

# Clinical Heme Panel Optimisation

Papaemmanuil Lab | Pierre Guilmin

December 2018

## 1 The process

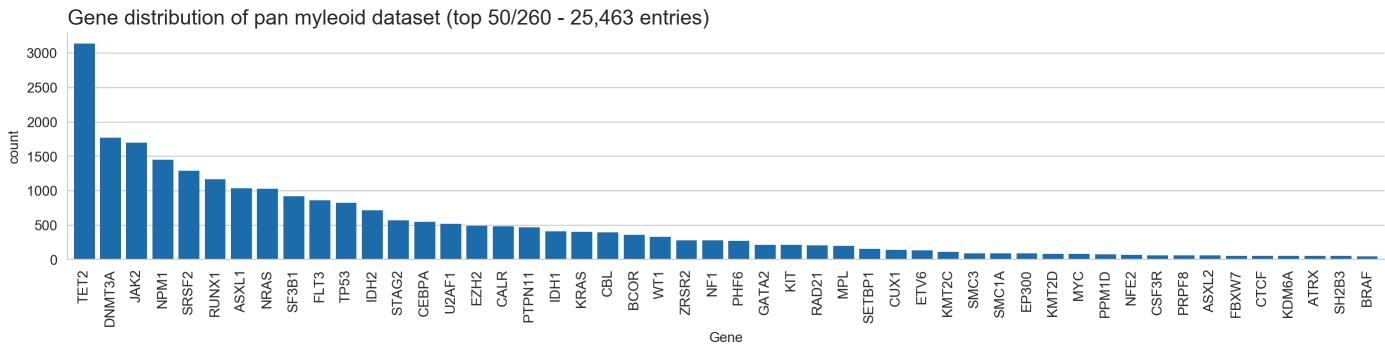
### 1.1 The pan-myeloid dataset

Brief overview of the pan-myeloid dataset:

- 25,463 mutations
- 8,966 patients
- only oncogenic mutations
- 3 different diseases:

disease	count	frequency	patient
AML	13,656	54%	4,348
MDS	8,179	32%	2,586
MPN	3,628	14%	2,063

- 260 genes:



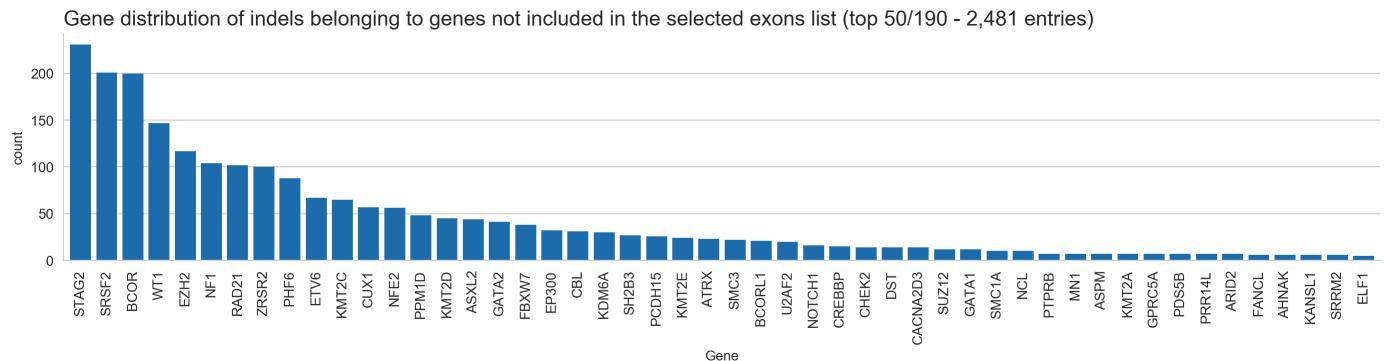
### 1.2 Methodology

- Filter the indels between reported and unreported following the selected exons list. An indel is considered as reported if its beginning or end is within one of the exons of the selected exons list.
- Filter the substitutions between reported and unreported following the selected hotspots list. A substitution is considered as reported if it hits one of the three nucleotides of an amino acid present in the selected hotspots list.

## 2 Unreported indels

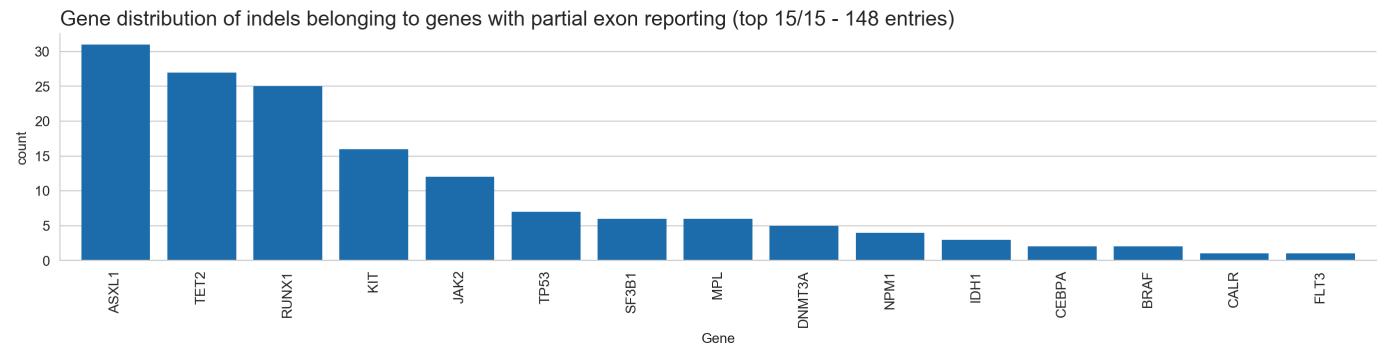
### 2.1 Genes not included in the selected exons list

$\approx 31\%$  of the indels belong to genes not included in the selected exons list. The following plot shows the distribution of the main unreported genes for indels:



### 2.2 Indels identified in genes with partial exon reporting

Only  $\approx 2\%$  of the indels belong to genes with partial exon reporting. The following plot shows the distribution of the genes having some unreported exons:



Here is a table of the unreported exons for the genes above. The table lists for each gene the missed exon numbers and how many indels belonging to that specific exon were missed (NA  $\leftrightarrow$  no corresponding exon found in transcript list). For example the first line means: 30 indels belonging to exon 11 of ASXL1 were missed as well as 1 indel belonging to an unidentified part of ASXL1.

gene	missed exons
ASXL1	11 (n = 30), NA (n = 1)
TET2	NA (n = 27)
RUNX1	NA (n = 25)
KIT	8 (n = 8), 2 (n = 2), 9 (n = 1), 10 (n = 1), 5 (n = 1), 11 (n = 1), 21 (n = 1), NA (n = 1)
JAK2	12 (n = 5), 11 (n = 5), 19 (n = 1), NA (n = 1)
TP53	NA (n = 7)
SF3B1	16 (n = 5), NA (n = 1)
MPL	12 (n = 3), 3 (n = 1), 11 (n = 1), NA (n = 1)
DNMT3A	NA (n = 5)
NPM1	10 (n = 2), NA (n = 2)
IDH1	6 (n = 1), 3 (n = 1), NA (n = 1)
CEBPA	NA (n = 2)
BRAF	14 (n = 1), 3 (n = 1)
CALR	7 (n = 1)
FLT3	3 (n = 1)

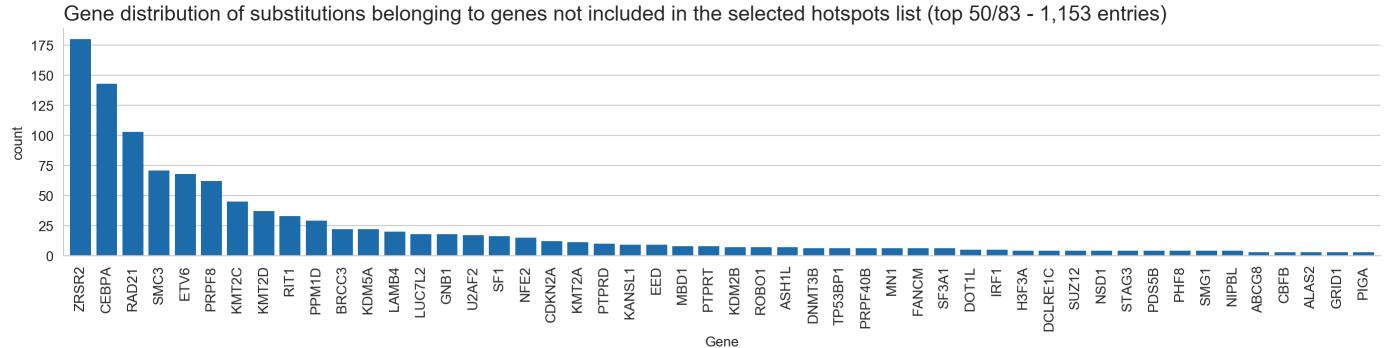
### 2.3 Summary table

indel status	count	frequency
in exons list	5,267	66,70%
belongs to gene not included in the selected exons list	2,481	31,42%
belongs to gene with partial exon reporting	148	1,87%

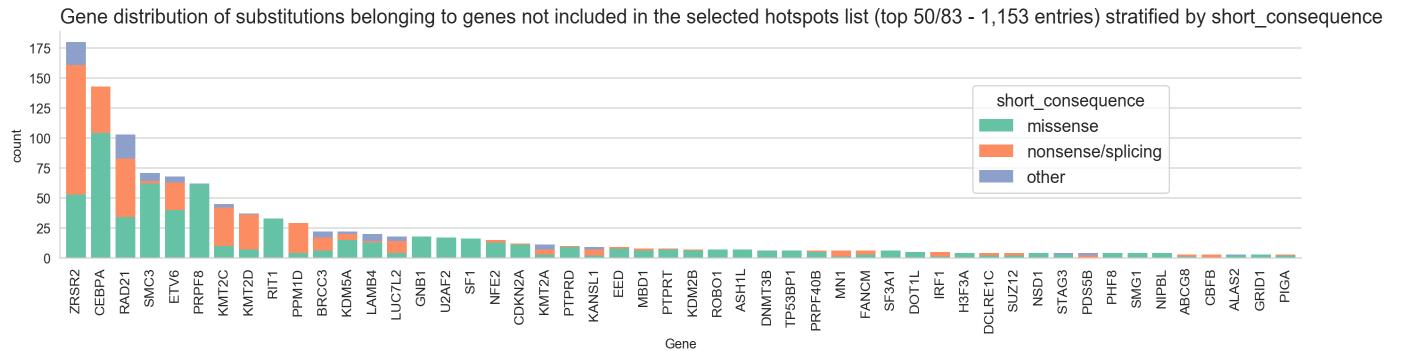
### 3 Unreported substitutions

#### 3.1 Genes not included in the selected hotspots list

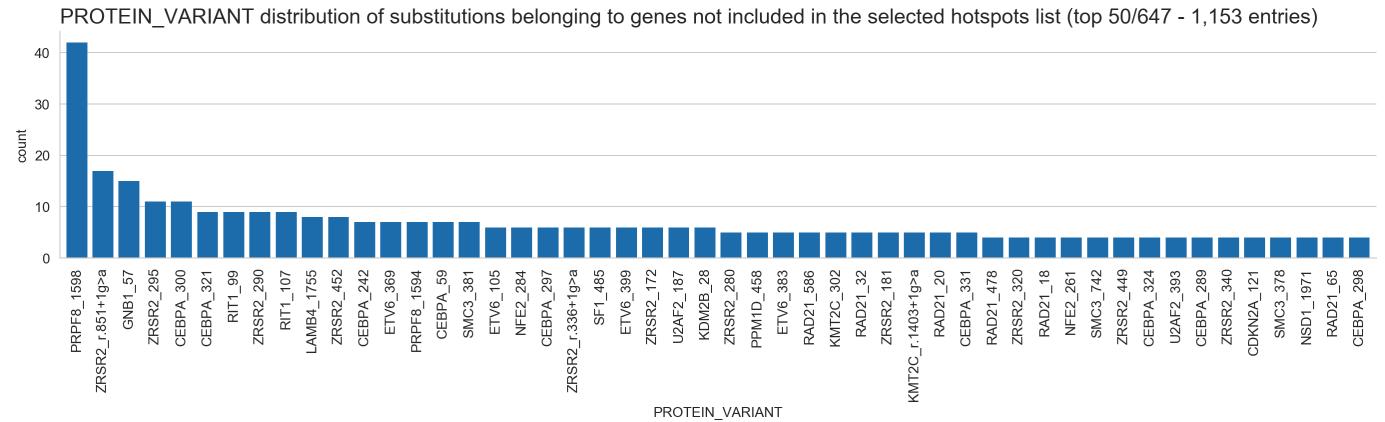
$\approx 7\%$  of the substitutions belong to genes not included in the selected hotspots list. The following plot shows the distribution of the main unreported genes for substitutions:



Here is the same plot stratified by mutation consequence:

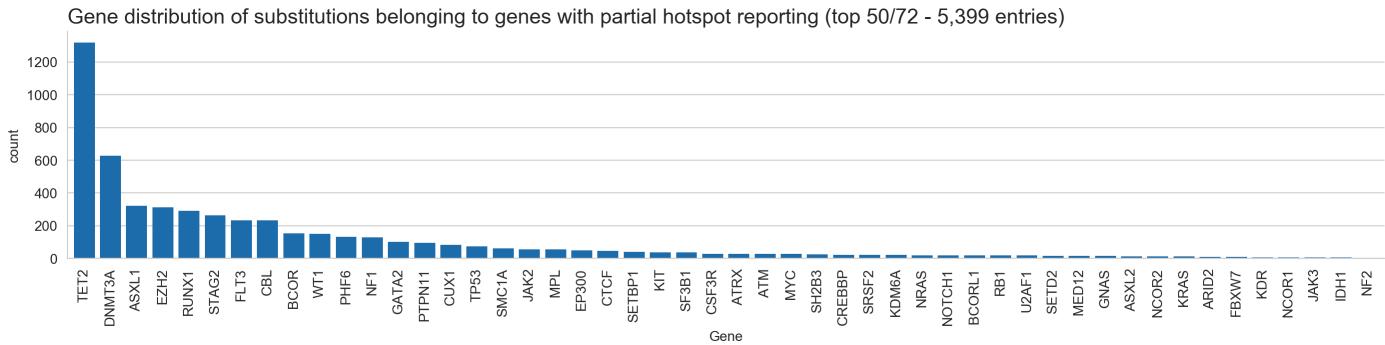


Here is a figure of the 50 most recurrent unreported substitutions belonging to genes not included in the selected hotspots list:

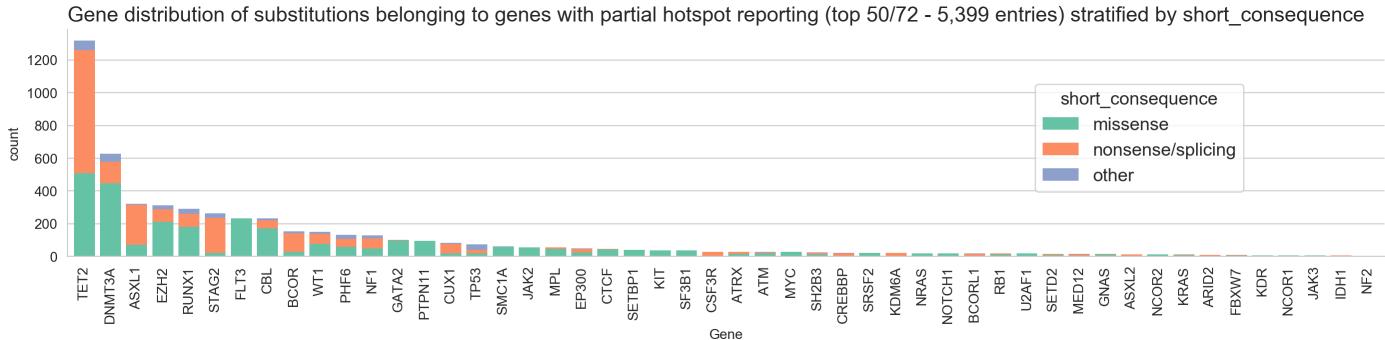


#### 3.2 Substitutions identified in genes with partial hotspot reporting

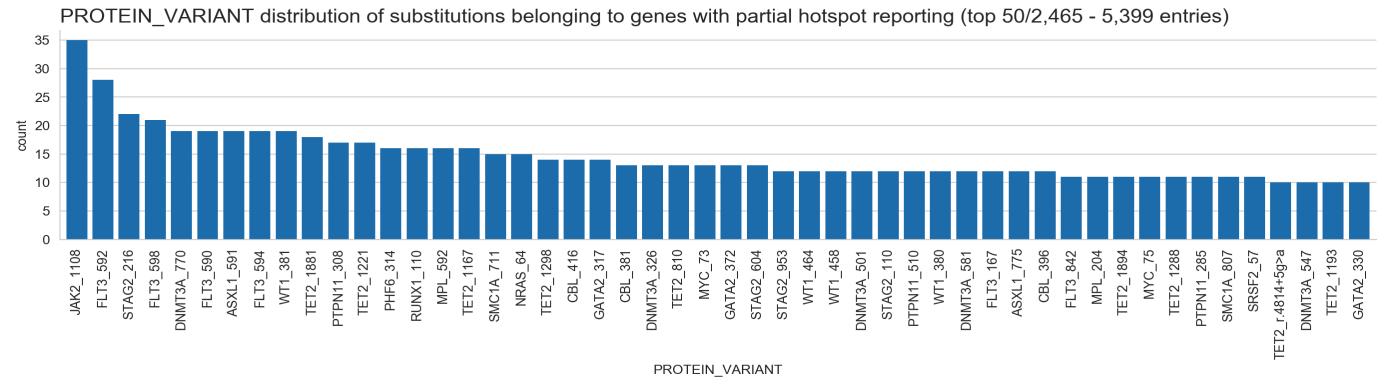
$\approx 31\%$  of the substitutions belong to genes with partial hotspot reporting. The following plot shows the distribution of the main genes having some unreported substitutions:



Here is the same plot stratified by mutation consequence:



Here is a figure of the 50 most recurrent unreported substitutions<sup>1</sup>:



### 3.3 Summary table

substitution status	count	frequency
in hotspots list	11,015	62,70%
belongs to gene not included in the selected hotspots list	1,153	6,56%
belongs to gene with partial hotspot reporting	5,399	30,73%

