969; 20826828


a
Chr11: 108098576 C>G, S49C
A T T G C A A
A T T G C A A
ATM WT
ATM MU
b
Chr11: 108123621 T>G, F627C
T T T T C C A
T T T T C C A
ATM WT
ATM MU
c
p=0.1544
% cells with Ki67 positive
0
1
2
3
4
ATM WT
ATM MU


a
Atm G2023R
5' GAGCCAGATAGTTTGTATGGCTGTGGTGGAGGGAAGATGTTACAACCC 3' Human DNA seq
5' GAGCCGGACAGCCTGTATGGCTGTGGAGGAGGGAAAATGTTACAACCC 3' Mouse DNA seq
E--P--D--S--L--Y--G--C--G--G--G--K--M--L--Q--P-- AA sequence
gRNA sequence GCCTGTATGGCTGTGGAGGA
Donor DNA sequence 3'--GCCTGTATGGTTGCGGGAGA---5' ~100nt total length
Atm S707P
5' AATTACTCATCTGAGATTACA 3' Human DNA seq
5' AATTACTCTTCTGAGATTACA 3' Mouse DNA seq
N--Y--S--S--E--I--T-- AA sequence
gRNA sequence TTTAAGTAATTACTCTTCTGAGG
Donor DNA sequence 3'--CTTAAGCAACTATTCTCCTGAGG---5' ~100nt total length
Variant site
PAM site
Green Text Wobbled DNA Sequence
---Homology Arm
Blue TextNext Exon
b
Human Immortalized SC
mRNA, fold change
1.2
0.8
0.4
0.0
shNT
shATM-1
shATM-2
WB: ATM
WB: ACTIN
shNT
shATM-1
shATM-2
c
Human Immortalized SC
Fold Change, Survival
1.5
1.0
0.5
0.0
ns
ns
WT iHSC
NF1-/- iHSC
shNT
shATM-1
shATM-2
d
Mouse Embryonic SC
mRNA, fold change
1.2
0.8
0.4
0.0
shNT
shAtm-43
shAtm-47
WB: ATM
WB: ACTIN
shNT
shAtm-43
shAtm-47
e
Mouse Embryonic SC
Fold Change, Survival
1.5
1.0
0.5
0.0
ns
ns
WT eSC
Nf1-/- eSC
shNT
shAtm-43
shAtm-47


a
H&E
S100
Toluidine blue
Nf1 fl/fl; DhhCre
Atm+/-;Nf1 fl/fl; DhhCre
50um
b
Toluidine blue+ cells/HPF
15
10
5
0
p=0.0019 **
Nf1 fl/fl; DhhCre
Atm+/-; Nf1 fl/fl; DhhCre

Supplementary Table 1. Each sample's depth of coverage.

GATK coverage info		
	Fibroblast	Schwann cell
Sample	*Depth of Coverage	*Depth of Coverage
T1	79.2	75.4
T2	82.8	100.8
T3	50.5	59.9
T4	98.4	80.2
T5a	118.6	117.2
T5c	101.9	143.8
T6	101.5	128.9
T7	122.2	111.0
T8	149.8	152.6
P9		97.4
Mean	100.6	106.7
Max.	149.8	152.6
Min.	50.5	59.9

*Sum of the sample specific depth in all loci divided by interval size

Supplementary Table 2 *NF1* gene Germline and somatic mutation status in different NF1 patients--Extended

sample	position (Chr: site)	Ref. Genotype	Mut. Genotype	Effect	HGMD ID	Cell type	Status	Reads with reference allele (FB/SC)	Reads with mutation allele (FB/SC)	% of mutation reads	GMAF/ ExAC MAF	1000Gp1_AC	COSMIC_ID	HGVSc	HGVSp	ClinVar disease	Variants with deleterious, damaging, and/or pathogenic effect (see Suppl. Table 9)			
																	ClinVar	SIFT	MetaLR	MetaSVM
T1	17:29652983	A	T	stop gained		FB/sc	germline	42/53	48/52	53/50		-	-	ENST00000358273.4:c.4981A>T	ENSP00000351015.4:p.Lys1661Ter		-		-	-
	17:29664575	C	G	missense		SC	somatic	79	37	47		-	-	ENST00000358273.4:c.6617C>G	ENSP00000351015.4:p.Thr2206Arg		-	deleterious	T	T
	17:29664862	G	T	missense		SC	somatic	41	4	9		-	-				-	deleterious	T	T
T2	17:29562746	C	G	missense	CM040785	FB/SC	germline	35/33	36/34	51/51		-	-	ENST00000358273.4:c.3826C>G	ENSP00000351015.4:p.Arg1276Gly	NF1	pathogenic	deleterious	D	D
	17:29509674	CATGAATAAG GTA	C	splicing		SC	somatic	35	7	16		-	-	ENST00000358273.4:c.881_888+4delTGAATAAGGTAA			-		-	-
	17:29677252	G	T	missense		SC	somatic	117	6	5		-	-				-	deleterious	T	T
T3	17:29497003	C	T	stop gained	CM000774	SC	somatic	52	13	25	0.000008238	-	COSM42794	ENST00000358273.4:c.574C>T	ENSP00000351015.4:p.Arg192Ter	NF1	pathogenic		-	-
T4	17:29585518	A	G	missense		FB	germline	79/44	61/39	44/47	8.24E-06	-	COSM24576	ENST00000358273.4:c.4330A>G	ENSP00000351015.4:p.Lys1444Glu	NF1	pathogenic	deleterious	D	D
	17:29556042	G	A	splicing		SC	somatic	35	6	15		-	-	ENST00000358273.4:c.2410-1G>A			likely_pathogenic		-	-
T5a	17:29483145	G	A	splicing		FB	germline	109/136	116/109	51/44		-	-	ENST00000358273.4:c.204+1G>A			pathogenic		-	-
	17:29533315	C	T	stop gained	CM000780	SC	somatic	192	85	44	8.24E-06	-	COSM977403	ENST00000358273.4:c.1318C>T	ENSP00000351015.4:p.Arg440Ter	NF1	pathogenic		-	-
T5c	17:29483145	G	A	splicing		FB	germline	118/210	130/118	52/36		-	-	ENST00000358273.4:c.204+1G>A			pathogenic		-	-
	17:29562641	C	T	stop gained	CM000799	SC	somatic	191	32	14		-	COSM24441	ENST00000358273.4:c.3721C>T	ENSP00000351015.4:p.Arg1241Ter	NF1	pathogenic		-	-
T6	17:29560021	C	T	plicing/synonamous		FB	germline	41/12	29/46	41/79	0.0044/1.244E-03	-	-	ENST00000358273.4:c.3498C>T	ENST00000358273.4:c.3498C>T(p.%3D)	NF1	-		-	-
	17:29546122	C	T	stop gained	CM020463	SC	somatic	74	20	21		-	-	ENST00000358273.4:c.1627C>T	ENSP00000351015.4:p.Gln543Ter		-		-	-
	17:29556143	G	A	stop gained	CM076345	SC	somatic	41	8	16		-	-	ENST00000358273.4:c.2510G>A	ENSP00000351015.4:p.Trp837Ter		-		-	-
T7	17:29557401	G	A	splicing		FB	germline	42/0	34/67	45/100		-	-	ENST00000358273.4:c.3113+1G>A		NF1	pathogenic		-	-
	17:29557401	G	A	splicing		SC	germline and somatic	0	67	100						NF1	pathogenic			
T8	17:29677208	T	TA	frame shift		FB	germline	272/276	104/116	28/30		-	-	ENST00000358273.4:c.7330dupA	ENSP00000351015.4:p.Thr2444AsnfsTer4		-		-	-
	17:29562709	AG	A	frame shift		SC	somatic	147	73	50		-	-	ENST00000358273.4:c.3790delG	ENSP00000351015.4:p.Glu1264LysfsTer2		-		-	-

End note: Highlighted sites are recurrent mutation with same site on same gene in different PNF (Ratner), DNF (Gosline) and PNF (Pemov) samples

Yellow: Overlapped mutated sites between PNF (Ratner) and DNF (Gosline)

Green: Overlapped mutated sites between PNF (Ratner) and PNF (Pemov)

Supplementary Table 3: Top 22 germline variant genes

Variants with deleterious, damaging, and/or pathogenic effect (see Suppl. Table 6)

