



# TP53 Allelic State in MSK-Impact Cohort

07/10/2020

# Outline

---

1. [Motivation](#)
2. [Workflow](#)
3. [General Distributions](#)
4. [Genome Instability](#)
5. [Cancer Study](#)
  - a. [Breast Cancer](#)
  - b. [Non-Small Cell Lung Cancer](#)
  - c. [Glioma](#)
6. [Conclusion & Next Steps](#)
7. [Appendix](#)

---

# Motivation

---

# TP53 Allelic State in MDS → Pan Cancer Analysis

---

## TP53 Allelic State in MDS

- TP53 Allelic State has important consequences on clinical outcomes
  - Higher risk of death for bi-allelic hits compared to mono-allelic
  - Bi-allelic hits is a driver of disease progression (MDS → AML)

# TP53 Allelic State in MDS → Pan Cancer Analysis

---

## TP53 Allelic State in MDS

- TP53 Allelic State has important consequences on clinical outcomes
    - Higher risk of death for bi-allelic hits compared to mono-allelic
    - Bi-allelic hits is a driver of disease progression (MDS → AML)
- 

## Pan Cancer Analysis: MSK-Impact Cohort

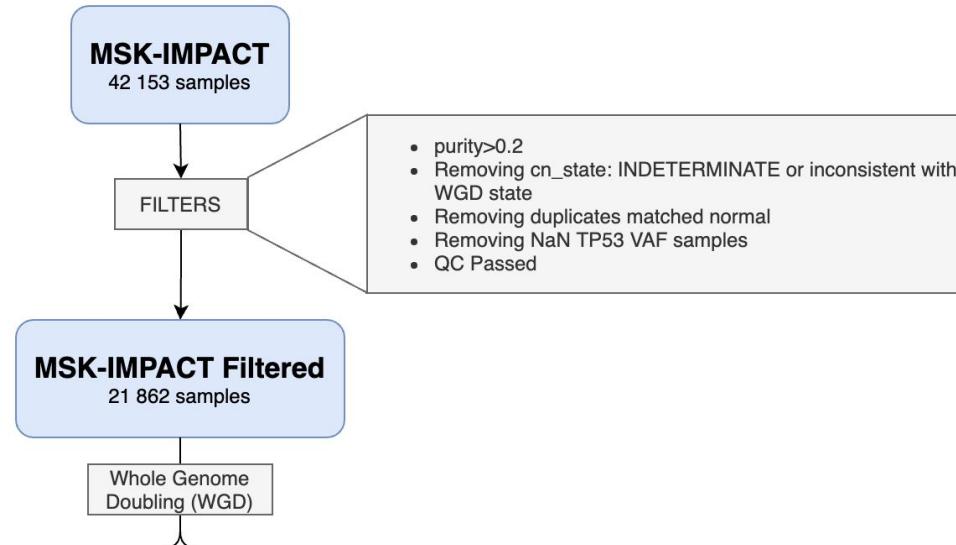
- What are the representations of TP53 allelic states across cancer types?
- Is there significant clinical / phenotypic differences between mono-allelic and bi-allelic states ?
  - What are the cancers where the differences are most significant?
  - The one where there are no differences?

---

# Workflow

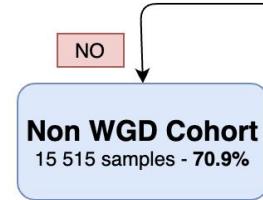
---

# Filtering the MSK-Impact Dataset

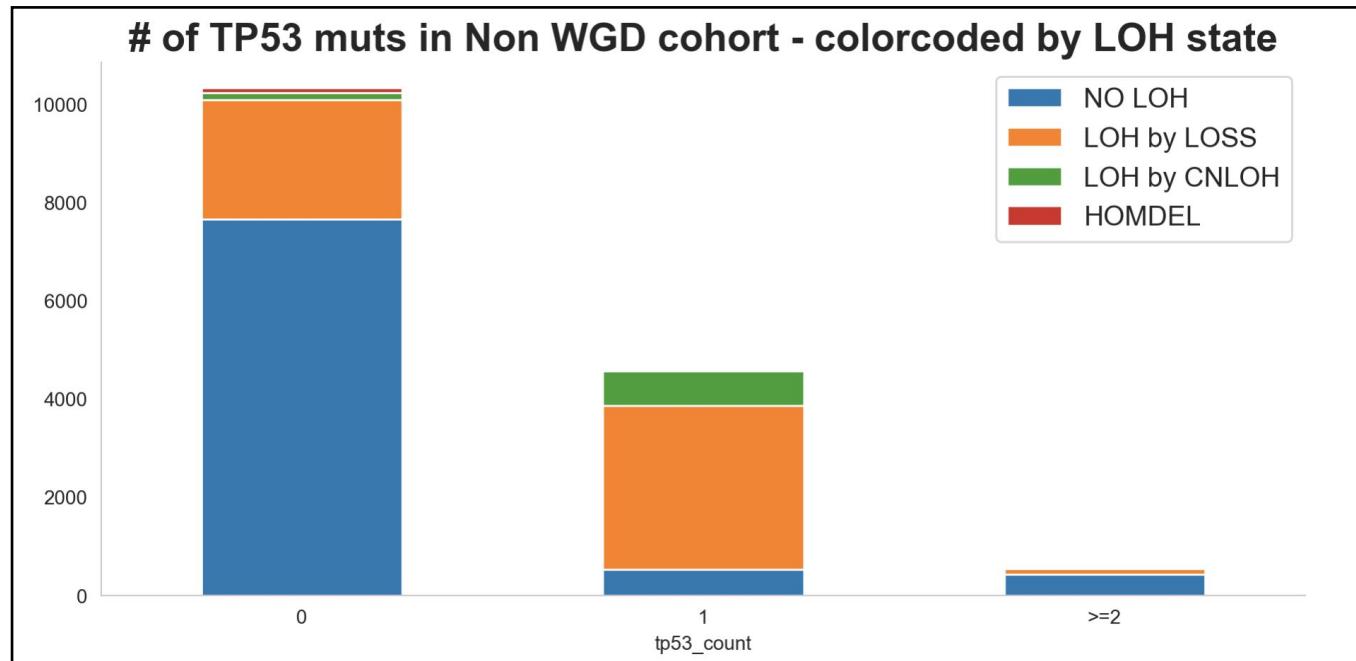


# Non-WGD Cohort

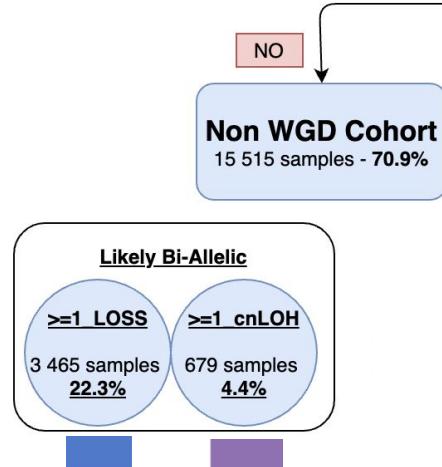
---



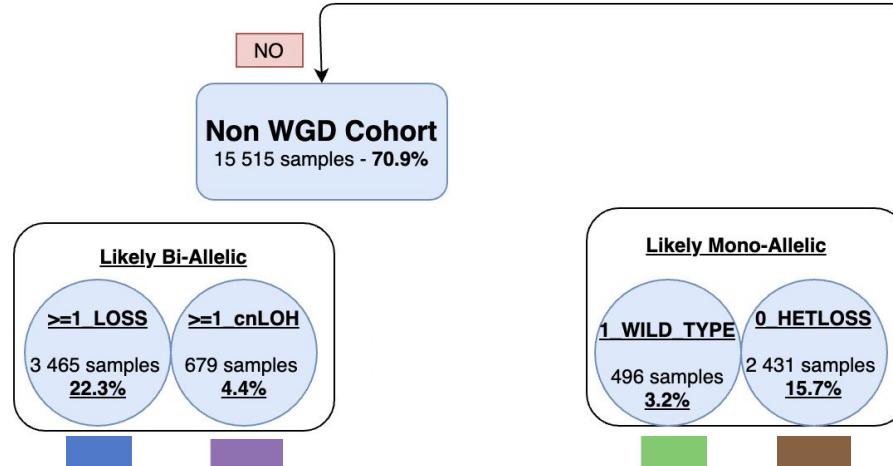
# LOH Status - Non WGD



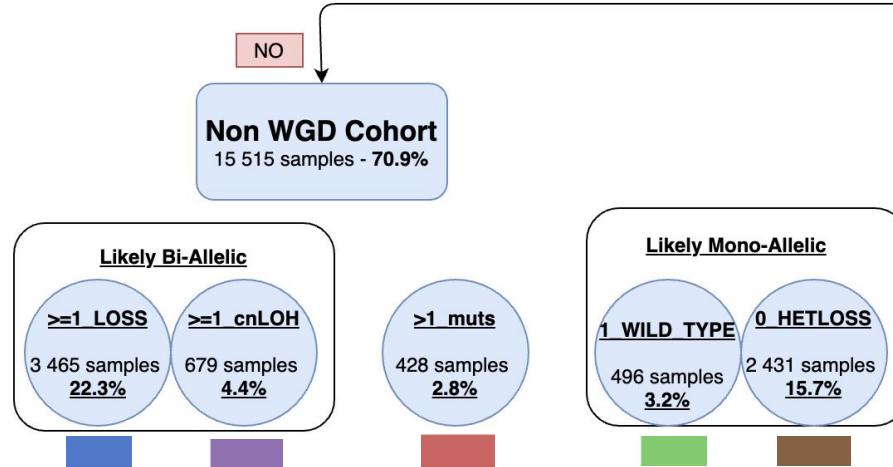
# Non-WGD Cohort



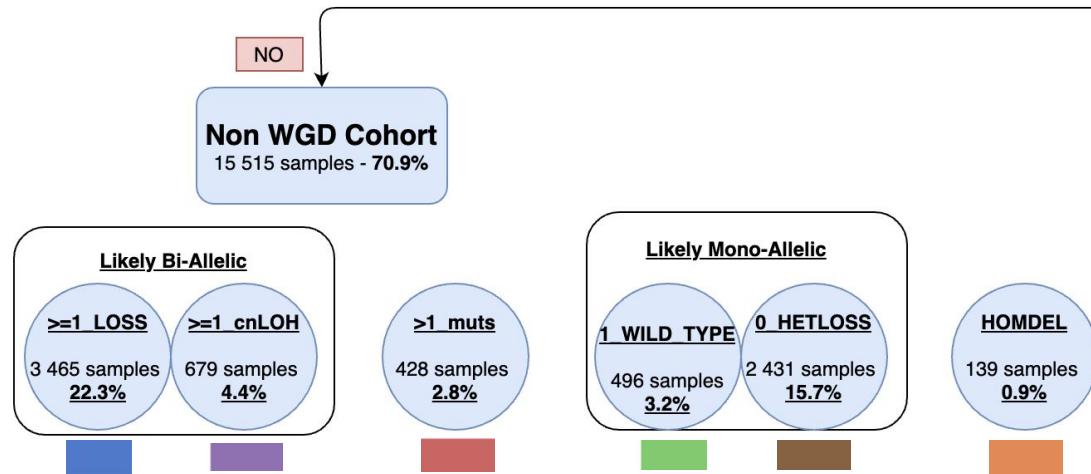
# Non-WGD Cohort



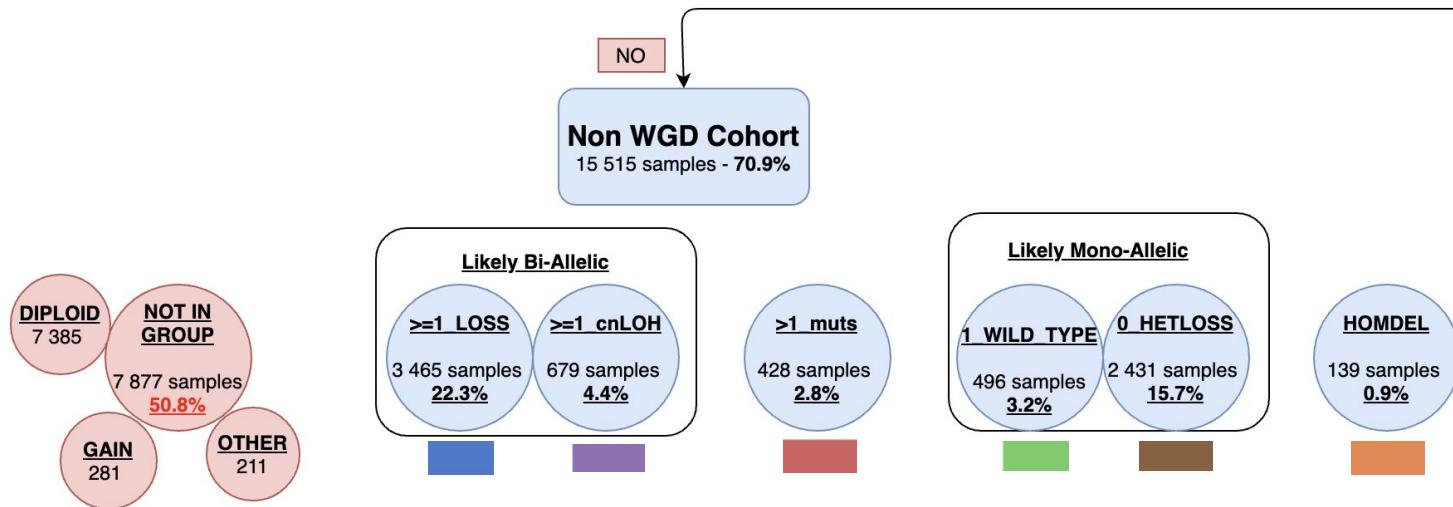
# Non-WGD Cohort



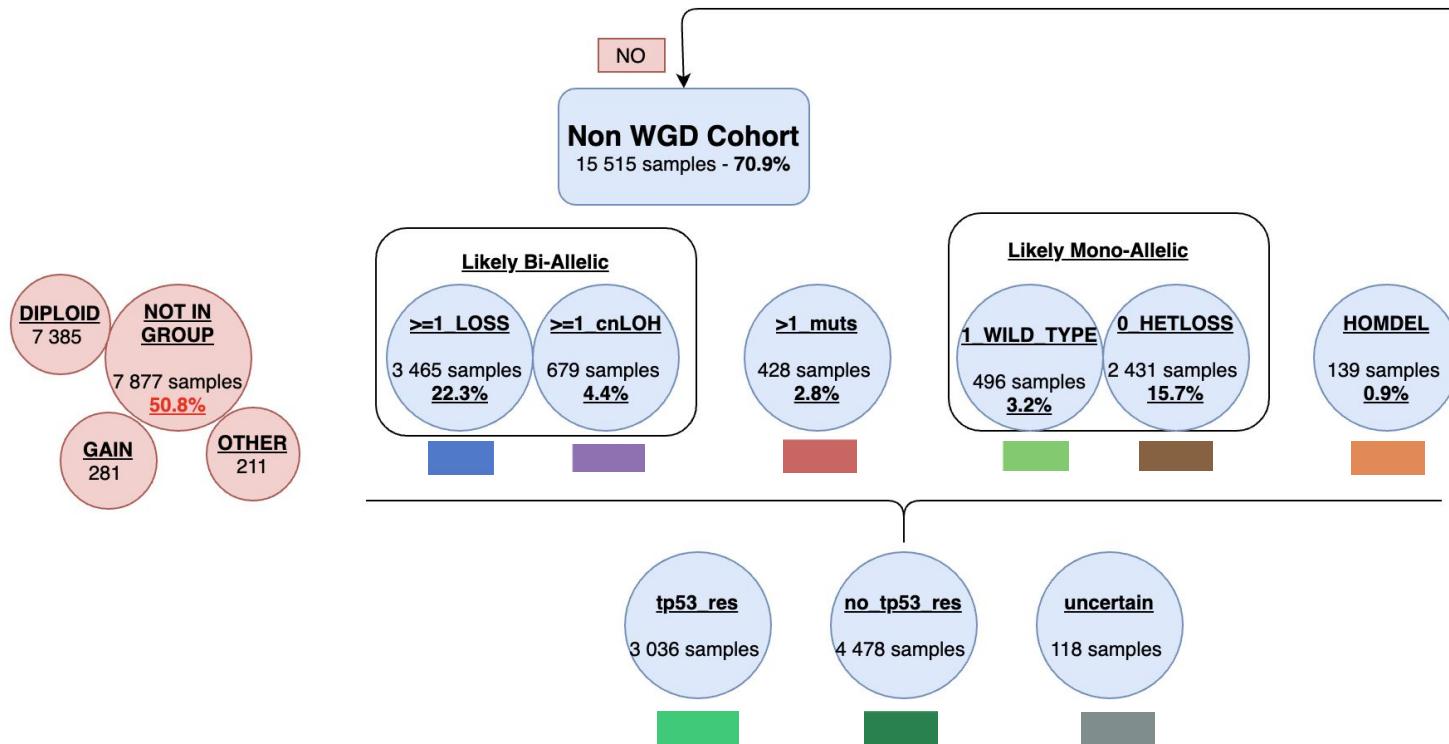
# Non-WGD Cohort



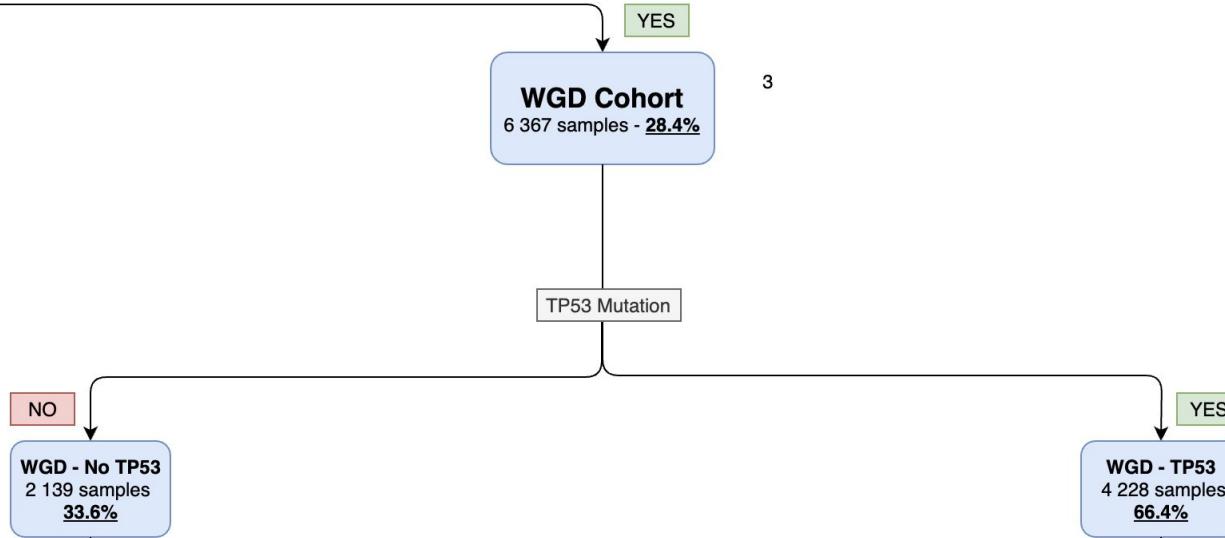
# Non-WGD Cohort



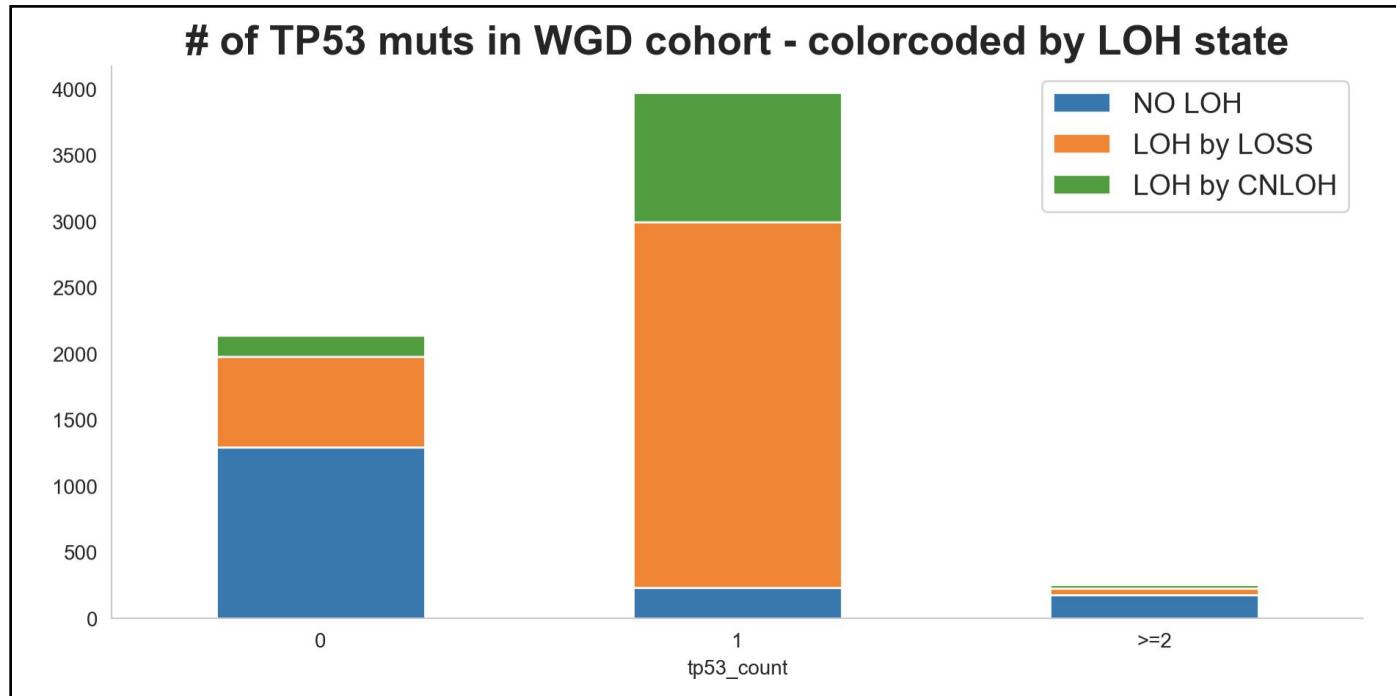
# Non-WGD Cohort



# WGD Cohort



# LOH Status - WGD



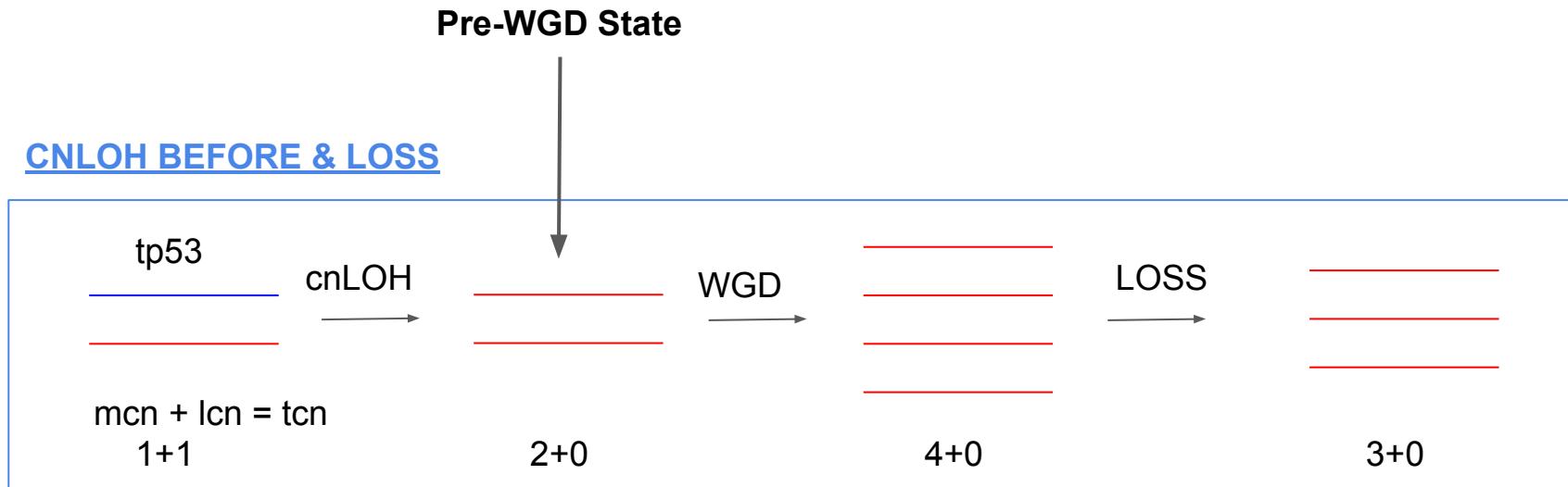
# TP53 State Study: Allelic Representation