<sup>1</sup> see supplementary material 1 for the complete table

## 4 Summary by patient

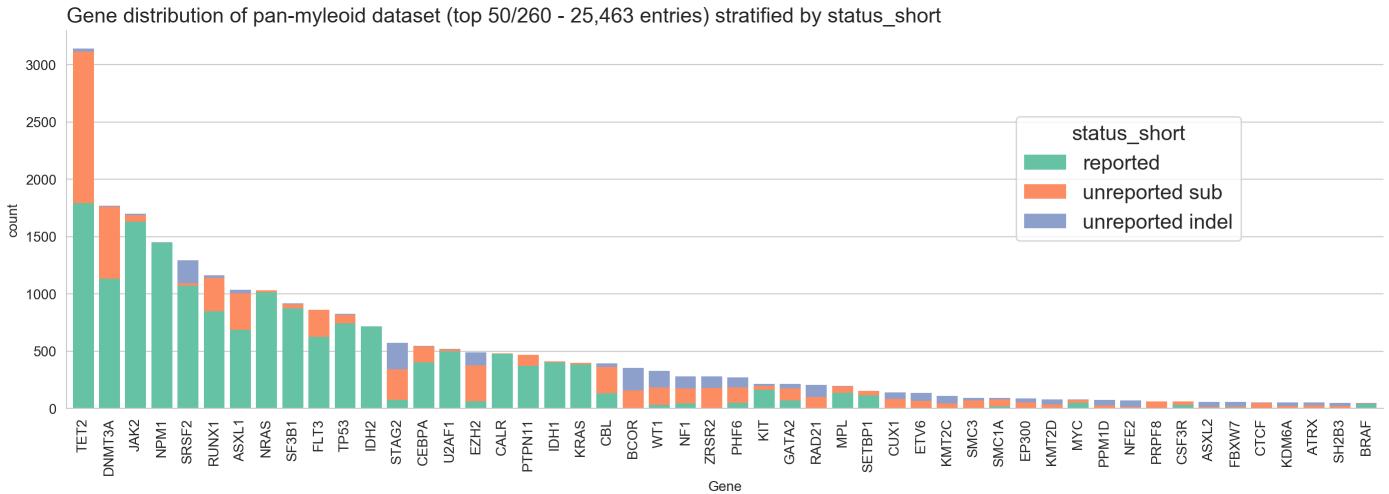
At the patient level, we have:

- a mean of  $\boxed{31\%}$  oncogenic mutations unreported by patient
- for  $\boxed{8\%}$  ( $711/8,966$ ) of the patients, **not a single oncogenic mutation reported**

## 5 Summary by gene

### 5.1 Overview

The following plot shows the proportion of unreported mutation for the most recurrent mutated genes<sup>2</sup>:



In the next part we propose an analysis gene by gene for the 33 most recurrent mutated genes (see lookup table under) gathering:

- The OS (overall survival) plot between:
  - wildtype**: patients without any mutation in the studied gene
  - gene name - reported**: patients with at least one reported mutation in the studied gene
  - gene name - unreported**: patients with mutations in the studied gene but not a single reported mutation

Warning, the number indicated in each category represents the number of patient used to plot the curve (modulo the missing values).

- A pivot table showing the count of mutation reported (✓) and unreported (✗) stratified by disease (AML, MDS or MPN) and mutation consequence (inframe, missense, nonsense/splicing, truncating or other).
- 4 lollipop plots of the gene, two for indels (reported vs unreported) and two for substitutions (reported vs unreported). Warning, these lollipop plots don't take into account the splicing mutations. The color of the mutation circle represents its consequence, the size of the mutation circle represents its occurrence in the dataset, see legend:

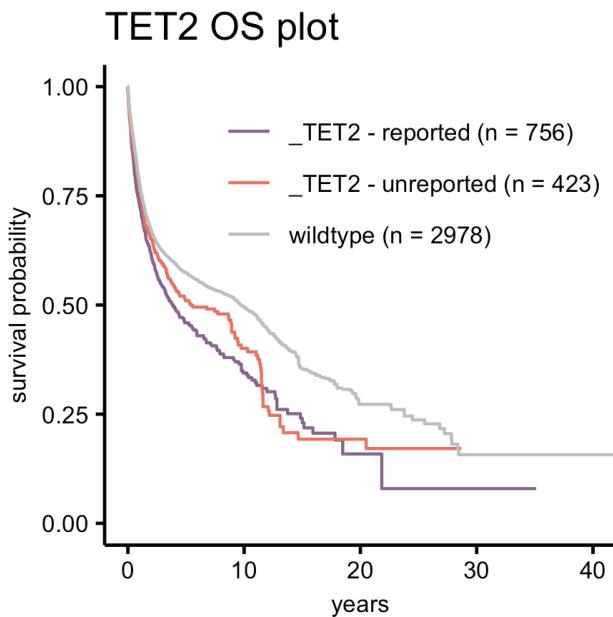


Lookup table gene ↔ page:

gene	page	gene	page	gene	page	gene	page	gene	page	gene	page
TET2	8	ASXL1	14	STAG2	20	IDH1	26	ZRSR2	32	SETBP1	38
DNMT3A	9	NRAS	15	CEBPA	21	KRAS	27	PHF6	33	CUX1	39
JAK2	10	SF3B1	16	U2AF1	22	CBL	28	KIT	34	ETV6	40
NPM1	11	FLT3	17	EZH2	23	BCOR	29	GATA2	35		
SRSF2	12	TP53	18	CALR	24	WT1	30	RAD21	36		
RUNX1	13	IDH2	19	PTPN11	25	NF1	31	MPL	37		

<sup>2</sup> see supplementary material 2 for the complete table

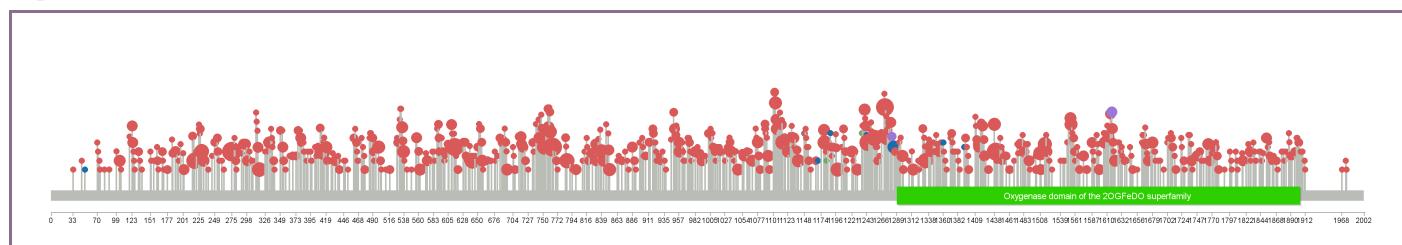
## 5.2 TET2



disease	status consequence	× indel	× sub	✓
AML	inframe	0	0	6
	missense	0	189	86
	nonsense/splicing	4	266	92
	other	10	31	4
	truncating	0	0	466
MDS	inframe	0	0	6
	missense	0	245	110
	nonsense/splicing	5	422	114
	other	2	15	3
	truncating	2	0	714
MPN	inframe	0	0	1
	missense	0	73	35
	nonsense/splicing	0	66	17
	other	3	13	0
	truncating	1	0	139
<b>TOTAL</b>		27	1320	1793
		0.01%	0.42%	0.57%

### Indels

#### reported

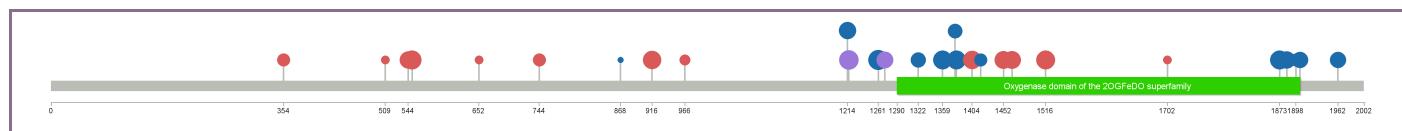


#### unreported

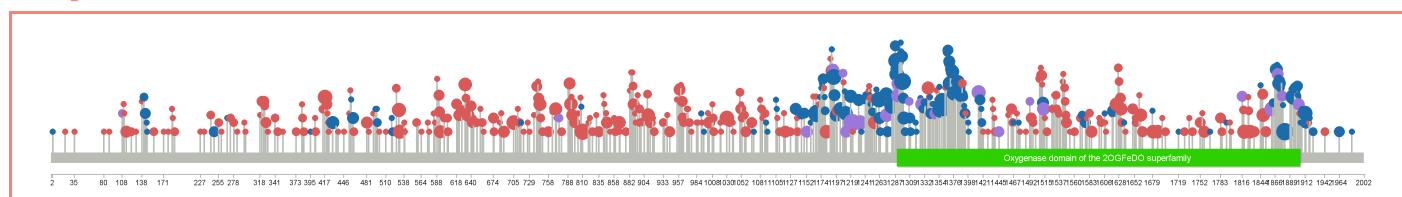
No unreported indel found.

### Substitutions

#### reported

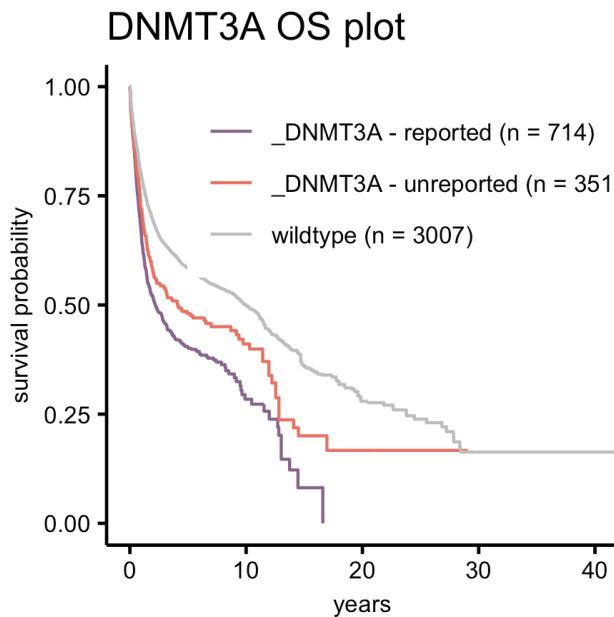


#### unreported



\* likely germline infiltration

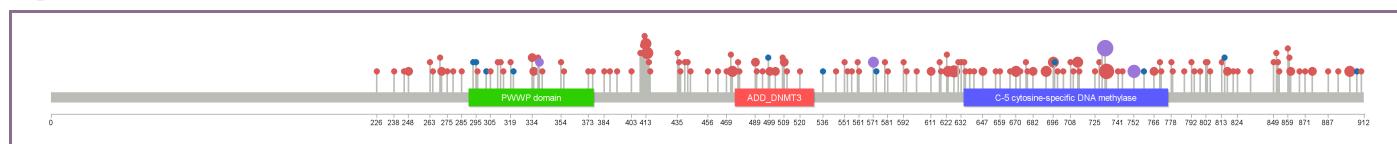
### 5.3 DNMT3A



disease	status consequence	× indel	× sub	✓
AML	inframe	1	0	14
	missense	0	297	733
	nonsense/splicing	1	70	6
	other	1	27	2
	truncating	0	0	119
MDS	inframe	0	0	11
	missense	0	105	126
	nonsense/splicing	0	48	4
	other	0	12	1
	truncating	0	0	51
MPN	inframe	0	0	1
	missense	0	44	45
	nonsense/splicing	0	16	3
	other	2	10	0
	truncating	0	0	18
<b>TOTAL</b>		5	629	1134
		0.00%	0.36%	0.64%

#### Indels

##### reported



##### unreported

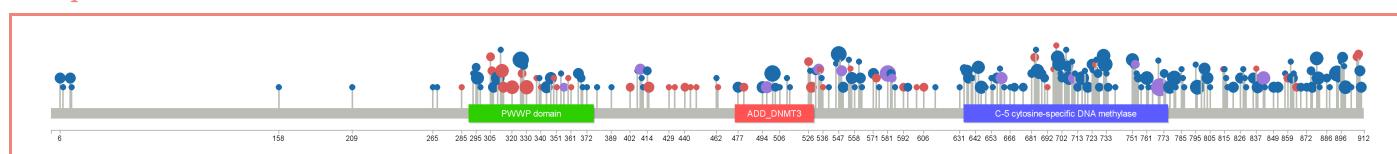
No unreported indel found.

#### Substitutions

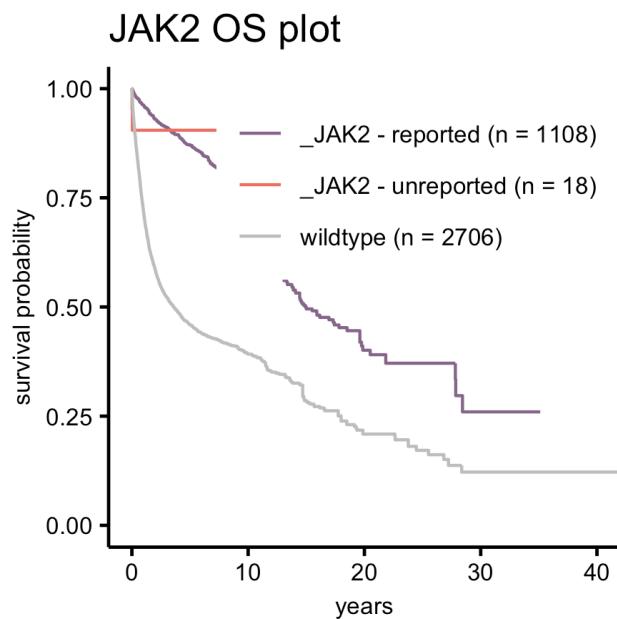
##### reported



##### unreported



## 5.4 JAK2



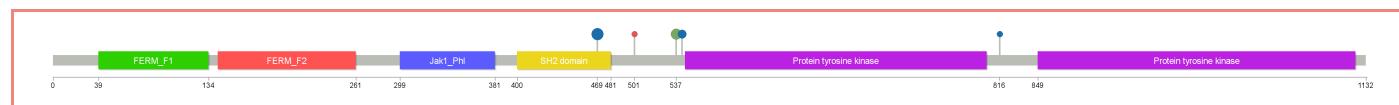
disease	status consequence	× indel	× sub	✓
<b>AML</b>	missense	0	8	134
	truncating	1	0	0
<b>MDS</b>	inframe	1	0	0
	missense	0	20	109
<b>MPN</b>	inframe	6	0	0
	missense	0	28	1388
	other	4	1	0
<b>TOTAL</b>		12	57	1631
		0.01%	0.03%	0.96%

### Indels

#### reported

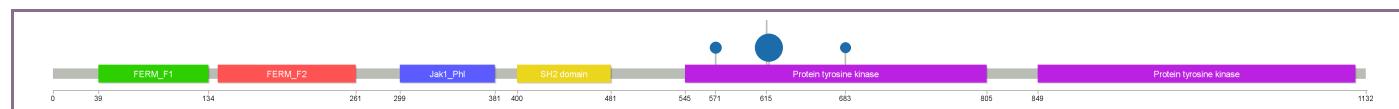
This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

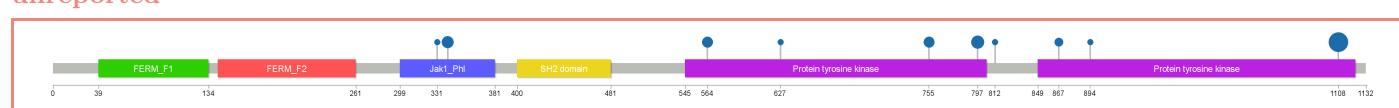


### Substitutions

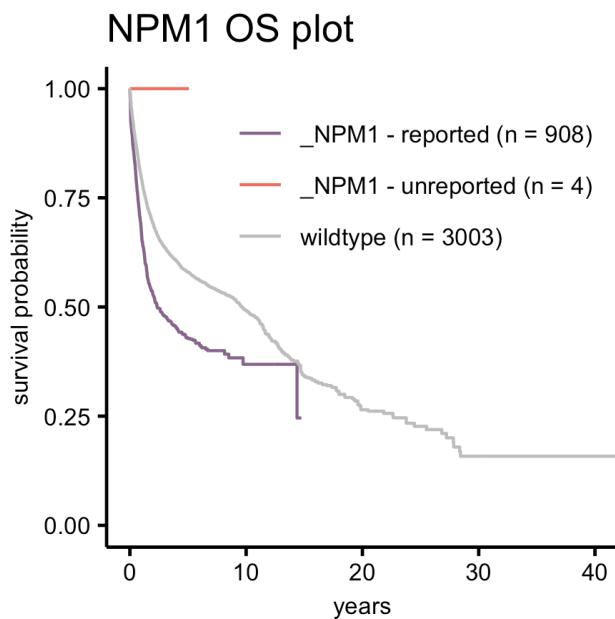
#### reported



#### unreported



## 5.5 NPM1



disease	status consequence	× indel	✓
AML	other	1	0
	truncating	1	1385
MDS	truncating	0	63
MPN	other	2	0
<b>TOTAL</b>		4	1448
		0.00%	1.00%

### Indels

#### reported



#### unreported

No unreported indel found.

### Substitutions

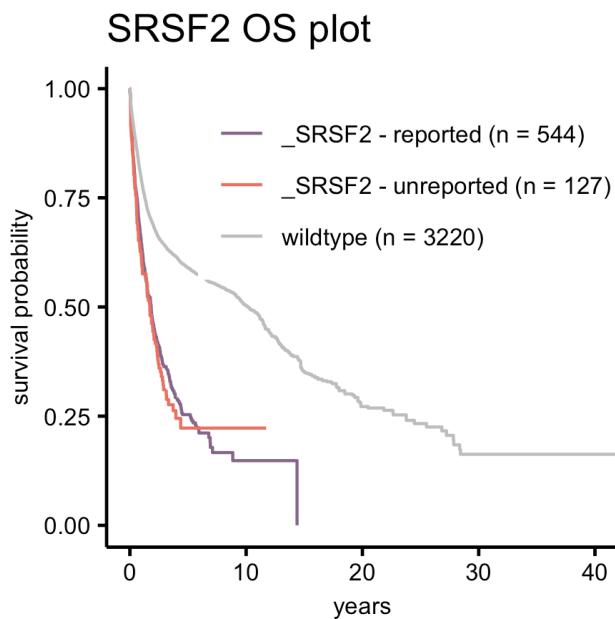
#### reported

This gene was not reported for hostpots (or no reported mutation was found in the dataset).

#### unreported

No unreported substitution found.