SYMBOL	CHR	POS	Samle #	Ref	Alt	Consequences	HGVSc	HGVSp	GMAF	ExAC_MAF	1000Gp1_AC allele counts in	ExAC_AC allele counts in	INFO	Reads with reference	Reads with mutation	Mutation reads Ratio(FB)	Reads with reference allele(SC)	Reads with mutation allele(SC)	Mutation reads Ratio(SC)	ClinVar (2019)	SIFT	MetaLR	MetaSVM
NF1	17	29652983	T1	A	T	stop_gained	ENST00000358273.4:c.4981A>T	ENSP00000351015.4:p.Lys1661Ter			-	-	.	42	48	53%	53	52	50%		-	-	-
		29562746	T2	C	G	missense_variant	ENST00000358273.4:c.3826C>G	ENSP00000351015.4:p.Arg1276Gly			-	-	.	35	36	51%	33	34	51%	Likely_pathogenic	deleterious	D	D
		29585518	T4	A	G	missense_variant	ENST00000358273.4:c.4330A>G	ENSP00000351015.4:p.Lys1444Glu		A:0.00000824	-	1	.	79	61	44%	44	39	47%	Pathogenic	deleterious	D	D
		29483145	T5a	G	A	splice_donor_variant	ENST00000358273.4:c.204+1G>A				-	-	.	116	109	48%	109	136	55%	Pathogenic	-	-	-
		29483145	T5c	G	A	splice_donor_variant	ENST00000358273.4:c.204+1G>A				-	-	.	116	109	48%	109	136	55%	Pathogenic	-	-	-
		29560021	T6	C	T	splicing/synonamous_variant	ENST00000358273.4:c.3498C>T		T:0.0044	T:1.244E-03	-	-	.	41	29	41%	12	46	79%	Conflicting_interpretations_	-	-	-
		29557401	T7	G	A	splice_donor_variant(LOH)	ENST00000358273.4:c.3113+1G>A				-	-	.	42	34	45%	0	67	100%	Pathogenic	-	-	-
		29677208	T8	T	TA	frameshift_variant	ENST00000358273.4:c.7330dupA	ENSP00000351015.4:p.Thr2444AsnfsTer4			-	-	.	272	104	28%	276	116	30%	Pathogenic	-	-	-
COL14A1	8	121354634	T2	C	G	missense_variant	ENST00000297848.3:c.4837C>G	ENSP00000297848.3:p.Gln1613Glu	G:0.0014	G:1.581e-03	3	192	rs140154122	95	123	56%	194	173	47%		deleterious	T	T
		121238881	T3	A	C	missense_variant	ENST00000297848.3:c.1880A>C	ENSP00000297848.3:p.Gln627Pro	C:0.0052	C:1.062e-03	17	129	rs116603414	36	22	38%	63	44	59%		tolerated	T	T
		121267490	T5a	G	A	missense_variant	ENST00000297848.3:c.2764G>A	ENSP00000297848.3:p.Val922Ile	A:0.0028	A:5.247e-03	10	637	rs11774228	58	53	48%	79	42	35%		tolerated	T	T
		121293179	T5a	G	A	missense_variant	ENST00000297848.3:c.3705G>A	ENSP00000297848.3:p.Met1235Ile		A:7.978e-03	-	968	.	175	102	35%	190	71	26%		deleterious	T	T
		121267490	T5c	G	A	missense_variant	ENST00000297848.3:c.2764G>A	ENSP00000297848.3:p.Val922Ile	A:0.0028	A:5.247e-03	10	637	rs11774228	57	72	56%	89	56	39%		deleterious	T	T
		121293179	T5c	G	A	missense_variant	ENST00000297848.3:c.3705G>A	ENSP00000297848.3:p.Met1235Ile		A:7.978e-03	-	968	.	258	117	30%	269	139	33%		deleterious	T	T
		121256190	T6	G	A	missense_variant	ENST00000297848.3:c.2422G>A	ENSP00000297848.3:p.Val808Met	A:0.0004	A:5.765e-05	-	7	rs115676614	40	78	66%	58	48	45%		deleterious	D	D
		121293179	T8	G	A	missense_variant	ENST00000297848.3:c.3705G>A	ENSP00000297848.3:p.Met1235Ile		A:7.978e-03	-	968	.	177	102	36%	322	204	38%		deleterious	T	T
COL4A2	13	111130496	T2	A	G	missense_variant	ENST00000360467.5:c.2572A>G	ENSP00000353654.5:p.Ile858Val		G:8.282e-06	-	1	.	7	9	56%	13	23	64%		tolerated	T	T
		111118382	T3	G	A	missense_variant	ENST00000360467.5:c.2011G>A	ENSP00000353654.5:p.Gly671Ser	A:0.0012	A:5.954e-04	4	72	rs143710874	99	95	49%	198	144	42%		tolerated	T	T
		111114695	T5a	T	A	missense_variant	ENST00000360467.5:c.1740T>A	ENSP00000353654.5:p.Asp580Glu		A:1.654e-05	-	2	.	30	30	50%	43	42	49%		tolerated	D	T
		111114695	T5c	T	A	missense_variant	ENST00000360467.5:c.1740T>A	ENSP00000353654.5:p.Asp580Glu		A:1.654e-05	-	2	.	54	54	50%	34	49	59%		tolerated	D	T
		111155473	T6	C	T	missense_variant	ENST00000360467.5:c.3883C>T	ENSP00000353654.5:p.Arg1295Trp	T:0.0006	T:1.821e-04	2	22	rs187526694	37	34	48%	54	38	41%		deleterious	D	D
		111164563	T6	T	C	3_prime_UTR_variant	ENST00000360467.5:c.*25T>C		C:0.0088	C:2.905e-03	-	-	rs113331483	27	10	27%	12	16	57%	Likely_benign	-	-	-
		111143601	T8	A	G	missense_variant	ENST00000360467.5:c.3368A>G	ENSP00000353654.5:p.Glu1123Gly	G:0.0058	G:7.829e-03	18	945	rs117412802	79	75	49%	87	59	40%	Benign/Likely_benign,_risk_	deleterious	D	D
MAST4	5	66441069	T2	G	A	missense_variant	ENST00000403625.2:c.2956G>A	ENSP00000385727.1:p.Glu986Lys	A:0.0036	A:9.205e-03	7	1114	rs55969676	33	25	43%	36	38	51%		deleterious		T
		66461391	T3	G	C	missense_variant	ENST00000403625.2:c.6384G>C	ENSP00000385727.1:p.Arg2128Ser	C:0.0052	A:8.259e-06&C:1.6	9	202	rs77191584	16	31	66%	53	28	35%		deleterious_low_confidenc	T	T
		66462143	T3	A	G	missense_variant	ENST00000403625.2:c.7136A>G	ENSP00000385727.1:p.Tyr2379Cys	G:0.0052	G:1.426e-03	9	172	rs138020214	39	17	30%	47	34	42%		deleterious_low_confidenc	T	T
		65892445	T5a	C	A	5_prime_UTR_variant	ENST00000403625.2:c.-39C>A		A:0.0002	A:4.756e-04&T:8.6	-	-	.	11	11	50%	4	6	60%		-	-	-
		65892445	T5c	C	A	5_prime_UTR_variant	ENST00000403625.2:c.-39C>A		A:0.0002	A:4.756e-04&T:8.6	-	-	.	8	3	27%	10	11	52%		-	-	-
		66460651	T7	G	A	missense_variant	ENST00000403625.2:c.5644G>A	ENSP00000385727.1:p.Gly1882Ser		A:3.310e-05	-	4	.	17	23	58%	27	15	36%		tolerated_low_confidence	T	T
		66458571	T8	C	G	missense_variant	ENST00000403625.2:c.3922C>G	ENSP00000385727.1:p.Arg1308Gly		T:2.482e-05&G:1.2	-	15	rs201605726	82	78	49%	63	48	43%		tolerated	T	T
MUC16	19	9066001	T1	T	C	missense_variant	ENST00000397910.4:c.21445A>G	ENSP00000381008.2:p.Thr7149Ala	C:0.0006	C:7.188e-04	3	87	rs151074528	469	390	45%	426	473	52%		-	T	T
		9065094	T2	A	C	missense_variant	ENST00000397910.4:c.22352T>G	ENSP00000381008.2:p.Leu7451Trp	C:0.0008	C:1.389e-03	-	168	rs201455232	208	184	47%	185	161	47%		-	T	T
		9009641	T3	G	A	missense_variant	ENST00000397910.4:c.3908S>T	ENSP00000381008.2:p.Arg13029Cys	A:0.0086	A:2.655e-03	11	321	rs115200029	381	222	37%	584	373	39%		-	T	T
		9073300	T3	C	A	missense_variant	ENST00000397910.4:c.14146G>T	ENSP00000381008.2:p.Ala4716Ser	A:0.0016	A:3.474e-04	3	42	rs116257308	47	35	43%	69	48	41%		-	T	T
		9086983	T5a	G	T	missense_variant	ENST00000397910.4:c.4832C>A	ENSP00000381008.2:p.Thr1611Lys	T:0.0010	T:2.696e-03	3	326	rs186319266	84	78	48%	90	88	49%		-	T	T
		9086983	T5c	G	T	missense_variant	ENST00000397910.4:c.4832C>A	ENSP00000381008.2:p.Thr1611Lys	T:0.0010	T:2.696e-03	3	326	rs186319266	156	134	46%	117	114	49%		-	T	T
		9020059	T8	T	C	missense_variant	ENST00000397910.4:c.37436A>G	ENSP00000381008.2:p.Glu12479Gly	C:0.0016	C:5.691e-03	-	688	rs201453310	117	117	50%	181	110	38%		-	T	T
		9068458	T8	A	T	missense_variant	ENST00000397910.4:c.18988T>A	ENSP00000381008.2:p.Ser6330Thr	T:0.0016	T:5.458e-03	4	660	rs150142305	164	148	47%	198	154	44%		-	T	T
OBSCN	1	228560779	T1	G	A	missense_variant	ENST00000570156.2:c.25171G>A	ENSP00000455507.2:p.Glu8391Lys	A:0.0002	A:3.439e-04	-	40	rs199507775	4	10	71%	5	5	50%		deleterious	T	T
		228557755	T2	G	A	missense_variant	ENST00000570156.2:c.22951G>A	ENSP00000455507.2:p.Ala7651Thr	A:0.0024	A:2.250e-03	5	272	rs145597580	29	34	54%	82	65	44%		deleterious	T	T
		228473833	T5a	C	T	missense_variant	ENST00000570156.2:c.10346C>T	ENSP00000455507.2:p.Ser3449Phe		T:1.653e-05	-	2	.	32	31	48%	39	35	47%		deleterious	T	T
		228473833	T5c	C	T	missense_variant	ENST00000570156.2:c.10346C>T	ENSP00000455507.2:p.Ser3449Phe		T:1.653e-05	-	2	.	47	36	43%	49	39	44%		deleterious	T	T
		228430967	T6	G	A	missense_variant	ENST00000570156.2:c.3289G>A	ENSP00000455507.2:p.Glu1097Lys	A:0.0002	A:1.323e-04	-	16	.	44	34	42%	50	54	52%		tolerated	T	T
		228407017	T7	C	T	ice_region_variant&intron_vari	ENST00000570156.2:c.2654-8C>T		T:0.0002	T:2.482e-05	-	-	.	14	11	44%	4	9	69%		-	-	-
ANKRD30A	10	37414855	T3	C	T	5_prime_UTR_variant	ENST00000361713.1:c.-29C>T		T:0.0084	T:2.358e-03	-	-	rs74414213	15	34	69%	49	43	47%		-	-	-
		37478440	T4	G	T	missense_variant	ENST00000361713.1:c.2299G>T	ENSP00000354433.1:p.Ala767Ser		T:4.116e-03	-	110	rs201628233	206	41	17%	196	44	18%		tolerated	T	T
		37422934	T6	A	G	missense_variant	ENST00000361713.1:c.540A>G	ENSP00000354432.1:p.Ile180Met	G:0.0032	G:9.432e-04	5	114	rs190023267	123	108	47%	99	121	55%		tolerated	T	T
		37508578	T1	CTG	C	frameshift_variant	eITG	ENSP00000354432.1:p.Glu1258ThrfsTer28		-:2.501e-03	-	-	.	49	26	35%	76	48	39%	Uncertain_significance	-	-	-
		37482182	T7	AG	A	frameshift_variant	ENST00000361713.1:c.2443delG	ENSP00000354432.1:p.Ala815GlnfsTer27	-:0.0066	-:0.021	-	-	.	180	176	49%	196	134	41%		-	-	-
CUBN	10	16893406	T3	G	C	missense_variant	ENST00000377833.4:c.9491C>G	ENSP00000367064.4:p.Ser3164Trp	C:0.0078	A:3.295e-05&C:1.9	10	242	rs57163243	13	20	61%	28	31	53%	Uncertain_significance	deleterious	T	T
		17113456	T4	C	T	missense_variant	ENST00000377833.4:c.2594G>A	ENSP00000367064.4:p.Ser865Asn	T:0.0032	A:8.237e-06&T:7.4	10	904	rs138083522	85	101	54%	86	78	48%	Likely_benign	tolerated	T	T
		16957036	T5a	A	G	missense_variant	ENST00000377833.4:c.7346T>C	ENSP00000367064.4:p.Met2449Thr	G:0.0034	C:3.542e-04&G:4.3	-	529	rs41301097	58	28	33%	68	38	36%	Benign	tolerated	T	T
		16957036	T5c	A	G	missense_variant	ENST00000377833.4:c.7346T>C	ENSP00000367064.4:p.Met2449Thr	G:0.0034	C:3.542e-04&G:4.3	-	529	rs41301097	47	40	46%	62	52	46%	Benign	tolerated	T	T
		16866929	T6	G	A	3_prime_UTR_variant	ENST00000377833.4:c.*45C>T		A:0.0098	A:2.883e-03	-	-	rs7085076	54	40	43%	73	42	37%	Likely_benign	-	-	-
EML6	2	55122149	T2	C	T	missense_variant	ENST00000356458.6:c.2840C>T	ENSP00000348842.6:p.Ser947Leu	T:0.0002	T:1.197e-04	-	3	.	23	31	57%	51	55	52%		deleterious	T	T
		55155592	T2	C	T	missense_variant	ENST00000356458.6:c.3818C>T	ENSP00000348842.6:p.Thr1273Ile	T:0.0004	T:5.026e-04	-	15	rs200306471	50	62	54%	77	85	52%		tolerated	T	T