---

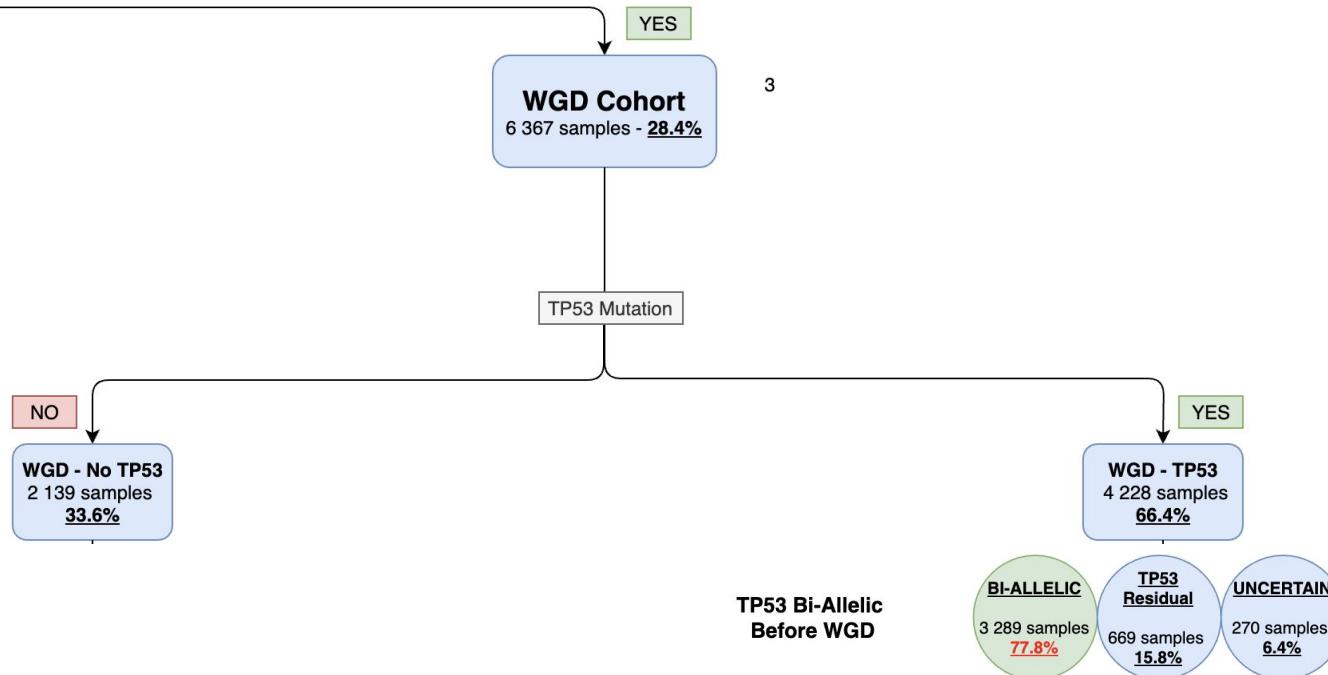
- Question:  
Is the association between TP53 and WGD driven by TP53 Allelic State ?

# TP53 State Study: Allelic Representation

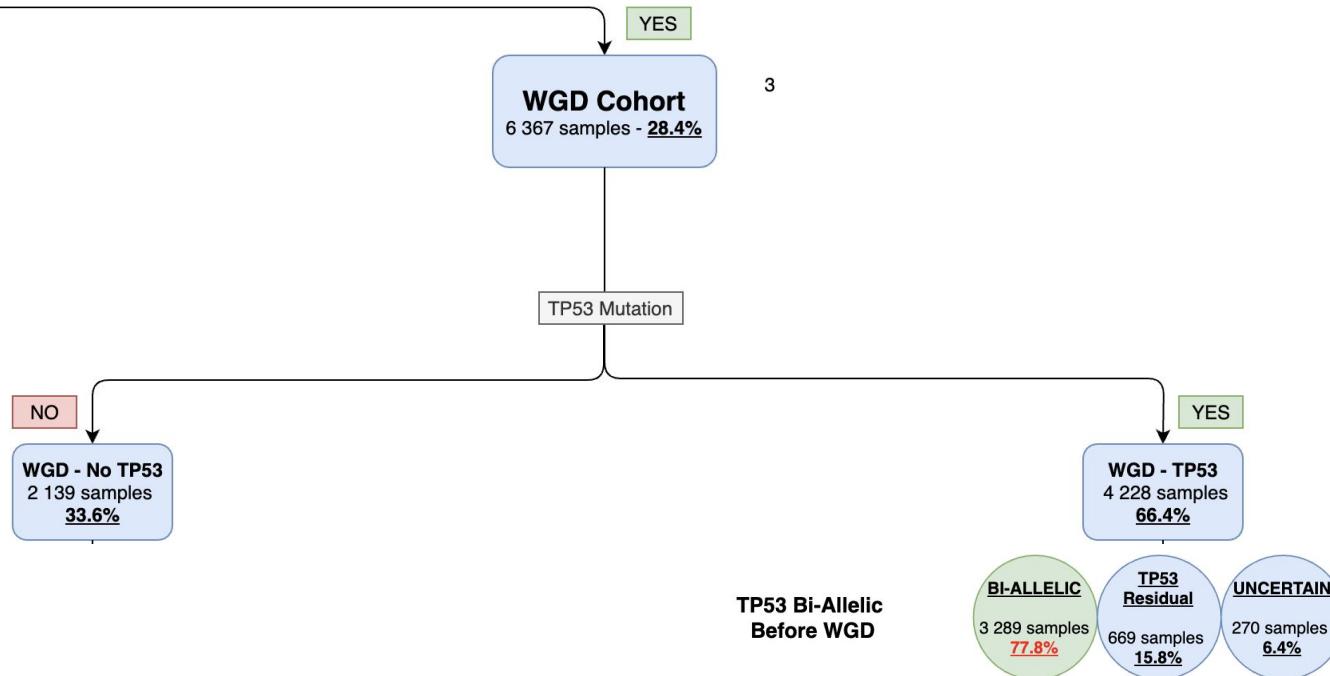
- Question:  
Is the association between TP53 and WGD driven by TP53 Allelic State ?



# WGD Cohort - Pre WGD Allelic State



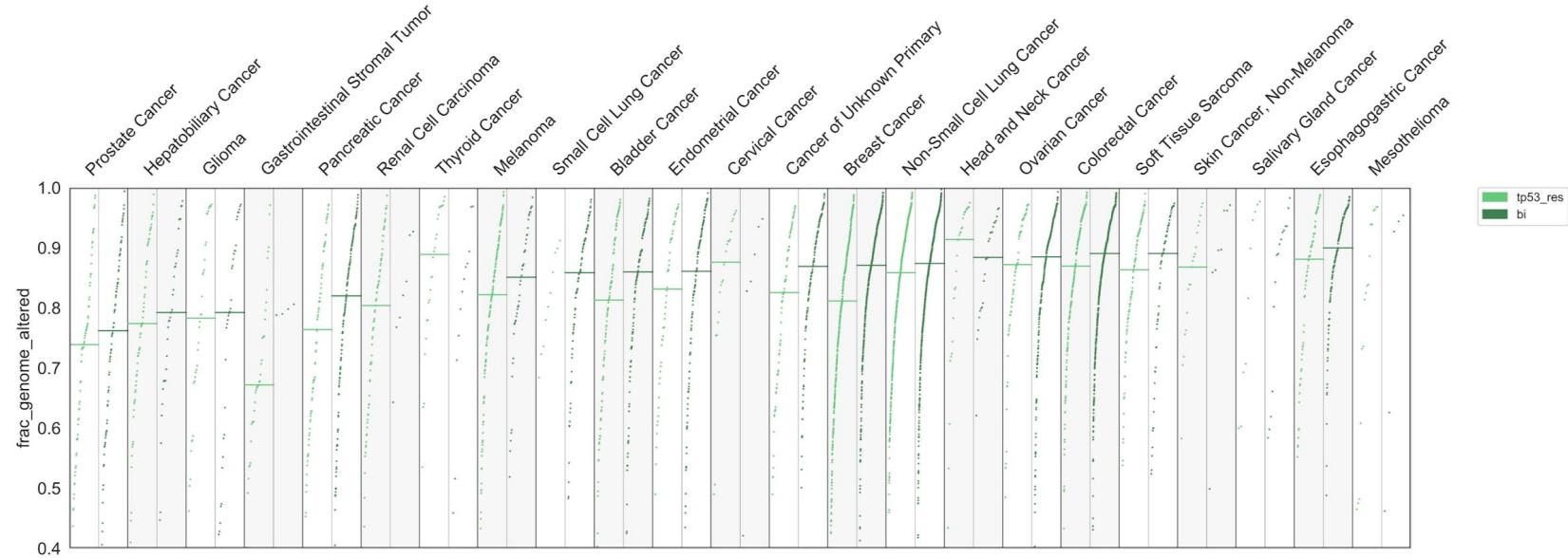
# WGD Cohort - Pre WGD Allelic State



- The known association between **TP53 mutant** and **WGD** (Bielski et al) is driven by **bi-allelic TP53** (at least 77.8%)

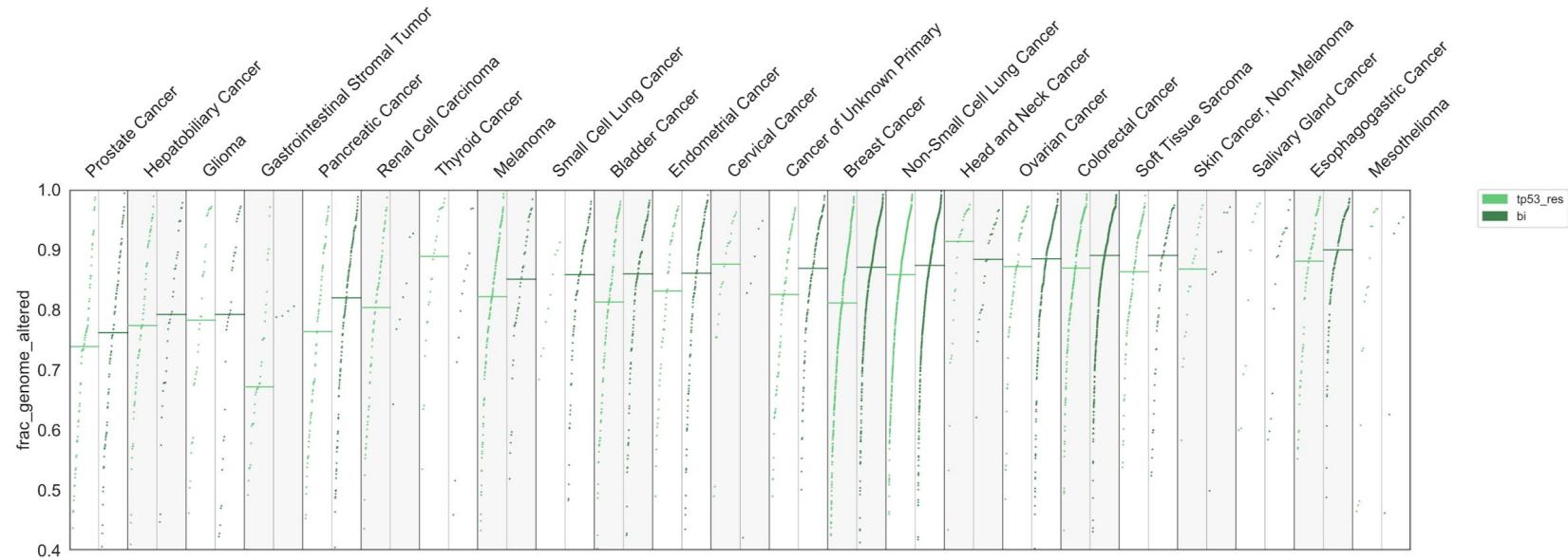
# WGD Cohort - Fraction of Genome Altered

## Fraction of Genome Altered - Pre WGD TP53 BI / TP53 RES - WGD



# WGD Cohort - Fraction of Genome Altered

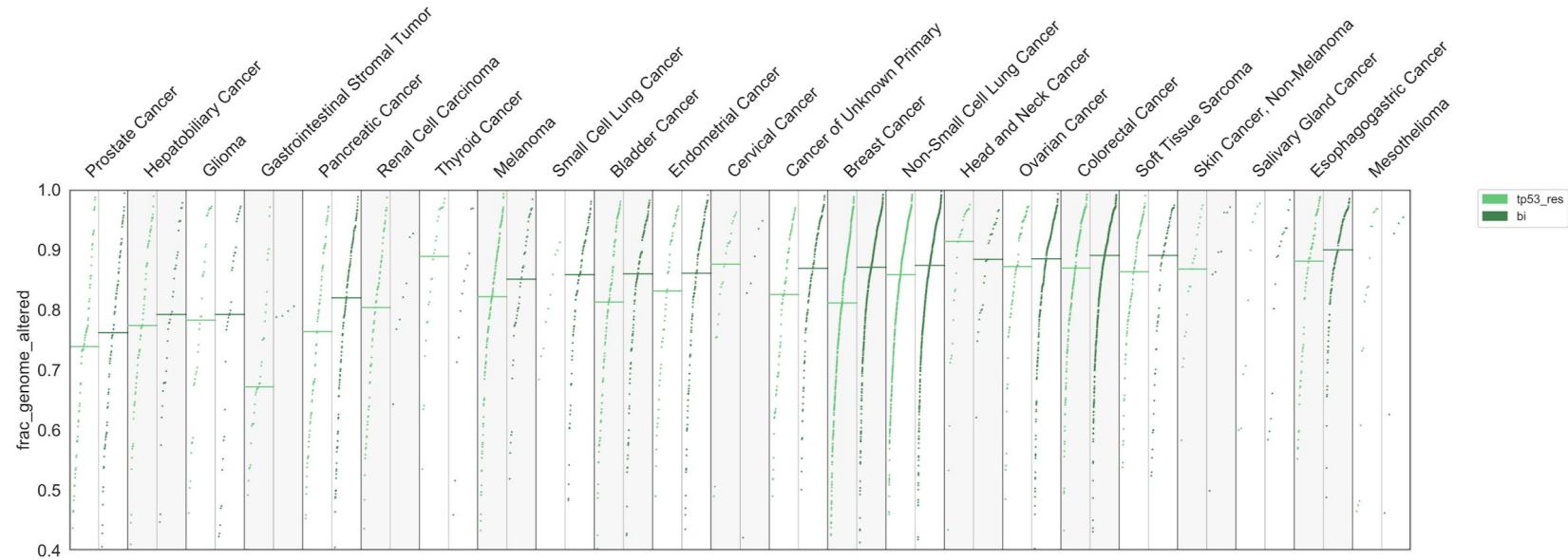
## Fraction of Genome Altered - Pre WGD TP53 BI / TP53 RES - WGD



- Minimal Fraction of Genome Altered Median: 67%
  - Difficulties to differentiate Mono-Allelic from Bi-Allelic Samples

# WGD Cohort - Fraction of Genome Altered

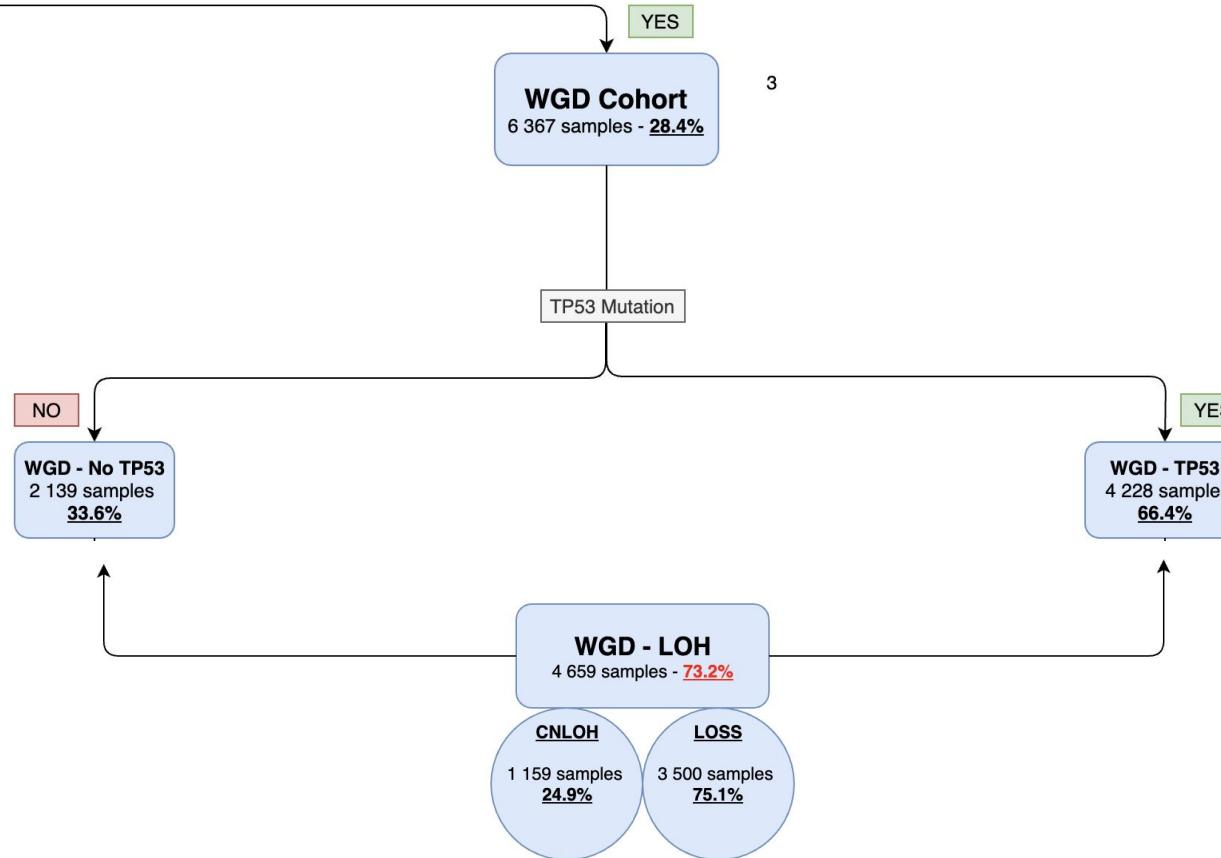
Fraction of Genome Altered - Pre WGD TP53 BI / TP53 RES - WGD



- Minimal Fraction of Genome Altered Median: 67%
  - Difficulties to differentiate Mono-Allelic from Bi-Allelic Samples