## 5.6 SRSF2



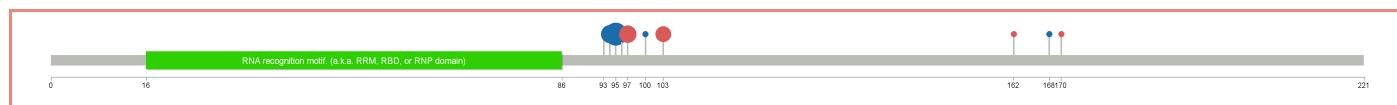
disease	status consequence	× indel	× sub	✓
AML	inframe	100	0	0
	missense	0	9	494
	other	1	0	0
	truncating	16	0	0
MDS	inframe	67	0	0
	missense	0	13	540
	truncating	15	0	0
MPN	inframe	2	0	0
	missense	0	1	36
<b>TOTAL</b>		201	23	1070
		0.16%	0.02%	0.83%

### Indels

#### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported



### Substitutions

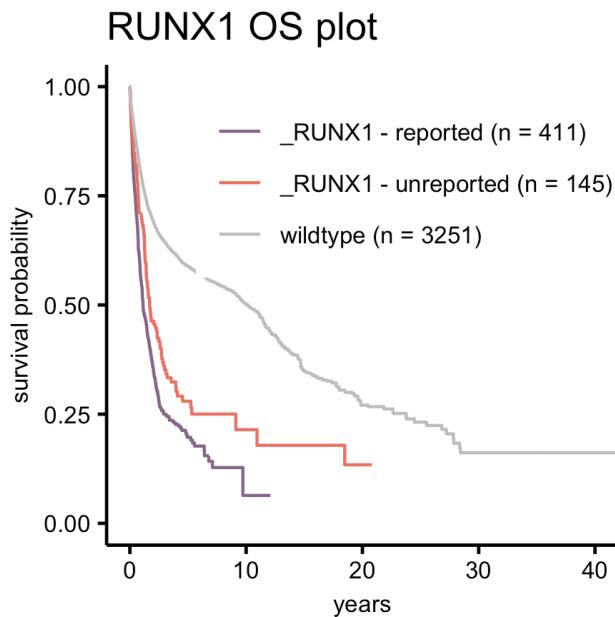
#### reported



#### unreported



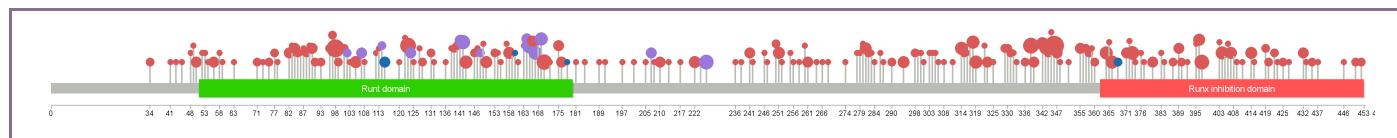
## 5.7 RUNX1



disease	status consequence	× indel	× sub	✓
<b>AML</b>	inframe	0	0	20
	missense	0	108	170
	nonsense/splicing	4	46	60
	other	12	23	6
<b>MDS</b>	truncating	2	0	296
	inframe	0	0	5
	missense	0	65	72
	nonsense/splicing	6	36	34
<b>MPN</b>	other	1	5	0
	truncating	0	0	180
	missense	0	7	2
	nonsense/splicing	0	0	2
<b>TOTAL</b>	other	0	2	0
	truncating	0	0	1
		25	292	848
		0.02%	0.25%	0.73%

### Indels

#### reported



#### unreported

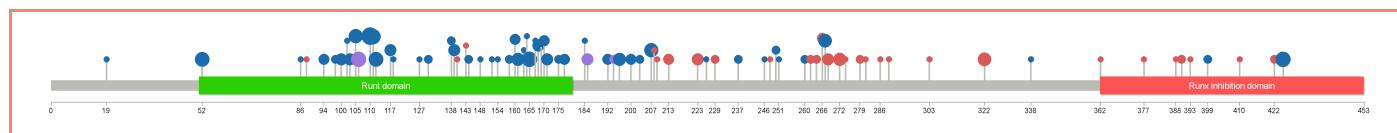
No unreported indel found.

### Substitutions

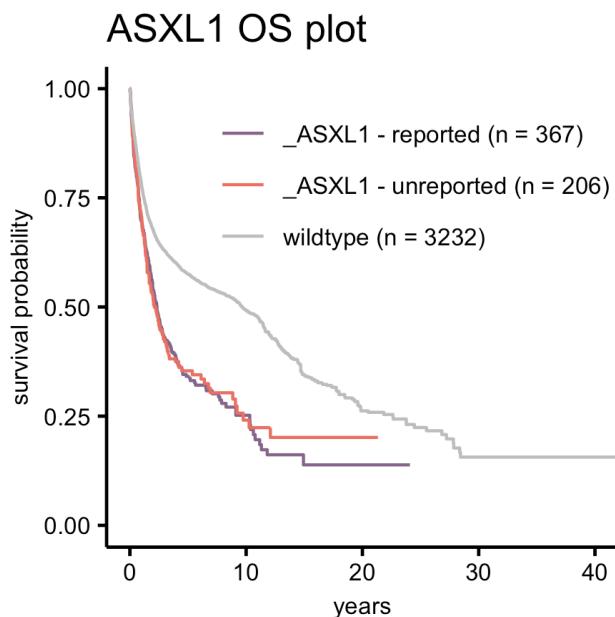
#### reported



#### unreported



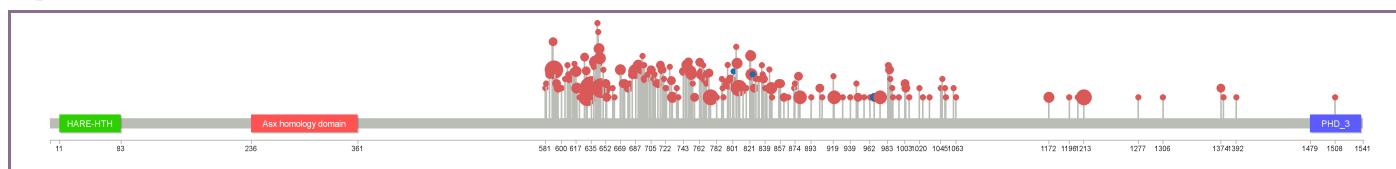
## 5.8 ASXL1



disease	status consequence	x indel	x sub	✓
AML	inframe	0	0	2
	missense	0	42	0
	nonsense/splicing	0	86	36
	other	1	0	0
	truncating	9	0	191
MDS	inframe	0	0	2
	missense	0	16	0
	nonsense/splicing	0	121	47
	other	0	1	0
	truncating	15	0	325
MPN	missense	0	14	0
	nonsense/splicing	0	37	9
	other	1	4	0
	truncating	5	0	73
<b>TOTAL</b>		31	321	685
		0.03%	0.31%	0.66%

### Indels

#### reported

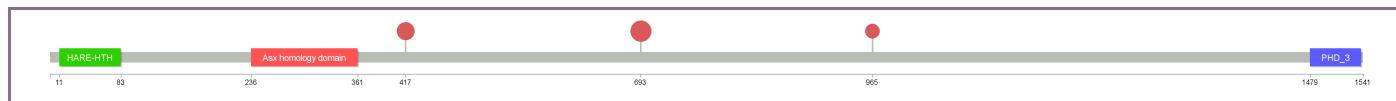


#### unreported

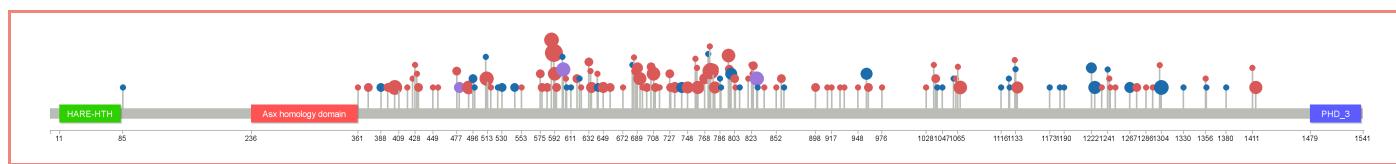


### Substitutions

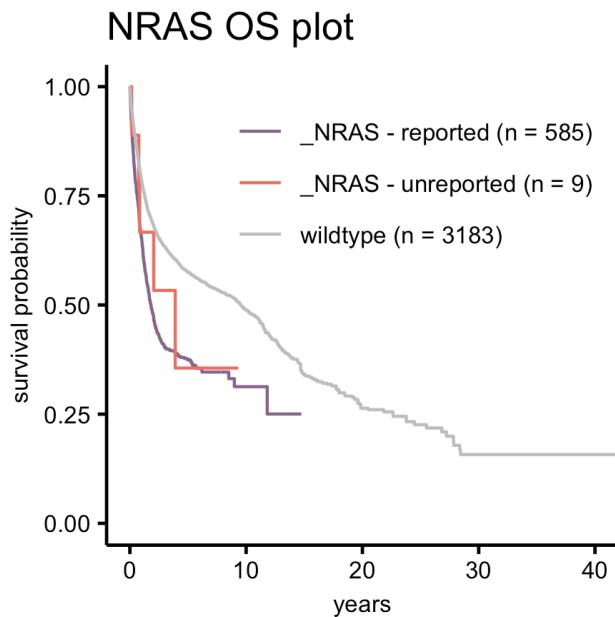
#### reported



#### unreported



## 5.9 NRAS



disease	status consequence	× sub	✓
AML	missense	6	823
MDS	missense	13	181
MPN	missense	0	9
<b>TOTAL</b>		19	1013
		0.02%	0.98%

### Indels

#### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

No unreported indel found.

### Substitutions

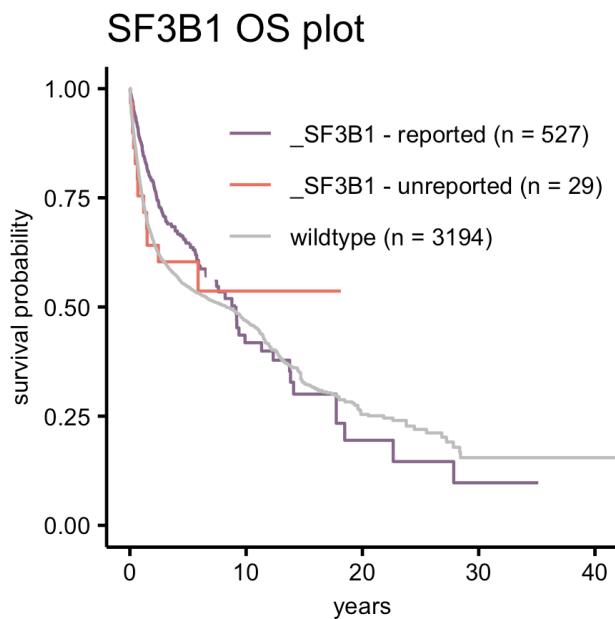
#### reported



#### unreported



## 5.10 SF3B1



disease	status consequence	× indel	× sub	✓
AML	missense	0	15	132
	other	0	1	0
MDS	missense	0	12	704
	other	4	0	0
MPN	inframe	1	0	0
	missense	0	9	37
	nonsense/splicing	0	1	0
	other	1	0	0
<b>TOTAL</b>		6	38	873
		0.01%	0.04%	0.95%

### Indels

#### reported

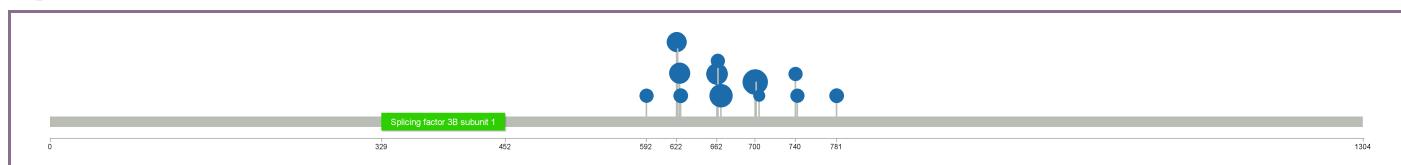
This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

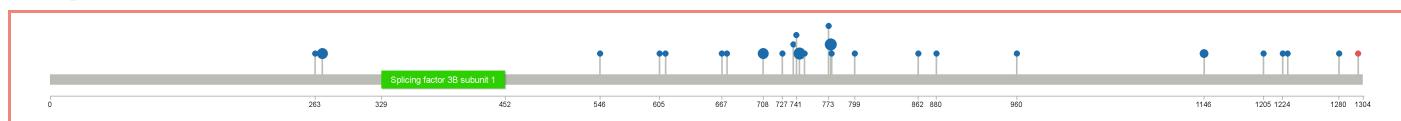


### Substitutions

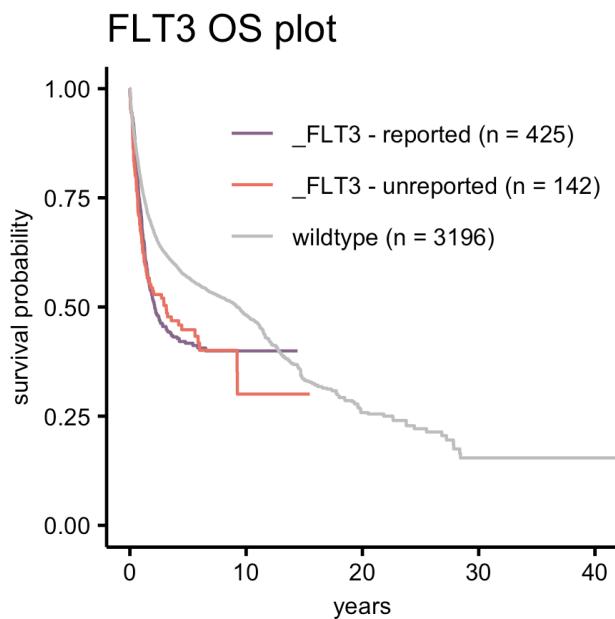
#### reported



#### unreported



## 5.11 FLT3



disease	status consequence	× indel	× sub	✓
AML	inframe	0	0	141
	missense	0	203	425
	nonsense/splicing	0	1	11
	other	0	0	26
MDS	inframe	0	0	9
	missense	0	20	16
MPN	missense	0	9	0
	truncating	1	0	0
<b>TOTAL</b>		1	233	628
		0.00%	0.27%	0.73%

### Indels

#### reported

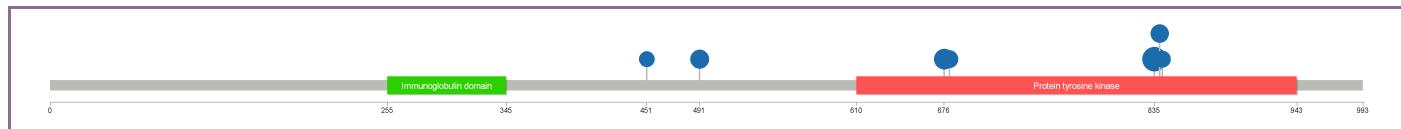


#### unreported

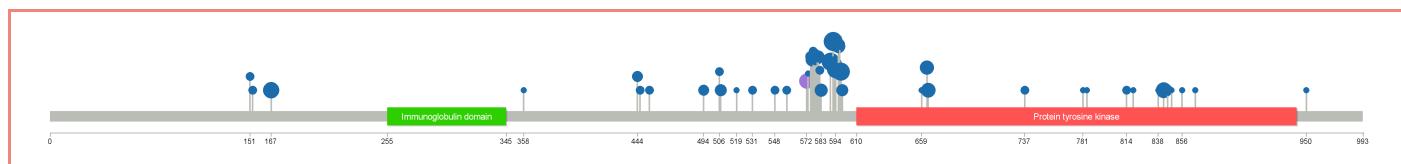


### Substitutions

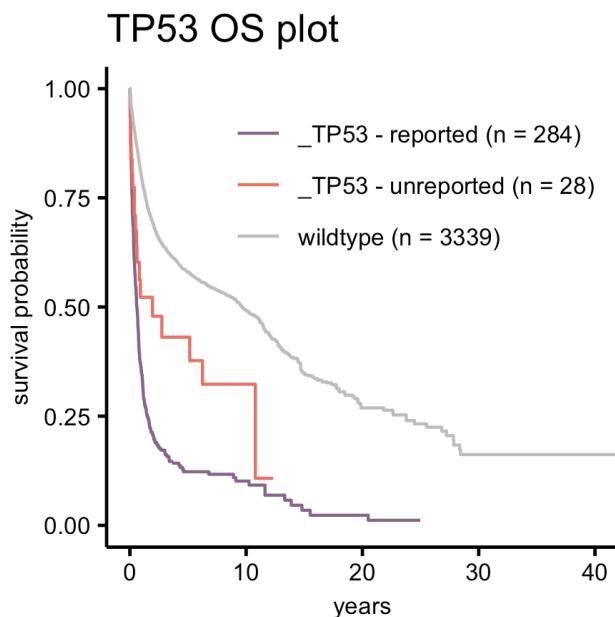
#### reported



#### unreported



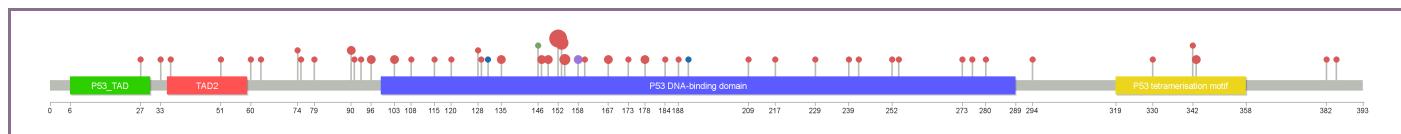
## 5.12 TP53



disease	status consequence	× indel	× sub	✓
AML	inframe	0	0	1
	missense	0	5	353
	nonsense/splicing	1	12	33
	other	2	23	7
	truncating	0	0	43
MDS	inframe	0	0	2
	missense	0	10	227
	nonsense/splicing	3	10	10
	other	0	5	2
	truncating	0	0	40
MPN	missense	0	4	23
	nonsense/splicing	0	0	2
	other	1	5	1
	truncating	0	0	1
TOTAL		7	74	745
		0.01%	0.09%	0.90%

### Indels

#### reported

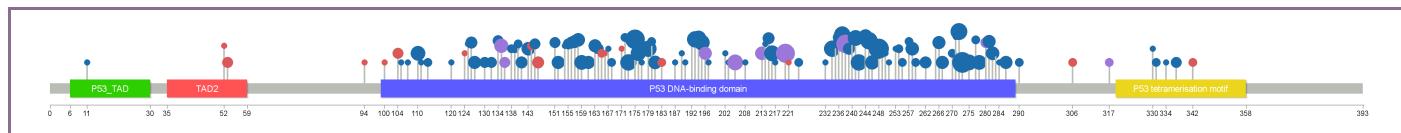


#### unreported

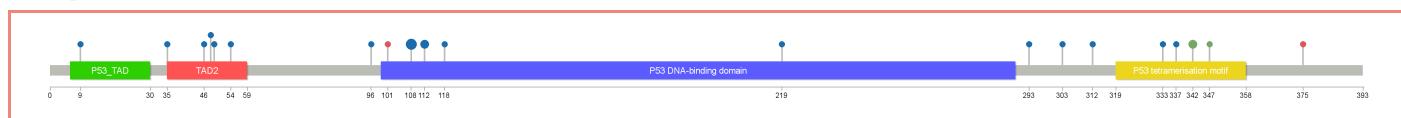
No unreported indel found.

