

## Supplementary Material

### Supplementary Tables

**Table S1. Cohorts analyzed in this study**

Samples	WGS	WES	RNAseq	Total
CVID	-	45	-	45
Argentinian trio/Mediterranean fever	-	23	-	23
Vall d'Hebron	12	13	5	13
Oxford Nanopore sample	1	-	-	1

**Table S2. MELT filter types**

Filter	Description
PASS	This means that the detected TE passed all quality filters and is likely to be a true positive.
ac0	This means that the alternate allele count is zero. In other words, none of the analyzed samples contained TE. These are most likely false positives and are typically filtered out.
hDP	This stands for a high level of coverage. It means that the region around the detected TE has an unusually large number of reads mapped to it. This could be due to sequencing errors or duplications in the genome, which are frequently excluded from further analysis.
lc	This represents low complexity. It indicates that the region surrounding the detected TE contains repetitive sequences. These regions can be difficult to accurately map reads to, so TEs detected in them are frequently treated with caution.

**Table S3. Exons distance program tags**

Tag	Distance	Description
Inside	The insertion is located within the nearest exon.	This implies that the insertion is very likely to influence the genome.
Close	The location of the insertion is less than 100 base pairs away from the nearest exon.	This suggests that the detected insertion is likely to influence the genome.
Near	The insertion is more than 100 base pairs but less than 1,000 base pairs away from the nearest exon.	This indicates that the detected insertion is still likely to influence the genome, although in a lesser way than the previous two tags.
Far	The insertion is found at more than 1,000 base pairs but less than 10,000 base pairs away from the nearest exon.	The detected insertion is unlikely to influence the genome.
Very far	The insertion is located at more than 10,000 base pairs away from the nearest exon.	This means that the detected insertion is highly unlikely to influence the genome.

**Table S4. Unique rate detection by the tool in each dataset**

Tool	CVID	Argentinian trio	Mediterranean fever	Vall d'Hebron WES	Vall d'Hebron WGS	Average detection
MELT	18.6%	23%	23.2%	16.3%	9.2%	18.06%
SCRAMble	29%	62%	39.1%	40.3%	41.4%	42.36%
mobster	44.7%	29.1%	27.6%	40.4%	10.5%	30.46%

**Table S5. Two-tool rate detection by the tool in each dataset**

Tool	CVID	Argentinian trio	Mediterranea n fever	Vall d'Hebron WES	Vall d'Hebron WGS	Average detection
MELT	26.6%	46.1%	50.7%	20%	20.9%	32.86%
SCRAMble	32.2%	24.1%	34.1%	28%	10.6%	25.8%
mobster	29.5%	54.2%	52.6%	32.2%	21.7%	38.04%

**Table S6. Three-tool rate detection by the tool in each dataset**

Tool	CVID	Argentinian trio	Mediterranean fever	Vall d'Hebron WES	Vall d'Hebron WGS	Average detection
MELT	54.8%	30.9%	26.1%	63.7%	69.9%	49.08%
SCRAMble	38.8%	13.9%	26.8%	31.7%	48%	31.84%
mobster	25.8%	16.7%	19.8%	27.4%	67.8%	31.5%