→ We did not push the Genome Instability Analysis further for the WGD Cohort

# WGD Cohort - TP53 LOH Enrichment

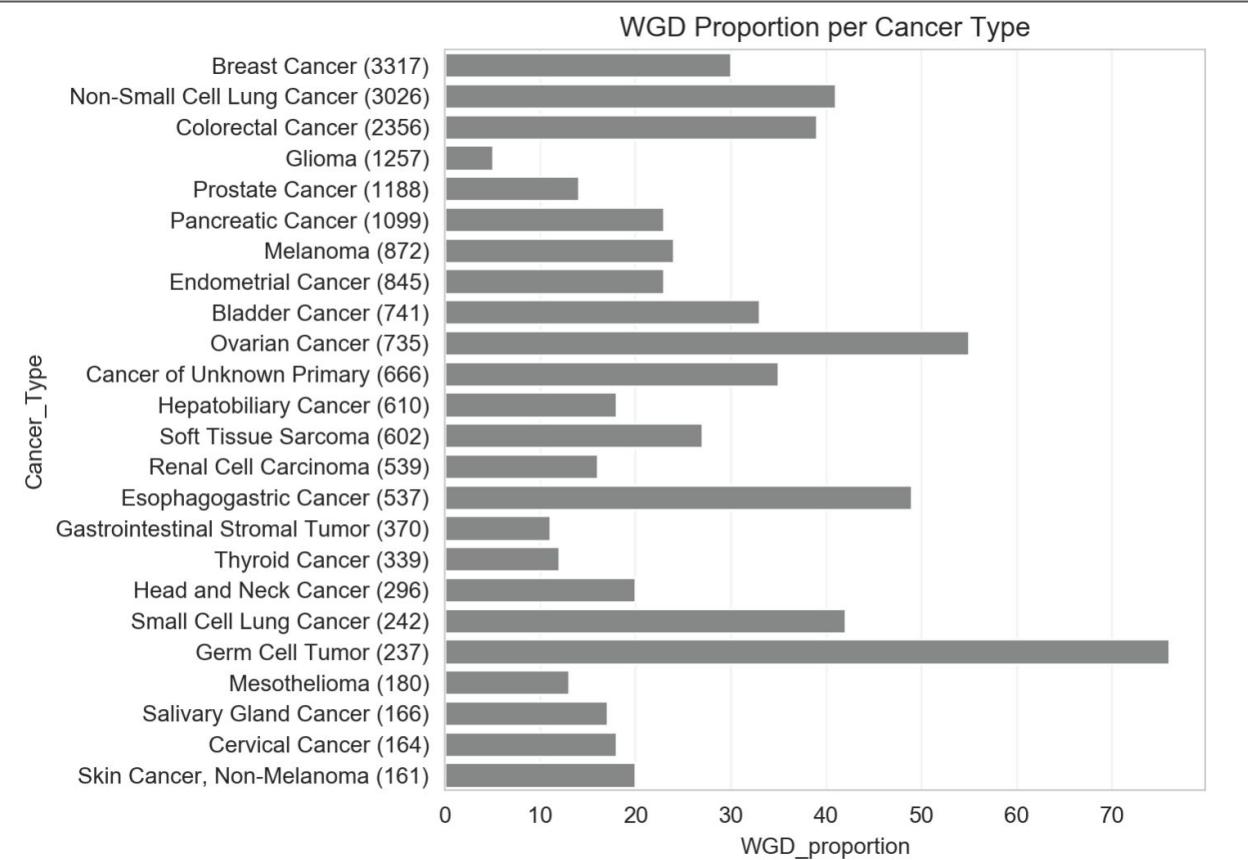


---

# General Distributions

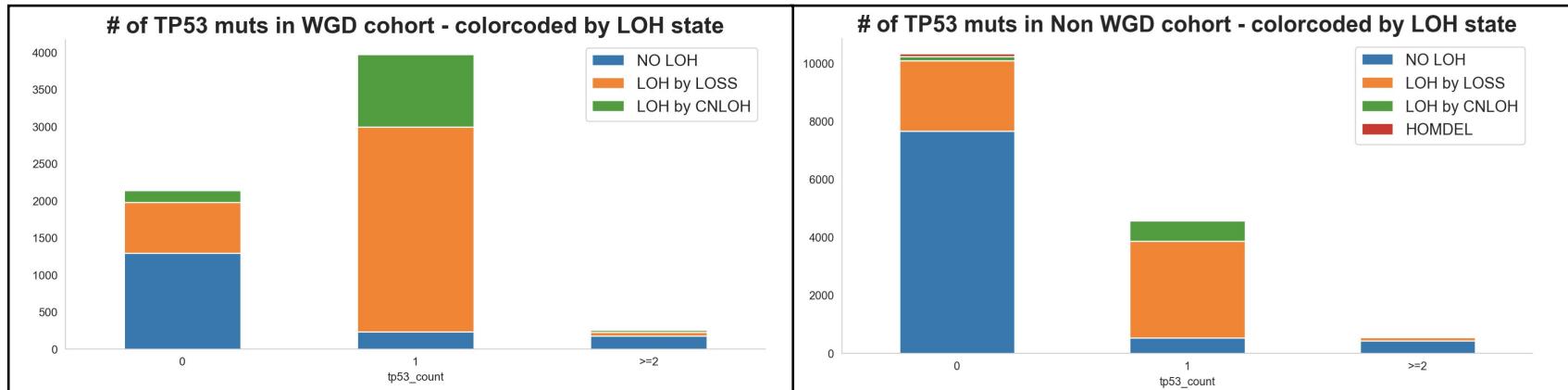
---

# WGD Proportion

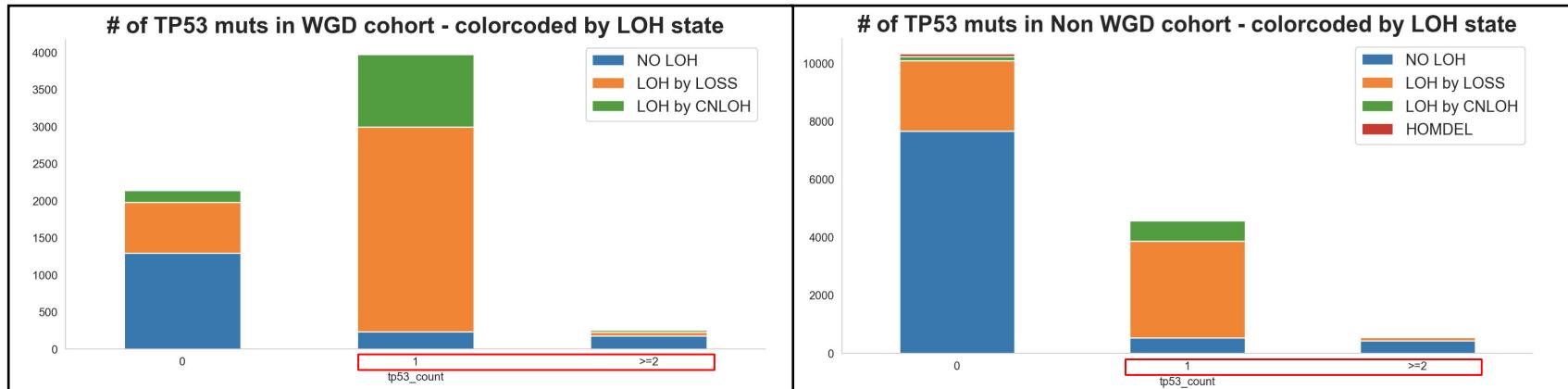


- Largest Cancers:
  - Breast
  - NSCLC
  - Colorectal
- WGD Outliers:
  - Glioma
  - Germ Cell Tumor

# LOH Status - Comparison

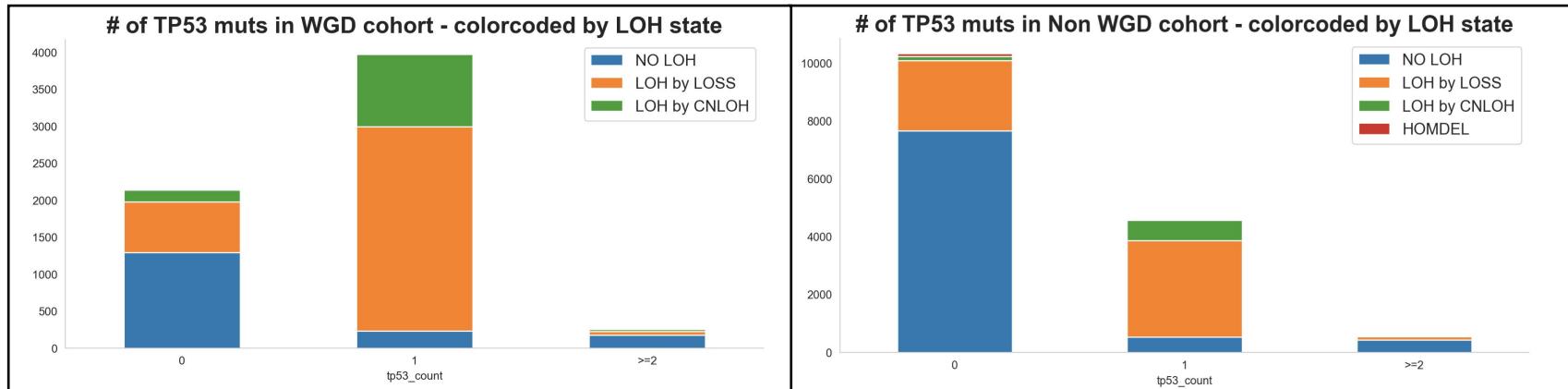


# LOH Status - Comparison



- TP53 mutations are enriched in WGD cohort vs. non-WGD (**66.4% vs. 33.4%**)

# LOH Status - Comparison



- TP53 mutations are enriched in WGD cohort vs. non-WGD (**66.4% vs. 33.4%**)
- Within TP53 hit samples more cnLOH in WGD vs. non-WGD (**24.4% vs 13.1%**)

# WGD - TP53 LOH/Mut Proportion



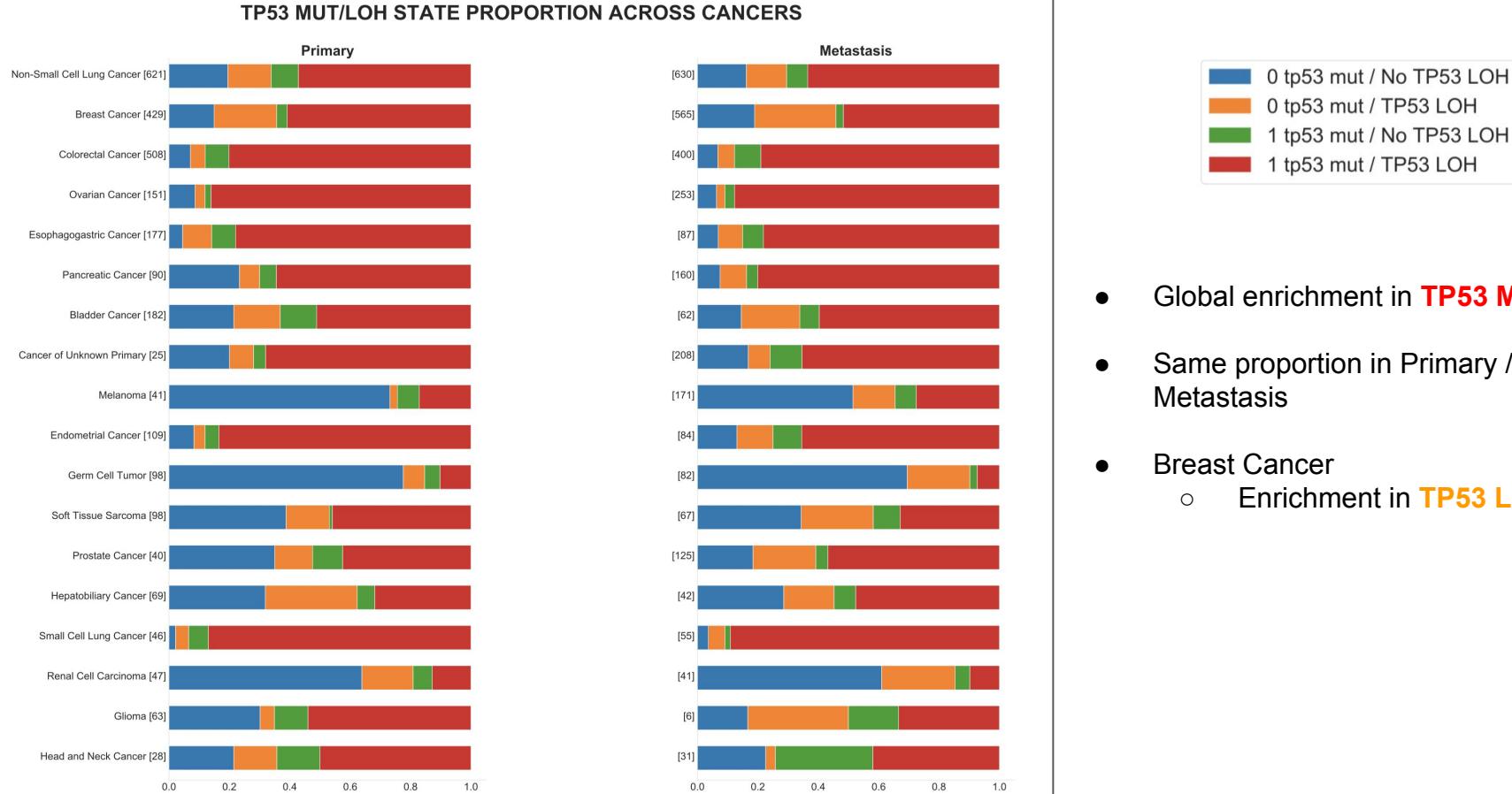
- Global enrichment in **TP53 Mut+LOH**

# WGD - TP53 LOH/Mut Proportion

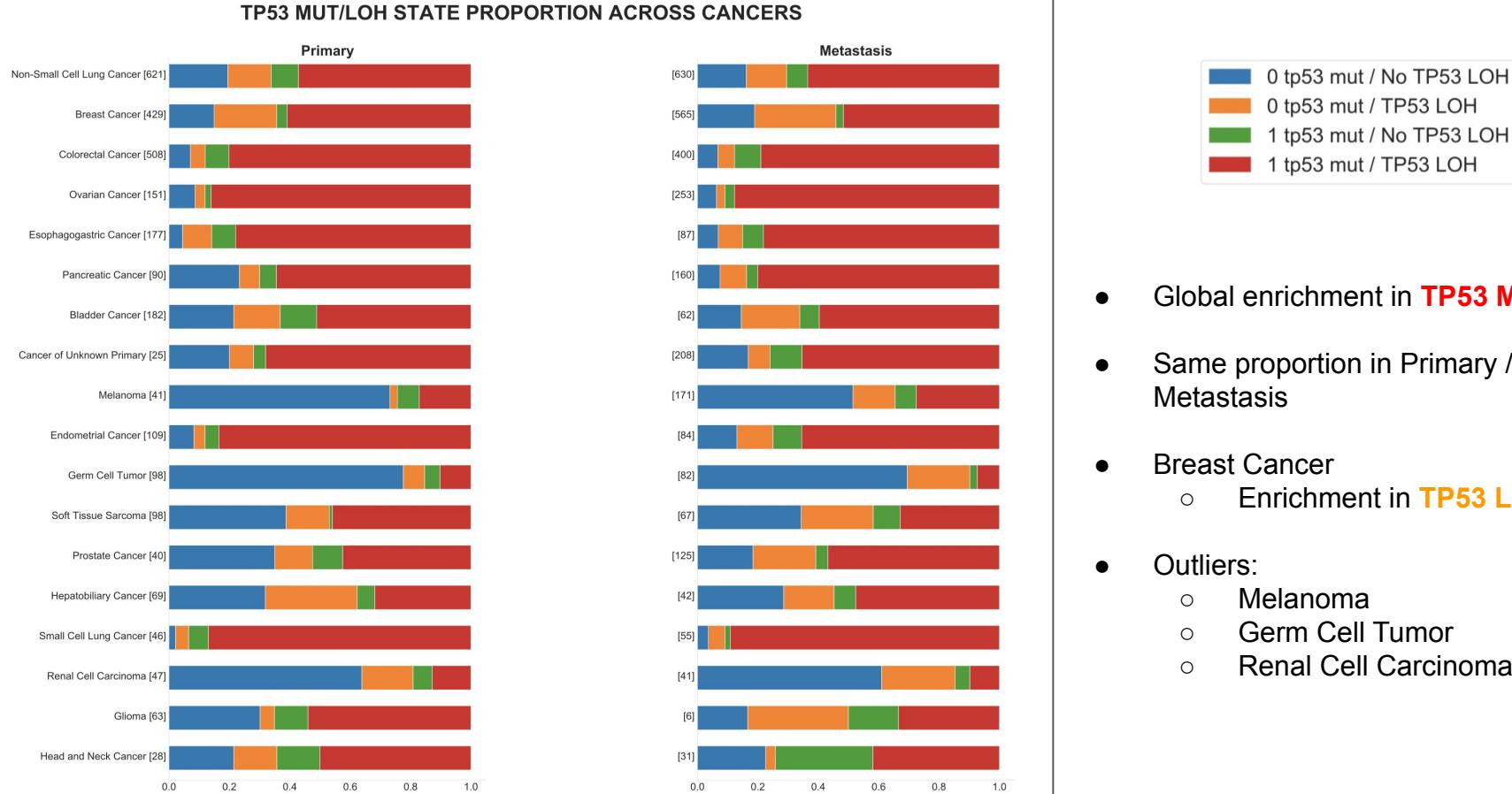


- Global enrichment in **TP53 Mut+LOH**
- Same proportion in Primary / Metastasis

# WGD - TP53 LOH/Mut Proportion



# WGD - TP53 LOH/Mut Proportion



- Global enrichment in **TP53 Mut+LOH**
- Same proportion in Primary / Metastasis
- Breast Cancer
  - Enrichment in **TP53 LOH**
- Outliers:
  - Melanoma
  - Germ Cell Tumor
  - Renal Cell Carcinoma

# Non WGD - TP53 Residual Proportion

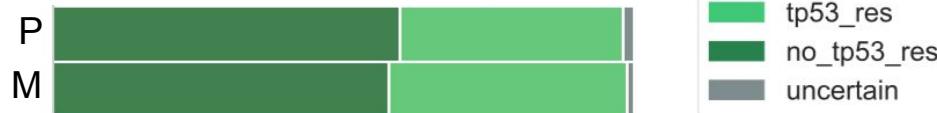
---

- In all proportion plot that are following, we only consider TP53 hit samples (Mutation or CN event)

# Non WGD - TP53 Residual Proportion

- In all proportion plot that are following, we only consider TP53 hit samples (Mutation or CN event)

Overall:

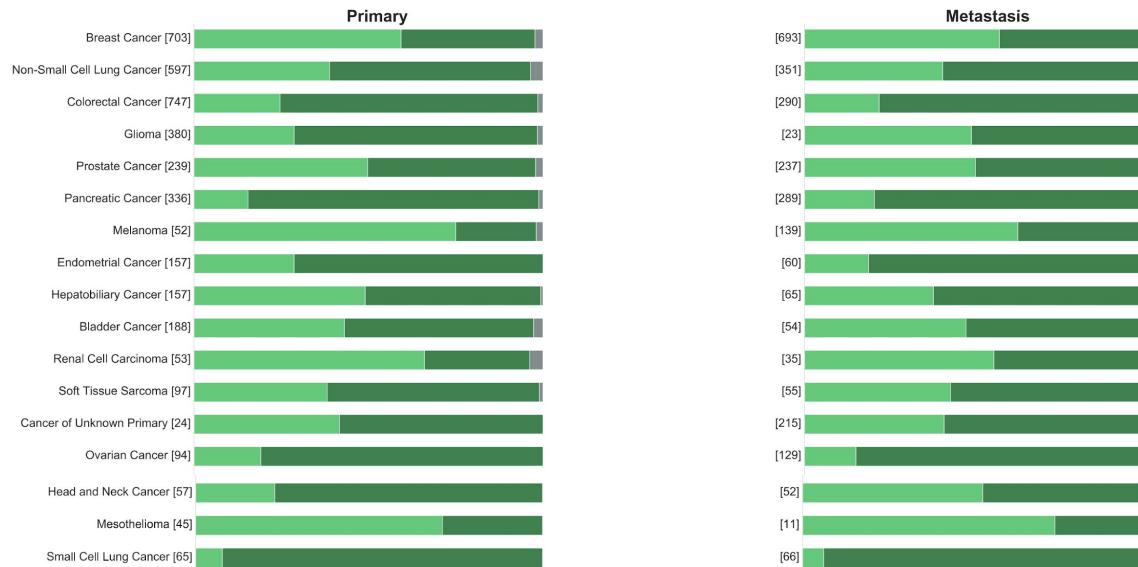


# Non WGD - TP53 Residual Proportion

Overall:



Per Cancer:

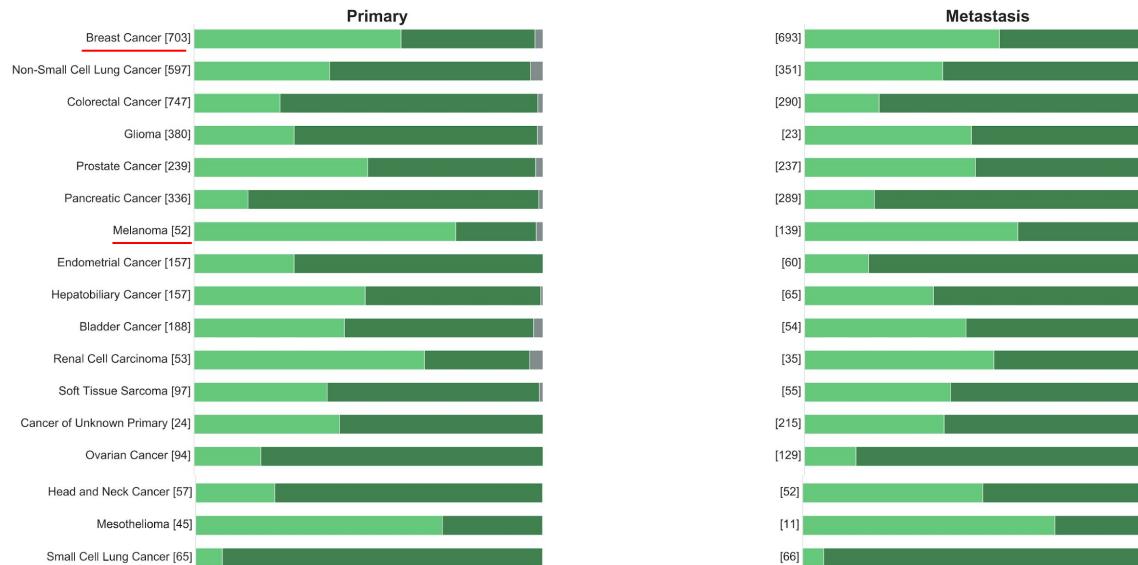


# Non WGD - TP53 Residual Proportion

Overall:



Per Cancer:



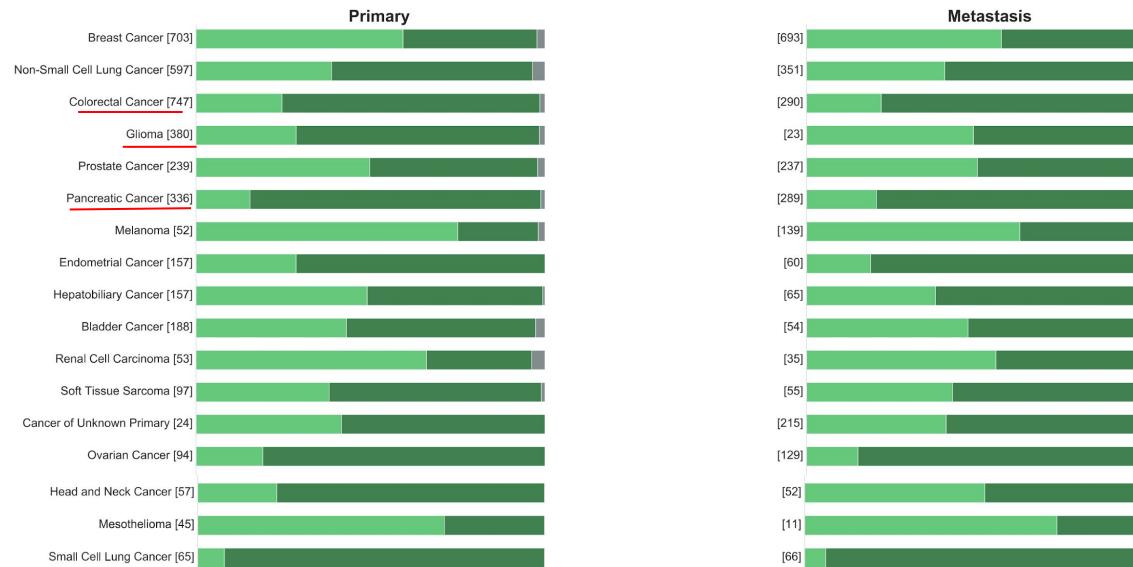
- ‘Mono-Allelic’ Cancers:
  - Breast Cancer
  - Melanoma

# Non WGD - TP53 Residual Proportion

Overall:



Per Cancer:



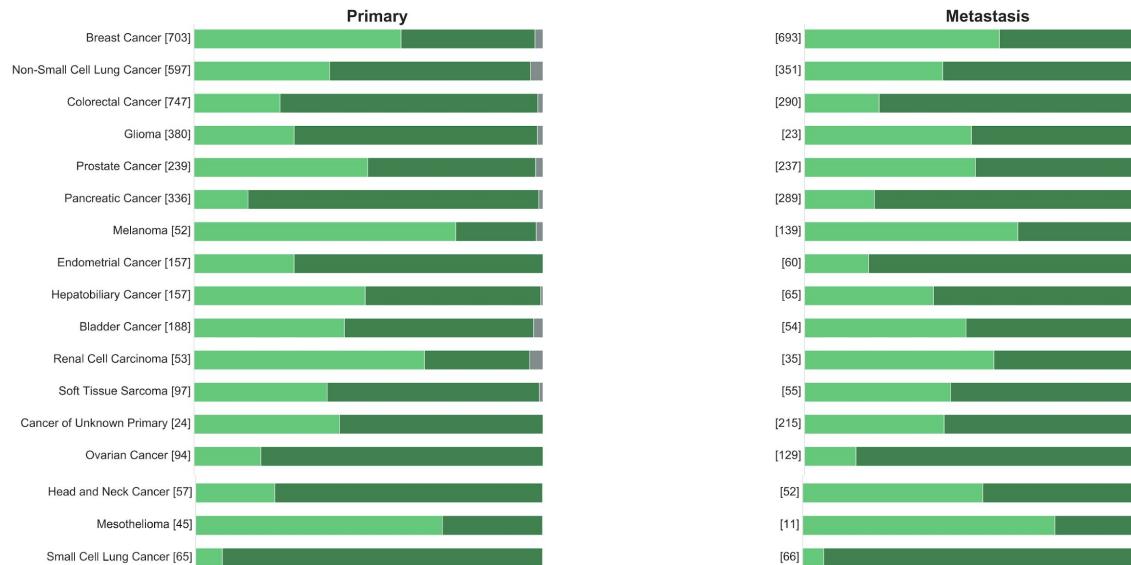
- ‘Mono-Allelic’ Cancers:
  - Breast Cancer
  - Melanoma
- ‘Bi-Allelic’ Cancers:
  - Colorectal
  - Glioma
  - Pancreatic

# Non WGD - TP53 Residual Proportion

Overall:



Per Cancer:



- ‘Mono-Allelic’ Cancers:
  - Breast Cancer
  - Melanoma
- ‘Bi-Allelic’ Cancers:
  - Colorectal
  - Glioma
  - Pancreatic
- Same proportions in Primary/Metastasis

# Non WGD - TP53 Subgroup Proportion

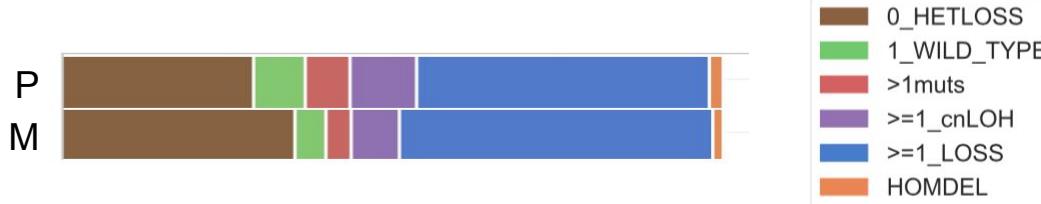
- In all proportion plot that are following, we only consider TP53 hit samples (Mutation or CN event)

Overall:

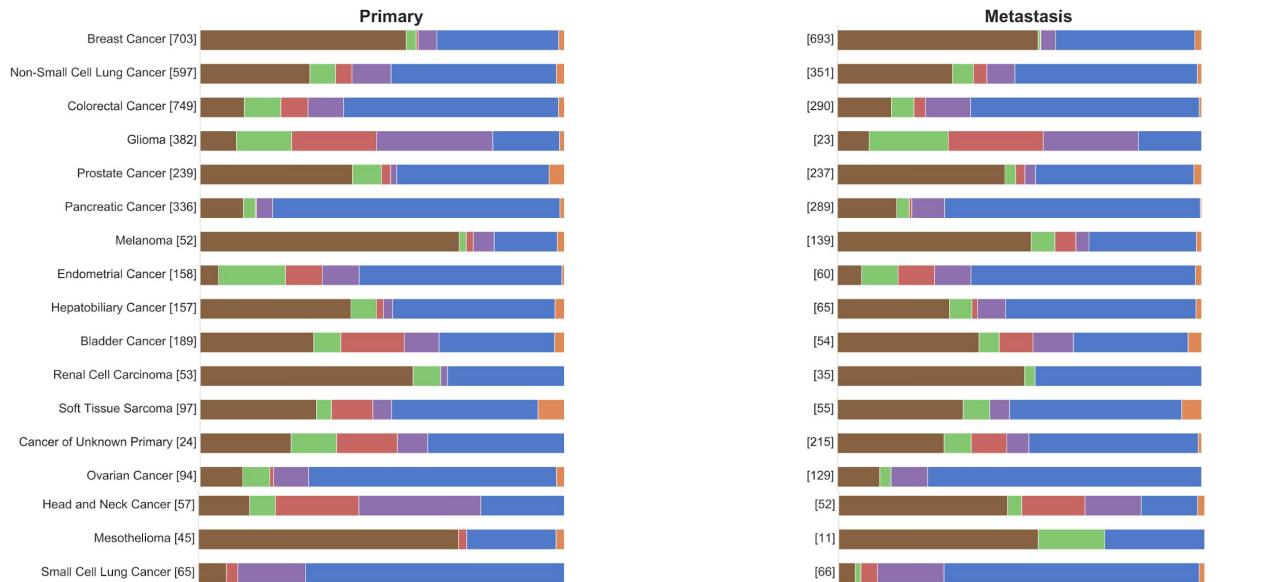


# Non WGD - TP53 Subgroup Proportion

Overall:



Per Cancer:

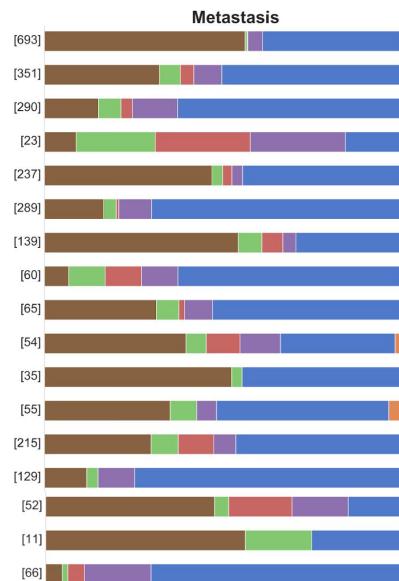
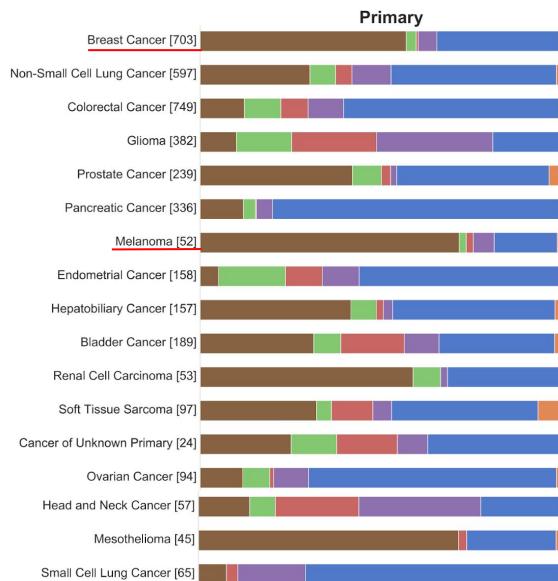


# Non WGD - TP53 Subgroup Proportion

Overall:



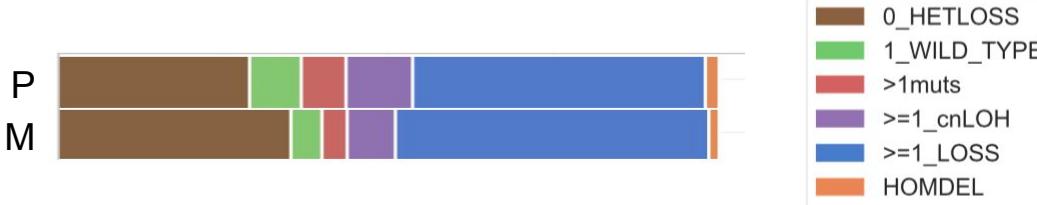
Per Cancer:



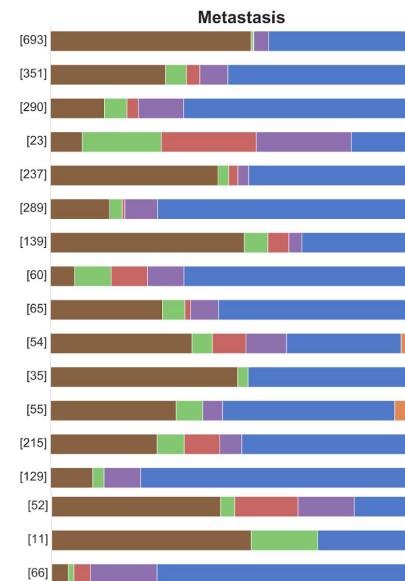
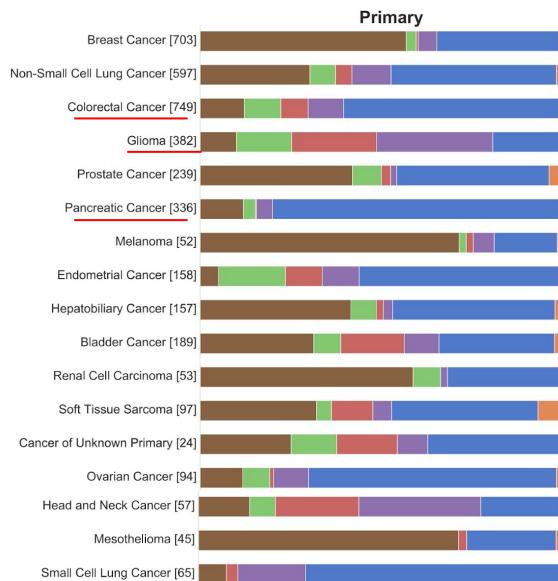
- 'Mono-Allelic' Cancers:
  - Breast Cancer
  - Melanoma

# Non WGD - TP53 Subgroup Proportion

Overall:



Per Cancer:



- ‘Mono-Allelic’ Cancers:
  - Breast Cancer
  - Melanoma
- ‘Bi-Allelic’ Cancers:
  - Colorectal
  - Glioma
  - Pancreatic

---

# Genome Instability

---

# Non WGD Cohort - Fraction of Genome Altered

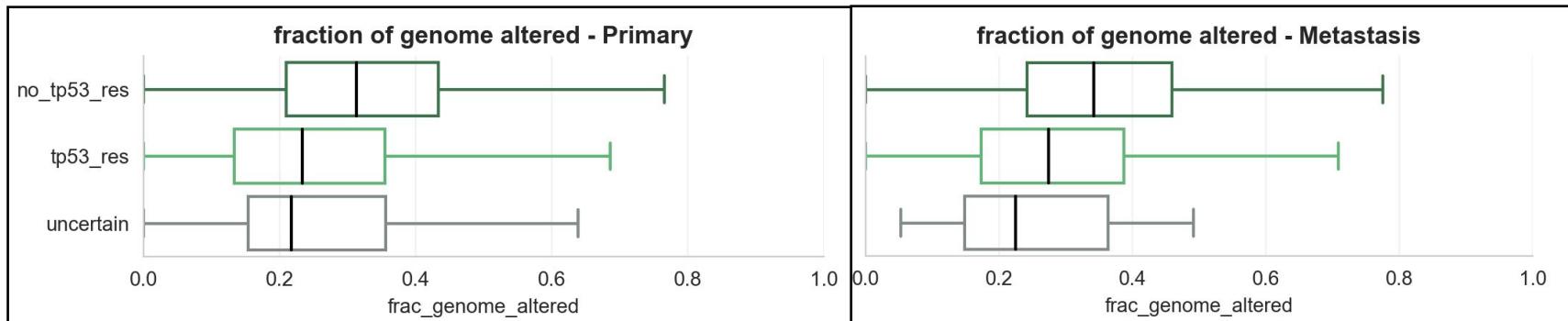
---

**Fraction of Genome Altered:** Proportion of the genome with a Copy Number State different from DIPLOID.

# Non WGD Cohort - Fraction of Genome Altered

**Fraction of Genome Altered:** Proportion of the genome with a Copy Number State different from DIPLOID.

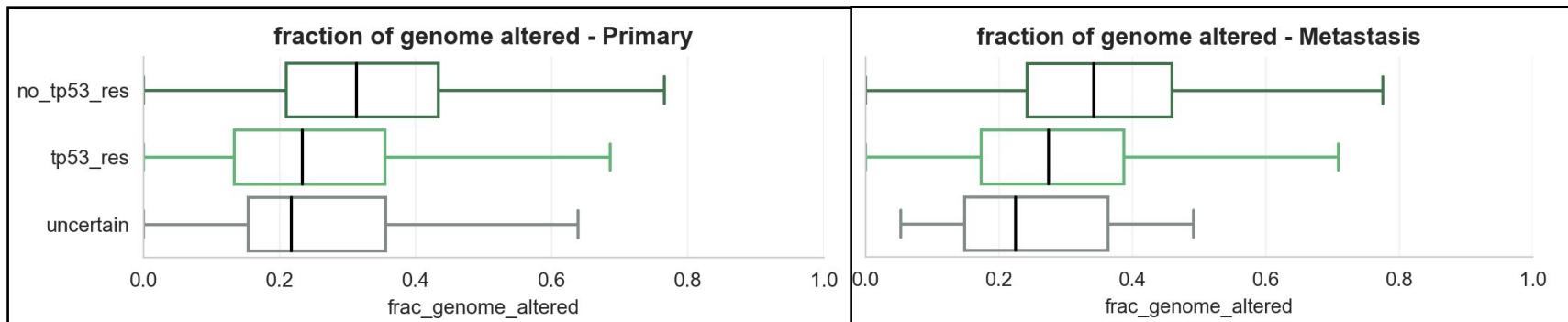
## TP53 Residual Subgroups (1/2)



# Non WGD Cohort - Fraction of Genome Altered

**Fraction of Genome Altered:** Proportion of the genome with a Copy Number State different from DIPLOID.

## TP53 Residual Subgroups (1/2)

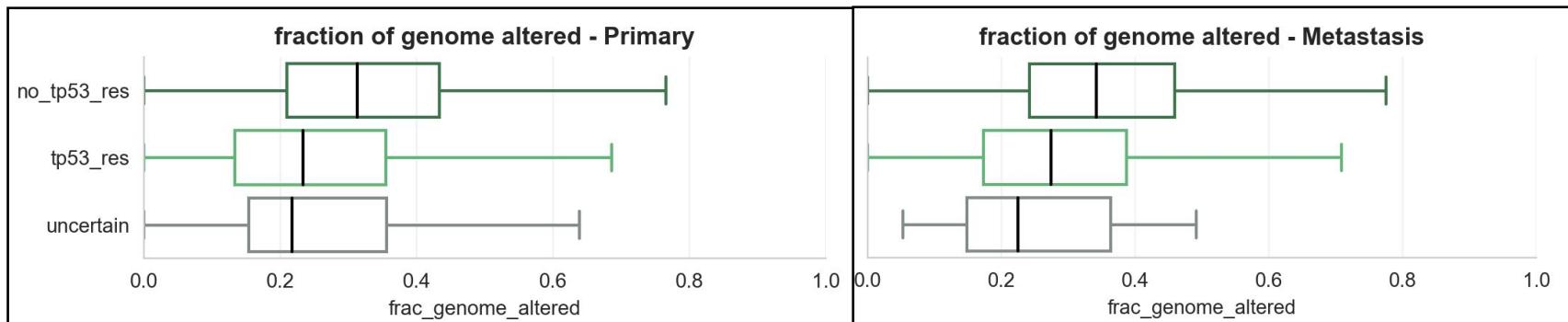


- In both Metastatic and Primary cohorts, Bi-allelic samples have higher Genome Instability than Mono-allelic samples
  - **Primary:** median 23.4% vs. 31.3%
  - **Metastasis:** median 27.4% vs. 34.2%

# Non WGD Cohort - Fraction of Genome Altered

**Fraction of Genome Altered:** Proportion of the genome with a Copy Number State different from DIPLOID.

## TP53 Residual Subgroups (1/2)

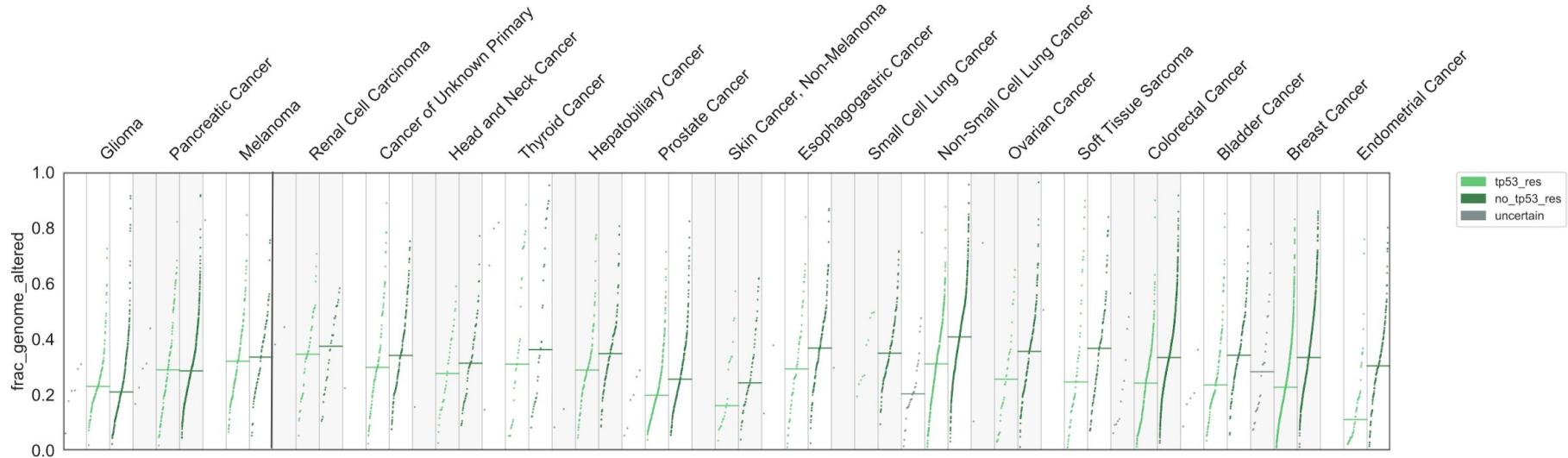


- In both Metastatic and Primary cohorts, Bi-allelic samples have higher Genome Instability than Mono-allelic samples
  - **Primary:** median 23.4% vs. 31.3%
  - **Metastasis:** median 27.4% vs. 34.2%
- No major difference of pattern between Primary and Metastatic cohorts.

# Non WGD Cohort - Fraction of Genome Altered

## TP53 Residual Subgroups (2/2)

### Fraction of Genome Altered - TP53 Residual Groups - NO WGD



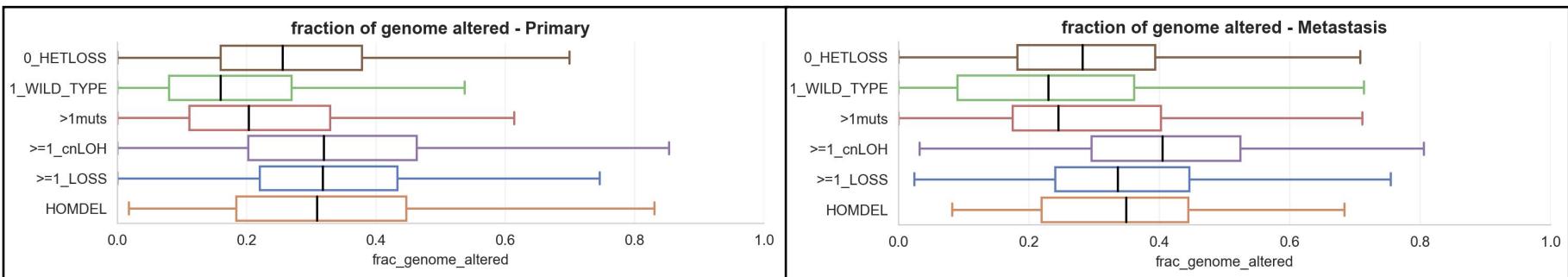
- **Outliers**  
→ No difference between Mono and Bi-Allelic

- **Average cases**  
→ Thyroid, Prostate, NSCLC

- **Textbook cases**  
→ Colorectal, Breast, Endometrial

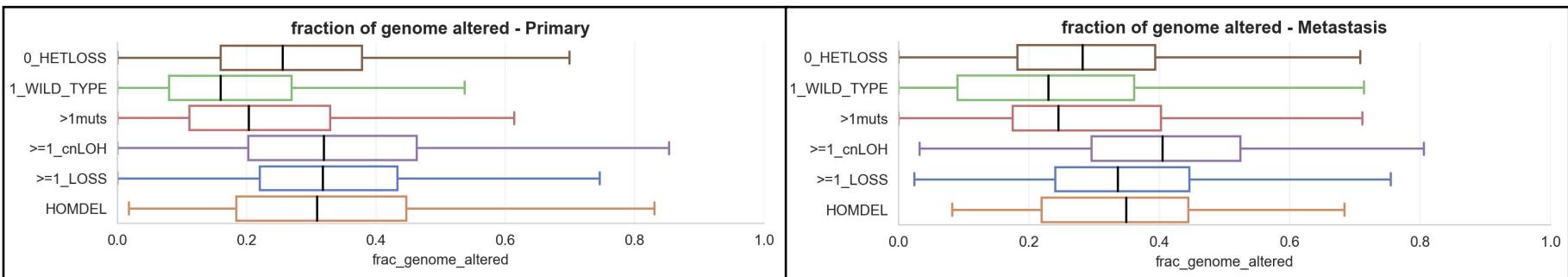
# Non WGD Cohort - Fraction of Genome Altered