### Substitutions

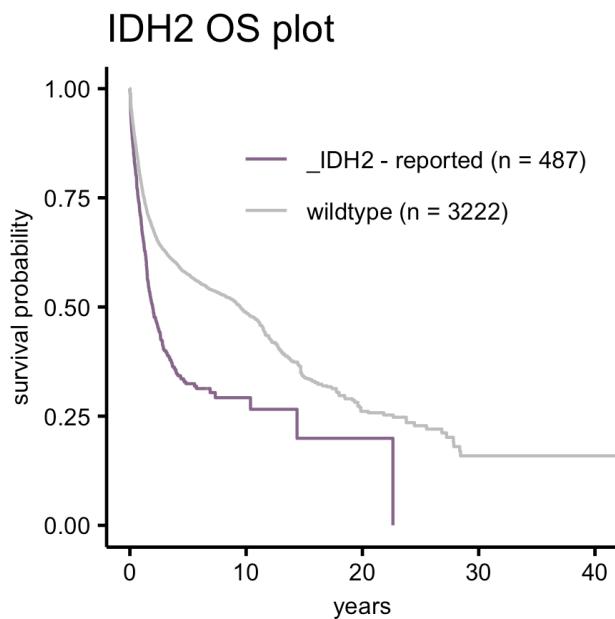
#### reported



#### unreported



## 5.13 IDH2



disease	status consequence	× sub	✓
AML	missense	2	578
MDS	missense	0	123
MPN	missense	0	16
<b>TOTAL</b>		2	717
		0.00%	1.00%

### Indels

#### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

No unreported indel found.

### Substitutions

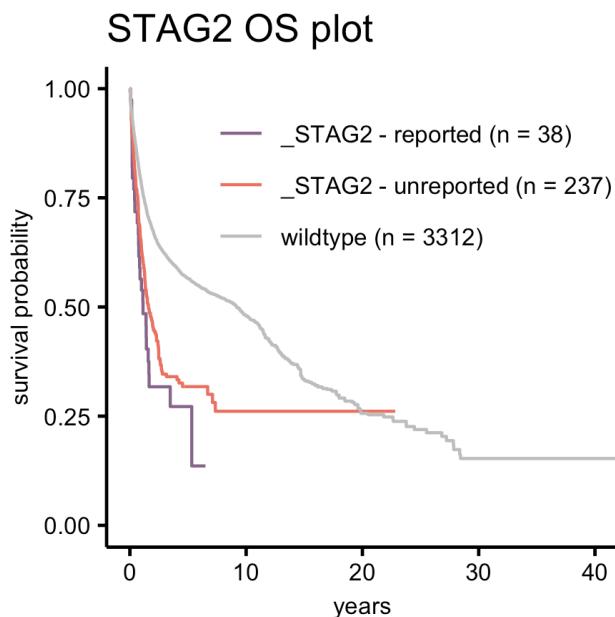
#### reported



#### unreported



## 5.14 STAG2



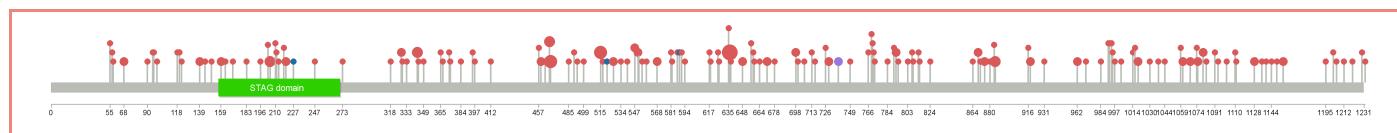
disease	status consequence	× indel	× sub	✓
AML	inframe	3	0	0
	missense	0	16	0
	nonsense/splicing	2	104	39
	other	7	19	0
MDS	truncating	105	0	0
	inframe	1	0	0
	missense	0	3	0
	nonsense/splicing	3	107	35
MPN	other	6	10	0
	truncating	102	0	0
	missense	0	4	0
	nonsense/splicing	0	1	2
TOTAL	other	1	1	0
	truncating	1	0	0
		231	265	76
		0.40%	0.46%	0.13%

### Indels

#### reported

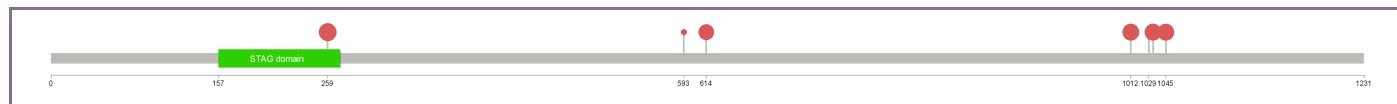
This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

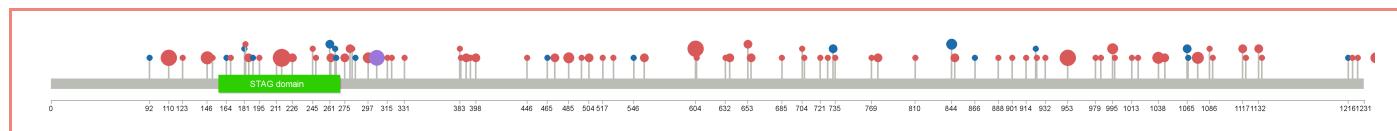


### Substitutions

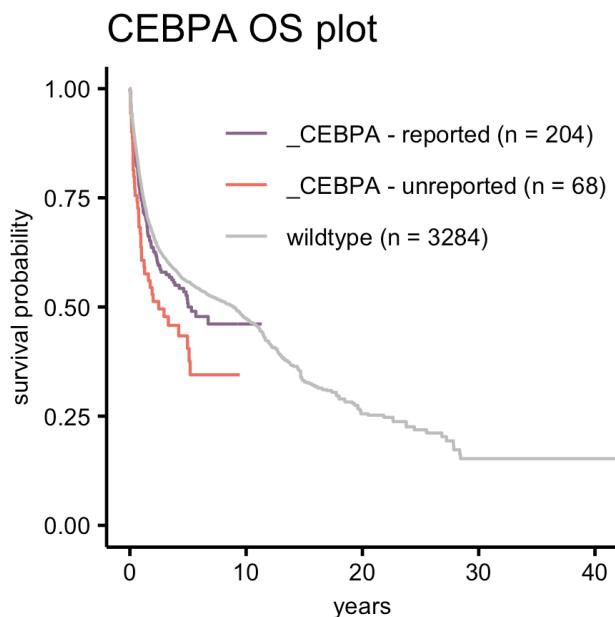
#### reported



#### unreported



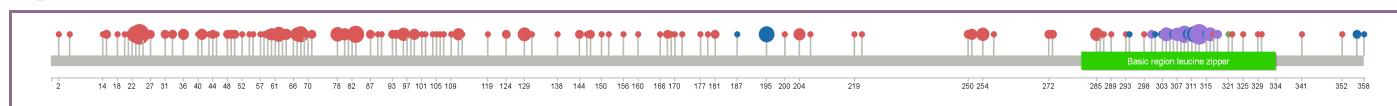
## 5.15 CEBPA



disease	status consequence	× indel	× sub	✓
AML	inframe	0	0	98
	missense	0	88	0
	nonsense/splicing	0	33	0
	other	2	0	17
	truncating	0	0	239
MDS	inframe	0	0	11
	missense	0	14	0
	nonsense/splicing	0	6	0
	truncating	0	0	36
MPN	missense	0	2	0
<b>TOTAL</b>		2	143	401
		0.00%	0.26%	0.73%

### Indels

#### reported



#### unreported

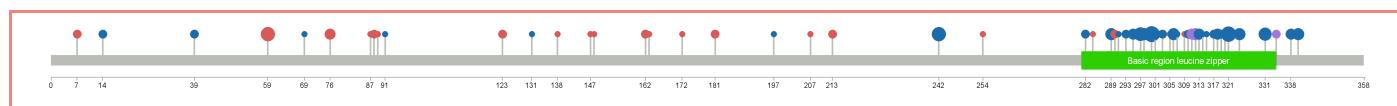
No unreported indel found.

### Substitutions

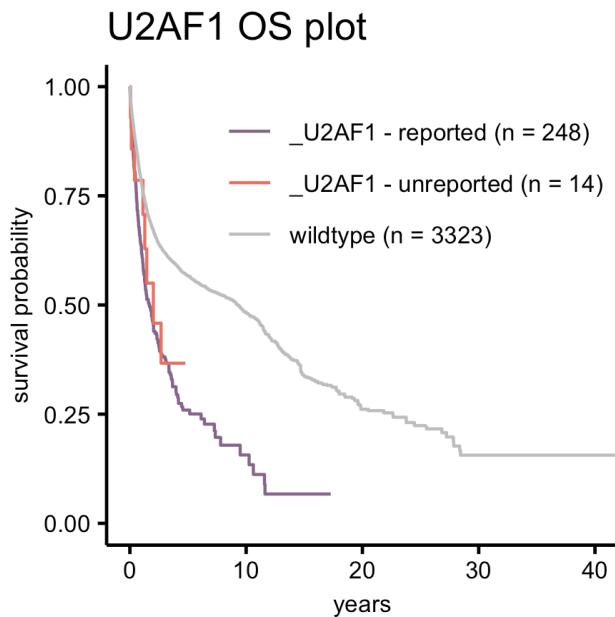
#### reported

This gene was not reported for hostpots (or no reported mutation was found in the dataset).

#### unreported



## 5.16 U2AF1



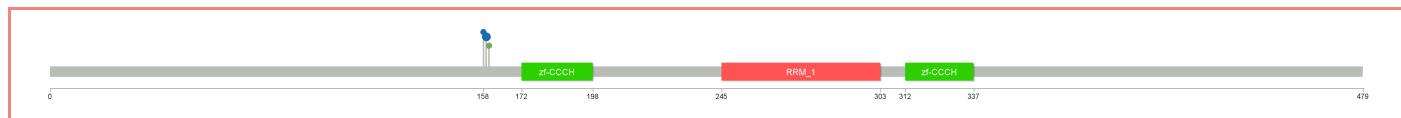
disease	status consequence	× indel	× sub	✓
<b>AML</b>	inframe	1	0	0
	missense	0	14	216
<b>MDS</b>	inframe	1	0	0
	missense	0	4	252
<b>MPN</b>	other	1	0	0
	inframe	1	0	0
	missense	0	0	30
<b>TOTAL</b>	other	1	0	0
		5	18	498
		0.01%	0.03%	0.96%

### Indels

#### reported

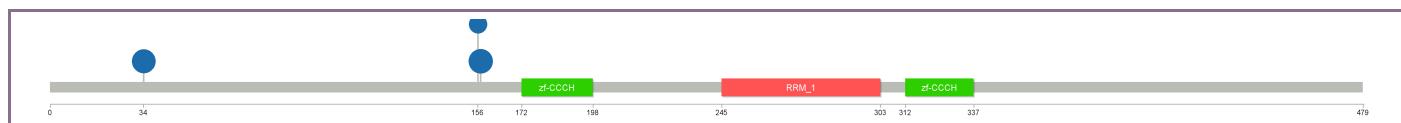
This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

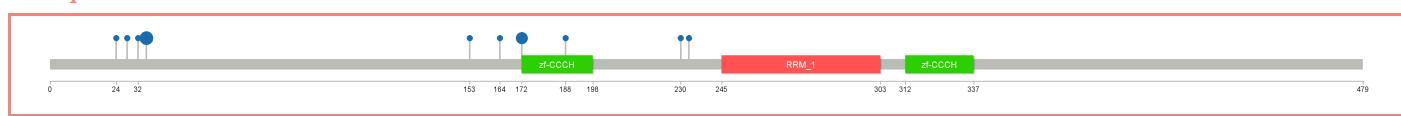


### Substitutions

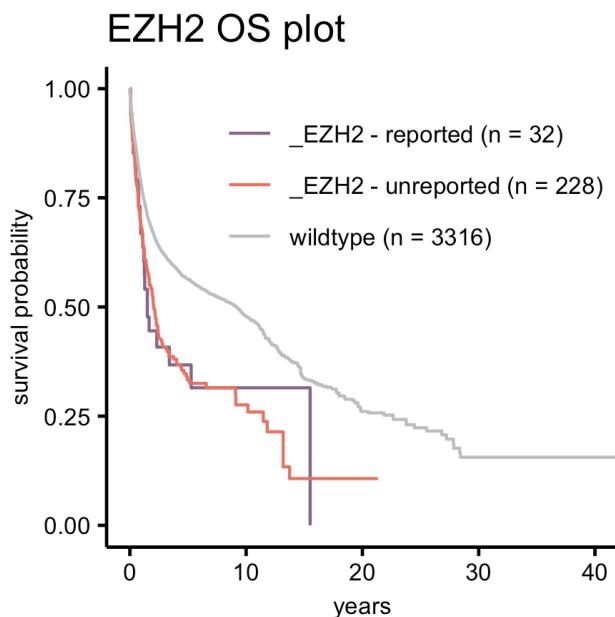
#### reported



#### unreported



## 5.17 EZH2



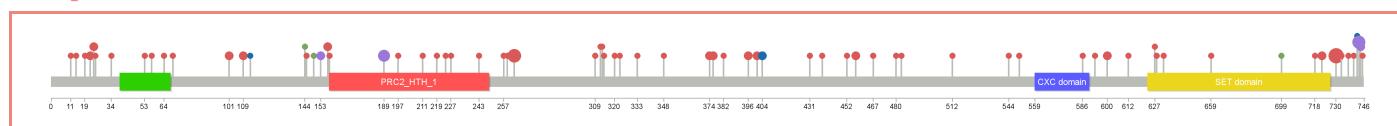
disease	status consequence	× indel	× sub	✓
AML	inframe	5	0	0
	missense	0	91	23
	nonsense/splicing	1	25	0
	other	4	12	0
	truncating	47	0	0
MDS	inframe	1	0	0
	missense	0	91	27
	nonsense/splicing	3	47	3
	other	2	5	0
	truncating	47	0	0
MPN	inframe	3	0	0
	missense	0	31	8
	nonsense/splicing	0	3	0
	other	2	9	0
	truncating	2	0	0
<b>TOTAL</b>		117	314	61
		0.24%	0.64%	0.12%

### Indels

#### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

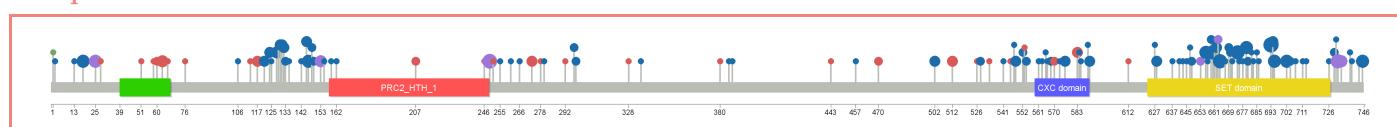


### Substitutions

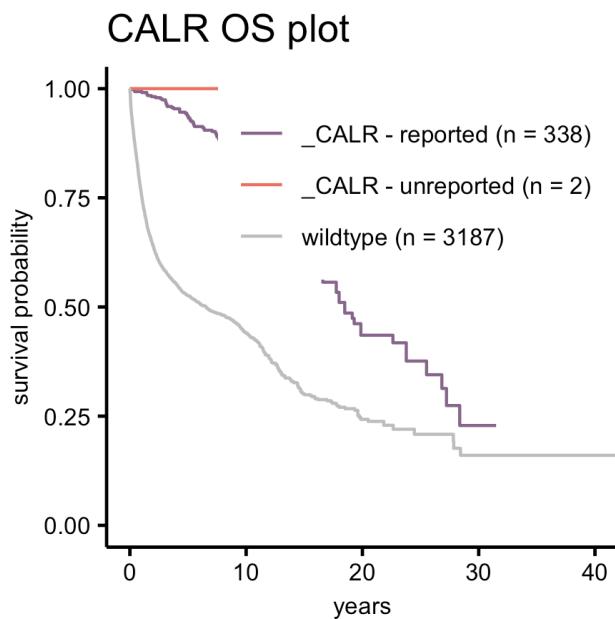
#### reported



#### unreported



## 5.18 CALR



disease	status consequence	× indel	× sub	✓
AML	truncating	0	0	28
MDS	truncating	0	0	4
MPN	nonsense/splicing	0	2	0
	truncating	1	0	446
<b>TOTAL</b>		1	2	478
		0.00%	0.00%	0.99%

### Indels

#### reported



#### unreported



### Substitutions

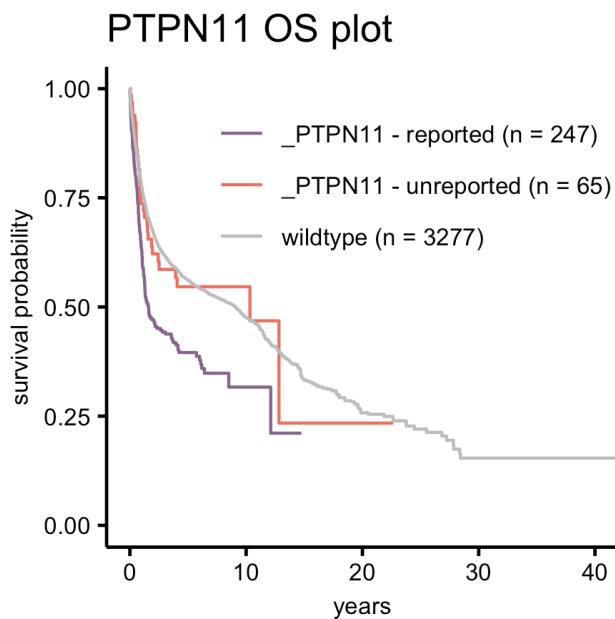
#### reported

This gene was not reported for hotspots (or no reported mutation was found in the dataset).