		55077136	T3	G	C	missense_variant	ENST00000356458.6:c.1225G>C	ENSP00000348842.6:p.Glu409Gln	C:0.0030	C:4.026e-04	5	10	rs146306448	10	13	57%	28	26	48%		tolerated	T	T
		55096447	T5a	T	A	missense_variant	ENST00000356458.6:c.2181T>A	ENSP00000348842.6:p.Asp727Glu		A:6.711e-04	-	17	.	75	56	43%	89	49	35%		deleterious	T	T
		55096447	T5c	T	A	missense_variant	ENST00000356458.6:c.2181T>A	ENSP00000348842.6:p.Asp727Glu		A:6.711e-04	-	17	.	95	72	43%	91	85	48%		deleterious	T	T
		55181176	T6	C	G	missense_variant	ENST00000356458.6:c.4366C>G	ENSP00000348842.6:p.Ser1455Cys	G:0.0072	G:2.328e-03	4	61	rs72910830	60	84	58%	84	79	48%		deleterious	T	T
FLNB	3	57994398	T3	G	A	missense_variant	ENST00000490882.1:c.4076G>A	ENSP00000420213.1:p.Arg1361His	A:0.0030	A:7.495e-04&T:2.4	4	91	rs142568031	89	117	57%	185	147	44%	Uncertain_significance	deleterious	D	D
		58145348	T5a	T	C	missense_variant	ENST00000490882.1:c.7049T>C	ENSP00000420213.1:p.Ile2350Thr	C:0.0024	C:6.482e-03	5	787	rs116826041	49	24	33%	49	45	48%	Benign/Likely_benign	deleterious	D	D
		58145348	T5c	T	C	missense_variant	ENST00000490882.1:c.7049T>C	ENSP00000420213.1:p.Ile2350Thr	C:0.0024	C:6.482e-03	5	787	rs116826041	47	59	56%	60	52	46%	Benign/Likely_benign	deleterious	D	D
		58108862	T6	A	C	missense_variant	ENST00000490882.1:c.3169A>C	ENSP00000420213.1:p.Lys1057Gln	C:0.0016	C:4.201e-04	3	51	rs13321615	56	70	56%	47	46	49%	Likely_benign	tolerated	T	T
		58139249	T6	G	A	missense_variant	ENST00000490882.1:c.6608G>A	ENSP00000420213.1:p.Arg2203His	A:0.0002	A:1.071e-04	2	13	rs139846706	19	19	50%	30	22	42%	Uncertain_significance	deleterious	T	T
		58134505	T7	A	G	missense_variant	ENST00000490882.1:c.6110A>G	ENSP00000420213.1:p.Lys2037Arg	G:0.0070	G:9.118e-03	19	1107	rs62621996	29	32	52%	34	34	50%	Benign/Likely_benign	tolerated	T	T
KIFC1	6	33372768	T1	G	A	missense_variant	ENST00000428849.2:c.896G>A	ENSP00000393963.2:p.Arg299Gln		A:3.295e-05	-	4	.	39	28	42%	24	45	65%		tolerated	T	T
		33359672	T3	G	A	5_prime_UTR_variant	ENST00000428849.2:c.-91G>A		A:0.0026		-	-	rs146617967	19	16	46%	29	25	46%		-	-	-
		33371806	T5a	G	A	missense_variant	ENST00000428849.2:c.656G>A	ENSP00000393963.2:p.Arg219Gln	A:0.0090	A:8.945e-03	13	1086	rs61736175	53	40	43%	55	48	47%		tolerated	T	T
		33371806	T5c	G	A	missense_variant	ENST00000428849.2:c.656G>A	ENSP00000393963.2:p.Arg219Gln	A:0.0090	A:8.945e-03	13	1086	rs61736175	71	70	50%	72		57%		tolerated	T	T
		33372840	T6	G	C	missense_variant	ENST00000428849.2:c.968G>C	ENSP00000393963.2:p.Gly323Ala	C:0.0020	A:8.236e-06&C:8.2	3	100	rs145666693	45	22	33%	32	30	48%		tolerated	T	T
PKHD1L1	8	110464510	T1	G	A	splice_donor_variant	ENST00000378402.5:c.6507+1G>A		A:0.0050	A:7.811e-03	-	-	rs72687022	38	41	51%	37	48	56%		-	-	-
		110477162	T2	G	A	missense_variant	ENST00000378402.5:c.8101G>A	ENSP00000367655.5:p.Gly2701Arg	A:0.0002	A:5.297e-04	1	64	rs200166927	79	77	49%	65	52	44%		deleterious	D	D
		110492352	T5a	C	T	missense_variant	ENST00000378402.5:c.9311C>T	ENSP00000367655.5:p.Thr3104Ile	T:0.0062	T:9.206e-03	17	1099	rs117952538	109	73	40%	89	76	46%		deleterious	D	D
		110492352	T5c	C	T	missense_variant	ENST00000378402.5:c.9311C>T	ENSP00000367655.5:p.Thr3104Ile	T:0.0062	T:9.206e-03	17	1099	rs117952538	160	100	38%	122	118	49%		deleterious	D	D
		110450593	T8	G	T	missense_variant	ENST00000378402.5:c.366B>T	ENSP00000367655.5:p.Gly1223Val	T:0.0054	T:8.477e-03	18	908	rs117037399	56	40	43%	65	54	45%		deleterious	T	T
SLC14A2	18	43224125	T5a	G	A	ense_variant&splice_region_va	ENST00000255226.6:c.1351G>A	ENSP00000255226.5:p.Gly451Arg	A:0.0032	A:1.244e-03	6	151	rs61738671	68	47	41%	44	33	43%		tolerated_low_confidence	T	T
		43224125	T5c	G	A	ense_variant&splice_region_va	ENST00000255226.6:c.1351G>A	ENSP00000255226.5:p.Gly451Arg	A:0.0032	A:1.244e-03	6	151	rs61738671	39	63	62%	55	50	48%		tolerated_low_confidence	T	T
		43248372	T6	G	T	missense_variant	ENST00000255226.6:c.1966G>T	ENSP00000255226.5:p.Ala656Ser	T:0.0006	T:6.589e-05	1	8	rs145751449	205	158	43%	172	230	57%		deleterious	T	T
		43262408	T7	G	A	missense_variant	ENST00000255226.6:c.2687G>A	ENSP00000255226.5:p.Arg896His	A:0.0052	A:8.195e-03&T:8.2	12	995	rs41301139	57	60	51%	73	69	48%		deleterious	T	T
		43262408	T8	G	A	missense_variant	ENST00000255226.6:c.2687G>A	ENSP00000255226.5:p.Arg896His	A:0.0052	A:8.195e-03&T:8.2	12	995	rs41301139	100	105	51%	113	87	43%		deleterious	T	T
AC079612.1	2	240500117	T3	G	A	5_prime_UTR_variant	ENST00000358775.1:c.-3G>A		A:0.0004	A:8.485e-05	-	-	.	3	19	86%	21	23	52%		-	-	-
		240504816	T5a	T	C	3_prime_UTR_variant	ENST00000358775.1:c.*12T>C		C:0.0018	C:5.145e-03	-	-	rs185537591	168	157	48%	148	133	47%		-	-	-
		240504816	T5c	T	C	3_prime_UTR_variant	ENST00000358775.1:c.*12T>C		C:0.0018	C:5.145e-03	-	-	rs185537591	200	137	41%	201	160	44%		-	-	-
		240500231	T6	G	A	missense_variant	ENST00000358775.1:c.1126G>A	ENSP00000351625.1:p.Glu38Lys	A:0.0028	A:1.113e-03	4	27	rs78492176	59	48	45%	47	27	36%		-	T	T
ANKK1	11	113264267	T3	G	C	missense_variant	ENST00000303941.3:c.250G>C	ENSP00000306678.3:p.Val84Leu	C:0.0010	C:9.096e-05	-	11	.	61	59	49%	85	98	54%		tolerated	T	T
		113264382	T5a	G	A	missense_variant	ENST00000303941.3:c.365G>A	ENSP00000306678.3:p.Arg122His	A:0.0042	A:8.983e-03&T:2.4	12	1089	rs35877321	54	54	50%	44	42	49%		tolerated	T	T
		113264382	T5c	G	A	missense_variant	ENST00000303941.3:c.365G>A	ENSP00000306678.3:p.Arg122His	A:0.0042	A:8.983e-03&T:2.4	12	1089	rs35877321	68	53	44%	69	55	44%		tolerated	T	T
		113266933	T6	C	T	missense_variant	ENST00000303941.3:c.827C>T	ENSP00000306678.3:p.Pro276Leu	T:0.0082	A:8.274e-06&T:1.8	12	219	rs35488601	40	41	51%	57	44	44%		deleterious	D	D
		113270450	T6	G	T	stop_gained	ENST00000303941.3:c.1759G>T	ENSP00000306678.3:p.Glu587Ter	T:0.0060	T:1.448e-03	12	175	rs113005509	15	13	46%	9	9	50%		-	-	-
ATM	11	108098576	T1	C	G	missense_variant	ENST00000278616.4:c.146C>G	ENSP00000278616.4:p.Ser49Cys	G:0.0042	G:7.364e-03	12	894	rs1800054	19	24	56%	47	26	35%	Benign/Likely_benign,_risk_	tolerated	T	T
		108138003	T4	T	C	missense_variant	ENST00000278616.4:c.2572T>C	ENSP00000278616.4:p.Phe858Leu	C:0.0046	C:9.142e-03	15	1110	rs1800056	81	71	47%	54	50	48%	Conflicting_interpretations_	tolerated	T	T
		108186610	T5a	G	A	missense_variant	ENST00000278616.4:c.6067G>A	ENSP00000278616.4:p.Gly2023Arg	A:0.0012	A:1.573e-03	3	191	rs11212587	159	126	44%	180	108	38%	Conflicting_interpretations_	deleterious	T	T
		108186610	T5c	G	A	missense_variant	ENST00000278616.4:c.6067G>A	ENSP00000278616.4:p.Gly2023Arg	A:0.0012	A:1.573e-03	3	191	rs11212587	204	153	43%	211	170	45%	Conflicting_interpretations_	deleterious	T	T
FCGBP	19	40368764	T1	T	C	missense_variant	ENST00000221347.6:c.12584A>G	ENSP00000221347.5:p.Tyr4195Cys	C:0.0004	C:4.942e-04	-	60	.	307	112	27%	284	83	23%		deleterious	T	T
		40408811	T5a	G	A	missense_variant	ENST00000221347.6:c.4028C>T	ENSP00000221347.5:p.Ala1343Val	A:0.0094	A:9.922e-03&A:9.9	8	1204	rs141750442	28	38	58%	43	24	36%		tolerated	T	T
		40408811	T5c	G	A	missense_variant	ENST00000221347.6:c.4028C>T	ENSP00000221347.5:p.Ala1343Val	A:0.0094	A:9.925e-03&A:9.9	8	1204	rs141750442	37	41	53%	45	52	54%		tolerated	T	T
		40376654	T6	G	T	missense_variant	ENST00000221347.6:c.11768C>A	ENSP00000221347.5:p.Pro3923His		A:1.682e-05&T:1.1	-	139	rs140951083	282	65	19%	306	67	18%		tolerated_low_confidence	T	T
RGPD4	2	108487440	T5a	A	C	missense_variant	ENST00000408999.3:c.2980A>C	ENSP00000386810.3:p.Lys994Gln	C:0.0034	C:1.527e-03	-	257	.	187	66	26%	140	96	41%		tolerated	T	T
		108487440	T5c	A	C	missense_variant	ENST00000408999.3:c.2980A>C	ENSP00000386810.3:p.Lys994Gln	C:0.0034	C:1.527e-03	-	257	.	154	53	26%	225	137	38%		tolerated	T	T
		108443503	T7	G	C	missense_variant	ENST00000408999.3:c.34G>C	ENSP00000386810.3:p.Val12Leu	C:0.0078	A:8.326e-06&C:7.8	-	937	rs200269994	18	20	53%	20	23	53%		deleterious	T	T
		108489225	T8	G	T	missense_variant	ENST00000408999.3:c.4765G>T	ENSP00000386810.3:p.Ala1589Ser	T:0.0014	T:2.770e-03	2	336	rs190948032	280	220	44%	455	397	46%		tolerated_low_confidence	T	T
CELSR2	1	109793499	T1	G	C	missense_variant	ENST00000271332.3:c.798G>C	ENSP00000271332.3:p.Met266Ile		C:7.083e-04	-	86	.	51	51	50%	46	58	56%		tolerated	T	T
		109794593	T5a	C	T	missense_variant	ENST00000271332.3:c.1892C>T	ENSP00000271332.3:p.Thr631Met	T:0.0012	T:3.904e-03	4	474	rs41279706	60	61	50%	57	38	40%		deleterious	T	T
		109794593	T5c	C	T	missense_variant	ENST00000271332.3:c.1892C>T	ENSP00000271332.3:p.Thr631Met	T:0.0012	T:3.904e-03	4	474	rs41279706	60	51	46%	54	59	52%		deleterious	T	T
ATR	3	142272098	T2	A	G	missense_variant	ENST00000350721.4:c.2776T>C	ENSP00000343741.4:p.Phe926Leu	G:0.0004	G:1.203e-03	2	146	rs141783863	28	35	56%	30	24	44%	Conflicting_interpretations_	tolerated	T	T
		142231134	T6	C	T	missense_variant	ENST00000350721.4:c.4820G>A	ENSP00000343741.4:p.Ser1607Asn	T:0.0048	T:1.367e-03	8	166	rs55724025	71	58	45%	90	67	43%	Conflicting_interpretations_	tolerated	T	T
MROH2B	5	40998261	T1	C	T	splice_acceptor_variant	ENST00000399564.4:c.4652-1G>A		T:0.0004	T:9.925e-04	-	-	rs111906143	30	30	50%	37	38	51%		-	-	-
		41047867	T3	C	A	splice_acceptor_variant	ENST00000399564.4:c.1685-1G>T				-	-	.	14	22	61%	26	34	57%		-	-	-
PASK	2	242051654	T2	C	T	splice_donor_variant	ENST00000358649.4:c.3554+1G>A		T:0.0010	T:1.235e-03	-	-	rs140030739	9	6	40%	22	27	55%	Uncertain_significance	-	-	-
		242072401	T8	G	A	stop_gained	ENST00000358649.4:c.1351C>T	ENSP00000351475.4:p.Arg451Ter	A:0.0010	A:3.212e-04	1	39	rs140101325	56	47	46%	44	33	43%		-	-	-