## TP53 Groups (1/2)



# Non WGD Cohort - Fraction of Genome Altered

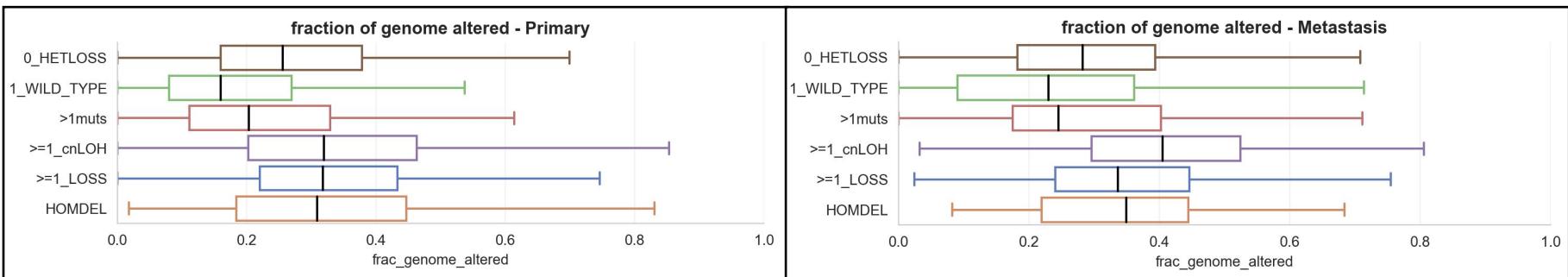
## TP53 Groups (1/2)



- Significant difference between likely Mono Allelic subgroups (Green, Brown) and likely Bi-Allelic subgroups (Purple, Blue)

# Non WGD Cohort - Fraction of Genome Altered

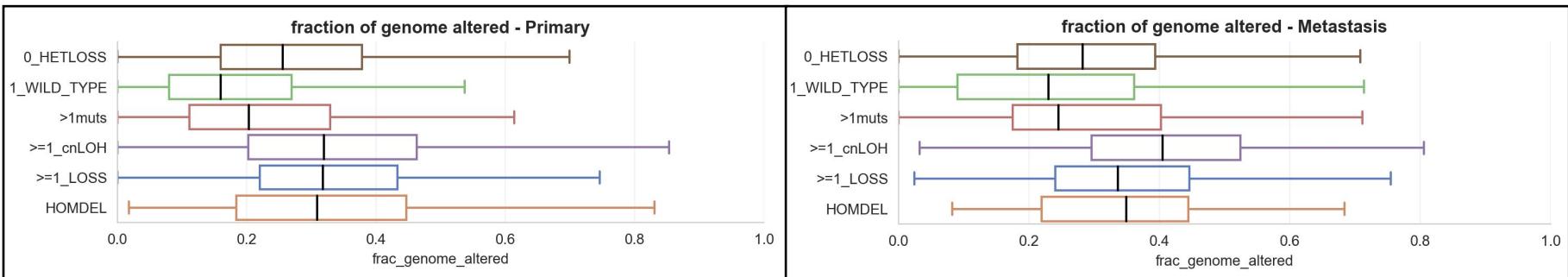
## TP53 Groups (1/2)



- Significant difference between likely Mono Allelic subgroups (Green, Brown) and likely Bi-Allelic subgroups (Purple, Blue)
- Significant difference between the 2 Mono-Allelic subgroups
  - **Why are Mono Allelic by LOSS more unstable than Mono Allelic by mutation?**

# Non WGD Cohort - Fraction of Genome Altered

## TP53 Groups (1/2)

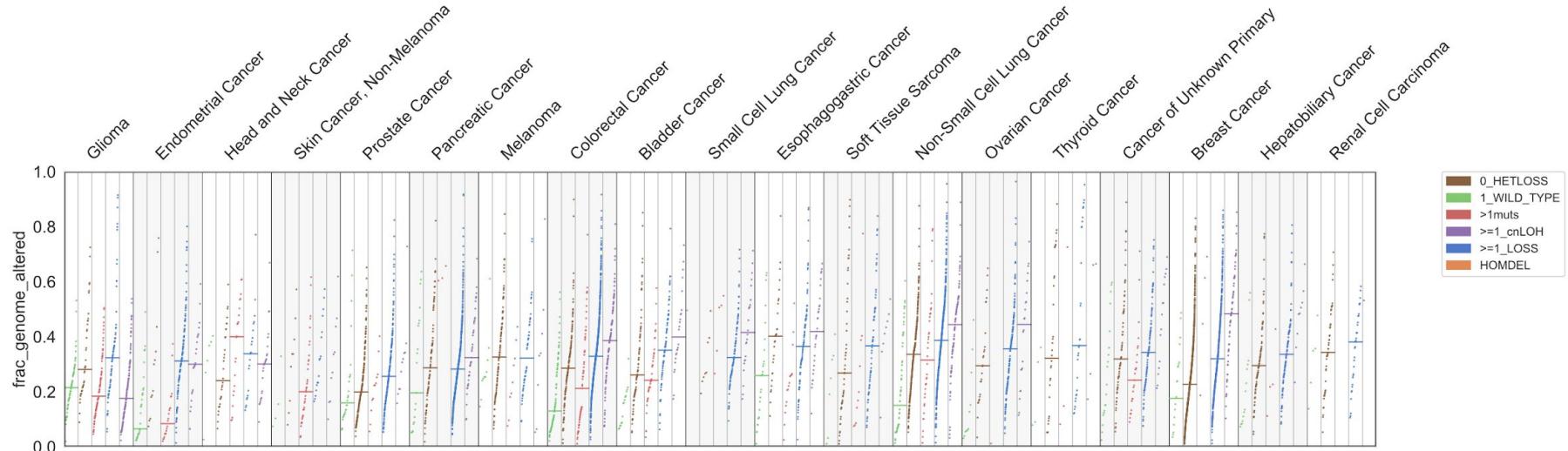


- Significant difference between likely Mono Allelic subgroups (Green, Brown) and likely Bi-Allelic subgroups (Purple, Blue)
- significant difference between the 2 Mono-Allelic subgroups
  - Why are Mono Allelic by LOSS more unstable than Mono Allelic by mutation?
- Same overall pattern
  - Difference between Mono and Bi
  - Difference between 0\_HETLOSS and 1\_WT

# Non WGD Cohort - Fraction of Genome Altered

## TP53 Groups (2/2)

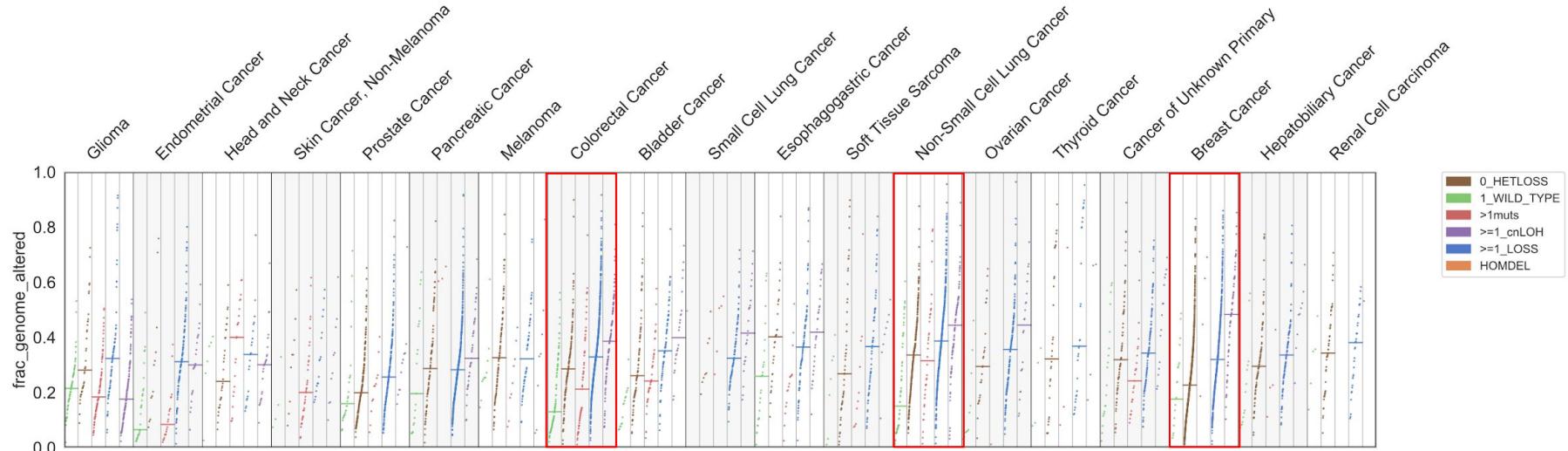
### Fraction of Genome Altered - TP53 Groups Detailed - NO WGD



# Non WGD Cohort - Fraction of Genome Altered

## TP53 Groups (2/2)

### Fraction of Genome Altered - TP53 Groups Detailed - NO WGD



1.

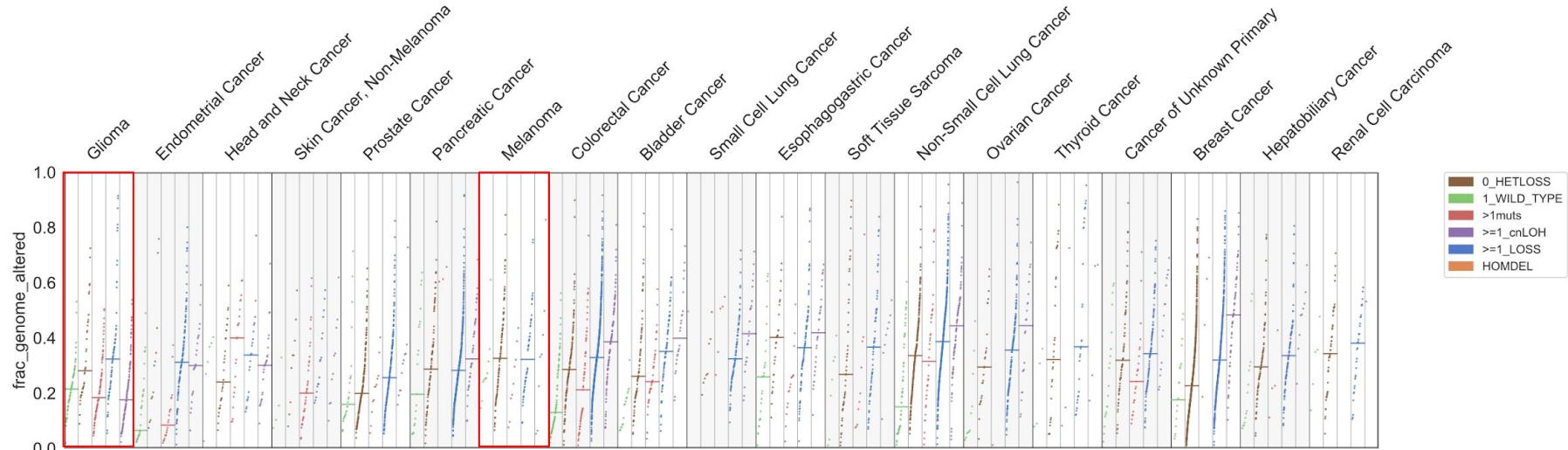
TextBook  
Cases

- Colorectal
- NSCLC
- Breast Cancer

# Non WGD Cohort - Fraction of Genome Altered

## TP53 Groups (2/2)

### Fraction of Genome Altered - TP53 Groups Detailed - NO WGD



1.

TextBook  
Cases

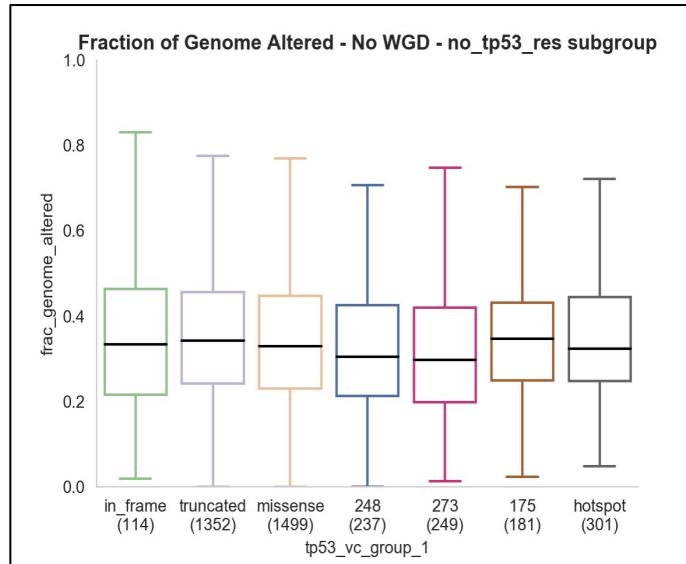
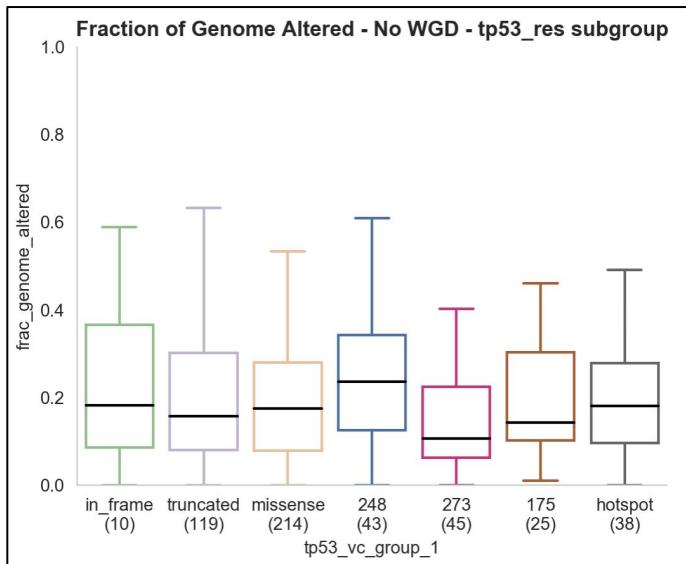
- Colorectal
- NSCLC
- Breast Cancer

2.

Outliers

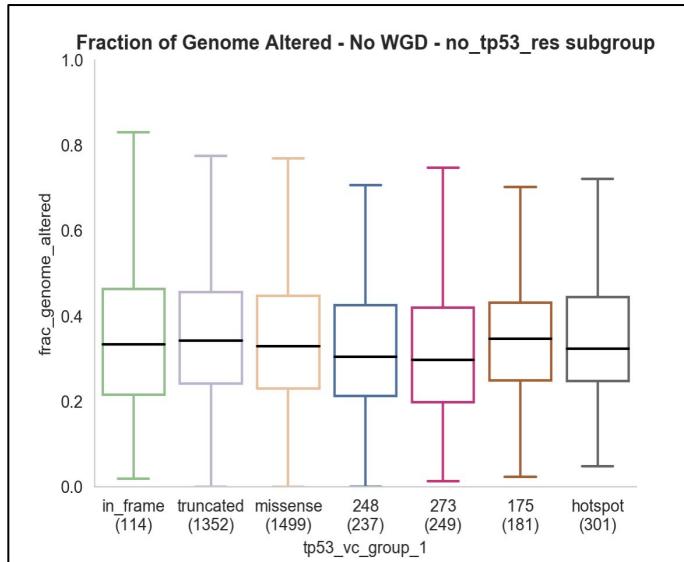
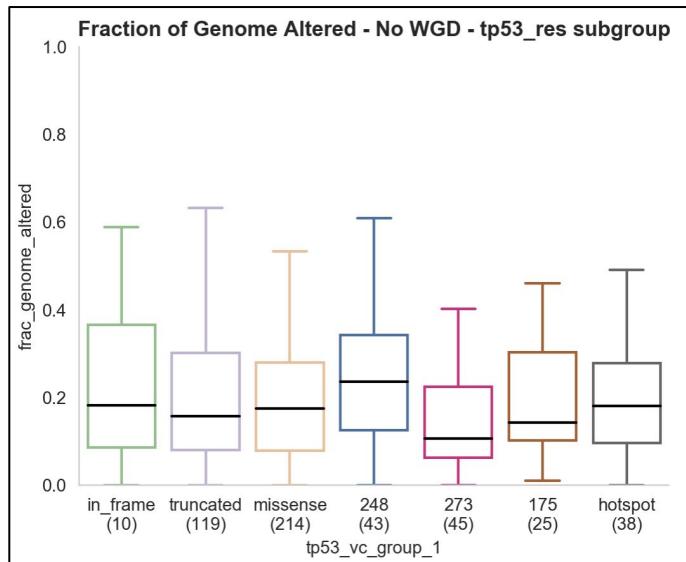
- Glioma
- Melanoma

# Non WGD Cohort - Hotspot TP53



- *hotspot contains 245, 282, 213, 352, 220, 196*

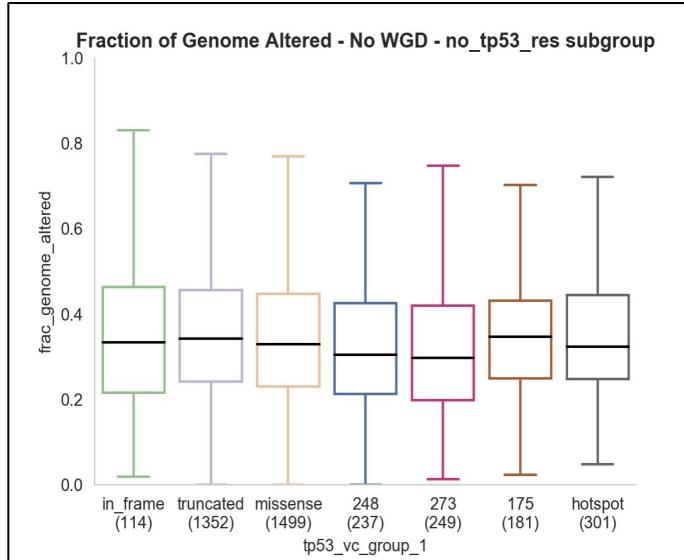
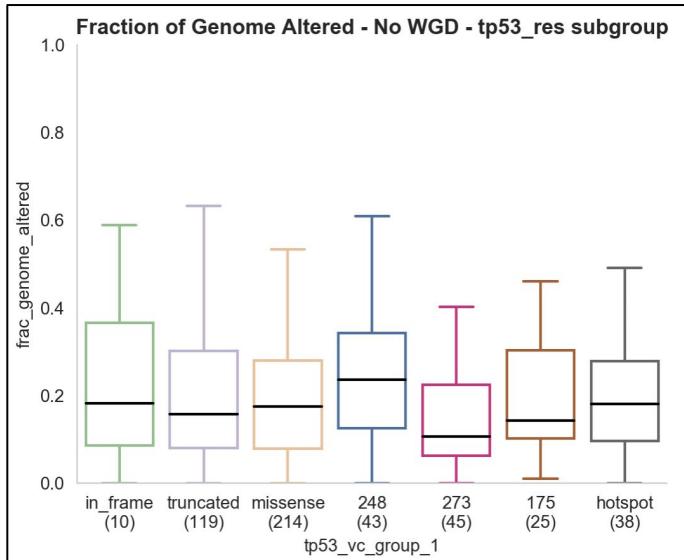
# Non WGD Cohort - Hotspot TP53



- In Mono-Allelic samples, Hotspot 248 leads to a higher Genome Instability
  - Hotspot 248 is known to have a dominant negative effect

- *hotspot contains*  
245, 282, 213,  
352, 220, 196

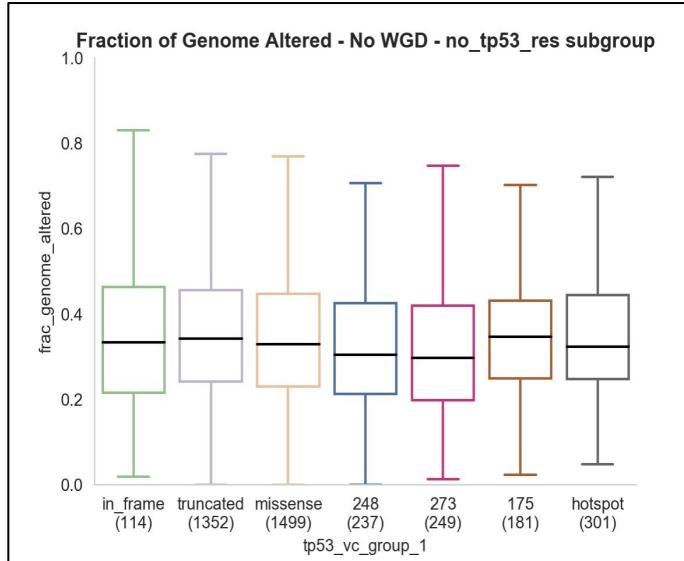
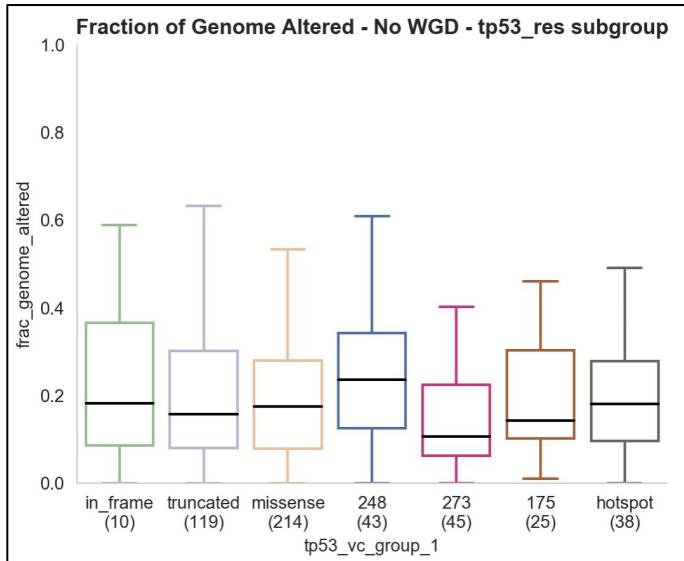
# Non WGD Cohort - Hotspot TP53



- hotspot contains 245, 282, 213, 352, 220, 196

- In Mono-Allelic samples, Hotspot 248 leads to a higher Genome Instability
  - Hotspot 248 is known to have a dominant negative effect
- In Bi-Allelic Samples, all hotspots have overall the same Genome Instability distribution
  - Selection pressure for a second mutation on all hotspots

# Non WGD Cohort - Hotspot TP53



- hotspot contains 245, 282, 213, 352, 220, 196

- In Mono-Allelic samples, Hotspot 248 leads to a higher Genome Instability
    - Hotspot 248 is known to have a dominant negative effect
  - In Bi-Allelic Samples, all hotspots have overall the same Genome Instability distribution
    - Selection pressure for a second mutation on all hotspots
- Difficulties to do this analysis per cancer type, not enough hotspots

---

# Cancer Exploration

---

---

# Cancer Exploration

---

- NSCLC

- Breast Cancer

- Glioma

---

# Cancer Exploration

---

- NSCLC
- Breast Cancer
- Glioma

In this section, we only analyse Non-WGD Cohort.