#### unreported



## 5.19 PTPN11



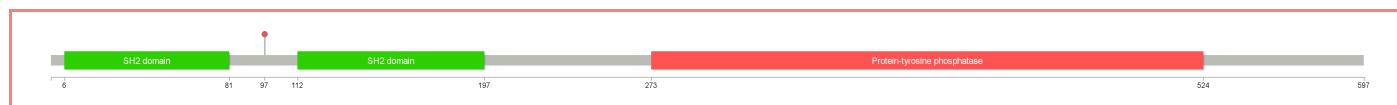
disease	status consequence	× indel	× sub	✓
AML	missense	0	67	329
MDS	missense	0	20	38
MPN	missense	0	9	4
	other	1	0	0
	truncating	1	0	0
<b>TOTAL</b>		2	96	371
		0.00%	0.20%	0.79%

### Indels

#### reported

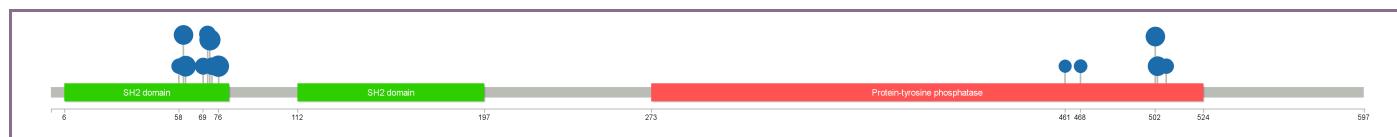
This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

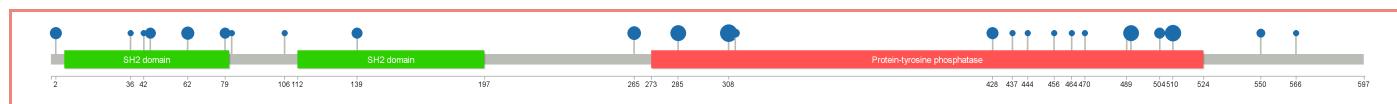


### Substitutions

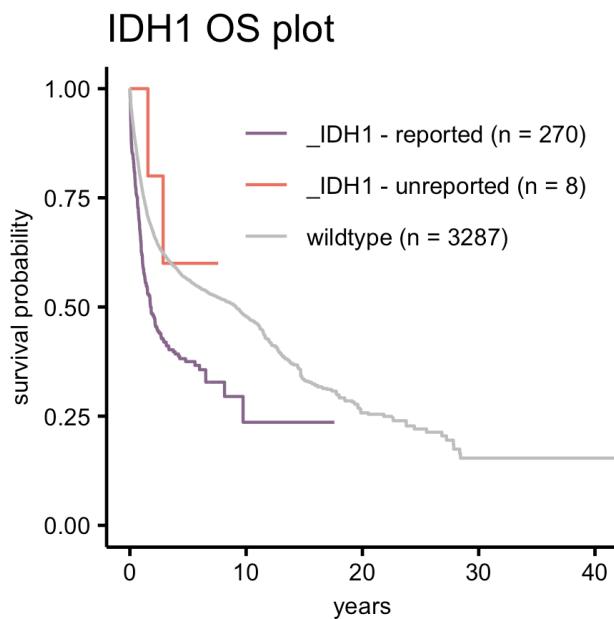
#### reported



#### unreported



## 5.20 IDH1



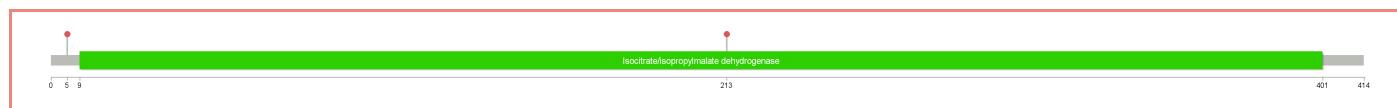
disease	status consequence	× indel	× sub	✓
AML	missense	0	1	337
MDS	missense	0	3	60
	nonsense/splicing	0	1	0
	truncating	1	0	0
MPN	missense	0	1	7
	other	1	0	0
	truncating	1	0	0
<b>TOTAL</b>		3	6	404
		0.01%	0.01%	0.98%

### Indels

#### reported

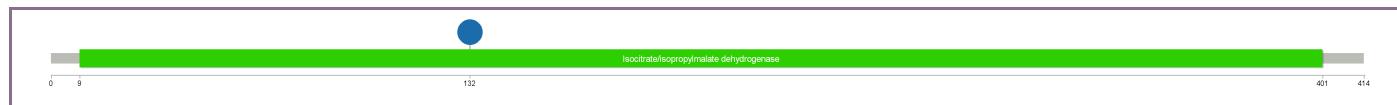
This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported



### Substitutions

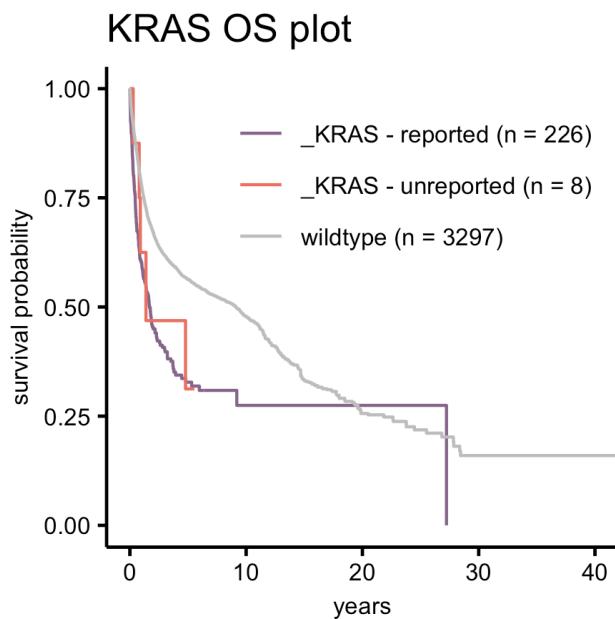
#### reported



#### unreported



## 5.21 KRAS



disease	status consequence	× indel	× sub	✓
AML	inframe	2	0	0
	missense	0	1	258
	nonsense/splicing	0	1	1
MDS	missense	0	8	116
MPN	missense	0	2	10
<b>TOTAL</b>		2	12	385
		0.01%	0.03%	0.96%

### Indels

#### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

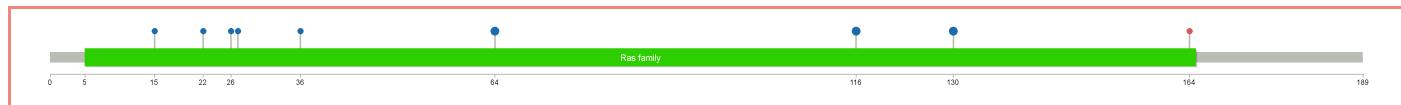


### Substitutions

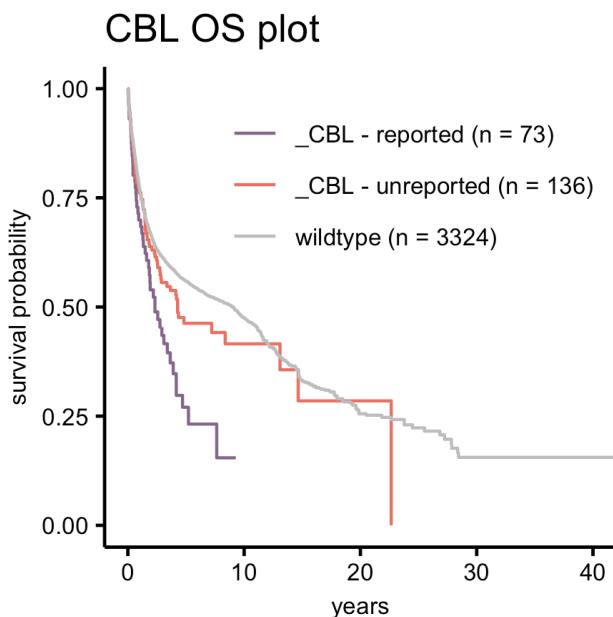
#### reported



#### unreported



## 5.22 CBL



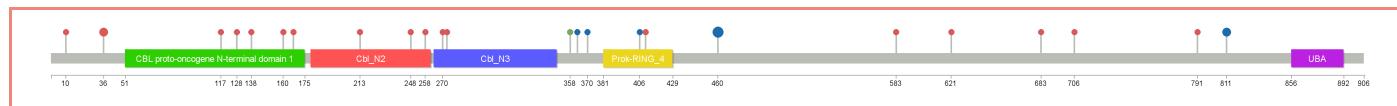
disease	status consequence	× indel	× sub	✓
AML	inframe	1	0	0
	missense	0	65	54
	nonsense/splicing	0	10	0
	other	3	6	0
	truncating	2	0	0
	inframe	2	0	0
MDS	missense	0	84	69
	nonsense/splicing	0	35	0
	other	1	3	0
	truncating	14	0	0
	inframe	5	0	0
	missense	0	24	8
MPN	nonsense/splicing	0	4	0
	other	0	1	0
	truncating	3	0	0
	TOTAL	31	232	131
		0.08%	0.59%	0.33%

### Indels

#### reported

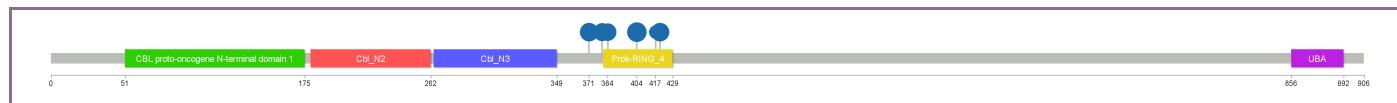
This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

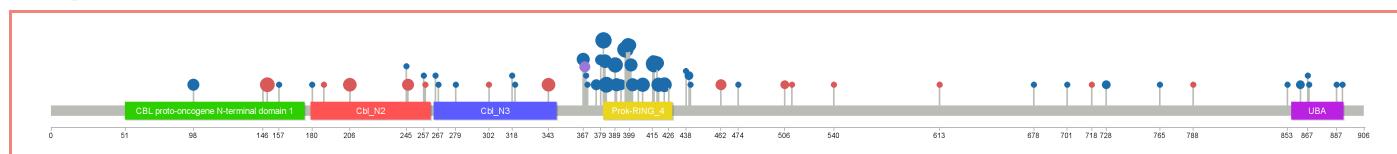


### Substitutions

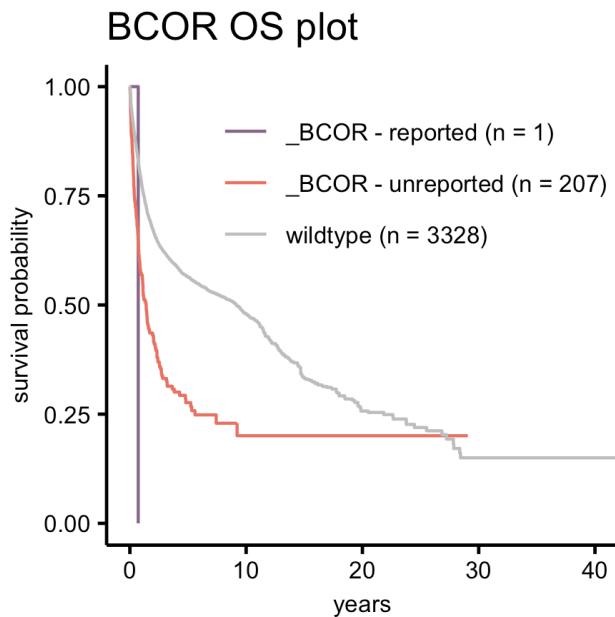
#### reported



#### unreported



## 5.23 BCOR



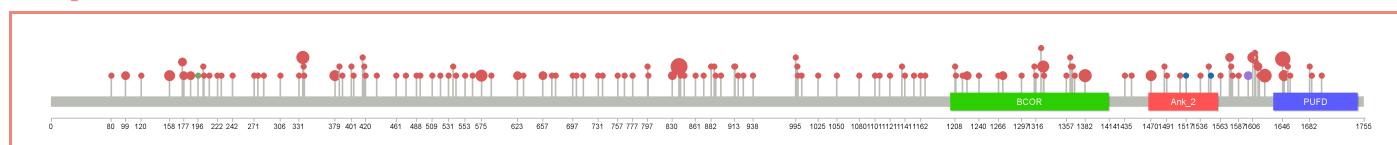
disease	status consequence	x indel	x sub	✓
AML	inframe	2	0	0
	missense	0	14	1
	nonsense/splicing	0	76	0
	other	0	8	0
	truncating	136	0	0
MDS	inframe	1	0	0
	missense	0	8	1
	nonsense/splicing	0	37	0
	other	2	6	0
	truncating	54	0	0
MPN	missense	0	6	0
	truncating	5	0	0
<b>TOTAL</b>		200	155	2
		0.56%	0.43%	0.01%

### Indels

#### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

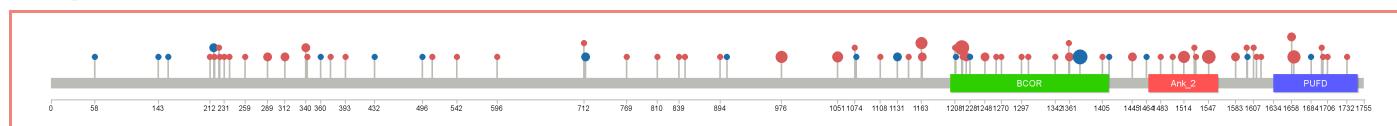


### Substitutions

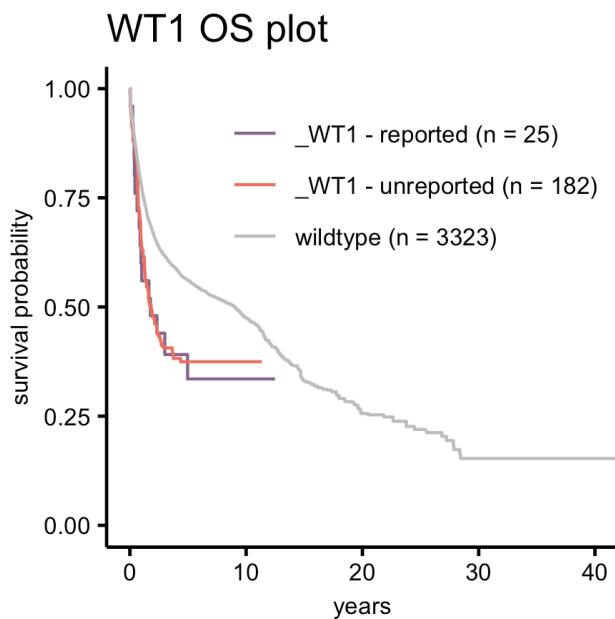
#### reported



#### unreported



## 5.24 WT1



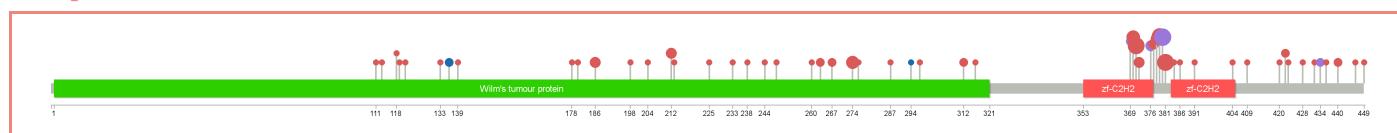
disease	status consequence	× indel	× sub	✓
AML	inframe	6	0	0
	missense	0	67	30
	nonsense/splicing	2	58	0
	other	4	13	0
MDS	truncating	121	0	0
	missense	0	10	1
	nonsense/splicing	0	4	0
TOTAL	truncating	14	0	0
		147	152	31
		0.45%	0.46%	0.09%

### Indels

#### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

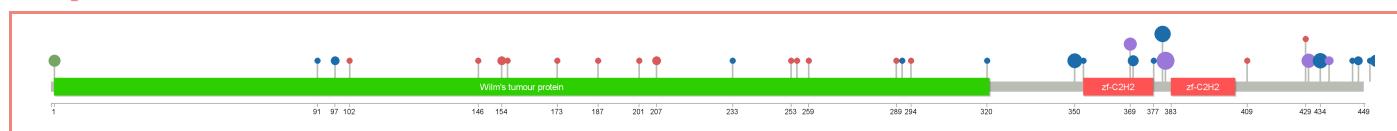


### Substitutions

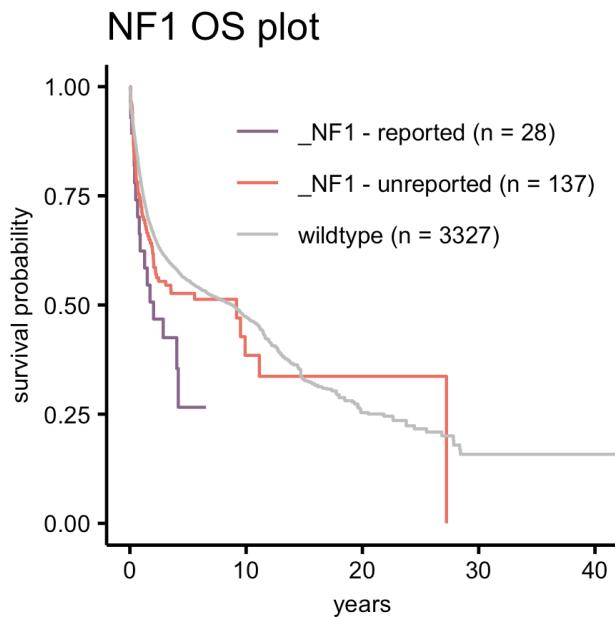
#### reported



#### unreported



## 5.25 NF1



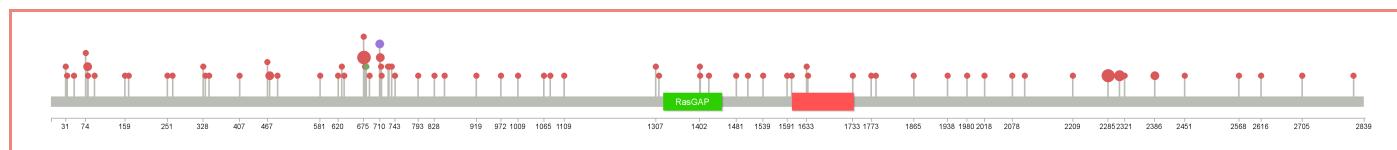
disease	status consequence	× indel	× sub	✓
AML	inframe	1	0	0
	missense	0	24	7
	nonsense/splicing	1	42	19
	other	3	12	0
	truncating	59	0	0
MDS	missense	0	10	12
	nonsense/splicing	1	17	7
	other	1	3	0
	truncating	25	0	0
MPN	missense	0	16	0
	nonsense/splicing	1	1	1
	other	9	5	0
	truncating	3	0	0
<b>TOTAL</b>		104	130	46
		0.37%	0.46%	0.16%