**Table S7. The detection rate of WES insertions also found in WGS data**

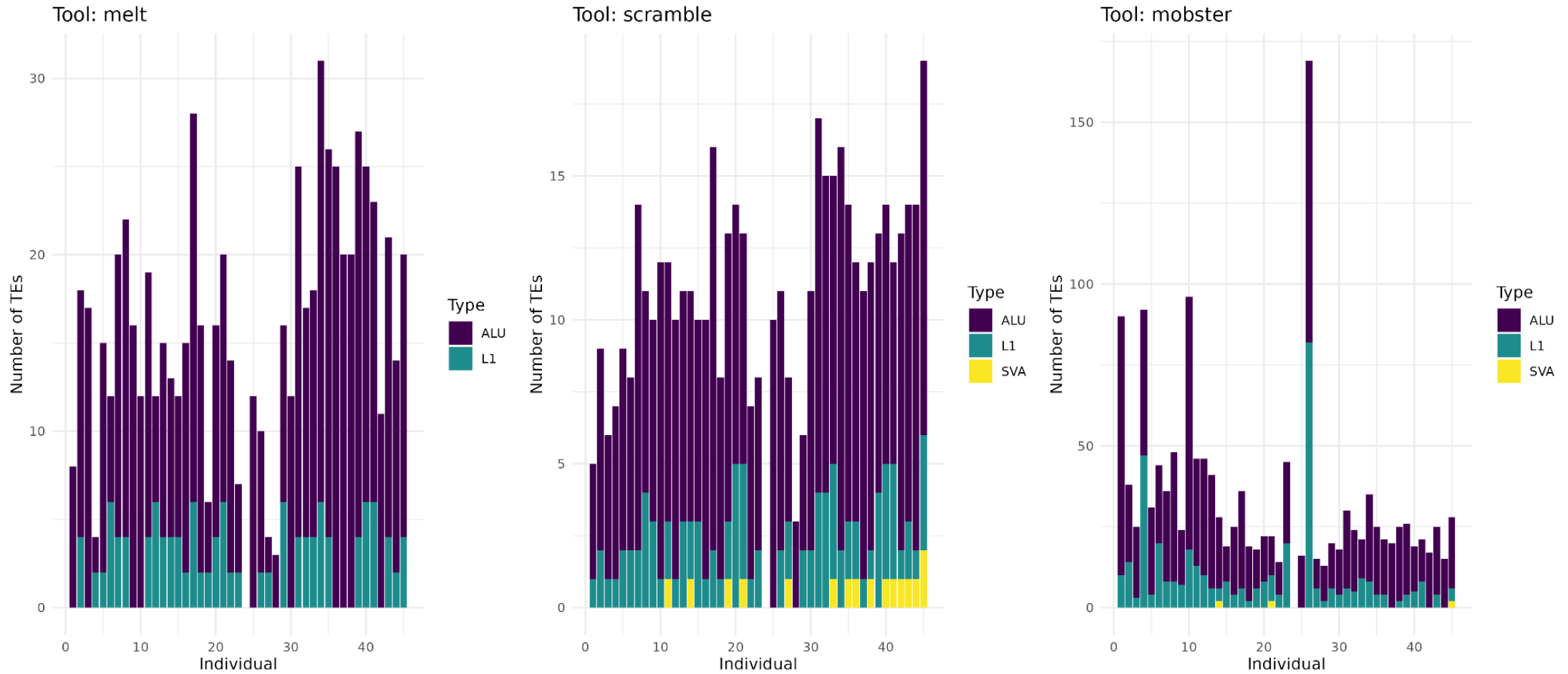
Tool	MELT	SCRAMble	mobster	Average detection
S58/VH17i	68.7%	57.1%	62.9%	62.9%
S59/VH17f	69.2%	63.1%	60%	64.1%
S60/VH17m	72.7%	58.8%	51.8%	61.1%
VH60/VH11s	47.3%	46.1%	40.3%	44.56%
VH61/VH11i	42.8%	55%	28.6%	42.13%
VH63/VH11f	50%	80%	47.8%	59.26%
VH71/VH16i	66.6%	60%	66.6%	64.4%
VH72/VH16m	100%	50%	50%	66.66%
VH73/VH16f	-	33.3%	100%	66.65%
VH77/VH01i	-	66.6%	100%	83.3%
VH78/VH01m	50%	40.9%	35.1%	42%
VH79/VH01f	62.5%	47.1%	38.5%	49.36%
<b>Average detection (tool)</b>	62.98%	54.83%	56.8%	

**Table S8. *De novo* insertions detected by each tool**

Family	MELT	SCRAMble	mobster	Total
VH01	47	303	34	384
VH016	456	255	38	749
VH017	430	277	39	746