End note: Highlighted sites are recurrent mutation with same site on same gene in different PNF (Ratner), DNF (Gosline) and PNF (Pemov) samples.

Yellow: Overlapped mutated sites between PNF (Ratner) and DNF (Gosline)
Green: Overlapped mutated sites between PNF (Ratner) and PNF (Pemov)
Orange: Overlapped mutated sites between three datasets.

Supplementary Table 4: Top 25 somatic variant genes

Symbol	Chr#	Position	Ref	Alt	mutation type	HGVSc	HGVSp	Patient #	Reads with reference allele	Reads with mutation allele	Mutation reads ratio	GMAF/ExAC MAF	1000Gp1_AC allele counts in 1000	ExAC_AC allele counts in ~60,706	DOMAIN	COSMIC_ID	Variants with deleterious, damaging, and/or pathogenic effect (see Suppl. Table 9) ClinVar (2019)	SIFT	MetaLR	MetaSVM
NF1	17	29664575	C	T	missense	ENST00000358273.4:c.6617C>G	ENSP00000351015.4:p.Thr2206Arg	T1	79	37	47%		-	-	Superfamily_domains:SSF4837	-	Uncertain	deleterious	T	T
	17	29664862	G	T	missense	ENST00000358273.4:c.6668G>T	ENSP00000351015.4:p.Cys2223Phe	T1	41	4	9%		-	-	Superfamily_domains:SSF4837	-		deleterious	T	T
	17	29509674	CATGAATA AGGTA	C	splicing	ENST00000358273.4:c.881_888+4 delTGAATAAGGTAA		T2	35	7	16%		-	-		-		-	-	-
	17	29677252	G	T	missense	ENST00000358273.4:c.7373G>T	ENSP00000351015.4:p.Arg2458Ile	T2	117	6	5%		-	-	hmmpanther:PTHR10194:SF608	-		deleterious	T	T
	17	29497003	C	T	stop gained	ENST00000358273.4:c.574C>T	ENSP00000351015.4:p.Arg192Ter	T3	52	13	25%	0.000008238	-	1	hmmpanther:PTHR10194:SF608	COSM42794	Pathogen	-	-	-
	17	29556042	G	A	splicing	ENST00000358273.4:c.2410-1G>A		T4	35	6	15%		-	-		-	Pathogen	-	-	-
	17	29533315	C	T	stop gained	ENST00000358273.4:c.1318C>T	ENSP00000351015.4:p.Arg440Ter	T5a	192	85	44%	8.24E-06	-	-	hmmpanther:PTHR10194:SF608	COSM977403	Pathogen	-	-	-
	17	29562641	C	T	stop gained		ENSP00000351015.4:p.Arg1241Ter	T5c	191	32	14%		-	-	Superfamily_domains:SSF4835	COSM24441	Pathogen	-	-	-
	17	29556143	G	A	stop gained	ENST00000358273.4:c.2510G>A	ENSP00000351015.4:p.Trp837Ter	T6	41	8	16%		-	-	hmmpanther:PTHR10194:SF608	-	Pathogen	-	-	-
	17	29546122	C	T	stop gained	ENST00000358273.4:c.1627C>T	ENSP00000351015.4:p.Gln543Ter	T6	74	20	21%		-	-	hmmpanther:PTHR10194:SF608	-	Pathogen	-	-	-
	17	29557401	G	A	splicing(LOH)	ENST00000358273.4:c.3113+1G>A		T7	0	67	100%		-	-		-	Pathogen	-	-	-
	17	29562709	AG	A	frame shift	ENST00000358273.4:c.3790delG	ENSP00000351015.4:p.Glu1264LysfsTe	T8	147	73	50%		-	-	Superfamily_domains:SSF4835	-		-	-	-
DMXL2	15	51756891	G	T	missense	ENST00000543779.2:c.7789C>A	ENSP00000441858.2:p.Pro2597Thr	T2	53	6	10%		-	-	hmmpanther:PTHR13950:SF118	-		deleterious	T	T
	15	51741351	T	A	stop gained	ENST00000543779.2:c.8944A>T	ENSP00000441858.2:p.Lys2982Ter	T3	61	6	9%		-	-	Superfamily_domains:SSF5097	-		-	-	-
	15	51792221	A	T	missense	ENST00000543779.2:c.3200T>A	ENSP00000441858.2:p.Ile1067Asn	T4	110	6	5%		-	-	hmmpanther:PTHR13950:SF118	-		deleterious	T	T
	15	51829806	A	G	missense	ENST00000543779.2:c.1496T>C	ENSP00000441858.2:p.Leu499Pro	T4	139	7	5%		-	-	Superfamily_domains:SSF5097	-		deleterious	T	T
FREM2	13	39262630	C	A	missense	ENST00000280481.7:c.1149C>A	ENSP00000280481.7:p.Phe383Leu	T1	114	8	7%		-	-	hmmpanther:PTHR11878&hmn	-		deleterious	T	T
	13	39422643	C	A	missense	ENST00000280481.7:c.6215C>A	ENSP00000280481.7:p.Pro2072His	T1	108	6	5%		-	-	hmmpanther:PTHR11878&hmn	-		deleterious	T	T
	13	39422747	A	T	missense	ENST00000280481.7:c.6319A>T	ENSP00000280481.7:p.Asn2107Tyr	T3	137	8	6%		-	-	hmmpanther:PTHR11878&hmn	-		deleterious	T	T
	13	39452288	C	T	missense	ENST00000280481.7:c.8689C>T	ENSP00000280481.7:p.Arg2897Cys	T4	114	12	10%		-	-	hmmpanther:PTHR11878&hmn	-		deleterious	D	D
AMOT	X	112035168	C	A	missense	ENST00000371959.3:c.1818G>T	ENSP00000361027.3:p.Met606Ile	T1	64	4	6%		-	-	Coiled-coils_(Ncoils):Coil&hmn	-		-	T	T
	X	112022711	C	A	missense	ENST00000371959.3:c.2671G>T	ENSP00000361027.3:p.Ala891Ser	T3	190	14	7%		-	-	Low_complexity_(Seg):seg&hr	-		-	T	T
	X	112054600	C	T	missense	ENST00000371959.3:c.1414G>A	ENSP00000361027.3:p.Val472Ile	T5a	167	11	6%		-	5	Coiled-coils_(Ncoils):Coil&hmn	-		-	T	T
DOCK1	10	129237467	A	G	missense	ENST00000280333.6:c.5174A>G	ENSP00000280333.6:p.Glu1725Gly	T5a	146	8	5%		-	-		-		tolerated	T	T
	10	129137332	C	T	missense	ENST00000280333.6:c.3044C>T	ENSP00000280333.6:p.Ala1015Val	T7	86	10	10%		-	-	hmmpanther:PTHR23317&hmn	-		tolerated	T	T
	10	129242478	C	T	missense	ENST00000280333.6:c.5285C>T	ENSP00000280333.6:p.Pro1762Leu	T8	198	17	8%		-	1	Low_complexity_(Seg):seg	-		tolerated	T	T
DOCK10	2	225710266	C	A	missense	ENST00000258390.7:c.2329G>T	ENSP00000258390.7:p.Ala777Ser	T1	159	8	5%		-	-	Low_complexity_(Seg):seg&PF	-		tolerated	T	T
	2	225727425	G	T	stop_gained	ENST00000258390.7:c.1641C>A	ENSP00000258390.7:p.Cys547Ter	T2	64	6	9%		-	-	hmmpanther:PTHR23317:SF718	-		-	-	-
	2	225738724	G	C	missense	ENST00000258390.7:c.1246C>G	ENSP00000258390.7:p.Pro416Ala	T3	156	9	5%		-	-	hmmpanther:PTHR23317:SF718	-		tolerated	T	T
ESRRB	14	76905924	C	G	stop_gained	ENST00000380887.2:c.228C>G	ENSP00000370270.2:p.Tyr76Ter	T1	34	4	11%		-	-	hmmpanther:PTHR24084&hmn	-		-	-	-
	14	76957895	A	G	missense	ENST00000380887.2:c.893A>G	ENSP00000370270.2:p.Lys298Arg	T4	73	6	8%		-	-	hmmpanther:PTHR24084&hmn	-		tolerated	D	D
	14	76964736	G	A	missense	ENST00000380887.2:c.1237G>A	ENSP00000370270.2:p.Val413Ile	T8	27	6	18%		12	1336	hmmpanther:PTHR24084&hmn	-	Benign/Li	tolerated	T	T
GPC4	X	132437211	C	A	missense	ENST00000370828.3:c.216G>T	ENSP00000359864.3:p.Glu72Asp	T1	82	6	7%		-	-	hmmpanther:PTHR10822:SF258	-		deleterious	T	T
	X	132440114	A	G	missense	ENST00000370828.3:c.1451T>C	ENSP00000359864.3:p.Val484Ala	T3	116	6	5%		-	-	hmmpanther:PTHR10822:SF258	-		deleterious	T	T
	X	132473314	C	A	missense	ENST00000370828.3:c.946G>T	ENSP00000359864.3:p.Asp316Tyr	T8	75	39	34%		-	-	hmmpanther:PTHR10822:SF258	-		deleterious	T	T
HMCN1	1	185953439	G	T	missense	ENST00000271588.4:c.2929G>T	ENSP00000271588.4:p.Val977Leu	T1	151	11	7%		-	-	hmmpanther:PTHR19897:SF158	-		deleterious	D	D
	1	185946996	C	T	stop gained	ENST00000271588.4:c.2449C>T	ENSP00000271588.4:p.Gln817Ter	T4	137	13	9%		-	-	PROSITE_profiles:PS50835&hm	-		-	-	-
	1	185891512	C	T	missense	ENST00000271588.4:c.902C>T	ENSP00000271588.4:p.Thr301Ile	T5a	65	6	8%		-	-	hmmpanther:PTHR19897:SF158	-		tolerated	T	T
MDN1	6	90497672	G	T	missense	ENST00000369393.3:c.1235C>A	ENSP00000358400.3:p.Ser412Tyr	T2	71	5	7%		-	-	hmmpanther:PTHR22908:SF588	-		deleterious	T	T
	6	90472182	G	C	missense	ENST00000369393.3:c.2212C>G	ENSP00000358400.3:p.Gln738Glu	T3	93	5	5%		-	-	hmmpanther:PTHR22908:SF588	-		tolerated	T	T
	6	90402239	G	A	missense	ENST00000369393.3:c.10510C>T	ENSP00000358400.3:p.Leu3504Phe	T7	39	4	9%		-	-	hmmpanther:PTHR22908:SF588	-		deleterious	T	T
PKHD1L1	8	110432873	A	T	missense	ENST00000378402.5:c.2651A>T	ENSP00000367655.5:p.Tyr884Phe	T1	127	8	6%		-	-	hmmpanther:PTHR11915&hmn	-		tolerated	T	T
	8	110535503	G	T	missense	ENST00000378402.5:c.12372G>T	ENSP00000367655.5:p.Arg4124Ser	T2	57	3	5%		-	-		-		tolerated	T	T
	8	110445445	G	T	stop gained	ENST00000378402.5:c.3340G>T	ENSP00000367655.5:p.Glu1114Ter	T3	81	6	7%		-	-	hmmpanther:PTHR11915&hmn	-		-	-	-
PPP2R3A	3	135722018	G	T	missense	ENST00000264977.3:c.1678G>T	ENSP00000264977.3:p.Val560Phe	T1	28	4	13%		-	-	hmmpanther:PTHR14095&hmn	-	tolerated_low_confi		T	T
	3	135821011	G	T	missense	ENST00000264977.3:c.3090G>T	ENSP00000264977.3:p.Lys1030Asn	T3	252	12	5%		-	-	hmmpanther:PTHR14095&hmn	-		deleterious	T	T
	3	135721626	A	G	missense	ENST00000264977.3:c.1286A>G	ENSP00000264977.3:p.Lys429Arg	T4	36	5	12%		-	-	hmmpanther:PTHR14095&hmn	-	tolerated_low_confi		T	T
PTPRS	19	5212020	C	A	missense	ENST00000357368.4:c.5011G>T	ENSP00000349932.4:p.Val1671Leu	T1	65	6	8%		-	-	Superfamily_domains:SSF5279	-		tolerated	T	T
	19	5220096	G	T	missense	ENST00000357368.4:c.3619C>A	ENSP00000349932.4:p.Pro1207Thr	T3	55	6	10%		-	-	hmmpanther:PTHR19134:SF204	-		deleterious	T	T
	19	5286108	G	A	missense	ENST00000357368.4:c.44C>T	ENSP00000349932.4:p.Pro15Leu	T8	31	8	21%		-	-	Low_complexity_(Seg):seg&Tr	-	tolerated_low_confi		T	T
SLC4A7	3	27436157	A	T	missense	ENST00000295736.5:c.2942T>A	ENSP00000295736.5:p.Ile981Lys	T3	131	7	5%		-	-	hmmpanther:PTHR11453&hmn	-		deleterious	D	D
	3	27462214	T	C	missense	ENST00000295736.5:c.1462A>G	ENSP00000295736.5:p.Ile488Val	T4	139	8	6%	C:4.118e-05	-	5	hmmpanther:PTHR11453&hmn	-		tolerated	D	D
	3	27444771	T	C	missense	ENST00000295736.5:c.2153A>G	ENSP00000295736.5:p.Tyr718Cys	T5c	41	3	7%		-	-	Transmembrane_helices:TMhe	-		deleterious	D	D
AGAP2	12	58126185	C	T	splicing	ENST00000547588.1:c.1794+1G>A	ENST00000547588.1:c.1794+1G>A	T3	94	8	8%		-	-		-		-	-	-
	12	58125635	C	T	missense	ENST00000547588.1:c.1910G>A	ENSP00000449241.1:p.Gly637Glu	T7	64	5	7%		-	-	hmmpanther:PTHR23180&hmn	-		tolerated	T	T
DENND2C	1	115166127	C	A	splicing	ENST00000393274.1:c.943+1G>T	ENST00000393274.1:c.943+1G>T	T1	93	8	8%		-	-		-		-	-	-
	1	115130518	C	A	missense	ENST00000357232.4:c.2234G>T	ENSP00000349768.3:p.Gly745Val	T3	113	8	7%		-	-	PROSITE_profiles:PS50268&hm	-		deleterious	D	T
ITGB8	7	20418832	G	T	missense	ENST00000222573.4:c.547G>T	ENSP00000222573.3:p.Asp183Tyr	T1	108	10	8%		-	-	hmmpanther:PTHR10082:SF9&	-		deleterious	D	D
	7	20403328	G	T	stop_gained	ENST00000222573.4:c.196G>T	ENSP00000222573.3:p.Gly66Ter	T3	59	10	14%		-	-	hmmpanther:PTHR10082:SF9&	-		-	-	-
KIAA1432	9	5770197	C	A	missense	ENST00000414202.2:c.3535C>A	ENSP00000416696.2:p.Pro1179Thr	T3	130	10	7%		-	-	hmmpanther:PTHR22746&hmn	-		tolerated	T	T
	9	5747307	C	G	missense	ENST00000414202.2:c.1254C>G	ENSP00000416696.2:p.Asn418Lys	T4	114	7	6%		2	75	hmmpanther:PTHR22746&hmn	-		deleterious	T	T
MACF1	1	39748002	C	A	stop_gained	ENST00000545844.1:c.666C>A	ENSP00000439537.1:p.Cys222Ter	T3	63	7	10%		-	-	Superfamily_domains:SSF4757	-		-	-	-
	1	39913744	G	T	missense	ENST00000545844.1:c.13831G>T	ENSP00000439537.1:p.Asp4611Tyr	T4	113	6	5%		-	-	SMART_domains:SM00150&Ge	-		deleterious	T	T
MYH9	22	36684296	A	T	splicing	ENST00000216181.5:c.4932+2T>A		T1	98	7	7%		-	-		-		-	-	-
	22	36697073	G	A	stop_gained	ENST00000216181.5:c.2662C>T	ENSP00000216181.5:p.Gln888Ter	T5c	63	6	9%		-	-	Low_complexity_(Seg):seg&Cc	-		-	-	-
OBSCN	1	228527747	G	T	missense	ENST00000570156.2:c.20231G>T	ENSP00000455507.2:p.Gly6744Val	T1	53	3	5%		-	-	PROSITE_profiles:PS50010&hm	-		deleterious	T	T
	1	228539135	G	A	missense	ENST00000570156.2:c.21404G>A	ENSP00000455507.2:p.Gly7135Glu	T8	40	12	23%		-	1	PROSITE_profiles:PS50835&hm	-		deleterious	D	D
PHOSPHO1	17	47304109	C	T	5'UTR	ENST00000413580.1:c.-108G>A		T4	79	9	10%		-	-		-		-	-	-
	17	47304058	G	A	5'UTR	ENST00000413580.1:c.-57C>T		T8	132	15	10%		-	-		-		-	-	-
RNF213	17	78306369	C	A	missense	ENST00000582970.1:c.4081C>A	ENSP00000464087.1:p.Gln1361Lys	T1	62	8	11%		-	-	hmmpanther:PTHR22605&hmn	-		-	T	T
	17	78337082	C	T	missense	ENST00000582970.1:c.11536C>T	ENSP00000464087.1:p.Arg3846Cys	T5a	37	4	10%		-	1	hmmpanther:PTHR22605&hmn	-		-	T	T
RREB1	6	7229237	G	T	missense	ENST00000379938.2:c.905G>T	ENSP00000369270.2:p.Cys302Phe	T3	135	4	10%		-	-	hmmpanther:PTHR23233&hmn	-		deleterious	T	T
	6	7240820	T	C	missense	ENST00000379938.2:c.3958T>C	ENSP00000369270.2:p.Ser1320Pro	T6	109	7	6%		-	-	hmmpanther:PTHR23233&hmn	-		tolerated	T	T
SCN10A	3	38755567	G	T	stop_gained	ENST00000449082.2:c.3686C>A	ENSP00000390600.2:p.Ser1229Ter	T1	54	6	10%		-	-	Transmembrane_helices:TMhe	-		-	-	-
	3	38766806	C	A	splicing	ENST00000449082.2:c.3088-1G>T		T3	68	6	11%		-	-		-		-	-	-