### Indels

#### reported

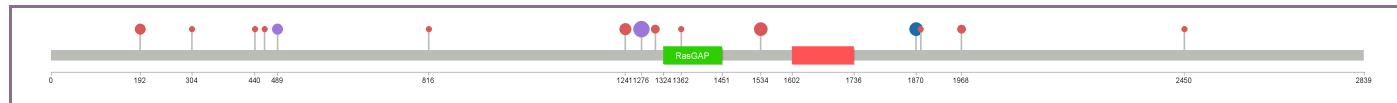
This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

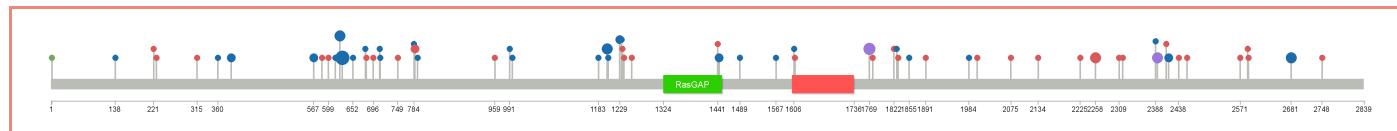


### Substitutions

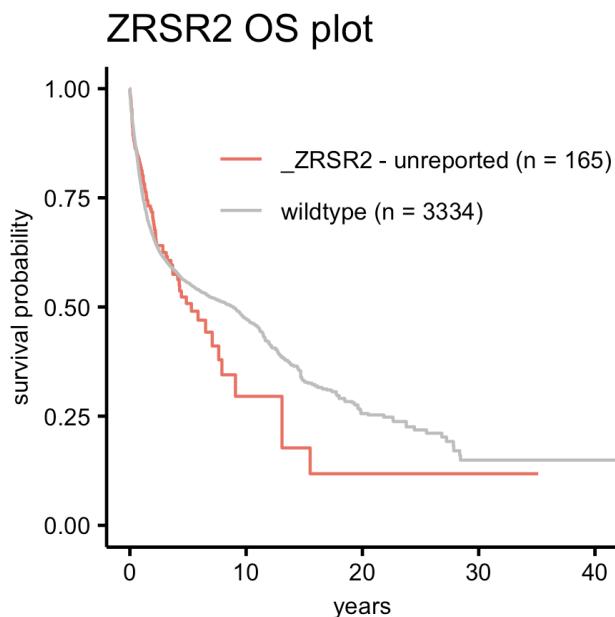
#### reported



#### unreported



## 5.26 ZRSR2



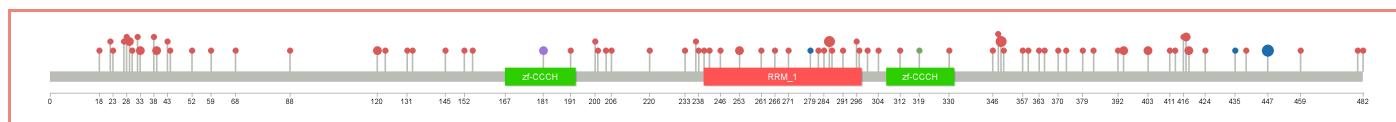
disease	status consequence	× indel	× sub
AML	inframe	1	0
	missense	0	9
	nonsense/splicing	0	22
	other	0	2
	truncating	24	0
MDS	inframe	1	0
	missense	0	43
	nonsense/splicing	2	80
	other	1	15
	truncating	57	0
MPN	inframe	5	0
	missense	0	1
	nonsense/splicing	0	6
	other	0	2
	truncating	9	0
<b>TOTAL</b>		100	180
		0.36%	0.64%

### Indels

#### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

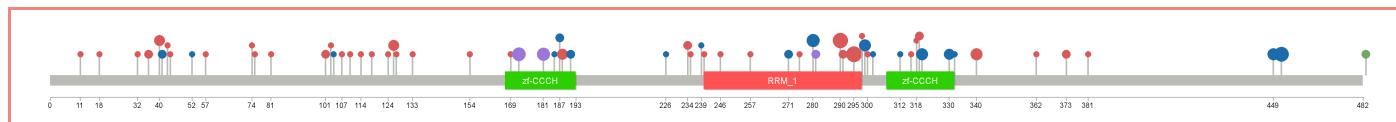


### Substitutions

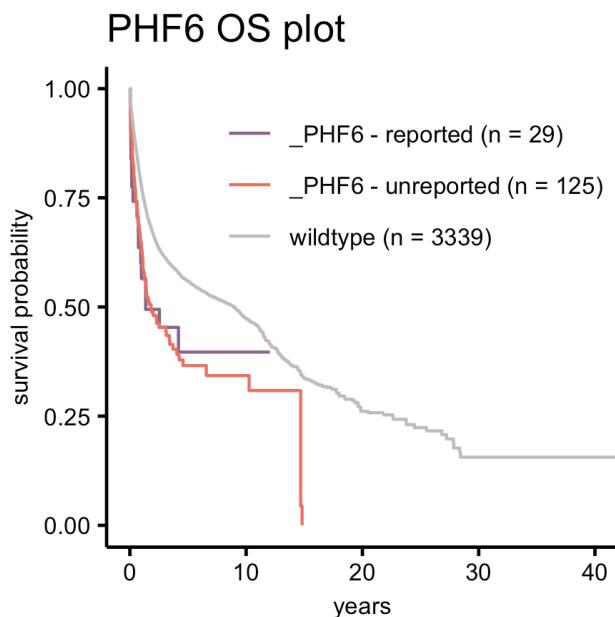
#### reported

This gene was not reported for hotspots (or no reported mutation was found in the dataset).

#### unreported



## 5.27 PHF6



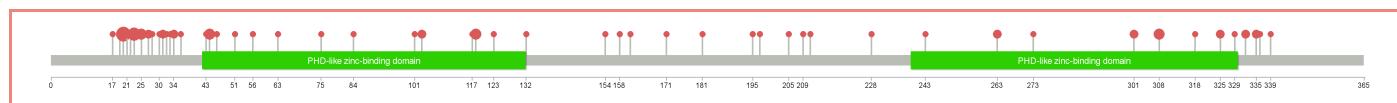
disease	status consequence	× indel	× sub	✓
AML	missense	0	35	12
	nonsense/splicing	1	26	19
	other	2	20	0
	truncating	47	0	0
MDS	missense	0	22	6
	nonsense/splicing	0	21	13
	other	2	5	0
	truncating	27	0	0
MPN	missense	0	3	0
	nonsense/splicing	0	1	0
	other	1	0	0
	truncating	8	0	0
<b>TOTAL</b>		88	133	50
		0.32%	0.49%	0.18%

### Indels

#### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

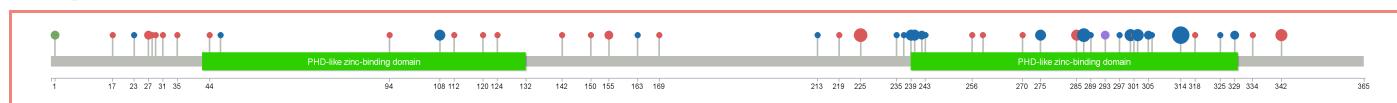


### Substitutions

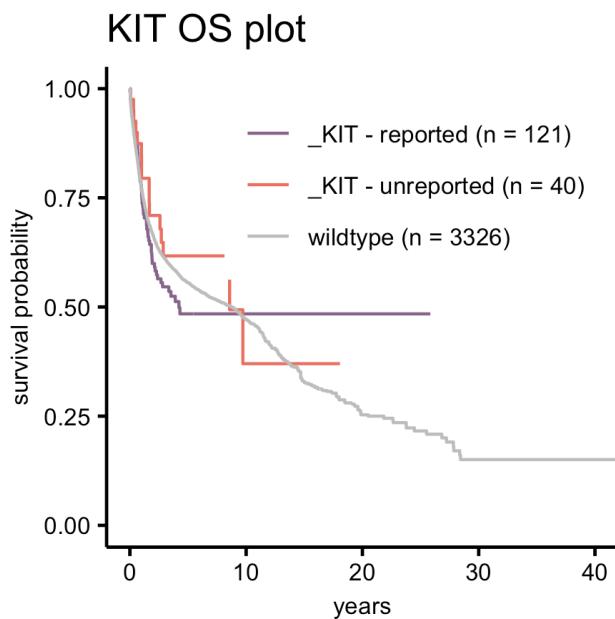
#### reported



#### unreported



## 5.28 KIT



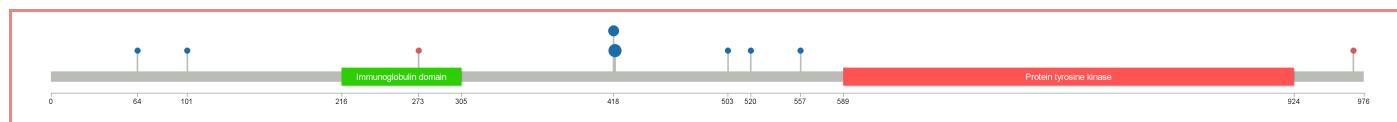
disease	status consequence	× indel	× sub	✓
AML	inframe	13	0	0
	missense	0	18	141
	truncating	1	0	0
MDS	missense	0	12	17
	other	1	0	0
MPN	missense	0	8	3
	truncating	1	0	0
TOTAL		16	38	161
		0.07%	0.18%	0.75%

### Indels

#### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

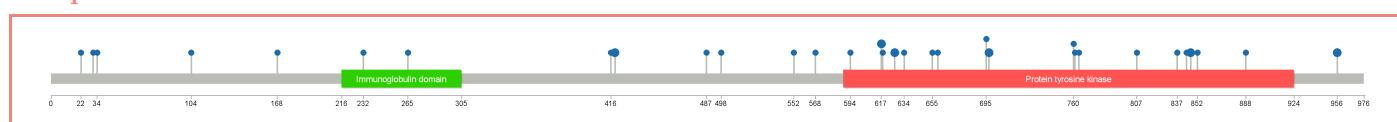


### Substitutions

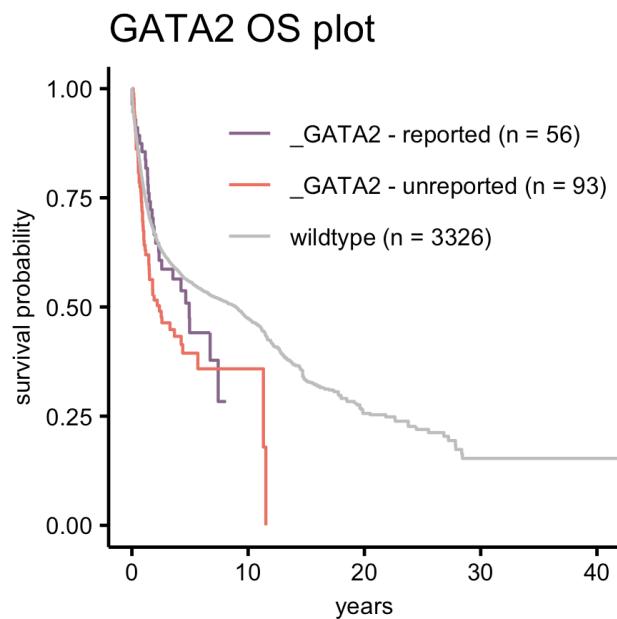
#### reported



#### unreported



## 5.29 GATA2



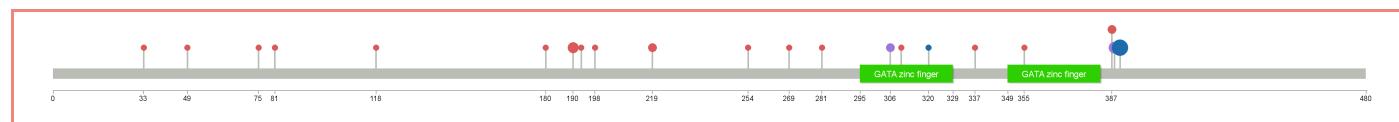
disease	status consequence	× indel	× sub	✓
AML	inframe	7	0	0
	missense	0	78	66
	other	2	0	0
	truncating	15	0	0
MDS	inframe	8	0	0
	missense	0	18	6
	nonsense/splicing	0	1	0
	truncating	6	0	0
MPN	inframe	2	0	0
	missense	0	4	1
	truncating	1	0	0
<b>TOTAL</b>		41	101	73
		0.19%	0.47%	0.34%

### Indels

#### reported

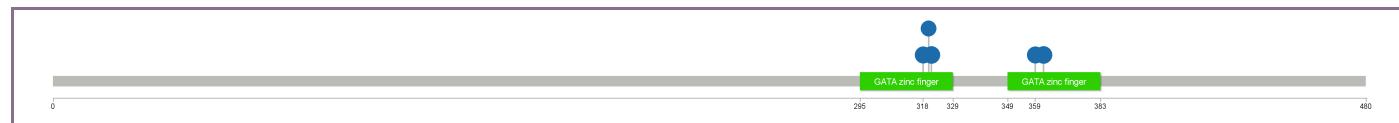
This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

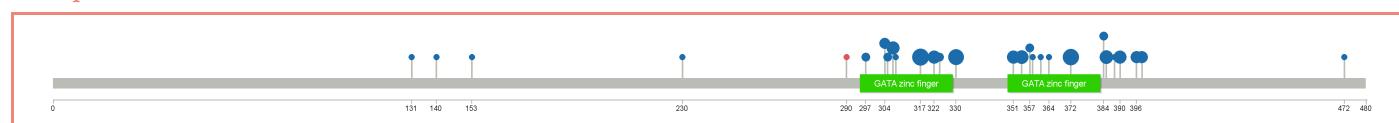


### Substitutions

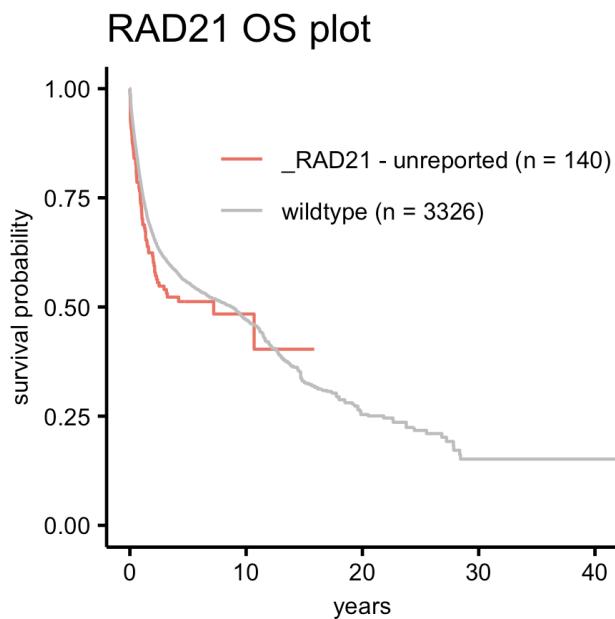
#### reported



#### unreported



### 5.30 RAD21



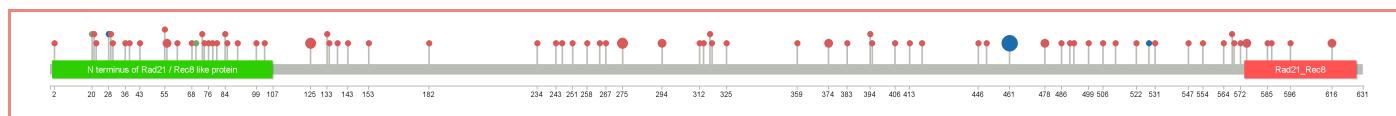
disease	status consequence	× indel	× sub
AML	inframe	8	0
	missense	0	26
	nonsense/splicing	0	40
	other	2	15
	truncating	68	0
MDS	inframe	6	0
	missense	0	8
	nonsense/splicing	0	9
	other	1	4
	truncating	12	0
MPN	other	2	1
	truncating	3	0
<b>TOTAL</b>		102	103
		0.50%	0.50%

#### Indels

##### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

##### unreported

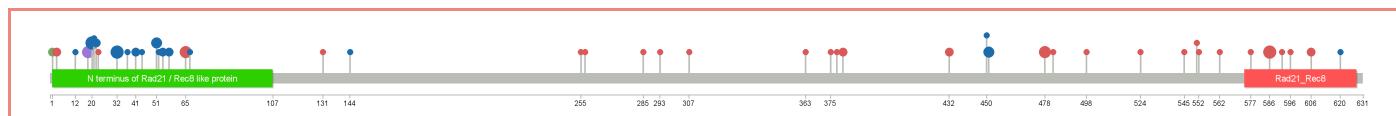


#### Substitutions

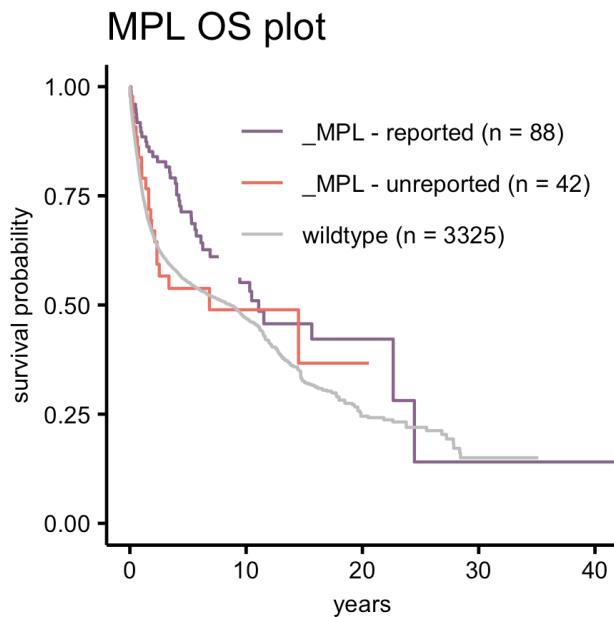
##### reported

This gene was not reported for hostpots (or no reported mutation was found in the dataset).