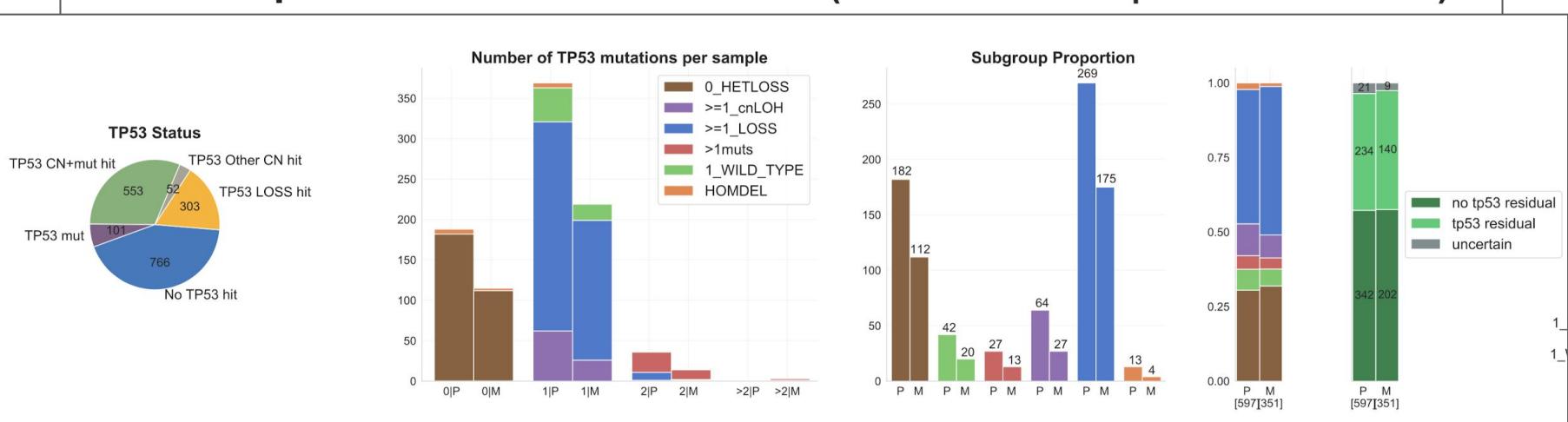
---

# **Non-Small Cell Lung Cancer**

# NSCLC - Proportions

## NON-SMALL CELL LUNG CANCER - NO WGD

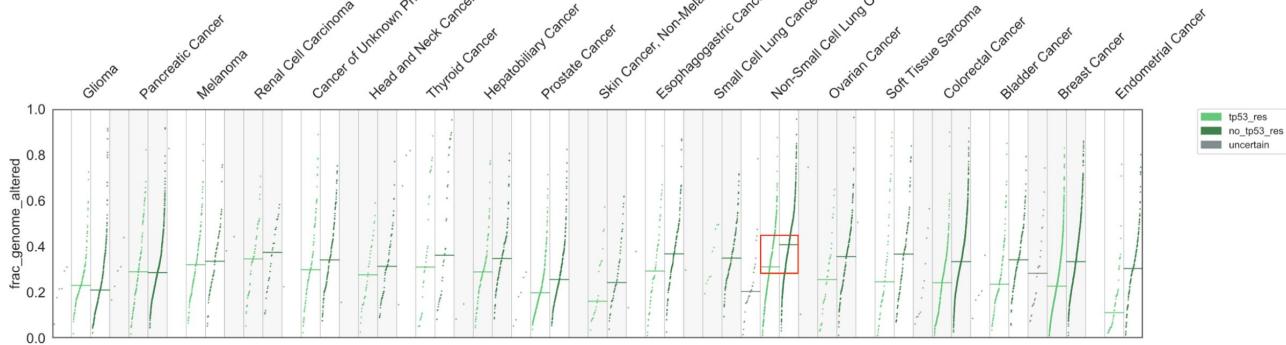
1775 samples - 11.44 % of the cohort (P: 1232 - 69.4 % | M: 543 - 30.6 %)  
1009 samples with TP53 hits - 56.85 % (P: 643 - 63.7 % | M: 366 - 36.3 %)



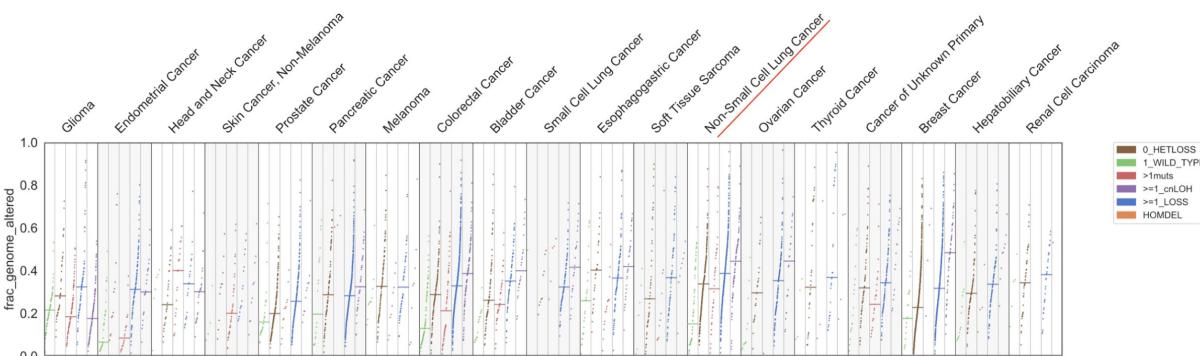
- High proportion of >=1\_LOSS
- Depletion of Metastatic Samples

# NSCLC - TextBook Case

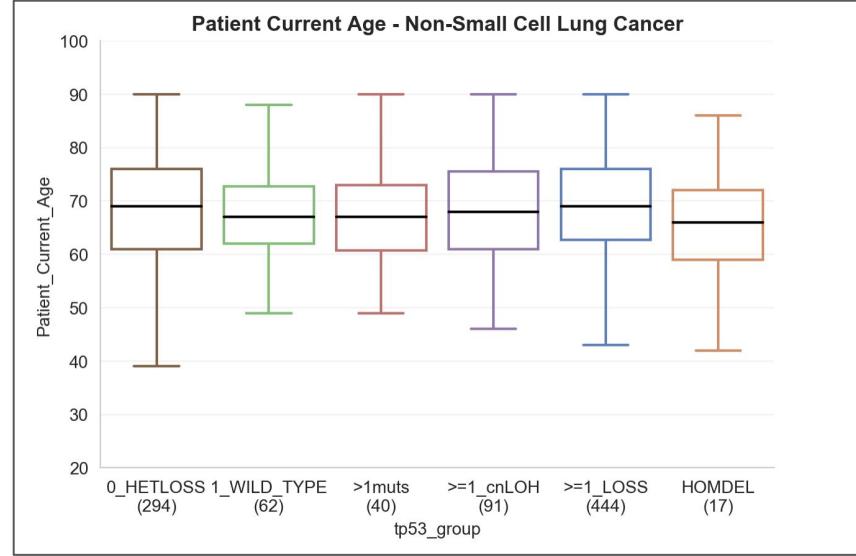
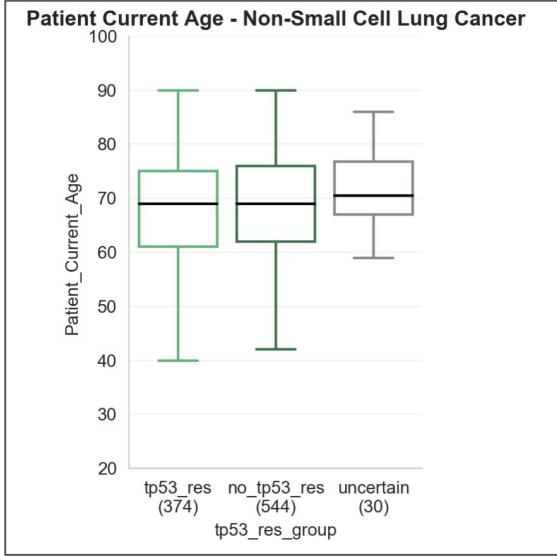
## Fraction of Genome Altered - TP53 Residual Groups - NO WGD



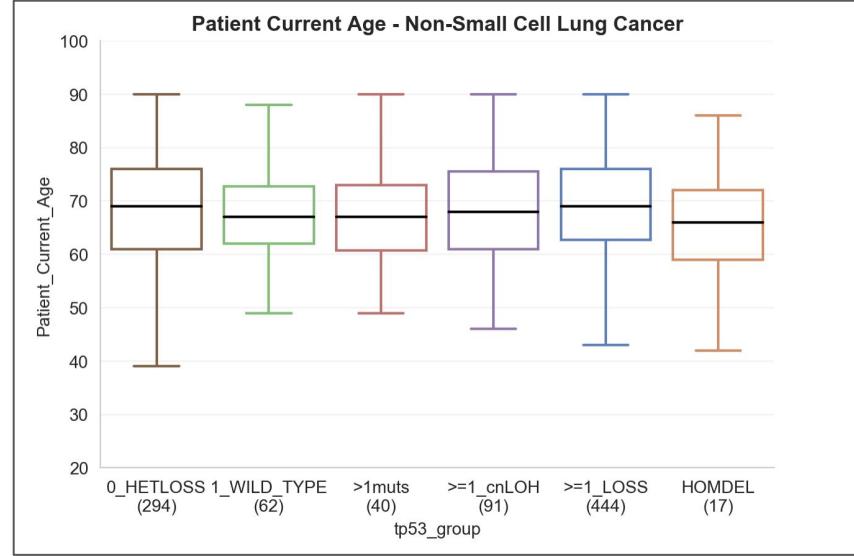
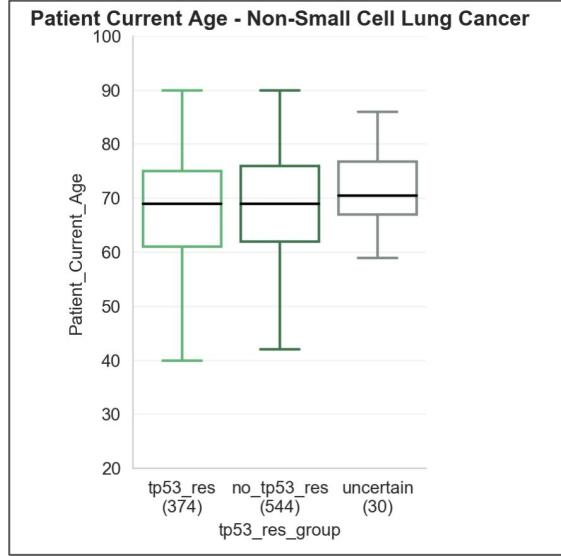
## Fraction of Genome Altered - TP53 Groups Detailed - NO WGD



# NSCLC - Age

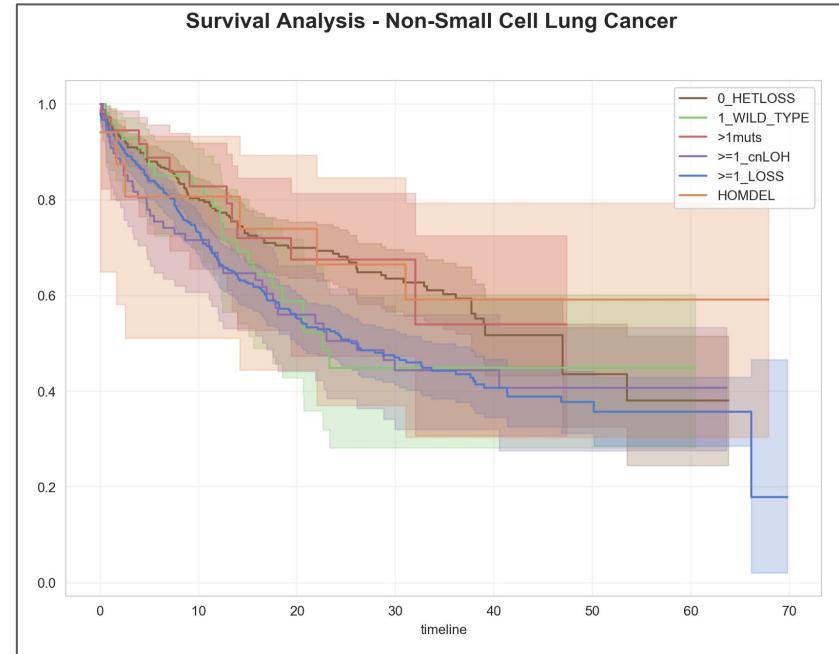
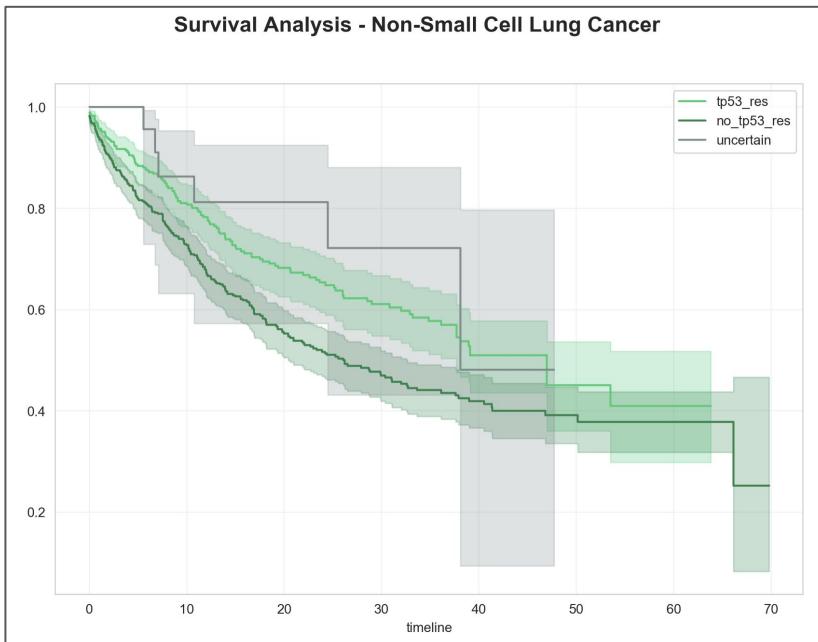


# NSCLC - Age



- No difference between subgroups in term of age

# NSCLC - Survival Analysis



- Mono-Allelic samples have a better survival than Bi-Allelic samples
- The signal is led by 0\_HETLOSS and >=1\_LOSS
- 1\_WILD\_TYPE has the same survival as >=1\_LOSS

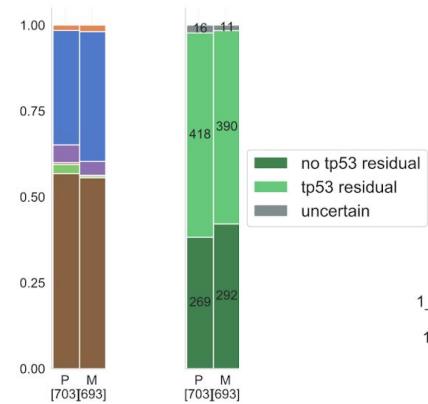
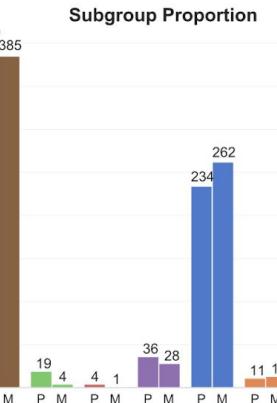
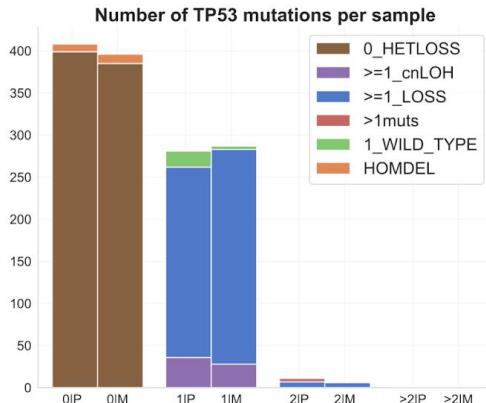
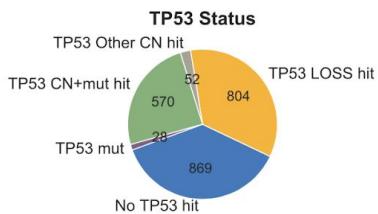
---

# Breast Cancer

# Breast Cancer - Proportions

## BREAST CANCER - NO WGD

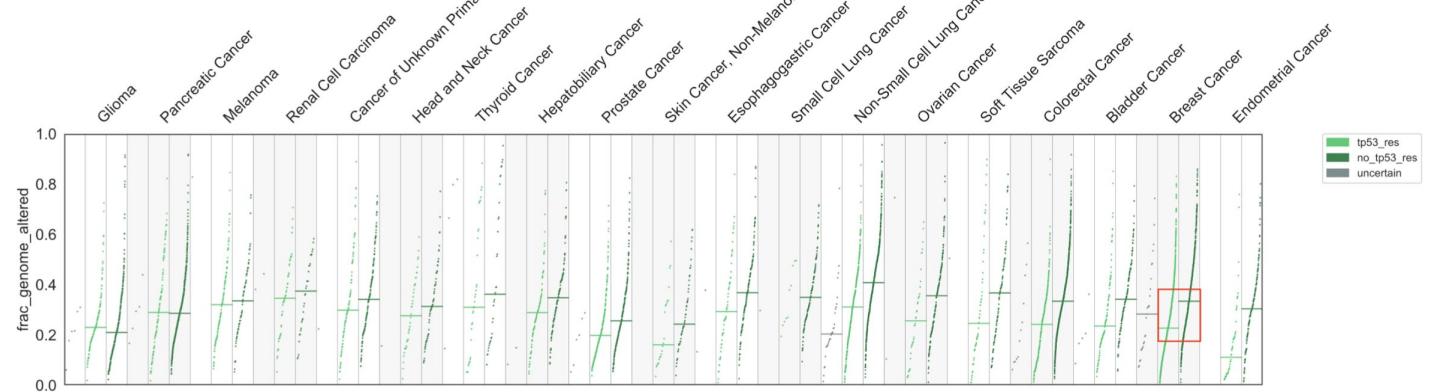
2323 samples - 14.97 % of the cohort (P: 1256 - 54.1 % | M: 1067 - 45.9 %)  
1454 samples with TP53 hits - 62.59 % (P: 740 - 50.9 % | M: 714 - 49.1 %)



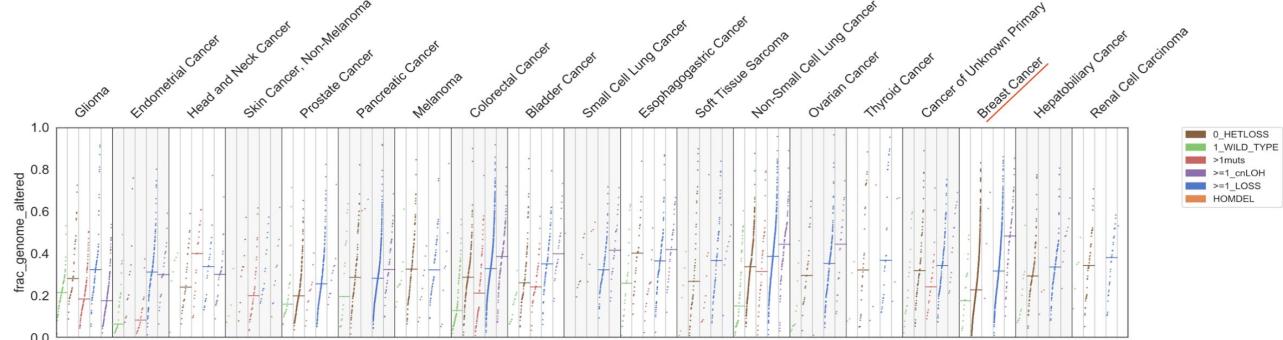
- High proportion of TP53 LOSS Hit
- Enrichment in 0\_HETLOSS
- Few TP53 composite samples

# Breast Cancer - TextBook Case

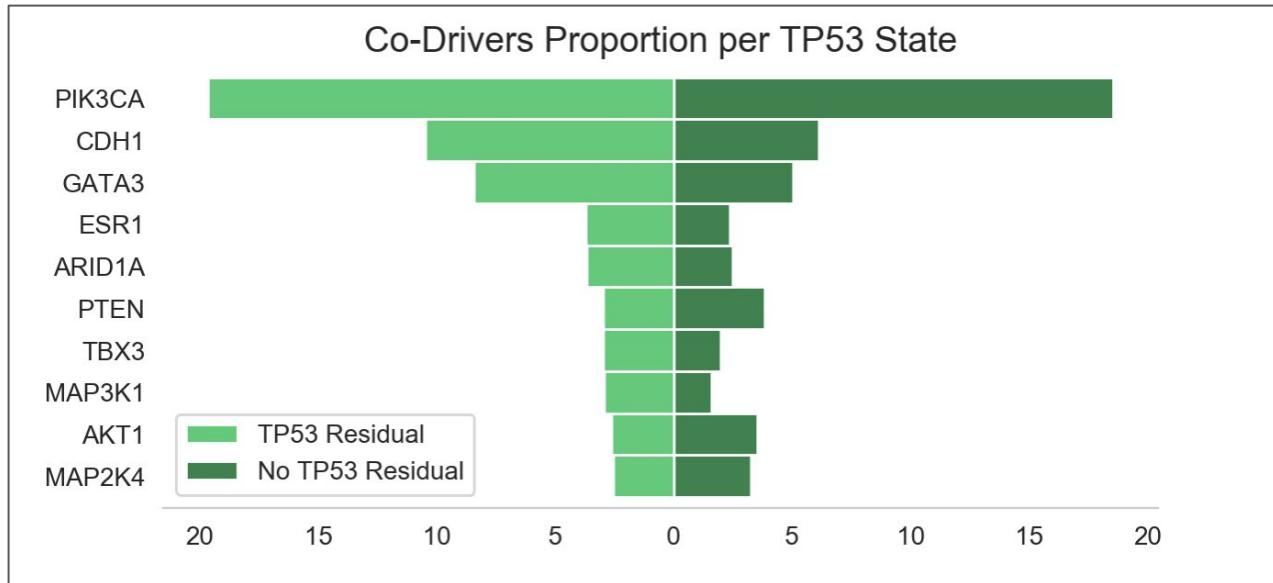
## Fraction of Genome Altered - TP53 Residual Groups - NO WGD