End note: Highlighted sites are recurrent mutation with same site on same gene in different PNF (Ratner), DNF (Gosline) and PNF (Pemov) samples.

Yellow: Overlapped mutated sites between PNF (Ratner) and DNF (Gosline)

Supplementary Table 5: Overlapped gemline variant genes in DNF data

Variants with deleterious, damaging, and/or pathogenic effect (see Suppl. Table 9)

Symbol	chr	LOC	Ref	Alt	Consequences	Impact	GMAF	ExAC_AC	Patient	HGVSc	HGVSp	ClinVar (2019)	SIFT	MetaLR	MetaSVM
AC079612.1	chr2	240498659	C	T	upstream_gene_varian	MODIFIEI	T:0.0008		patient8_tumor_syn4985	-	-		-	-	-
AC079612.1	chr2	240505863	G	A	3_prime_UTR_variant	MODIFIEI	A:0.0008		patient6_tumor_syn4985	ENST00000358775.1:c.*1059G>A	-		-	-	-
ANKK1	chr11	113270124	G	A	missense_variant	MODERA	A:0.0072	A:3.253e-03	patient2_tumor_syn4985	ENST00000303941.3:c.1433G>A	ENSP00000306678.3:p.Arg478Gln		tolerated	T	T
ANKK1	chr11	113270178	C	A	missense_variant	MODERA	A:0.0072	A:3.272e-03	patient2_tumor_syn4985	ENST00000303941.3:c.1487C>A	ENSP00000306678.3:p.Thr496Asn		deleterious	D	D
ANKRD30A	chr10	37455581	G	A	missense_variant	MODERATE		A:7.236e-05	patient6_tumor_syn4985	ENST00000361713.1:c.1945G>A	ENSP00000354432.1:p.Ala649Thr		tolerated	T	T
ANKRD30A	chr10	37481992	A	G	missense_variant&spli	MODERA	G:0.0042	G:1.838e-03	patient2_tumor_syn4985	ENST00000361713.1:c.2345A>G	ENSP00000354432.1:p.Glu782Gly		deleterious	T	T
ATM	chr11	108153488	A	G	missense_variant	MODERATE		G:3.295e-05	patient11_tumor_syn498	ENST00000278616.4:c.3628A>G	ENSP00000278616.4:p.Met1210Val	Uncertain_significance	tolerated	T	T
ATM	chr11	108188136	G	A	missense_variant	MODERA	A:0.0028	A:2.570e-03	patient8_tumor_syn4984	ENST00000278616.4:c.6235G>A	ENSP00000278616.4:p.Val2079Ile	Benign/Likely_benign	tolerated	T	T
ATM	chr11	108198391	T	C	missense_variant	MODERA	C:0.0072		patient8_tumor_syn4985	ENST00000278616.4:c.6995T>C	ENSP00000278616.4:p.Leu2332Pro	Benign/Likely_benign	tolerated	T	T
ATM	chr11	108236279	A	G	3_prime_UTR_variant	MODIFIEI	G:0.0028	G:2.398e-03	patient8_tumor_syn4985	ENST00000278616.4:c.*44A>G	-	Uncertain_significance	-	-	-
ATM	chr11	108236471	C	T	3_prime_UTR_variant	MODIFIEI	T:0.0028		patient8_tumor_syn4984	ENST00000278616.4:c.*236C>T	-	Uncertain_significance	-	-	-
CELSR2	chr1	109794593	C	T	missense_variant	MODERA	T:0.0012	T:3.904e-03	patient1_tumor_syn4985	ENST00000271332.3:c.1892C>T	ENSP00000271332.3:p.Thr631Met		deleterious	T	T
CELSR2	chr1	109807571	C	T	missense_variant	MODERA	T:0.0056	T:1.540e-03	patient11_tumor_syn498	ENST00000271332.3:c.5546C>T	ENSP00000271332.3:p.Ala1849Val		tolerated	D	T
CELSR2	chr1	109816898	C	G	3_prime_UTR_variant	MODIFIEI	G:0.0058		patient3_tumor_syn4987	ENST00000271332.3:c.*227C>G	-		-	-	-
CELSR2	chr1	109817633	A	G	3_prime_UTR_variant	MODIFIEI	G:0.0054		patient3_tumor_syn4987	ENST00000271332.3:c.*962A>G	-		-	-	-
COL14A1	chr8	121267490	G	A	missense_variant	MODERA	A:0.0028	A:5.247e-03	patient3_tumor_syn4987	ENST00000297848.3:c.2764G>A	ENSP00000297848.3:p.Val922Ile		tolerated	T	T
COL4A2	chr13	111114519	C	A	missense_variant	MODERA	A:0.0054	A:1.506e-03	patient8_tumor_syn4985	ENST00000360467.5:c.1655C>A	ENSP00000353654.5:p.Thr552Lys	Likely_benign	tolerated	T	T
COL4A2	chr13	111164467	G	A	missense_variant	MODERA	A:0.0010	A:5.370e-04&T:8.26	patient8_tumor_syn4985	ENST00000360467.5:c.5068G>A	ENSP00000353654.5:p.Ala1690Thr	Likely_benign	tolerated	D	T
COL4A2	chr13	111164563	T	C	3_prime_UTR_variant	MODIFIEI	C:0.0088	C:2.905e-03	patient8_tumor_syn4984	ENST00000360467.5:c.*25T>C	-	Likely_benign	-	-	-
CUBN	chr10	16893406	G	C	missense_variant	MODERA	C:0.0078	A:3.295e-05&C:1.99	patient8_tumor_syn4985	ENST00000377833.4:c.9491C>G	ENSP00000367064.4:p.Ser3164Trp	Uncertain_significance	deleterious	T	T
CUBN	chr10	16960624	T	C	missense_variant	MODERA	C:0.0028	C:1.063e-03	patient11_tumor_syn498	ENST00000377833.4:c.6997A>G	ENSP00000367064.4:p.Ile2333Val	Uncertain_significance	tolerated	T	T
CUBN	chr10	17153023	C	T	missense_variant	MODERA	T:0.0092	T:4.110e-03	patient2_tumor_syn4987	ENST00000377833.4:c.910G>A	ENSP00000367064.4:p.Glu304Lys	Uncertain_significance	deleterious	D	T
EML6	chr2	55197958	T	C	3_prime_UTR_variant	MODIFIEI	C:0.0036		patient3_tumor_syn4987	ENST00000356458.6:c.*727T>C	-		-	-	-
EML6	chr2	55198728	A	G	3_prime_UTR_variant	MODIFIEI	G:0.0030		patient9_tumor_syn4985	ENST00000356458.6:c.*1497A>G	-		-	-	-
FCGBP	chr19	40367555	C	T	missense_variant	MODERA	T:0.0098	T:3.417e-03	patient3_tumor_syn4987	ENST00000221347.6:c.13405G>A	ENSP00000221347.5:p.Val4469Ile		tolerated	T	T
FCGBP	chr19	40430459	G	A	missense_variant	MODERA	A:0.0002	A:2.638e-04	patient9_tumor_syn4985	ENST00000221347.6:c.1484C>T	ENSP00000221347.5:p.Ser495Leu		tolerated	T	T
KIFC1	chr6	33359736	C	G	5_prime_UTR_variant	MODIFIEI	G:0.0056	G:8.602e-03	patient4_tumor_syn4985	ENST00000428849.2:c.-27C>G	-		-	-	-
KIFC1	chr6	33371806	G	A	missense_variant	MODERA	A:0.0090	A:8.945e-03	patient6_tumor_syn4985	ENST00000428849.2:c.656G>A	ENSP00000393963.2:p.Arg219Gln		tolerated	T	T
KIFC1	chr6	33371806	G	A	missense_variant	MODERA	A:0.0090	A:8.945e-03	patient4_tumor_syn4985	ENST00000428849.2:c.656G>A	ENSP00000393963.2:p.Arg219Gln		tolerated	T	T
KIFC1	chr6	33372689	C	T	missense_variant	MODERA	T:0.0016	T:2.883e-03	patient4_tumor_syn4985	ENST00000428849.2:c.817C>T	ENSP00000393963.2:p.Arg273Trp		deleterious	T	T
MUC16	chr19	9014194	A	T	missense_variant	MODERATE		T:3.307e-05	patient11_tumor_syn498	ENST00000397910.4:c.38454T>A	ENSP00000381008.2:p.His12818Gln		-	T	T
MUC16	chr19	9010991	C	T	missense_variant	MODERATE		T:1.655e-05&G:8.27	patient6_tumor_syn4985	ENST00000397910.4:c.38927G>A	ENSP00000381008.2:p.Ser12976Asn		-	T	T
MUC16	chr19	9010991	C	T	missense_variant	MODERATE		T:1.655e-05&G:8.27	patient2_tumor_syn4985	ENST00000397910.4:c.38927G>A	ENSP00000381008.2:p.Ser12976Asn		-	T	T
MUC16	chr19	9009614	T	C	missense_variant	MODERATE		C:4.961e-05&G:8.26	patient4_tumor_syn4985	ENST00000397910.4:c.39112A>G	ENSP00000381008.2:p.Thr13038Ala		-	T	T
MUC16	chr19	8993008	G	C	missense_variant&spli	MODERA	C:0.0092	C:2.970e-03	patient8_tumor_syn4985	ENST00000397910.4:c.41751C>G	ENSP00000381008.2:p.Ser13917Arg		-	T	T
NF1	chr17	29486054	A	T	missense_variant	MODERA	G:0.0002	T:3.295e-05&G:8.23	patient8_tumor_syn4985	ENST00000358273.4:c.231A>T	ENSP00000351015.4:p.Lys77Asn	Uncertain_significance	deleterious	T	T
NF1	chr17	29497003	C	T	stop_gained	HIGH		T:8.238e-06	patient8_tumor_syn4985	ENST00000358273.4:c.574C>T	ENSP00000351015.4:p.Arg192Ter	Pathogenic	-	-	-
NF1	chr17	29541542	A	G	missense_variant	MODERATE		G:1.647e-05	patient5_tumor_syn4985	ENST00000358273.4:c.1466A>G	ENSP00000351015.4:p.Tyr489Cys	Pathogenic	tolerated	T	T
NF1	chr17	29588751	C	T	stop_gained	HIGH		T:8.236e-06	patient9_tumor_syn4985	ENST00000358273.4:c.4600C>T	ENSP00000351015.4:p.Arg1534Ter	Pathogenic/Likely_path	-	-	-
NF1	chr17	29702631	A	G	3_prime_UTR_variant	MODIFIEI	G:0.0094		patient11_tumor_syn498	ENST00000358273.4:c.*1458A>G	-		-	-	-
OBSCN	chr1	228404401	A	G	missense_variant&spli	MODERA	G:0.0002	G:9.381e-05	patient11_tumor_syn498	ENST00000570156.2:c.2375A>G	ENSP00000455507.2:p.Gln792Arg		tolerated	T	T
OBSCN	chr1	228468437	G	A	missense_variant	MODERA	A:0.0034	A:1.034e-03	patient11_tumor_syn498	ENST00000570156.2:c.9424G>A	ENSP00000455507.2:p.Ala3142Thr		tolerated	T	T
OBSCN	chr1	228470764	C	T	missense_variant	MODERA	T:0.0092	T:2.089e-03	patient11_tumor_syn498	ENST00000570156.2:c.9803C>T	ENSP00000455507.2:p.Ala3268Val		tolerated	T	T
OBSCN	chr1	228481916	C	T	missense_variant	MODERATE		A:8.256e-06&T:2.47	patient10_tumor_syn498	ENST00000570156.2:c.12482C>T	ENSP00000455507.2:p.Thr4161Met		tolerated	T	T
OBSCN	chr1	228509861	C	T	missense_variant	MODERA	T:0.0002	T:1.652e-04	patient11_tumor_syn498	ENST00000570156.2:c.18190C>T	ENSP00000455507.2:p.Pro6064Ser		tolerated	T	T
OBSCN	chr1	228511241	G	A	missense_variant	MODERATE		A:4.948e-05	patient3_tumor_syn4987	ENST00000570156.2:c.18457G>A	ENSP00000455507.2:p.Val6153Ile		deleterious	T	T
OBSCN	chr1	228557680	C	A	missense_variant	MODERA	A:0.0010	A:1.901e-04	patient2_tumor_syn4987	ENST00000570156.2:c.22876C>A	ENSP00000455507.2:p.Arg7626Ser		deleterious	T	T
OBSCN	chr1	228560317	G	A	missense_variant	MODERATE		A:8.567e-04	patient4_tumor_syn4985	ENST00000570156.2:c.24709G>A	ENSP00000455507.2:p.Val8237Met		deleterious	T	T
PASK	chr2	242063458	C	T	missense_variant	MODERA	T:0.0032	A:8.236e-06&T:9.44	patient4_tumor_syn4985	ENST00000358649.4:c.2810G>A	ENSP00000351475.4:p.Arg937His		deleterious	T	T
PKHD1L1	chr8	110418593	A	G	missense_variant	MODERA	G:0.0020	G:3.450e-04	patient2_tumor_syn4987	ENST00000378402.5:c.1699A>G	ENSP00000367655.5:p.Ile567Val		tolerated	T	T
PKHD1L1	chr8	110432873	A	G	missense_variant	MODERA	G:0.0014	G:2.185e-03	patient1_tumor_syn4985	ENST00000378402.5:c.2651A>G	ENSP00000367655.5:p.Tyr884Cys		tolerated	T	T
PKHD1L1	chr8	110450593	G	T	missense_variant	MODERA	T:0.0054	T:8.477e-03	patient4_tumor_syn4985	ENST00000378402.5:c.3668G>T	ENSP00000367655.5:p.Gly1223Val		deleterious	T	T
PKHD1L1	chr8	110450593	G	T	missense_variant	MODERA	T:0.0054	T:8.477e-03	patient3_tumor_syn4987	ENST00000378402.5:c.3668G>T	ENSP00000367655.5:p.Gly1223Val		deleterious	T	T
PKHD1L1	chr8	110477066	C	T	stop_gained	HIGH	T:0.0008	T:1.333e-03	patient9_tumor_syn4985	ENST00000378402.5:c.8005C>T	ENSP00000367655.5:p.Gln2669Ter		-	-	-
RGPD4	chr2	108443503	G	C	missense_variant	MODERA	C:0.0078	A:8.326e-06&C:7.80	patient4_tumor_syn4985	ENST00000408999.3:c.34G>C	ENSP00000386810.3:p.Val12Leu		deleterious	T	T
SLC14A2	chr18	43195469	G	A	5_prime_UTR_variant	MODIFIEI	A:0.0028		patient9_tumor_syn4985	ENST00000255226.6:c.-113G>A	-		-	-	-
SLC14A2	chr18	43262408	G	A	missense_variant	MODERA	A:0.0052	A:8.195e-03&T:8.23	patient8_tumor_syn4985	ENST00000255226.6:c.2687G>A	ENSP00000255226.5:p.Arg896His		deleterious	T	T

End note: Highlighted sites are recurrent mutations showed in different PNF and DNF samples

End note: Highlighted sites are recurrent mutation with same site on same gene in different PNF (Ratner), DNF (Gosline) and PNF (Pemov) samples.

Yellow: Overlapped mutated sites between PNF (Ratner) and DNF (Gosline)

Orange: Overlapped mutated sites between three datasets.

Supplementary Table 6: Overlapped somatic variant genes in DNF data

												Variants with deleterious, damaging, and/or pathogenic effect			
Symbol	chr	LOC	Ref	Alt	Consequences	Impact	GMAF	Patient	HGVSc	HGVSp	COSMIC	ClinVar	SIFT	MetaLR	MetaSVM
DMXL2	chr15	51798638	G	T	intron_variant	MODIFIER	T:0.0092	patient6_	ENST00000543779.2:c.2764+693C>A		-	-	-	-	-
MACF1	chr1	39597015	T	A	intron_variant	MODIFIER	A:0.0034	patient10	ENST00000545844.1:c.220+46905T>A		-	-	-	-	-
MACF1	chr1	39597015	T	A	intron_variant	MODIFIER	A:0.0034	patient9_	ENST00000545844.1:c.220+46905T>A		-	-	-	-	-
MACF1	chr1	39597015	T	A	intron_variant	MODIFIER	A:0.0034	patient9_	ENST00000545844.1:c.220+46905T>A		-	-	-	-	-
MYH9	chr22	36673329	C	T	downstream_g	MODIFIER	T:0.0054	patient8_	-		-	-	-	-	-

Supplemental Table7: Overlapped germline variant genes in PNF (Pemov et al, Re-analyzed with VarScan2)