##### unreported



### 5.31 MPL



disease	status	indel	sub	✓
	consequence			
AML	missense	0	10	20
	nonsense/splicing	0	1	2
	other	1	0	0
MDS	missense	0	15	29
	nonsense/splicing	0	8	1
	truncating	4	0	0
MPN	missense	0	22	75
	nonsense/splicing	0	0	11
	truncating	1	0	0
<b>TOTAL</b>		6	56	138
		0.03%	0.28%	0.69%

#### Indels

##### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

##### unreported

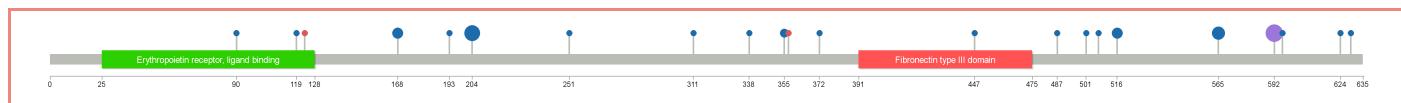


#### Substitutions

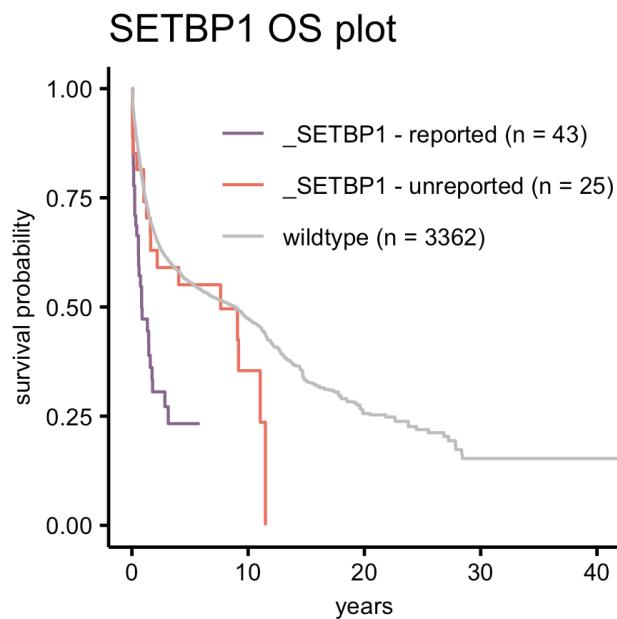
##### reported



##### unreported



## 5.32 SETBP1



disease	status consequence	× sub	✓
AML	missense	20	58
MDS	missense	6	59
MPN	missense	13	0
<b>TOTAL</b>		39	117
		0.25%	0.75%

### Indels

#### reported

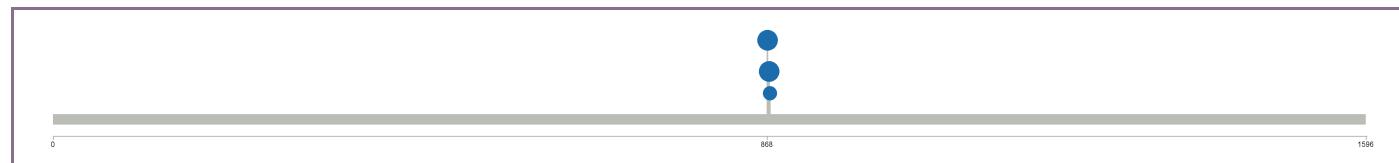
This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

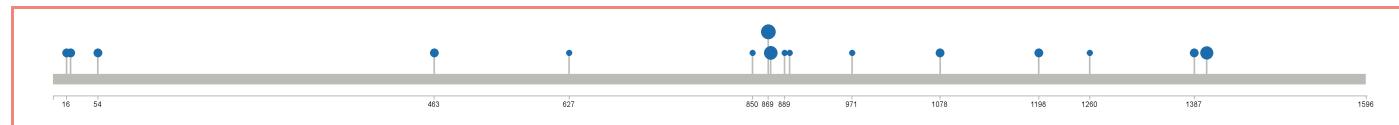
No unreported indel found.

### Substitutions

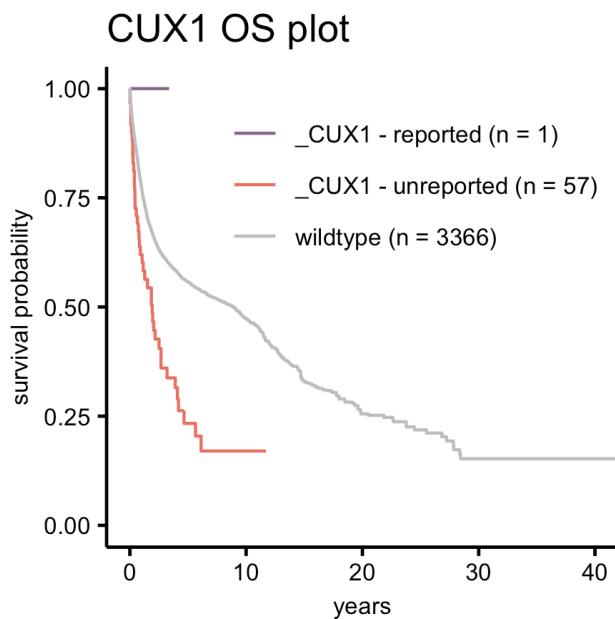
#### reported



#### unreported



### 5.33 CUX1



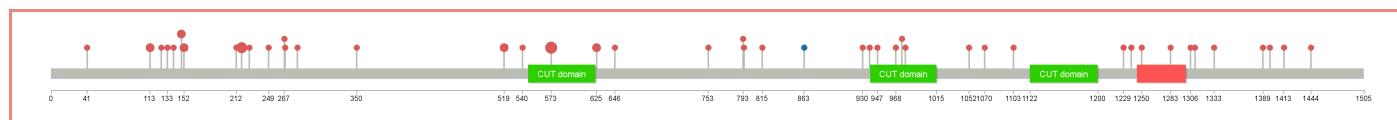
disease	status consequence	× indel	× sub	✓
AML	missense	0	13	0
	nonsense/splicing	0	12	1
	other	1	3	0
MDS	truncating	23	0	0
	inframe	1	0	0
	missense	0	1	0
MPN	nonsense/splicing	1	46	0
	other	0	2	0
	truncating	28	0	0
TOTAL	missense	0	4	0
	nonsense/splicing	0	1	0
	truncating	3	0	0
TOTAL		57	82	1
		0.41%	0.59%	0.01%

#### Indels

##### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

##### unreported

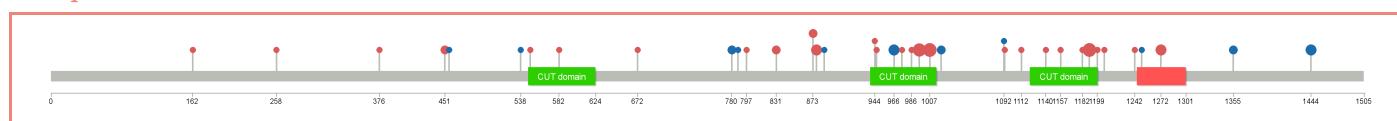


#### Substitutions

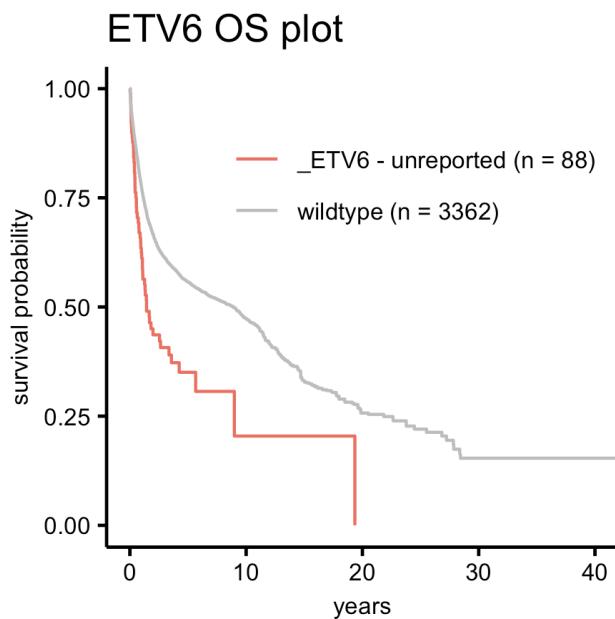
##### reported



##### unreported



## 5.34 ETV6



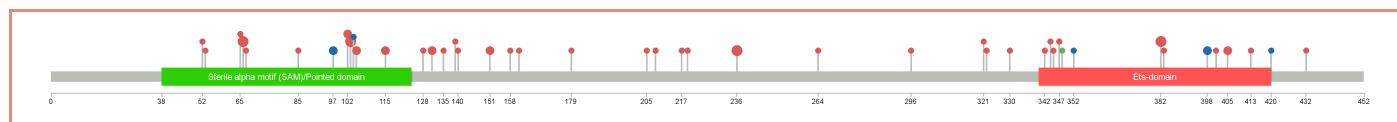
disease	status consequence	× indel	× sub
AML	inframe	5	0
	missense	0	21
	nonsense/splicing	0	13
	other	4	3
	truncating	29	0
MDS	inframe	2	0
	missense	0	19
	nonsense/splicing	1	10
	other	2	2
	truncating	23	0
MPN	truncating	1	0
<b>TOTAL</b>		67	68
		0.50%	0.50%

### Indels

#### reported

This gene was not reported for exons (or no reported mutation was found in the dataset).

#### unreported

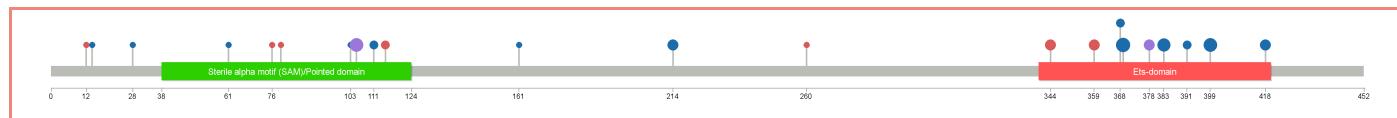


### Substitutions

#### reported

This gene was not reported for hotspots (or no reported mutation was found in the dataset).

#### unreported



## 6 Supplementary material

### 6.1 Unreported substitutions table (with count $\geq 5$ )

Table of the unreported substitutions belonging to genes with partial hotspot reporting.

protein variant	count
JAK2_1108	35
FLT3_592	28
STAG2_216	22
FLT3_598	21
FLT3_590	19
FLT3_594	19
DNMT3A_770	19
WT1_381	19
ASXL1_591	19
TET2_1881	18
TET2_1221	17
PTPN11_308	17
TET2_1167	16
MPL_592	16
PHF6_314	16
RUNX1_110	16
NRAS_64	15
SMC1A_711	15
CBL_416	14
GATA2_317	14
TET2_1298	14
TET2_810	13
STAG2_604	13
GATA2_372	13
DNMT3A_326	13
MYC_73	13
CBL_381	13
WT1_464	12
CBL_396	12
WT1_380	12
STAG2_953	12
PTPN11_510	12
FLT3_167	12
DNMT3A_501	12
ASXL1_775	12
STAG2_110	12
WT1_458	12
DNMT3A_581	12
PTPN11_285	11
MYC_75	11
TET2_1894	11
MPL_204	11
SMC1A_807	11
FLT3_842	11
TET2_1288	11
SRSF2_57	11
DNMT3A_893	10
TET2_1282	10
PTPN11_491	10

Continued on next page

protein variant	count
TET2_1193	10
TET2_1262	10
GATA2_330	10
CBL_383	10
STAG2_305	10
TET2_r.4814+5g>a	10
DNMT3A_547	10
CBL_389	9
TP53_r.973-1g>a	9
WT1_434	9
RUNX1_106	9
FLT3_572	9
RUNX1_111	9
RUNX1_165	9
EZH2_693	9
RUNX1_425	9
TET2_1355	9
FLT3_664	8
RB1_786	8
WT1_465	8
FLT3_579	8
TET2_1912	8
DNMT3A_796	8
CTCF_339	8
EZH2_249	8
TET2_1289	8
DNMT3A_732	8
RUNX1_112	8
TET2_1861	8
DNMT3A_r.2005+1g>a	8
BCOR_1217	8
EZH2_r.1029+5g>a	8
RUNX1_52	8
TET2_r.4663+5g>a	8
ASXL1_1304	8
SETBP1_869	8
CBL_423	8
TET2_1233	8
TET2_417	8
TET2_1361	8
ASXL1_588	8
TET2_1030	8
ASXL1_404	8
CBL_149	8
BCOR_1375	8
WT1_350	8
DNMT3A_780	8
TET2_705	8
EP300_2266	8
WT1_430	7
ASXL1_829	7
DNMT3A_330	7
TET2_1819	7
TET2_1211	7

Continued on next page

protein variant	count
TET2_r.4904+1g>a	7
ASXL1_512	7
TET2_1414	7
CBL_398	7
SETD2_629	7
TET2_1358	7
FLT3_663	7
TET2_1370	7
RUNX1_105	7
EZH2_674	7
TET2_1376	7
TET2_1191	7
NF1_629	7
CBL_408	7
CBL_419	7
ASXL1_602	7
EZH2_664	7
PHF6_r.932-2a>g	7
EZH2_731	7
TET2_531	7
RUNX1_207	7
SMC1A_1131	7
TET2_530	7
FLT3_593	7
EZH2_660	7
GATA2_354	6
TET2_1680	6
DNMT3A_310	6
SETBP1_872	6
DNMT3A_699	6
RUNX1_322	6
DNMT3A_707	6
DNMT3A_899	6
TET2_r.4905-1g>a	6
TET2_1687	6
TET2_1273	6
CBL_382	6
DNMT3A_716	6
CBL_343	6
TP53_r.749+1g>a	6
SRSF2_96	6
FLT3_596	6
TET2_1382	6
TET2_910	6
CUX1_1007	6
CBL_418	6
TET2_792	6
RUNX1_267	6
CREBBP_1745	6
ASXL1_491	6
ASXL1_708	6
TET2_r.4269+1g>a	6
GATA2_385	6
DNMT3A_731	6

Continued on next page

protein variant	count
PHF6_225	6
TET2_1848	6
TET2_1182	6
DNMT3A_313	6
TET2_1291	6
TET2_1203	6
U2AF1_35	6
TET2_631	6
BCOR_1547	6
DNMT3A_842	6
CUX1_1190	6
TET2_1387	6
DNMT3A_879	6
DNMT3A_761	6
DNMT3A_320	6
TET2_1828	6
PHF6_287	6
CTCF_371	6
CBL_206	6
DNMT3A_r.1767+1g>a	6
DNMT3A_685	6
TET2_1374	6
PTPN11_265	6
ASXL1_592	6
BCOR_r.4657+1g>a	6
ASXL1_796	6
TET2_1872	6
ASXL1_1068	6
TET2_1274	6
DNMT3A_700	6
TET2_1624	6
SH2B3_566	5
TET2_1384	5
TET2_1377	5
CUX1_995	5
CBL_400	5
TET2_1300	5
EZH2_695	5
TET2_593	5
EZH2_679	5
ASXL1_1226	5
FLT3_576	5
EZH2_146	5
CBL_367	5
GATA2_351	5
RUNX1_r.796+1g>a	5
JAK2_797	5
RUNX1_r.953+2u>c	5
RUNX1_196	5
TET2_769	5
CSF3R_776	5
RUNX1_161	5
DNMT3A_751	5
TET2_429	5

Continued on next page

protein variant	count
TET2_1661	5
TET2_745	5
EZH2_702	5
ASXL1_760	5
TET2_1868	5
EZH2_745	5
ASXL1_692	5
ASXL1_1415	5
ASXL2_635	5
TET2_1197	5
EZH2_656	5
SETBP1_1402	5
DNMT3A_880	5
EZH2_18	5
TET2_1219	5
DNMT3A_296	5
EP300_638	5
TET2_635	5
ASXL2_614	5
STAG2_146	5
RUNX1_272	5
CBL_390	5
CBL_401	5
PTPN11_62	5
TET2_1327	5
TET2_r.4814+1g>u	5
BCOR_1661	5
DNMT3A_777	5
MPL_565	5
FLT3_583	5
DNMT3A_644	5
CTCF_342	5
TET2_1337	5
CBL_399	5
TET2_1255	5
GATA2_390	5
GATA2_322	5
TET2_1631	5
CSF3R_770	5
WT1_369	5
GATA2_307	5
TET2_r.4454+1g>u	5
TET2_1513	5
SH2B3_395	5
EZH2_25	5
DNMT3A_646	5
TET2_866	5

## 6.2 Gene summary table

gene	count	freq	unreported	reported	unreported	sub	unreported	indel	% unreported
<b>TET2</b>	3140	12.33%		1347	1793		1320	27	42.9%
<b>DNMT3A</b>	1768	6.94%		634	1134		629	5	35.9%
<b>JAK2</b>	1700	6.68%		69	1631		57	12	4.1%
<b>NPM1</b>	1452	5.70%		4	1448		0	4	0.3%
<b>SRSF2</b>	1294	5.08%		224	1070		23	201	17.3%
<b>RUNX1</b>	1165	4.58%		317	848		292	25	27.2%
<b>ASXL1</b>	1037	4.07%		352	685		321	31	33.9%
<b>NRAS</b>	1032	4.05%		19	1013		19	0	1.8%
<b>SF3B1</b>	917	3.60%		44	873		38	6	4.8%
<b>FLT3</b>	862	3.39%		234	628		233	1	27.1%
<b>TP53</b>	826	3.24%		81	745		74	7	9.8%
<b>IDH2</b>	719	2.82%		2	717		2	0	0.3%
<b>STAG2</b>	572	2.25%		496	76		265	231	86.7%
<b>CEBPA</b>	546	2.14%		145	401		143	2	26.6%
<b>U2AF1</b>	521	2.05%		23	498		18	5	4.4%
<b>EZH2</b>	492	1.93%		431	61		314	117	87.6%
<b>CALR</b>	481	1.89%		3	478		2	1	0.6%
<b>PTPN11</b>	469	1.84%		98	371		96	2	20.9%
<b>IDH1</b>	413	1.62%		9	404		6	3	2.2%
<b>KRAS</b>	399	1.57%		14	385		12	2	3.5%
<b>CBL</b>	394	1.55%		263	131		232	31	66.8%
<b>BCOR</b>	357	1.40%		355	2		155	200	99.4%
<b>WT1</b>	330	1.30%		299	31		152	147	90.6%
<b>NF1</b>	280	1.10%		234	46		130	104	83.6%
<b>ZRSR2</b>	280	1.10%		280	0		180	100	100.0%
<b>PHF6</b>	271	1.06%		221	50		133	88	81.5%
<b>KIT</b>	215	0.84%		54	161		38	16	25.1%
<b>GATA2</b>	215	0.84%		142	73		101	41	66.0%
<b>RAD21</b>	205	0.81%		205	0		103	102	100.0%
<b>MPL</b>	200	0.79%		62	138		56	6	31.0%
<b>SETBP1</b>	156	0.61%		39	117		39	0	25.0%
<b>CUX1</b>	140	0.55%		139	1		82	57	99.3%
<b>ETV6</b>	135	0.53%		135	0		68	67	100.0%
<b>KMT2C</b>	110	0.43%		110	0		45	65	100.0%
<b>SMC3</b>	93	0.37%		93	0		71	22	100.0%
<b>SMC1A</b>	91	0.36%		71	20		61	10	78.0%
<b>EP300</b>	87	0.34%		81	6		49	32	93.1%
<b>KMT2D</b>	82	0.32%		82	0		37	45	100.0%
<b>MYC</b>	80	0.31%		32	48		27	5	40.0%
<b>PPM1D</b>	77	0.30%		77	0		29	48	100.0%
<b>NFE2</b>	71	0.28%		71	0		15	56	100.0%
<b>PRPF8</b>	63	0.25%		63	0		62	1	100.0%
<b>CSF3R</b>	63	0.25%		31	32		29	2	49.2%
<b>ASXL2</b>	58	0.23%		58	0		14	44	100.0%
<b>FBXW7</b>	56	0.22%		47	9		9	38	83.9%
<b>CTCF</b>	54	0.21%		52	2		48	4	96.3%
<b>KDM6A</b>	53	0.21%		52	1		22	30	98.1%
<b>ATRX</b>	52	0.20%		51	1		28	23	98.1%
<b>SH2B3</b>	51	0.20%		51	0		24	27	100.0%
<b>BRAF</b>	48	0.19%		6	42		4	2	12.5%
<b>GNAS</b>	47	0.18%		19	28		15	4	40.4%