## Supplementary Figures

Figure S1: Number of TEs per Individual detected by each tool



Source: CVID patients

Figure S2: Number of TEs per Individual detected by each tool

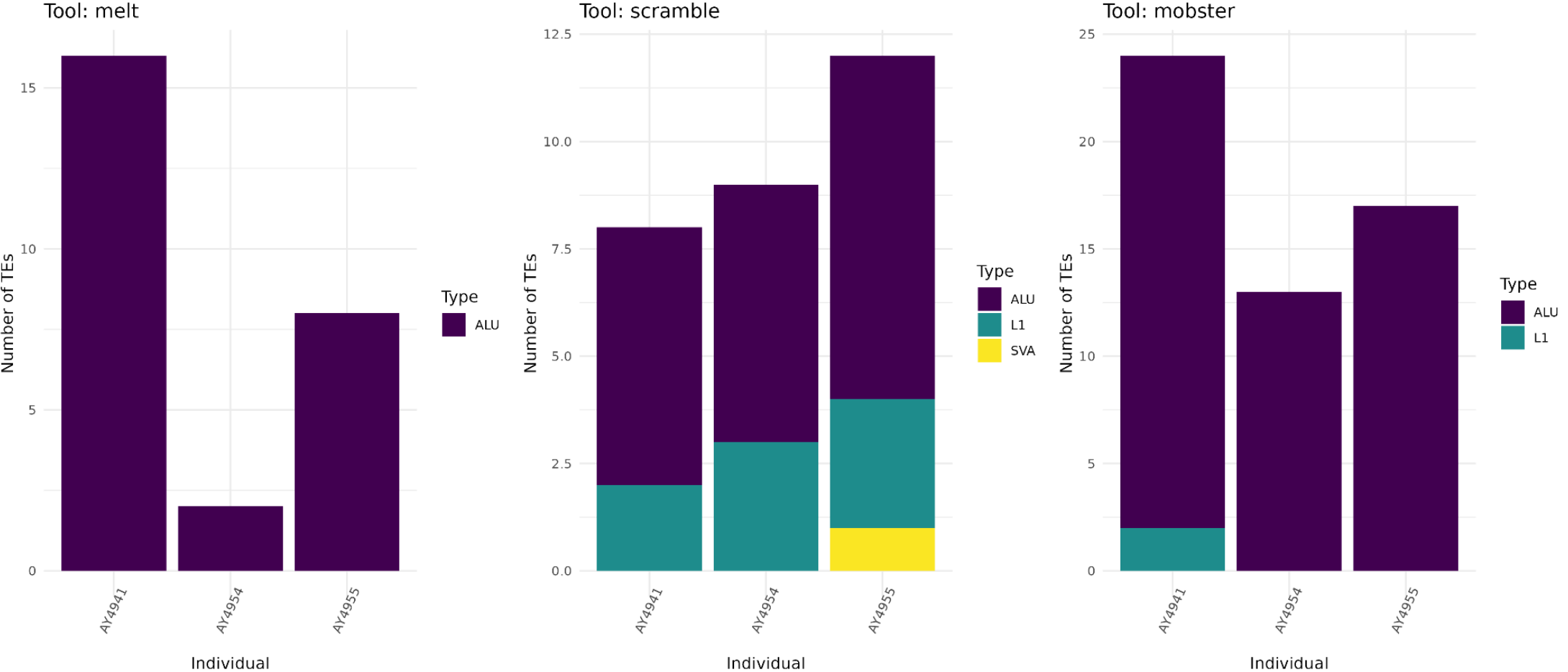