SYMBOL	Patient	CHROM	POS	REF	ALT	Consequence	IMPACT	HGVSc	HGVSp	GMAF	ExAC_MAF	Existing_variation	ClinVar (2019)	SIFT	MetaLR	MetaSVM
ANKK1	p9	chr11	113266080	G	T	splice_acceptor_vari	HIGH	ENST00000303941.3:c.633-1G>T		T:0.0096	T:2.440e-03	rs139841238				
ANKRD30A	p23	chr10	37419292	G	T	stop_gained	HIGH	ENST00000361713.1:c.328G>T	ENSP00000354432.1:p.Glu110Ter	T:0.0060	A:1.654e-05&T:0.012	rs116939015				
	p24	chr10	37419292	G	T	stop_gained	HIGH	ENST00000361713.1:c.328G>T	ENSP00000354432.1:p.Glu110Ter	T:0.0060	A:1.654e-05&T:0.012	rs116939015				
	p9	chr10	37433929	G	T	missense_variant	MODERA	ENST00000361713.1:c.1232G>T	ENSP00000354432.1:p.Arg411Met		T:8.032e-04	rs202149101		T&T	T	T
	p10	chr10	37478440	G	T	missense_variant	MODERA	ENST00000361713.1:c.2299G>T	ENSP00000354432.1:p.Ala767Ser	T:0.0040	T:4.118e-03	rs201628233&COSM4144666&COSM4144667		T&D	T	T
	p6	chr10	37508256	G	C	missense_variant	MODERA	ENST00000361713.1:c.3448G>C	ENSP00000354432.1:p.Val1150Leu			COSM3751773&COSM3751774		T&T	T	T
ATM	p27	chr11	108098576	C	G	missense_variant	MODERA	ENST00000278616.4:c.146C>G	ENSP00000278616.4:p.Ser49Cys	G:0.0042	G:7.364e-03	rs1800054	Benign/Likely_benign	D&D&D&D	T	T
	p5	chr11	108123621	T	G	missense_variant	MODERA	ENST00000278616.4:c.1880T>G	ENSP00000278616.4:p.Phe627Cys	G:0.0002	G:1.648e-05	rs546087885	Uncertain_significance	T&T&T	T	T
	p8	chr11	108127035	G	A	missense_variant	MODERA	ENST00000278616.4:c.2218G>A	ENSP00000278616.4:p.Ala740Thr				Uncertain_significance	D&D&D	D	D
	p16	chr11	108198394	C	A	missense_variant	MODERA	ENST00000278616.4:c.6998C>A	ENSP00000278616.4:p.Thr2333Lys		A:6.589e-05&T:4.118e-(	rs150503164	Conflicting_interpreta	D&D	T	T
	p8	chr11	108201023	T	C	missense_variant	MODERA	ENST00000278616.4:c.7390T>C	ENSP00000278616.4:p.Cys2464Arg		C:5.683e-04	rs55801750&COSM758329	Conflicting_interpreta	T&T	T	T
CELSR2	p22	chr1	109793126	G	A	missense_variant	MODERA	ENST00000271332.3:c.425G>A	ENSP00000271332.3:p.Arg142Gln					T	T	T
	p10	chr1	109803697	G	A	missense_variant	MODERA	ENST00000271332.3:c.3992G>A	ENSP00000271332.3:p.Arg1331His	A:0.0012	A:1.145e-03	rs115077620		D	T	T
	p8	chr1	109810200	G	A	missense_variant	MODERA	ENST00000271332.3:c.6044G>A	ENSP00000271332.3:p.Arg2015Lys	A:0.0096	A:0.022	rs72703203		T	T	T
	p7	chr1	109813223	G	T	splice_donor_varian	HIGH	ENST00000271332.3:c.7483+1G>T								
COL14A1	p29	chr8	121256206	C	T	missense_variant	MODERA	ENST00000297848.3:c.2438C>T	ENSP00000297848.3:p.Thr813Met		T:4.695e-04	rs148241340		T&T&T&T	T	T
COL4A2	p8	chr13	111118344	C	T	splice_region_variar	LOW	ENST00000360467.5:c.1979-6C>T		T:0.0026	T:4.532e-03	rs190632602	Likely_benign			
CUBN	p25	chr10	16870912	G	T	missense_variant	MODERA	ENST00000377833.4:c.10656C>A	ENSP00000367064.4:p.Asn3552Lys				Likely_benign	T&T	T	T
	p17	chr10	16882363	G	A	missense_variant	MODERA	ENST00000377833.4:c.9998C>T	ENSP00000367064.4:p.Thr3333Met		A:3.295e-05&T:1.647e-(	rs140637500		T&T	T	T
	p9	chr10	16893406	G	C	missense_variant	MODERA	ENST00000377833.4:c.9491C>G	ENSP00000367064.4:p.Ser3164Trp	C:0.0078	A:3.295e-05&C:1.993e-	rs57163243	Uncertain_significance	D	T	T
	p27	chr10	16932384	G	A	missense_variant	MODERA	ENST00000377833.4:c.8741C>T	ENSP00000367064.4:p.Ala2914Val	A:0.0064	A:0.012	rs45551835&COSM32707	Benign/Likely_benign	D	T	T
	p12	chr10	16942831	C	A	missense_variant	MODERA	ENST00000377833.4:c.8203G>T	ENSP00000367064.4:p.Asp2735Tyr		A:2.224e-04	rs149802222	Uncertain_significance	D	T	T
	p11	chr10	16955978	A	T	missense_variant	MODERA	ENST00000377833.4:c.7365T>A	ENSP00000367064.4:p.Asp2455Glu	T:0.0026	T:5.650e-03	rs117128556	Uncertain_significance	T	T	T
	p8	chr10	16967417	T	C	missense_variant	MODERA	ENST00000377833.4:c.6469A>G	ENSP00000367064.4:p.Asn2157Asp	C:0.0032	C:5.610e-03	rs144360241		T	T	T
	p9	chr10	16981015	T	C	missense_variant	MODERA	ENST00000377833.4:c.5680A>G	ENSP00000367064.4:p.Ile1894Val	C:0.0050	C:1.293e-03	rs75789390	Uncertain_significance	T	T	T
	p16	chr10	17085903	C	G	missense_variant	MODERA	ENST00000377833.4:c.3752G>C	ENSP00000367064.4:p.Ser1251Thr	G:0.0064	G:1.754e-03	rs115048360	Uncertain_significance	T	T	T
FCGBP	p23	chr19	40354279	G	A	missense_variant	MODERA	ENST00000221347.6:c.16190C>T	ENSP00000221347.5:p.Ala5397Val		A:1.483e-04	rs199845989		T	T	T
	p24	chr19	40354279	G	A	missense_variant	MODERA	ENST00000221347.6:c.16190C>T	ENSP00000221347.5:p.Ala5397Val		A:1.483e-04	rs199845989		T	T	T
	p13	chr19	40357636	C	T	missense_variant	MODERA	ENST00000221347.6:c.15677G>A	ENSP00000221347.5:p.Arg5226Gln	T:0.0002	T:8.236e-05	rs202029040		T	T	T
	p16	chr19	40376811	C	T	missense_variant	MODERA	ENST00000221347.6:c.11611G>A	ENSP00000221347.5:p.Gly3871Arg		T:7.692e-03	rs4802062		D	T	T
	p19	chr19	40376811	C	T	missense_variant	MODERA	ENST00000221347.6:c.11611G>A	ENSP00000221347.5:p.Gly3871Arg		T:7.692e-03	rs4802062		D	T	T
	p22	chr19	40376811	C	T	missense_variant	MODERA	ENST00000221347.6:c.11611G>A	ENSP00000221347.5:p.Gly3871Arg		T:7.692e-03	rs4802062		D	T	T
	p17	chr19	40377117	C	T	missense_variant	MODERA	ENST00000221347.6:c.11305G>A	ENSP00000221347.5:p.Ala3769Thr	T:0.0004	T:1.226e-04	rs534122657		T	T	T
	p22	chr19	40392008	G	A	missense_variant	MODERA	ENST00000221347.6:c.8378C>T	ENSP00000221347.5:p.Ala2793Val		A:6.611e-05	rs201020361&COSM3823077		T	T	T
	p25	chr19	40392008	G	A	missense_variant	MODERA	ENST00000221347.6:c.8378C>T	ENSP00000221347.5:p.Ala2793Val		A:6.611e-05	rs201020361&COSM3823077		T	T	T
	p27	chr19	40392008	G	A	missense_variant	MODERA	ENST00000221347.6:c.8378C>T	ENSP00000221347.5:p.Ala2793Val		A:6.611e-05	rs201020361&COSM3823077		T	T	T
	p29	chr19	40392008	G	A	missense_variant	MODERA	ENST00000221347.6:c.8378C>T	ENSP00000221347.5:p.Ala2793Val		A:6.611e-05	rs201020361&COSM3823077		T	T	T
	p11	chr19	40392339	G	T	missense_variant	MODERA	ENST00000221347.6:c.8165C>A	ENSP00000221347.5:p.Pro2722His		A:1.395e-04&T:2.092e-(	rs782699713		D	T	T
	p16	chr19	40399385	C	T	missense_variant	MODERA	ENST00000221347.6:c.6310G>A	ENSP00000221347.5:p.Gly2104Arg		T:9.734e-03	rs201168964&COSM1481055		T	T	T
	p24	chr19	40399385	C	T	missense_variant	MODERA	ENST00000221347.6:c.6310G>A	ENSP00000221347.5:p.Gly2104Arg		T:9.734e-03	rs201168964&COSM1481055		T	T	T
	p7	chr19	40399385	C	T	missense_variant	MODERA	ENST00000221347.6:c.6310G>A	ENSP00000221347.5:p.Gly2104Arg		T:9.734e-03	rs201168964&COSM1481055		T	T	T
	p25	chr19	40402235	G	C	missense_variant	MODERA	ENST00000221347.6:c.5164C>G	ENSP00000221347.5:p.Gln1722Glu		C:0.029	rs7247426		T	T	T
	p16	chr19	40408806	C	T	missense_variant	MODERA	ENST00000221347.6:c.4033G>A	ENSP00000221347.5:p.Gly13455er	T:0.0004	T:9.890e-05&T:9.890e-(	rs760171441&rs148703638		T	T	T
	p16	chr19	40411643	C	T	missense_variant	MODERA	ENST00000221347.6:c.3985G>A	ENSP00000221347.5:p.Val1329Ile	T:0.0012	T:6.919e-04	rs146825162		T	T	T
	p11	chr19	40420097	C	A	missense_variant	MODERA	ENST00000221347.6:c.2897G>T	ENSP00000221347.5:p.Gly966Val	A:0.0010	A:6.177e-04	rs144899113		D	T	T
KIFC1	p25	chr6	33371778	C	A	missense_variant	MODERA	ENST00000428849.2:c.628C>A	ENSP00000393963.2:p.Arg210Ser					T	T	T
MAST4	p6	chr5	66055573	C	T	missense_variant	MODERA	ENST00000403625.2:c.400C>T	ENSP00000385727.1:p.Arg134Trp		T:9.099e-04	rs200785678		D&D&D&D&D	D	D
	p16	chr5	66449319	G	T	stop_gained	HIGH	ENST00000403625.2:c.3559G>T	ENSP00000385727.1:p.Gly1187Ter							
	p10	chr5	66460352	G	A	missense_variant	MODERA	ENST00000403625.2:c.5345G>A	ENSP00000385727.1:p.Arg1782Lys	A:0.0016	A:4.791e-03	rs114551553		D&D&D&D&D	T	T