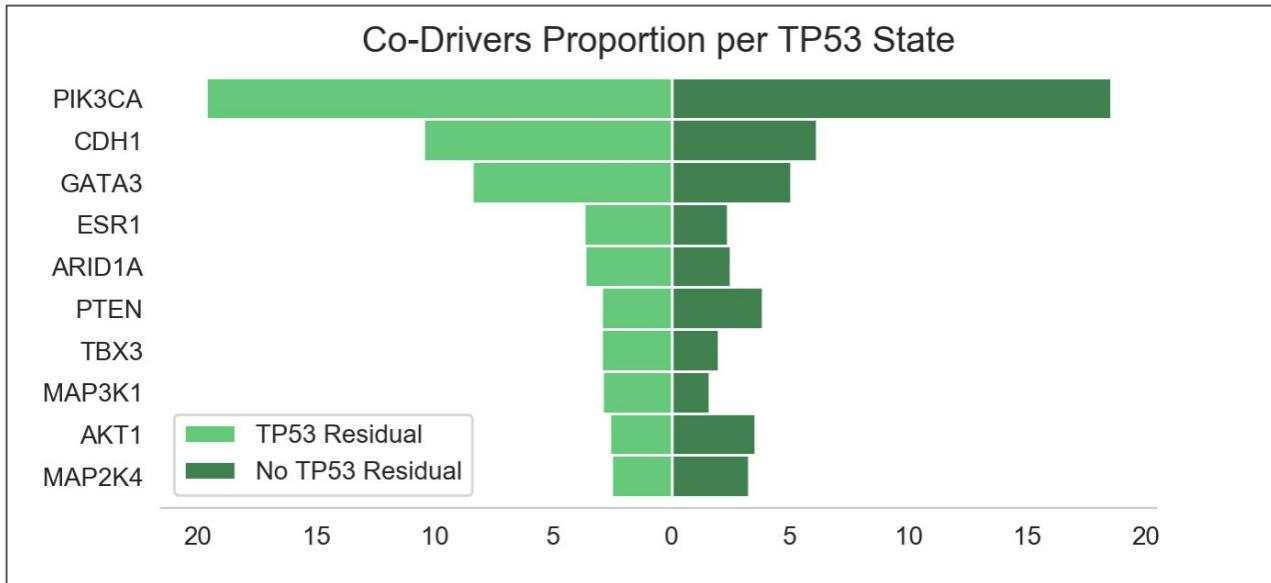
## Fraction of Genome Altered - TP53 Groups Detailed - NO WGD



# Breast Cancer - Co Driver Analysis



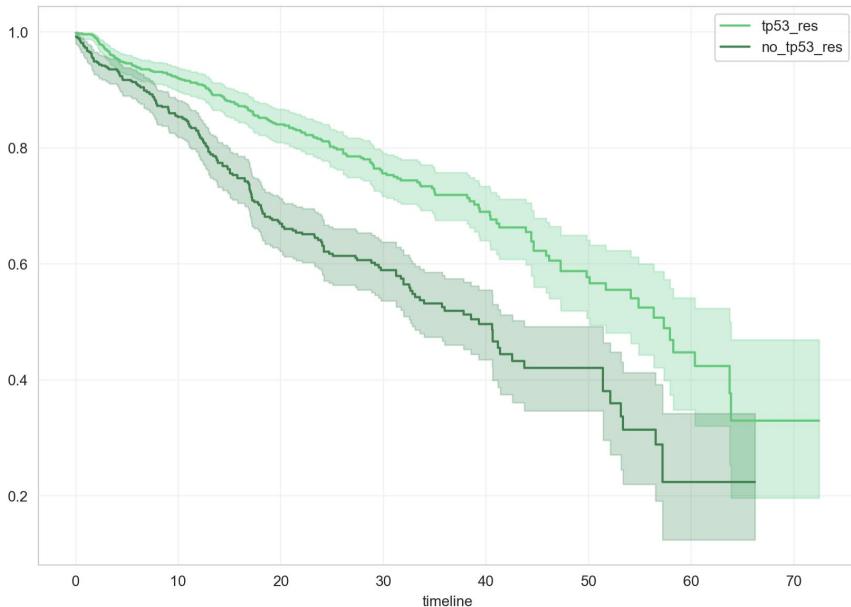
# Breast Cancer - Co Driver Analysis



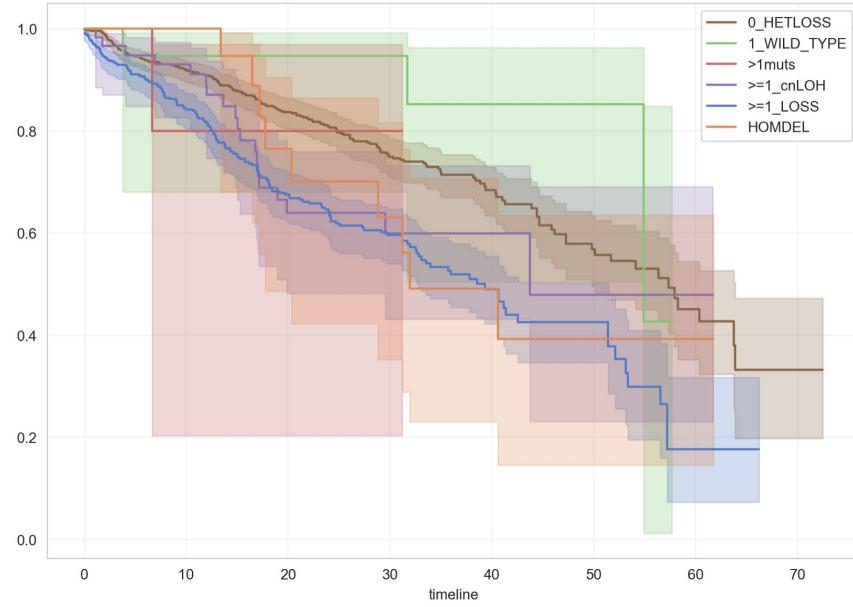
- Overall the same distribution between Mono-Allelic and Bi-Allelic samples
  - Co-drivers are not leading the signal if we are to observe one in Clinical Outcomes

# Breast Cancer - Survival Analysis

Survival Analysis - Non-WGD Cohort - Breast Cancer



Survival Analysis - Non-WGD Cohort - Breast Cancer

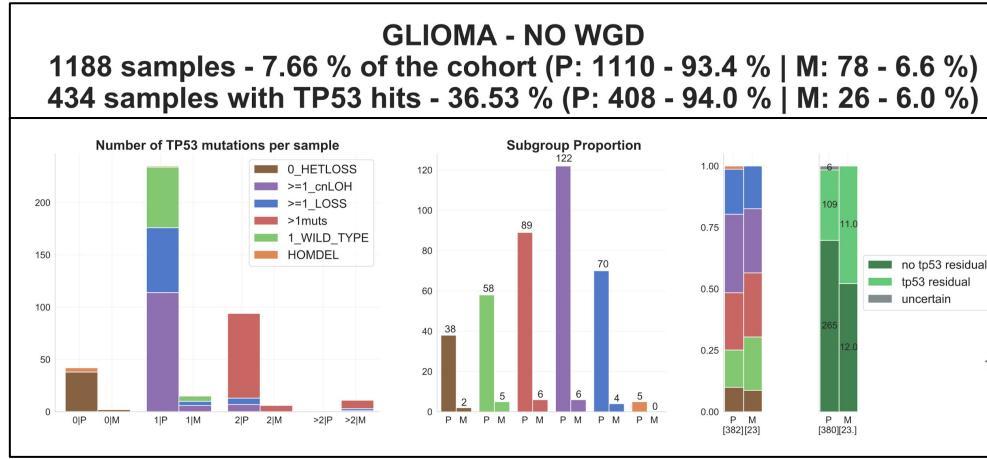


- Significant difference between Mono and Bi Allelic Survivals
- The signal is led by  $\geq 1\_LOSS$  and  $0\_HETLOSS$

---

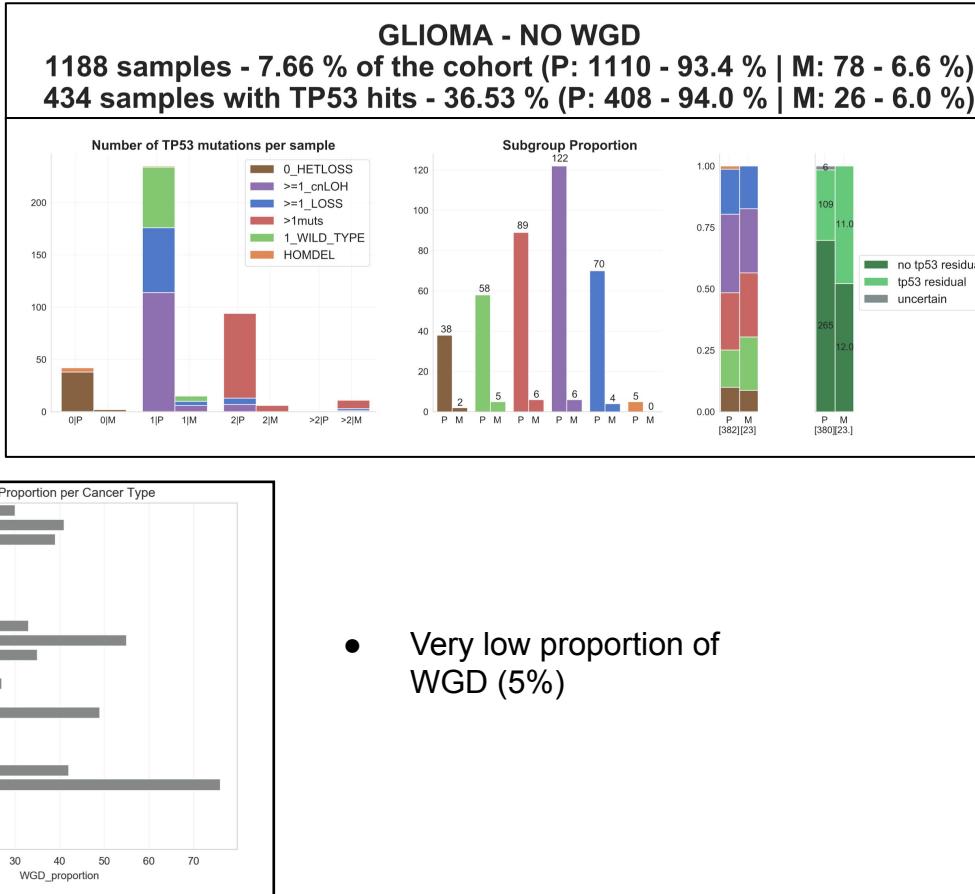
# **Glioma**

# Glioma - Outlying Signals



- Low proportion of Metastatic Samples
- High proportion of TP53 Composite samples

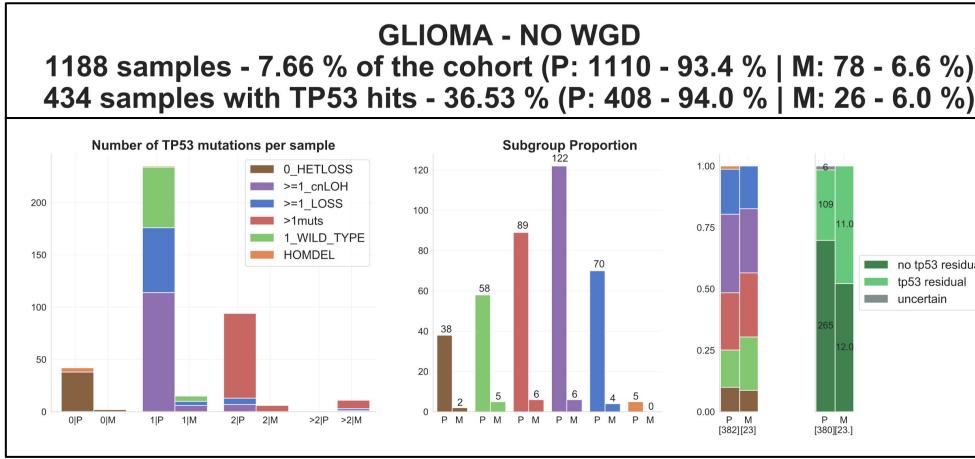
# Glioma - Outlying Signals



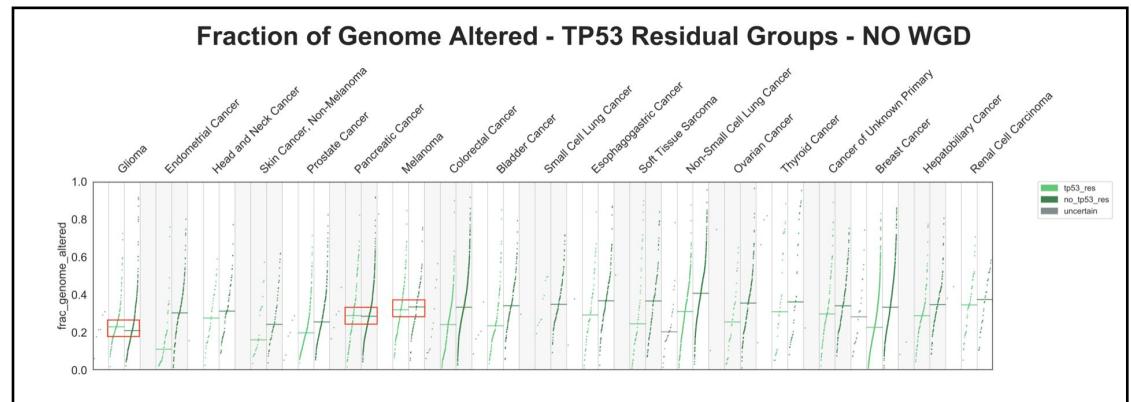
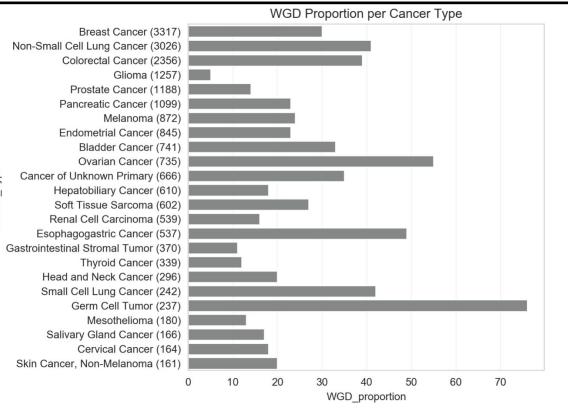
- Low proportion of Metastatic Samples
- High proportion of TP53 Composite samples

- Very low proportion of WGD (5%)

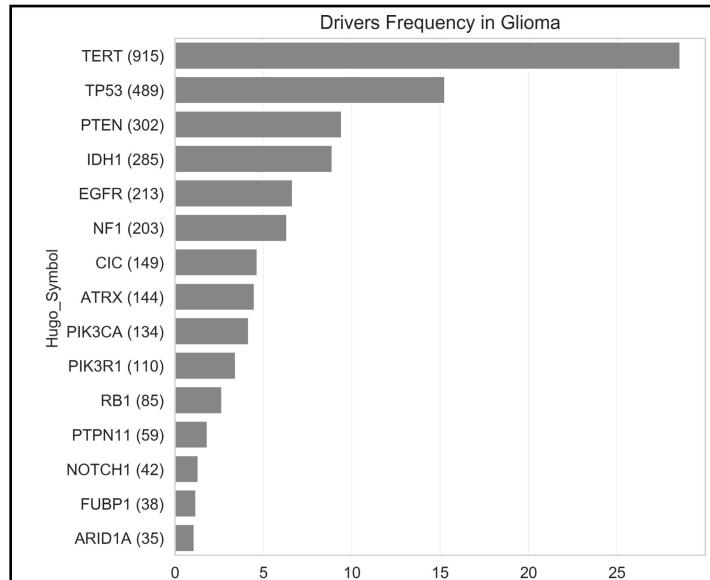
# Glioma - Outlying Signals



- Low proportion of Metastatic Samples
- High proportion of TP53 Composite samples



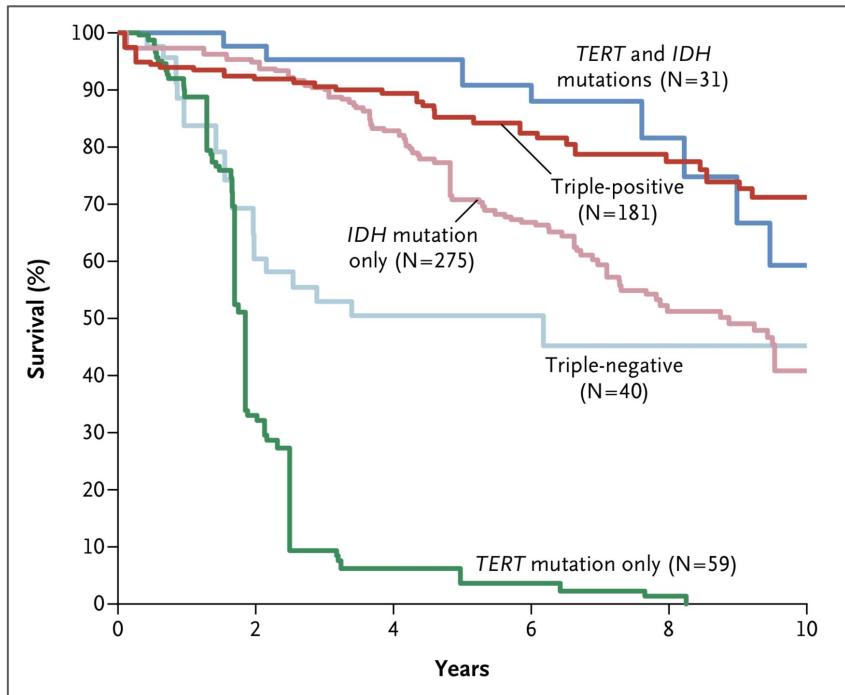
# Glioma - Driver Analysis



Major drivers in Glioma:

- TERT (29 %)
- TP53 (15%)
- PTEN (9.5%)
- IDH1 (9%)

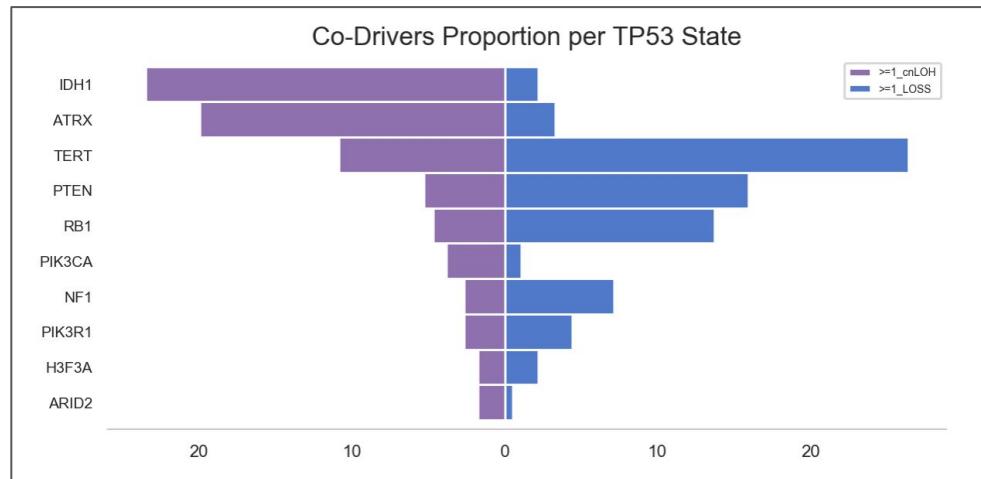
# Glioma - Driver Analysis



- TERT only is associated with ‘really bad’ Survival
- IDH only is associated with ‘normal’ Survival
- IDH+TERT is associated with good Survival

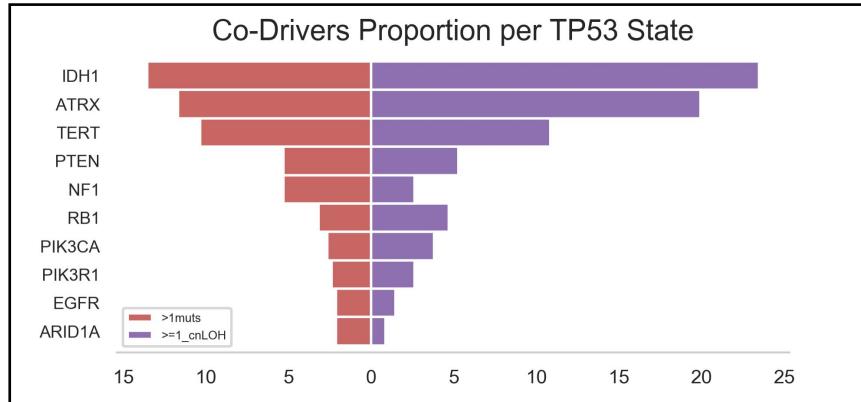
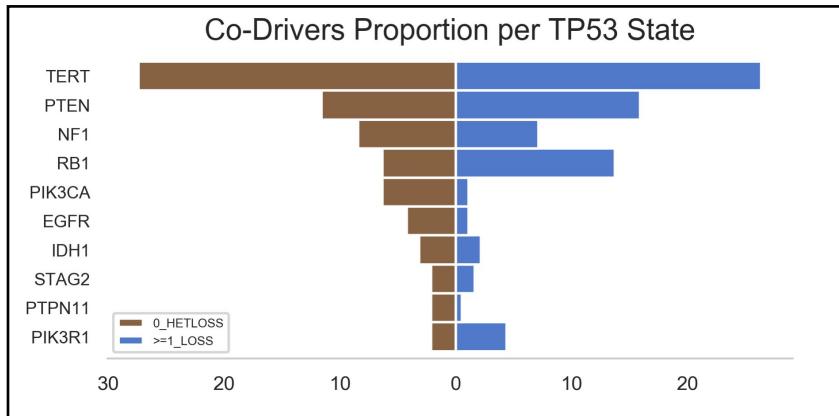
<https://www.nejm.org/doi/pdf/10.1056/NEJMoa1407279?articleTools=true>

# Glioma - Driver Analysis



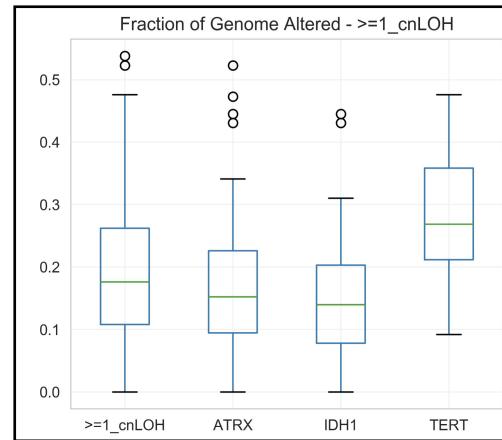
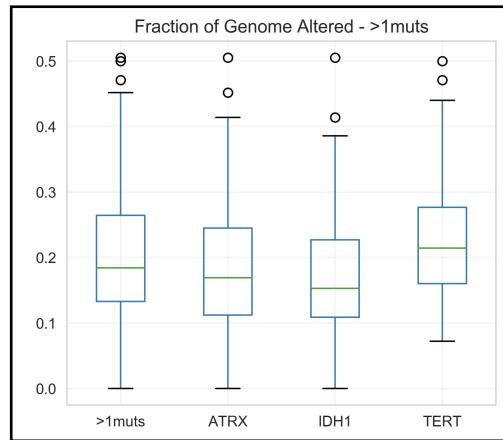
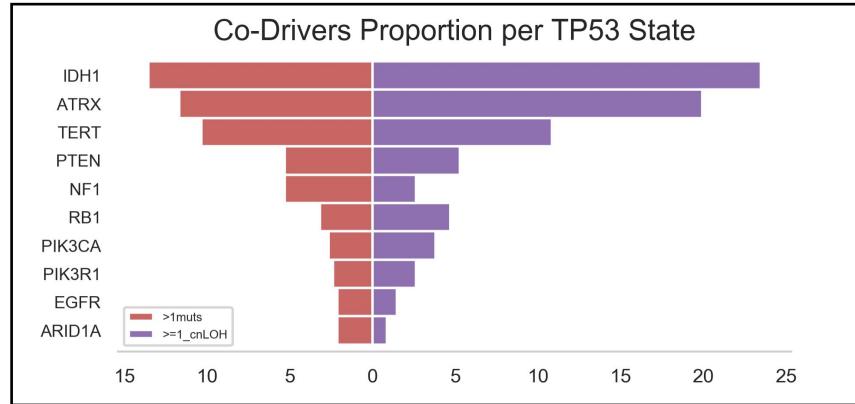
- Huge Difference in proportion between the two Bi-Allelic subgroups
  - **>=1\_cnLOH:** IDH1 and ATRX
  - **>=1 LOSS:** TERT and PTEN