Continued on next page

gene	count	freq	unreported	reported	unreported	sub	unreported	indel	%	unreported
<b>NOTCH1</b>	41	0.16%		35	6		19		16	85.4%
<b>BCORL1</b>	40	0.16%		40	0		19		21	100.0%
<b>CREBBP</b>	38	0.15%		38	0		23		15	100.0%
<b>U2AF2</b>	37	0.15%		37	0		17		20	100.0%
<b>ATM</b>	36	0.14%		32	4		27		5	88.9%
<b>RB1</b>	36	0.14%		20	16		18		2	55.6%
<b>RIT1</b>	33	0.13%		33	0		33		0	100.0%
<b>ETNK1</b>	30	0.12%		1	29		1		0	3.3%
<b>KMT2E</b>	27	0.11%		27	0		3		24	100.0%
<b>PCDH15</b>	26	0.10%		26	0		0		26	100.0%
<b>BRCC3</b>	26	0.10%		26	0		22		4	100.0%
<b>KDM5A</b>	23	0.09%		23	0		22		1	100.0%
<b>LAMB4</b>	21	0.08%		21	0		20		1	100.0%
<b>SF1</b>	21	0.08%		21	0		16		5	100.0%
<b>SETD2</b>	20	0.08%		20	0		17		3	100.0%
<b>LUC7L2</b>	19	0.07%		19	0		18		1	100.0%
<b>MED12</b>	19	0.07%		15	4		15		0	78.9%
<b>CHEK2</b>	19	0.07%		15	4		1		14	78.9%
<b>KMT2A</b>	18	0.07%		18	0		11		7	100.0%
<b>GNB1</b>	18	0.07%		18	0		18		0	100.0%
<b>ARID2</b>	18	0.07%		18	0		11		7	100.0%
<b>GATA1</b>	16	0.06%		16	0		4		12	100.0%
<b>SUZ12</b>	16	0.06%		16	0		4		12	100.0%
<b>NCOR2</b>	15	0.06%		15	0		13		2	100.0%
<b>KANSL1</b>	15	0.06%		15	0		9		6	100.0%
<b>JAK3</b>	15	0.06%		7	8		6		1	46.7%
<b>DST</b>	14	0.05%		14	0		0		14	100.0%
<b>CACNA2D3</b>	14	0.05%		14	0		0		14	100.0%
<b>CDKN2A</b>	13	0.05%		13	0		12		1	100.0%
<b>MN1</b>	13	0.05%		13	0		6		7	100.0%
<b>PTPRD</b>	12	0.05%		12	0		10		2	100.0%
<b>PDS5B</b>	11	0.04%		11	0		4		7	100.0%
<b>EED</b>	11	0.04%		11	0		9		2	100.0%
<b>NCL</b>	10	0.04%		10	0		0		10	100.0%
<b>ROBO1</b>	10	0.04%		10	0		7		3	100.0%
<b>KDM2B</b>	10	0.04%		10	0		7		3	100.0%
<b>IRF1</b>	9	0.04%		9	0		5		4	100.0%
<b>NCOR1</b>	9	0.04%		9	0		8		1	100.0%
<b>FANCA</b>	9	0.04%		9	0		4		5	100.0%
<b>FANCM</b>	8	0.03%		8	0		6		2	100.0%
<b>ARID1A</b>	8	0.03%		7	1		3		4	87.5%
<b>DCLRE1C</b>	8	0.03%		8	0		4		4	100.0%
<b>ARID1B</b>	8	0.03%		8	0		4		4	100.0%
<b>MBD1</b>	8	0.03%		8	0		8		0	100.0%
<b>PTPRT</b>	8	0.03%		8	0		8		0	100.0%
<b>NSD1</b>	8	0.03%		8	0		4		4	100.0%
<b>KDR</b>	8	0.03%		8	0		8		0	100.0%
<b>PTEN</b>	7	0.03%		6	1		4		2	85.7%
<b>FANCL</b>	7	0.03%		7	0		1		6	100.0%
<b>GPRC5A</b>	7	0.03%		7	0		0		7	100.0%
<b>CBFB</b>	7	0.03%		7	0		3		4	100.0%
<b>ASPM</b>	7	0.03%		7	0		0		7	100.0%
<b>DNMT3B</b>	7	0.03%		7	0		6		1	100.0%

Continued on next page

gene	count	freq	unreported	reported	unreported	sub	unreported	indel	%	unreported
<b>PHIP</b>	7	0.03%	7	0		2		5	100.0%	
<b>PTPRB</b>	7	0.03%	7	0		0		7	100.0%	
<b>PRR14L</b>	7	0.03%	7	0		0		7	100.0%	
<b>TP53BP1</b>	7	0.03%	7	0		6		1	100.0%	
<b>ASH1L</b>	7	0.03%	7	0		7		0	100.0%	
<b>AHNAK</b>	6	0.02%	6	0		0		6	100.0%	
<b>PIGA</b>	6	0.02%	6	0		3		3	100.0%	
<b>SF3A1</b>	6	0.02%	6	0		6		0	100.0%	
<b>PRPF40B</b>	6	0.02%	6	0		6		0	100.0%	
<b>STAT5B</b>	6	0.02%	1	5		1		0	16.7%	
<b>SRRM2</b>	6	0.02%	6	0		0		6	100.0%	
<b>STAG3</b>	5	0.02%	5	0		4		1	100.0%	
<b>DOT1L</b>	5	0.02%	5	0		5		0	100.0%	
<b>ZBTB33</b>	5	0.02%	5	0		0		5	100.0%	
<b>ELF1</b>	5	0.02%	5	0		0		5	100.0%	
<b>MRE11A</b>	5	0.02%	5	0		2		3	100.0%	
<b>HUWE1</b>	5	0.02%	5	0		0		5	100.0%	
<b>SRRM1</b>	4	0.02%	4	0		1		3	100.0%	
<b>SYNE2</b>	4	0.02%	4	0		0		4	100.0%	
<b>H3F3A</b>	4	0.02%	4	0		4		0	100.0%	
<b>SMG1</b>	4	0.02%	4	0		4		0	100.0%	
<b>STAT3</b>	4	0.02%	0	4		0		0	0.0%	
<b>TG</b>	4	0.02%	4	0		0		4	100.0%	
<b>WHSC1</b>	4	0.02%	4	0		4		0	100.0%	
<b>PHF8</b>	4	0.02%	4	0		4		0	100.0%	
<b>NIPBL</b>	4	0.02%	4	0		4		0	100.0%	
<b>NF2</b>	4	0.02%	4	0		4		0	100.0%	
<b>FANCD2</b>	4	0.02%	4	0		4		0	100.0%	
<b>CMYA5</b>	4	0.02%	4	0		0		4	100.0%	
<b>NOTCH2</b>	4	0.02%	4	0		3		1	100.0%	
<b>HNRNPK</b>	4	0.02%	4	0		0		4	100.0%	
<b>ZFHX4</b>	3	0.01%	3	0		0		3	100.0%	
<b>YLPM1</b>	3	0.01%	3	0		3		0	100.0%	
<b>CDKN1B</b>	3	0.01%	3	0		0		3	100.0%	
<b>ALAS2</b>	3	0.01%	3	0		3		0	100.0%	
<b>DICER1</b>	3	0.01%	3	0		0		3	100.0%	
<b>ABCG8</b>	3	0.01%	3	0		3		0	100.0%	
<b>ASXL3</b>	3	0.01%	3	0		2		1	100.0%	
<b>NRD1</b>	3	0.01%	3	0		0		3	100.0%	
<b>HIPK2</b>	3	0.01%	3	0		0		3	100.0%	
<b>ITPR1</b>	3	0.01%	3	0		0		3	100.0%	
<b>FARS2</b>	3	0.01%	3	0		0		3	100.0%	
<b>CSF2RB</b>	3	0.01%	3	0		2		1	100.0%	
<b>RAD50</b>	3	0.01%	3	0		1		2	100.0%	
<b>KIAA1549</b>	3	0.01%	3	0		0		3	100.0%	
<b>GRID1</b>	3	0.01%	3	0		3		0	100.0%	
<b>SARDH</b>	2	0.01%	2	0		0		2	100.0%	
<b>MLH3</b>	2	0.01%	2	0		1		1	100.0%	
<b>FANCB</b>	2	0.01%	2	0		0		2	100.0%	
<b>FAT4</b>	2	0.01%	2	0		0		2	100.0%	
<b>C17orf74</b>	2	0.01%	2	0		2		0	100.0%	
<b>CSMD1</b>	2	0.01%	2	0		0		2	100.0%	
<b>CTNNA1</b>	2	0.01%	2	0		1		1	100.0%	

Continued on next page

gene	count	freq	unreported	reported	unreported	sub	unreported	indel	%	unreported
<b>GPR116</b>	2	0.01%		2	0		0		2	100.0%
<b>PTPN14</b>	2	0.01%		2	0		0		2	100.0%
<b>NUP98</b>	2	0.01%		2	0		0		2	100.0%
<b>FAM175A</b>	2	0.01%		2	0		1		1	100.0%
<b>DNAH9</b>	2	0.01%		2	0		0		2	100.0%
<b>CSF1R</b>	2	0.01%		2	0		0		2	100.0%
<b>SRCAP</b>	2	0.01%		2	0		1		1	100.0%
<b>ZEB2</b>	2	0.01%		2	0		0		2	100.0%
<b>ERG</b>	2	0.01%		2	0		0		2	100.0%
<b>GPR112</b>	2	0.01%		2	0		0		2	100.0%
<b>PPFIA2</b>	2	0.01%		2	0		2		0	100.0%
<b>ABL1</b>	2	0.01%		1	1		1		0	50.0%
<b>BRCA1</b>	2	0.01%		2	0		2		0	100.0%
<b>SVEP1</b>	2	0.01%		2	0		0		2	100.0%
<b>BEND6</b>	2	0.01%		2	0		0		2	100.0%
<b>FAT2</b>	2	0.01%		2	0		1		1	100.0%
<b>RRAS</b>	2	0.01%		2	0		2		0	100.0%
<b>SMC1B</b>	2	0.01%		2	0		2		0	100.0%
<b>STAG1</b>	2	0.01%		2	0		2		0	100.0%
<b>CBLB</b>	2	0.01%		2	0		2		0	100.0%
<b>ME1</b>	2	0.01%		2	0		0		2	100.0%
<b>TCF4</b>	2	0.01%		2	0		0		2	100.0%
<b>MAP2K5</b>	2	0.01%		2	0		0		2	100.0%
<b>CD101</b>	2	0.01%		2	0		0		2	100.0%
<b>SPTA1</b>	2	0.01%		2	0		0		2	100.0%
<b>STK17B</b>	2	0.01%		2	0		0		2	100.0%
<b>RINT1</b>	2	0.01%		2	0		0		2	100.0%
<b>GRIN2B</b>	2	0.01%		2	0		0		2	100.0%
<b>DIAPH2</b>	2	0.01%		2	0		2		0	100.0%
<b>NXF1</b>	2	0.01%		2	0		2		0	100.0%
<b>ABCC9</b>	2	0.01%		2	0		0		2	100.0%
<b>CACNA1G</b>	1	0.00%		1	0		0		1	100.0%
<b>ARHGAP32</b>	1	0.00%		1	0		0		1	100.0%
<b>SPRED2</b>	1	0.00%		1	0		1		0	100.0%
<b>RASGRF1</b>	1	0.00%		1	0		1		0	100.0%
<b>FANCE</b>	1	0.00%		1	0		0		1	100.0%
<b>SF3A3</b>	1	0.00%		1	0		1		0	100.0%
<b>MYB</b>	1	0.00%		1	0		0		1	100.0%
<b>DRC1</b>	1	0.00%		1	0		0		1	100.0%
<b>SPP1</b>	1	0.00%		1	0		0		1	100.0%
<b>RP1L1</b>	1	0.00%		1	0		0		1	100.0%
<b>PRDX2</b>	1	0.00%		1	0		0		1	100.0%
<b>AHNAK2</b>	1	0.00%		1	0		0		1	100.0%
<b>CRLF2</b>	1	0.00%		1	0		1		0	100.0%
<b>MGA</b>	1	0.00%		1	0		1		0	100.0%
<b>CDCA5</b>	1	0.00%		1	0		0		1	100.0%
<b>BTNL8</b>	1	0.00%		1	0		0		1	100.0%
<b>SRPK2</b>	1	0.00%		1	0		0		1	100.0%
<b>IL31RA</b>	1	0.00%		1	0		0		1	100.0%
<b>HIF1A</b>	1	0.00%		1	0		0		1	100.0%
<b>CUL2</b>	1	0.00%		1	0		0		1	100.0%
<b>GATA3</b>	1	0.00%		1	0		1		0	100.0%
<b>EGFR</b>	1	0.00%		1	0		0		1	100.0%

Continued on next page

gene	count	freq	unreported	reported	unreported	sub	unreported	indel	%	unreported
<b>DDX18</b>	1	0.00%		1	0		0		1	100.0%
<b>OCA2</b>	1	0.00%		1	0		0		1	100.0%
<b>ATXN7L1</b>	1	0.00%		1	0		0		1	100.0%
<b>PHF12</b>	1	0.00%		1	0		0		1	100.0%
<b>NR3C1</b>	1	0.00%		1	0		0		1	100.0%
<b>SLC25A38</b>	1	0.00%		1	0		0		1	100.0%
<b>FANCI</b>	1	0.00%		1	0		1		0	100.0%
<b>HNRNPCL1</b>	1	0.00%		1	0		0		1	100.0%
<b>GRIN2A</b>	1	0.00%		1	0		0		1	100.0%
<b>SPAG9</b>	1	0.00%		1	0		0		1	100.0%
<b>PTPN21</b>	1	0.00%		1	0		0		1	100.0%
<b>FAT1</b>	1	0.00%		1	0		1		0	100.0%
<b>S100B</b>	1	0.00%		1	0		1		0	100.0%
<b>KSR2</b>	1	0.00%		1	0		0		1	100.0%
<b>PTPN2</b>	1	0.00%		1	0		0		1	100.0%
<b>OTOP1</b>	1	0.00%		1	0		0		1	100.0%
<b>KIAA1377</b>	1	0.00%		1	0		0		1	100.0%
<b>RPS6KA6</b>	1	0.00%		1	0		1		0	100.0%
<b>PTPRC</b>	1	0.00%		1	0		0		1	100.0%
<b>ZNF717</b>	1	0.00%		1	0		0		1	100.0%
<b>IL12B</b>	1	0.00%		1	0		0		1	100.0%
<b>EPHA2</b>	1	0.00%		1	0		0		1	100.0%
<b>NAAA</b>	1	0.00%		1	0		0		1	100.0%
<b>PIK3CA</b>	1	0.00%		0	1		0		0	0.0%
<b>WAPAL</b>	1	0.00%		1	0		1		0	100.0%
<b>IKZF1</b>	1	0.00%		1	0		1		0	100.0%
<b>NFKB2</b>	1	0.00%		1	0		0		1	100.0%
<b>SLC12A1</b>	1	0.00%		1	0		0		1	100.0%
<b>PTCH1</b>	1	0.00%		1	0		1		0	100.0%
<b>ESCO2</b>	1	0.00%		1	0		0		1	100.0%
<b>MSH6</b>	1	0.00%		1	0		0		1	100.0%
<b>CNTN5</b>	1	0.00%		1	0		0		1	100.0%
<b>DDX41</b>	1	0.00%		1	0		0		1	100.0%
<b>KCNMA1</b>	1	0.00%		1	0		0		1	100.0%
<b>IFNGR1</b>	1	0.00%		1	0		0		1	100.0%
<b>HLA-C</b>	1	0.00%		1	0		1		0	100.0%
<b>DTNA</b>	1	0.00%		1	0		0		1	100.0%
<b>TRPM4</b>	1	0.00%		1	0		0		1	100.0%
<b>FANCG</b>	1	0.00%		1	0		1		0	100.0%
<b>CDH1</b>	1	0.00%		1	0		0		1	100.0%
<b>FAM47C</b>	1	0.00%		1	0		0		1	100.0%
<b>CSNK1A1</b>	1	0.00%		1	0		1		0	100.0%
<b>RAD51</b>	1	0.00%		1	0		0		1	100.0%
<b>COLCA1</b>	1	0.00%		1	0		0		1	100.0%
<b>CTNNA2</b>	1	0.00%		1	0		0		1	100.0%
<b>SH3RF3</b>	1	0.00%		1	0		0		1	100.0%
<b>CSMD3</b>	1	0.00%		1	0		0		1	100.0%
<b>IL6ST</b>	1	0.00%		1	0		0		1	100.0%
<b>ALK</b>	1	0.00%		1	0		1		0	100.0%
<b>PTPRF</b>	1	0.00%		1	0		0		1	100.0%