	p17	chr5	66460390	G	A	missense_variant	MODERA	ENST00000403625.2:c.5383G>A	ENSP00000385727.1:p.Val1795Ile	A:0.0098	A:0.019&T:1.655e-05	rs17221458			T	T
	p21	chr5	66460390	G	A	missense_variant	MODERA	ENST00000403625.2:c.5383G>A	ENSP00000385727.1:p.Val1795Ile	A:0.0098	A:0.019&T:1.655e-05	rs17221458			T	T
	p22	chr5	66460390	G	A	missense_variant	MODERA	ENST00000403625.2:c.5383G>A	ENSP00000385727.1:p.Val1795Ile	A:0.0098	A:0.019&T:1.655e-05	rs17221458			T	T
	p9	chr5	66461713	A	G	missense_variant	MODERA	ENST00000403625.2:c.6706A>G	ENSP00000385727.1:p.Ser2236Gly	G:0.0058	G:1.680e-03	rs56027793		T&T&T&T&T	T	T
	p12	chr5	66462640	A	G	missense_variant	MODERA	ENST00000403625.2:c.7633A>G	ENSP00000385727.1:p.Ser2545Gly		G:8.289e-06	rs758737257			T	T
MROH2B	p9	chr5	41008777	G	A	missense_variant	MODERA	ENST00000399564.4:c.3539C>T	ENSP00000382476.4:p.Thr1180Ile	A:0.0086	A:2.000e-03	rs34245494		D&D&D	T	T
	p16	chr5	41047836	G	A	missense_variant	MODERA	ENST00000399564.4:c.1715C>T	ENSP00000382476.4:p.Thr572Ile		A:2.492e-05	rs375041927		D&D&D	T	T
NF1	p6	chr17	29496957	T	A	missense_variant	MODERA	ENST00000358273.4:c.528T>A	ENSP00000351015.4:p.Asp176Glu	A:0.0008	A:3.270e-03	rs112306990&COSM24498	Conflicting_interpreta	T&T&T	T	T
	p17	chr17	29496994	A	T	stop_gained	HIGH	ENST00000358273.4:c.565A>T	ENSP00000351015.4:p.Lys189Ter							
	p29	chr17	29528480	T	C	missense_variant	MODERA	ENST00000358273.4:c.1237T>C	ENSP00000351015.4:p.Ser413Pro				Pathogenic	D&D&D&D	D	D
	p6	chr17	29541596	T	TG	frameshift_variant	HIGH	ENST00000358273.4:c.1521dupG	ENSP00000351015.4:p.Leu508AlafsTer3							
	p12	chr17	29546122	C	T	stop_gained	HIGH	ENST00000358273.4:c.1627C>T	ENSP00000351015.4:p.Gln543Ter				Pathogenic			
	p13	chr17	29550488	A	G	missense_variant	MODERA	ENST00000358273.4:c.1748A>G	ENSP00000351015.4:p.Lys583Arg			rs199474760&COSM36863	Pathogenic/Likely_pat	T&T&T	T	T
	p7	chr17	29554235	G	C	splice_acceptor_vari	HIGH	ENST00000358273.4:c.2252-1G>C				rs587781577	Likely_pathogenic			
	p10	chr17	29556079	C	T	stop_gained	HIGH	ENST00000358273.4:c.2446C>T	ENSP00000351015.4:p.Arg816Ter			COSM24444	Pathogenic			
	p21	chr17	29556079	C	T	stop_gained	HIGH	ENST00000358273.4:c.2446C>T	ENSP00000351015.4:p.Arg816Ter			COSM24444	Pathogenic			
	p11	chr17	29556175	G	T	missense_variant	MODERA	ENST00000358273.4:c.2542G>T	ENSP00000351015.4:p.Gly848Trp					D&D&D	T	T
	p9	chr17	29662047	C	T	stop_gained&splice	HIGH	ENST00000358273.4:c.6004C>T	ENSP00000351015.4:p.Gln2002Ter				Pathogenic			
	p8	chr17	29665110	C	T	stop_gained	HIGH	ENST00000358273.4:c.6772C>T	ENSP00000351015.4:p.Arg2258Ter			COSM1324055&COSM215676	Pathogenic			
OBSCN	p19	chr1	228461097	C	G	missense_variant	MODERA	ENST00000570156.2:c.6111C>G	ENSP00000455507.2:p.Asp2037Glu	G:0.0022	G:5.037e-03	rs116320185		D	T	T
	p6	chr1	228461097	C	G	missense_variant	MODERA	ENST00000570156.2:c.6111C>G	ENSP00000455507.2:p.Asp2037Glu	G:0.0022	G:5.037e-03	rs116320185		D	T	T
	p14	chr1	228462255	T	C	missense_variant	MODERA	ENST00000570156.2:c.7450T>C	ENSP00000455507.2:p.Trp2484Arg	C:0.0026	C:7.053e-03	rs62621832		D&D&D	T	T
	p13	chr1	228464306	G	C	missense_variant	MODERA	ENST00000570156.2:c.7501G>C	ENSP00000455507.2:p.Val2501Leu		C:2.485e-05	rs7629283369		T&T&T	T	T
	p9	chr1	228467611	C	T	missense_variant	MODERA	ENST00000570156.2:c.8773C>T	ENSP00000455507.2:p.Pro2925Ser		T:5.806e-05	rs757897389		T&T&T&	D	D
	p11	chr1	228506704	C	T	missense_variant	MODERA	ENST00000570156.2:c.17122C>T	ENSP00000455507.2:p.Arg5708Cys		A:8.270e-06&T:1.654e-(	rs369347698		D&D&D&D	T	T
	p12	chr1	228558952	T	G	missense_variant	MODERA	ENST00000570156.2:c.23344T>G	ENSP00000455507.2:p.Phe7782Val	G:0.0022	G:6.994e-03	rs113876855		D&D	T	T
	p25	chr1	228560686	C	T	missense_variant	MODERA	ENST00000570156.2:c.25078C>T	ENSP00000455507.2:p.His8360Tyr		T:2.493e-05	rs774192971			T	T
	p15	chr1	228560783	G	A	splice_donor_varian	HIGH	ENST00000570156.2:c.25174+1G>A			A:3.447e-04	rs200271514				
PASK	p6	chr2	242076585	G	A	missense_variant	MODERA	ENST00000358649.4:c.971C>T	ENSP00000351475.4:p.Ala324Val		A:4.942e-05	rs556062263&COSM1406912&COSM1406913		T&T&T&T&T&	T	T
PKHD1L1	p9	chr8	110418573	C	T	missense_variant	MODERA	ENST00000378402.5:c.1679C>T	ENSP00000367655.5:p.Pro560Leu	T:0.0018	T:3.326e-04	rs142950253		T	D	D
	p24	chr8	110445418	G	A	missense_variant	MODERA	ENST00000378402.5:c.3313G>A	ENSP00000367655.5:p.Val1105Ile		A:3.477e-04	rs200879586		T	T	T
	p16	chr8	110468517	C	T	missense_variant	MODERA	ENST00000378402.5:c.6901C>T	ENSP00000367655.5:p.Arg2301Cys	T:0.0006	T:2.399e-04	rs73309320		T	D	D
	p6	chr8	110492352	C	T	missense_variant	MODERA	ENST00000378402.5:c.9311C>T	ENSP00000367655.5:p.Thr3104Ile	T:0.0062	T:9.266e-03	rs117952538		D	D	D
	p8	chr8	110492352	C	T	missense_variant	MODERA	ENST00000378402.5:c.9311C>T	ENSP00000367655.5:p.Thr3104Ile	T:0.0062	T:9.266e-03	rs117952538		D	D	D
	p6	chr8	110499005	C	T	missense_variant	MODERA	ENST00000378402.5:c.9835C>T	ENSP00000367655.5:p.Arg3279Cys	T:0.0012	T:2.308e-03	rs189682035&COSM453855		D&D	D	D
RGPD4	p25	chr2	108443503	G	C	missense_variant	MODERA	ENST00000408999.3:c.34G>C	ENSP00000386810.3:p.Val12Leu	C:0.0078	A:8.326e-06&C:7.802e-(	rs200269994		T&T	T	T
	p23	chr2	108455329	G	T	missense_variant	MODERA	ENST00000408999.3:c.314G>T	ENSP00000386810.3:p.Cys105Phe		T:6.063e-03	rs201257819		D&D&D	T	T
	p24	chr2	108455329	G	T	missense_variant	MODERA	ENST00000408999.3:c.314G>T	ENSP00000386810.3:p.Cys105Phe		T:6.063e-03	rs201257819		D&D&D	T	T
	p11	chr2	108455376	G	A	missense_variant	MODERA	ENST00000408999.3:c.361G>A	ENSP00000386810.3:p.Ala121Thr	A:0.0006	A:1.746e-03	rs201566550		T&T&D	T	T
	p13	chr2	108473216	G	A	missense_variant&s	MODERA	ENST00000408999.3:c.1067G>A	ENSP00000386810.3:p.Gly356Glu					D&D&D	T	T
	p14	chr2	108475609	C	G	missense_variant	MODERA	ENST00000408999.3:c.1316C>G	ENSP00000386810.3:p.Thr439Ser					T&T&T	T	T
	p11	chr2	108475628	G	A	stop_gained	HIGH	ENST00000408999.3:c.1335G>A	ENSP00000386810.3:p.Trp445Ter				not_provided			
	p13	chr2	108475628	G	A	stop_gained	HIGH	ENST00000408999.3:c.1335G>A	ENSP00000386810.3:p.Trp445Ter				not_provided			
	p9	chr2	108475698	A	T	missense_variant	MODERA	ENST00000408999.3:c.1405A>T	ENSP00000386810.3:p.Thr469Ser					T&T&D	D	T
	p10	chr2	108479198	A	G	missense_variant	MODERA	ENST00000408999.3:c.2266A>G	ENSP00000386810.3:p.Asn756Asp					T&T&T	T	T
	p19	chr2	108487440	A	T	missense_variant	MODERA	ENST00000408999.3:c.2980A>C	ENSP00000386810.3:p.Lys994Gln	C:0.0034	C:3.527e-03	rs199695825		T&T&T	T	T
	p23	chr2	108487440	A	C	missense_variant	MODERA	ENST00000408999.3:c.2980A>C	ENSP00000386810.3:p.Lys994Gln	C:0.0034	C:3.527e-03	rs199695825		T&T&T	T	T
	p24	chr2	108487440	A	C	missense_variant	MODERA	ENST00000408999.3:c.2980A>C	ENSP00000386810.3:p.Lys994Gln	C:0.0034	C:3.527e-03	rs199695825		T&T&T	T	T
	p27	chr2	108488017	G	T	missense_variant	MODERA	ENST00000408999.3:c.3557G>T	ENSP00000386810.3:p.Gly1186Val		T:3.024e-05	rs373812613		D&D&D	T	T
	p17	chr2	108499138	G	A	missense_variant	MODERA	ENST00000408999.3:c.5075G>A	ENSP00000386810.3:p.Ser1692Asn	A:0.0004	A:2.891e-03	rs201108061		D&D&D	T	T
SLC14A2	p16	chr18	43224101	G	A	missense_variant	MODERA	ENST00000255226.6:c.1327G>A	ENSP00000255226.5:p.Gly443Ser	A:0.0068	A:2.257e-03	rs35245152&COSM3937749		T	T	T
	p25	chr18	43243797	G	A	missense_variant	MODERA	ENST00000255226.6:c.1399G>A	ENSP00000255226.5:p.Gly467Ser		A:8.237e-06	rs767446154		T	T	T
	p22	chr18	43262408	G	A	missense_variant	MODERA	ENST00000255226.6:c.2687G>A	ENSP00000255226.5:p.Arg896His	A:0.0052	A:8.195e-03&T:8.236e-(	rs41301139		D	T	T