# Glioma - Driver Analysis

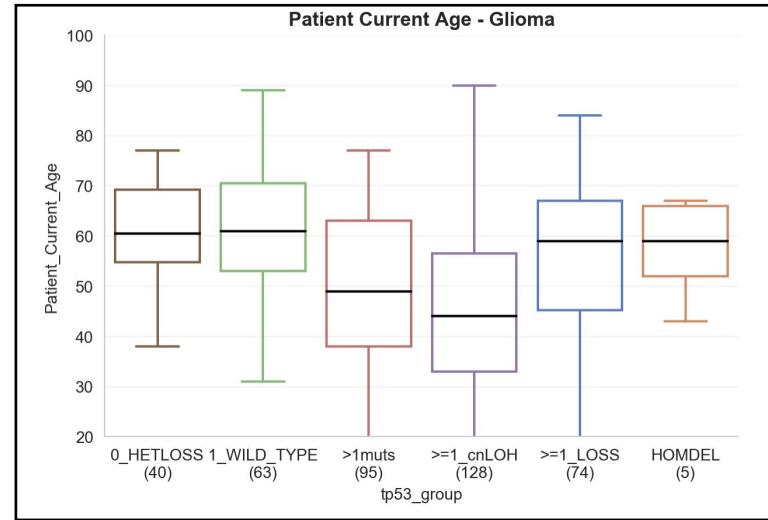
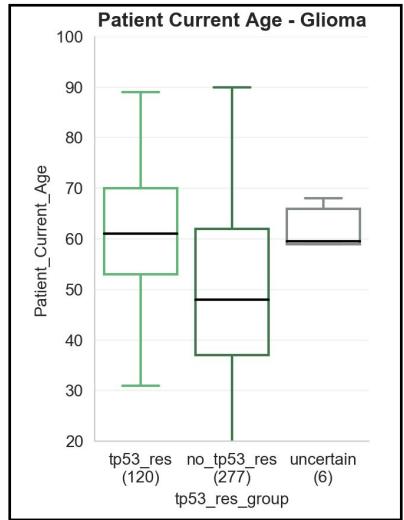


- Co-Driver distribution association between 0\_HETLOSS and >=1\_LOSS
- Co-Driver distribution association between >1muts and >=1\_cnLOH

# Glioma - Driver Analysis

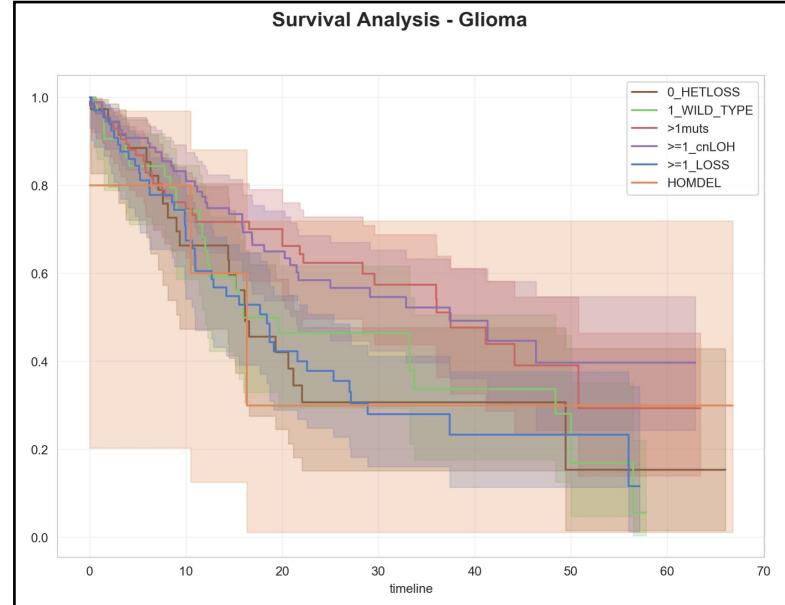
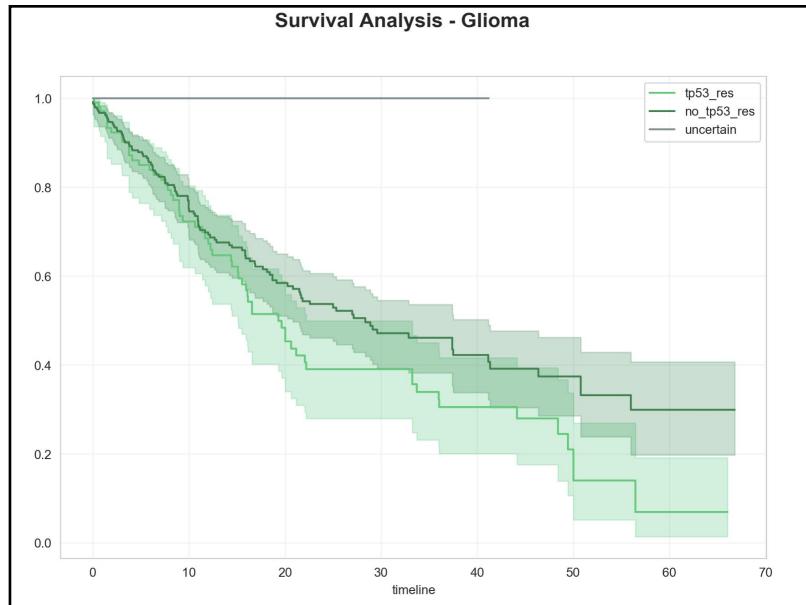


# Glioma - Age



- $\geq 1\_cnLOH$  subgroup shows a lower median for the Patient Age  
→ Knowing the proportion of  $\geq 1\_cnLOH$  in Glioma, it leads the signal in no\_tp53\_res Age

# Glioma - Survival Analysis



- Opposite pattern as for the 2 other cancers  
→ Bi-Allelic samples have a better Survival than Mono-Allelic samples

- Similarities observed in terms of Co-Driver distribution are still present:
  - 0\_HETLOSS and >=1\_LOSS - TERT-bad survival
  - >1mut ann >=1\_cnLOH- IDH1- good survival

---

## Conclusion & Next Steps

---

# Conclusion & Next Steps

---

## Conclusive Points

- TP53 Allelic State has not a uniform representation through all cancer types
  - While some cancers are more mono-allelic, others show a high proportion of Bi-allelic

# Conclusion & Next Steps

---

## Conclusive Points

- TP53 Allelic State has not a uniform representation through all cancer types
  - While some cancers are more mono-allelic, others show a high proportion of Bi-allelic
- We observe a recurrent pattern of Genome Instability change through a majority of cancer types
  - Bi-allelic samples are more genetically unstable than Mono-allelic, and have a worse Survival
  - Some outliers like Glioma, melanoma, Pancreatic Cancer

# Conclusion & Next Steps

---

## Conclusive Points

- TP53 Allelic State has not a uniform representation through all cancer types
  - While some cancers are more mono-allelic, others show a high proportion of Bi-allelic
- We observe a recurrent pattern of Genome Instability change through a majority of cancer types
  - Bi-allelic samples are more genetically unstable than Mono-allelic, and have a worse Survival
  - Some outliers like Glioma, melanoma, Pancreatic Cancer
- In WGD State, the association between TP53 mutant and WGD is driven by Bi-Alelic TP53

# Conclusion & Next Steps

---

## Conclusive Points

- TP53 Allelic State has not a uniform representation through all cancer types
  - While some cancers are more mono-allelic, others show a high proportion of Bi-allelic
- We observe a recurrent pattern of Genome Instability change through a majority of cancer types
  - Bi-allelic samples are more genetically unstable than Mono-allelic, and have a better Survival
  - Some outliers like Glioma, melanoma, Pancreatic Cancer
- In WGD State, the association between TP53 mutant and WGD is driven by Bi-Alelic TP53

# Conclusion & Next Steps

---

## Conclusive Points

- TP53 Allelic State has not a uniform representation through all cancer types
    - While some cancers are more mono-allelic, others show a high proportion of Bi-allelic
  - We observe a recurrent pattern of Genome Instability change through a majority of cancer types
    - Bi-allelic samples are more genetically unstable than Mono-allelic, and have a worse Survival
    - Some outliers like Glioma, melanoma, Pancreatic Cancer
  - In WGD State, the association between TP53 mutant and WGD is driven by Bi-Allelic TP53
- 

## Next steps

- Pair Analysis: understand the association between TP53 State and Disease Progression
- Co occurring event on 17p: understand the Genome Instability gap between 0\_HETLOSS and 1\_WILD\_TYPE
- Speaking with cancer specialists

---

# Appendix

---

# TP53 Residual Subgroups

We used the following formula to compute the expected number of WT tp53 residual:

$$wt\_residual = tcn - \frac{vaf}{purity} \times (tcn \times purity + 2 \times (1 - purity))$$

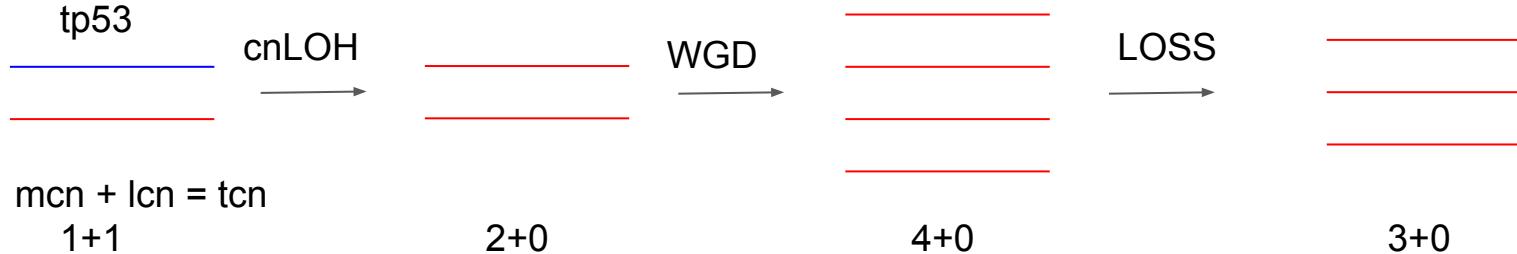
---

We tailored specific thresholds for each tp53\_cn\_state and defined the following groups:

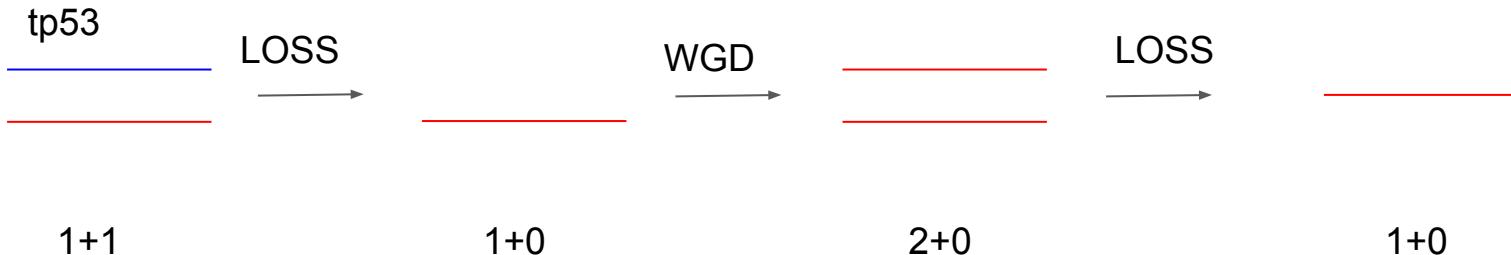
 tp53_res	tp53_res: 3 036
 no_tp53_res	no_tp53_res: 4 478
 uncertain	uncertain: 118

# TP53 State Study: Allelic Representation

## CNLOH BEFORE & LOSS



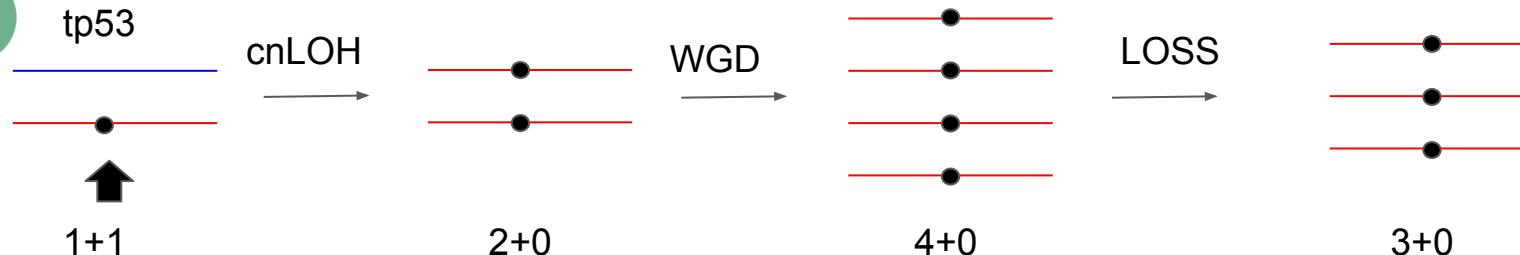
## LOSS BEFORE & AFTER



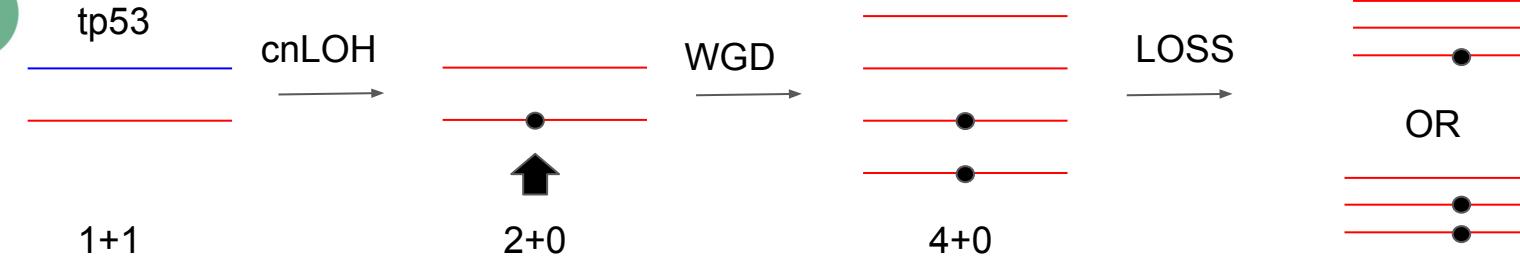
# TP53 State Study: With mutation

## CNLOH BEFORE & LOSS

1

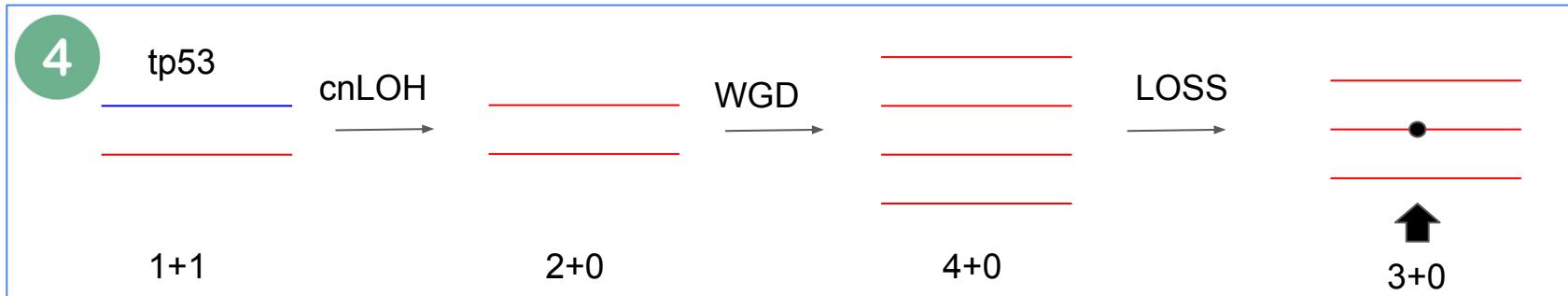
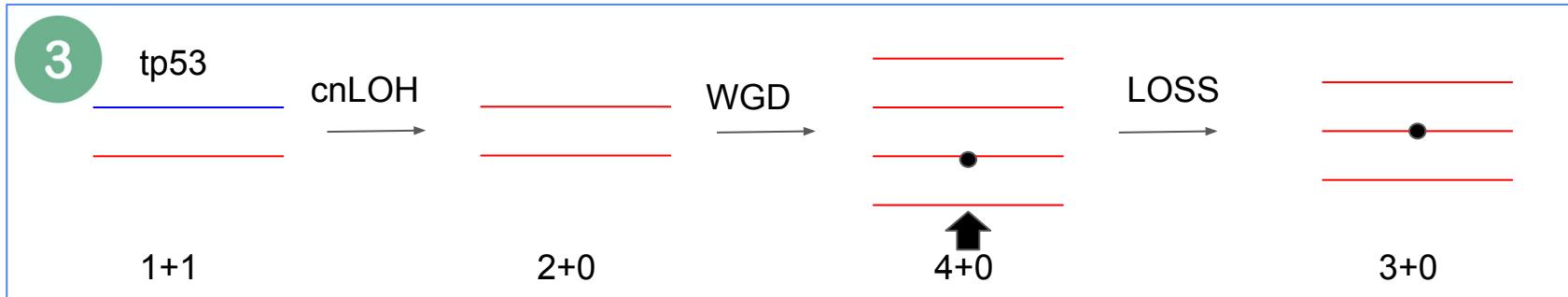


2



# TP53 State Study: With mutation

## CNLOH BEFORE & LOSS



# TP53 State Study: With mutation

CNLOH BEFORE & LOSS

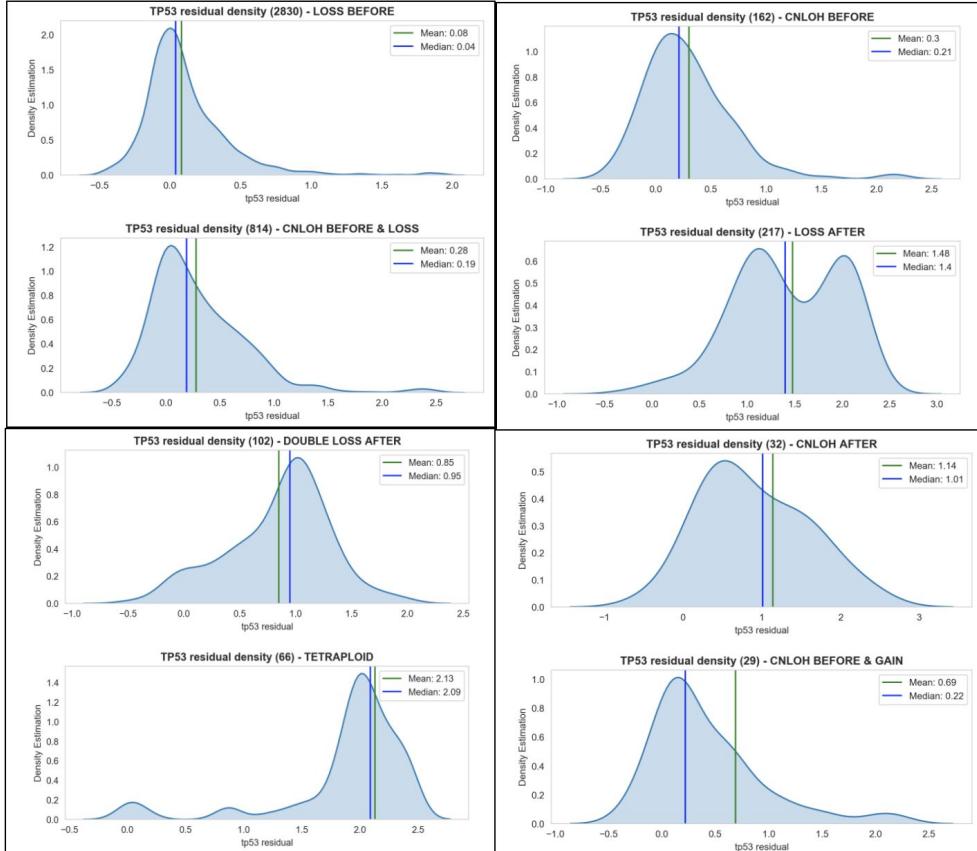
Mutation Timing	# of tp53 mutant copies	# WT Alleles
mut before CNLOH	3	0
mut after CNLOH, before WGD	1	2
mut after CNLOH, before WGD	2	1
mut after WGD, before LOSS	1	2
mut after LOSS	1	2

# Methods - Pre WGD TP53 Allelic State

NUMBER OF TP53 residual ASSOCIATED WITH ALLELIC STATE BEFORE WGD  
BI-ALLELIC | MONO - ALLEGIC | 2WT

LOSS BEFORE :	0		1		-
CNLOH BEFORE & LOSS:	0		1,2		2
<u>CNLOH</u> BEFORE:	0		2		3
LOSS AFTER:	-		1,2		2
DOUBLE LOSS AFTER:	-		0,1		1
TETRAPLOID:	-		2		3
<u>CNLOH</u> AFTER:	-		1,3		3
CNLOH BEFORE & GAIN:	0		2,3		3,4

# Methods - Pre WGD TP53 Allelic State



## THRESHOLDS BETWEEN ALLELIC STATES BEFORE WGD

### BI-ALLELIC | MONO - ALLEGIC | 2WT

LOSS BEFORE :	<0.4	>0.6	-
CNLOH BEFORE & LOSS:	<0.4	0.6<<1.5	uncertain
CNLOH BEFORE:	<1.5	1.5<<2.5	>2.5
LOSS AFTER:	-	<1.5	uncertain
DOUBLE LOSS AFTER:	-	<0.5	uncertain
TETRAPLOID:	-	<2.5	>2.5
CNLOH AFTER:	-	<1.5	uncertain
CNLOH BEFORE & GAIN:	<1.4	1.6<<2.5	uncertain