End note: Highlighted sites are recurrent mutation with same site on same gene in different PNF (Ratner) and PNF (Pemov) samples.

Green: Overlapped mutated sites between PNF (Ratner) and PNF (Pemov)

Orange: Overlapped mutated sites between three datasets.

Supplemental Table 8: Somatic variant genes overlapped with PNF (Pemov et al, Re-analyzed with VarScan2)

SYMBOL	Patient	#CHROM	POS	REF	ALT	Consequence	IMPACT	HGVSc	HGVSp	GMAF	ExAC_MAF	Existing_variation	ClinVar (2019)	SIFT	MetaLR	MetaSVM
DOCK10	p22	chr2	225750388	C	A	missense_variant	MODERATE	NST00000258390.7:c.747G>	ENSP00000258390.7:p.Gln249His					D&D	D	T
NF1	p13	chr17	29483036	CAA	C	frameshift_varian	HIGH	T00000358273.4:c.98_99de	ENSP00000351015.4:p.Lys33SerfsTer4							
	p7	chr17	29527590	C	T	stop_gained	HIGH	NST00000358273.4:c.1039C>	ENSP00000351015.4:p.Gln347Ter			COSM1479414&COSM1479413	Pathogenic			
	p9	chr17	29556992	G	A	missense_variant	MODERATE	NST00000358273.4:c.2990G>	ENSP00000351015.4:p.Arg997Lys					T&T&T	T	T
	p23	chr17	29654553	C	T	stop_gained	HIGH	NST00000358273.4:c.5305C>	ENSP00000351015.4:p.Arg1769Ter			COSM36883&COSM303874	Pathogenic			
	p10	chr17	29665110	C	T	stop_gained	HIGH	NST00000358273.4:c.6772C>	ENSP00000351015.4:p.Arg2258Ter			COSM1324055&COSM215676	Pathogenic			
	p15	chr17	29683983	C	T	stop_gained	HIGH	NST00000358273.4:c.7744C>	ENSP00000351015.4:p.Gln2582Ter							
OBSCN	p19	chr1	228465551	G	T	missense_variant	MODERATE	NST00000570156.2:c.8138G>	ENSP00000455507.2:p.Arg2713Leu					T&T&T	T	T
	p17	chr1	228486163	G	T	missense_variant	MODERATE	ST00000570156.2:c.12955G	ENSP00000455507.2:p.Ala4319Ser		T:8.361e-0	rs765653440		T&T	T	T
	p22	chr1	228504457	G	T	missense_variant	MODERATE	ST00000570156.2:c.16204G	ENSP00000455507.2:p.Ala5402Ser					T&T&T&T	T	T
	p25	chr1	228564931	C	A	missense_variant	MODERATE	ST00000570156.2:c.26089C	ENSP00000455507.2:p.His8697Asn						T	T
RREB1	p17	chr6	7246860	G	T	stop_gained	HIGH	NST00000379938.2:c.4177G>	ENSP00000369270.2:p.Glu1393Ter							

LOH variants (Pemov et al, Re-analyzed with VarScan2)

SYMBOL	Patient	#CHROM	POS	REF	ALT	Consequence	IMPACT	HGVSc	HGVSp	GMAF	ExAC_MAF	Existing_variation	ClinVar (2019)	SIFT	MetaLR	MetaSVM
ATM	p29	chr11	108124761	T	C	missense_variant	MODERATE	NST00000278616.4:c.2119T>	ENSP00000278616.4:p.Ser707Pro	C:0.0044	:7.808e-0	rs4986761&COSM41595	Benign/Li	T&T&T	T	T
FCGBP	p7	chr19	40374034	A	G	missense_variant	MODERATE	ST00000221347.6:c.12044T	ENSP00000221347.5:p.Val4015Ala	G:0.504&G:0.46009&rs138587194&COSM439495				T	T	T
MDN1	p10	chr6	90408618	T	C	missense_variant	MODERATE	NST00000369393.3:c.9134A>	ENSP00000358400.3:p.Glu3045Gly	C:0.0014	:1.095e-0	rs116103426		T&T	T	T
NF1	p5	chr17	29550493	TTAAC	T	frameshift_varian	HIGH	000358273.4:c.1756_1759d	NSP00000351015.4:p.Thr586ValfsTer18				Pathogenic			

Supplemental Table 9:Disease-causing

	Methods	Category	Description
ClinVar clinical significance	Repository	unknown, untested, non-pathogenic, probable-non-pathogenic, **probable-pathogenic, lkey-pathogenic, pathogenic**, drug-response,histocompatibility, other	asserted to be pathogenic or likely pathogenic by at least one submitter
SIFT	P(An amino acid at a position is tolerated | The most frequentest amino acid being tolerated)	**D: Deleterious (sift<=0.05)**; T: tolerated (sift>0.05)	protein function based on sequence homology and the physico-chemical similarity between the alternate amino acids
MetaLR	Logistic regression	**D: Deleterious**; T: Tolerated; higher scores are more deleterious	Ensemble score
MetaSVM	Support vector machine	**D: Deleterious**; T: Tolerated; higher scores are more deleterious	Ensemble score