Figure S3: Number of TEs per Individual detected by each tool

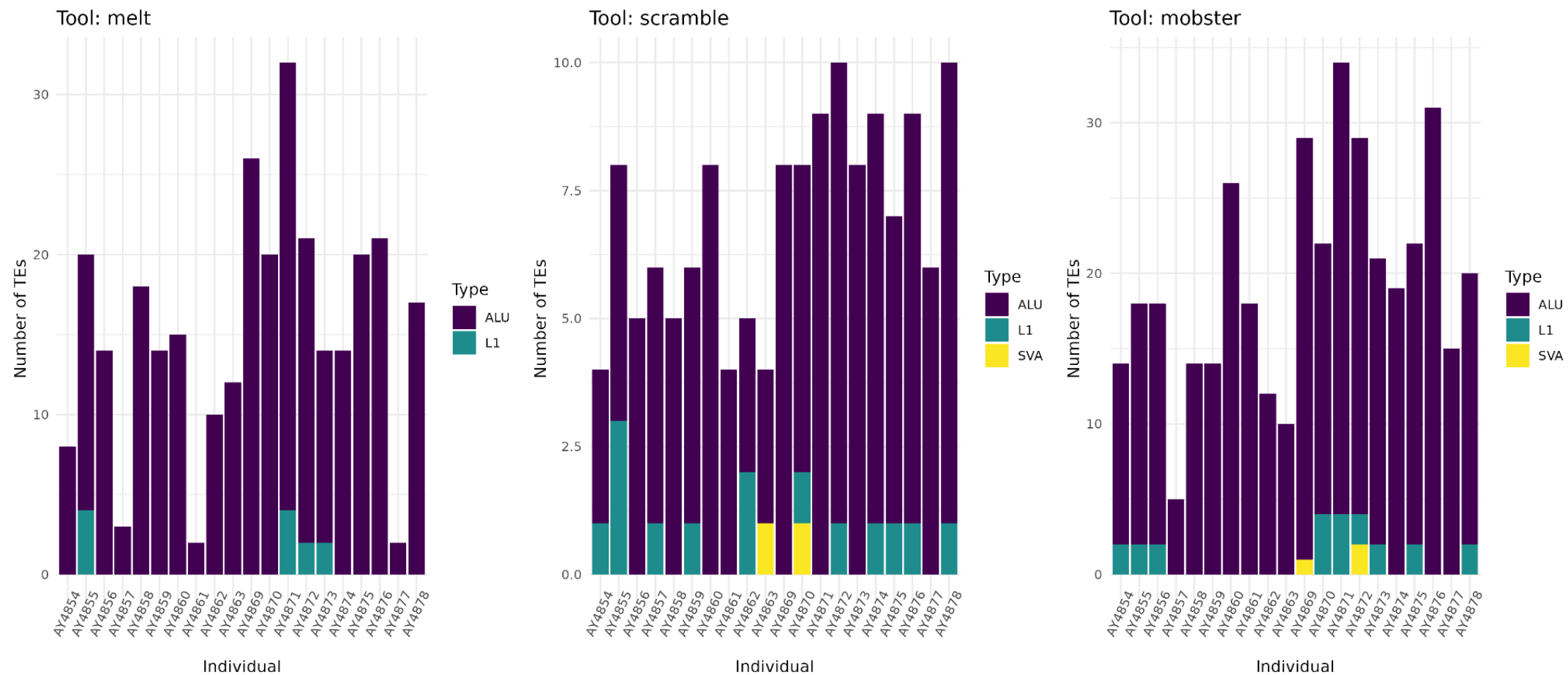


Figure S4: Number of TEs per Individual detected by each tool

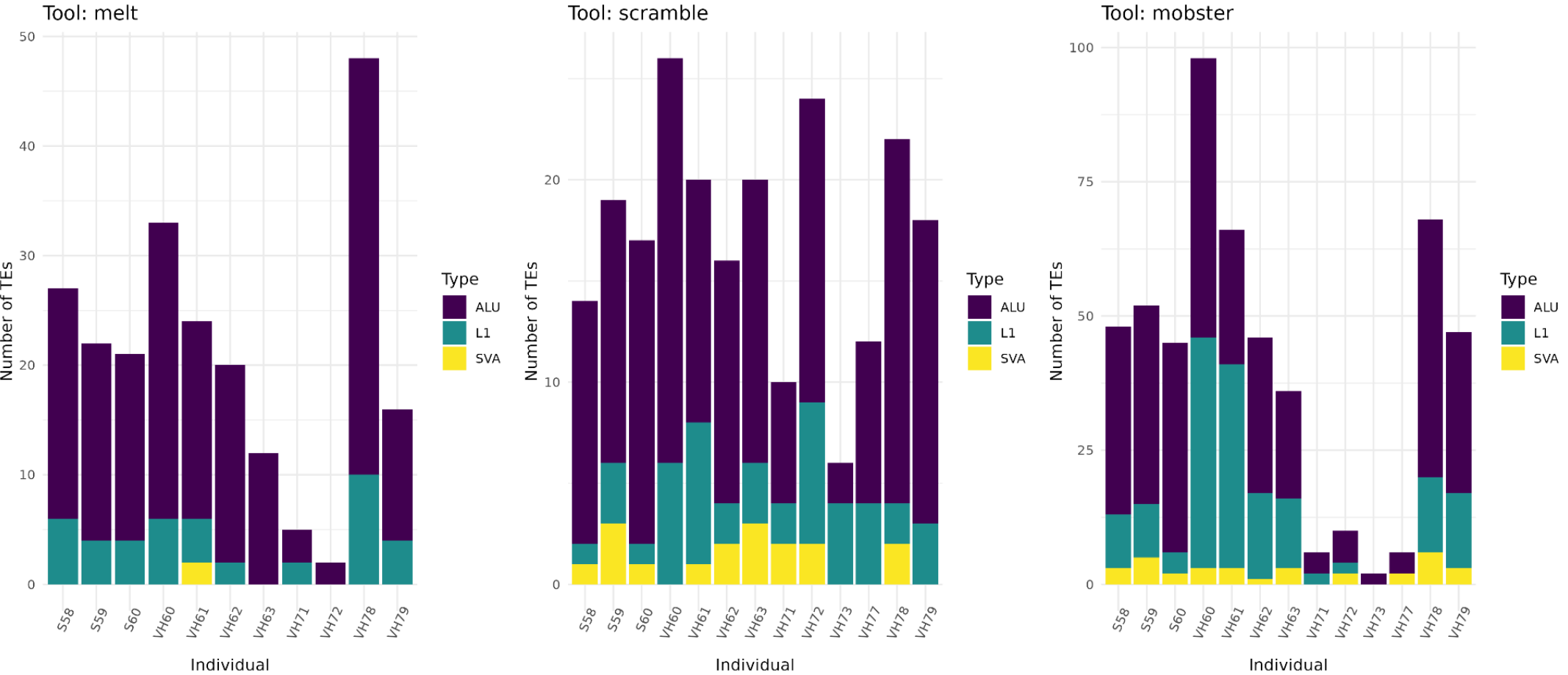
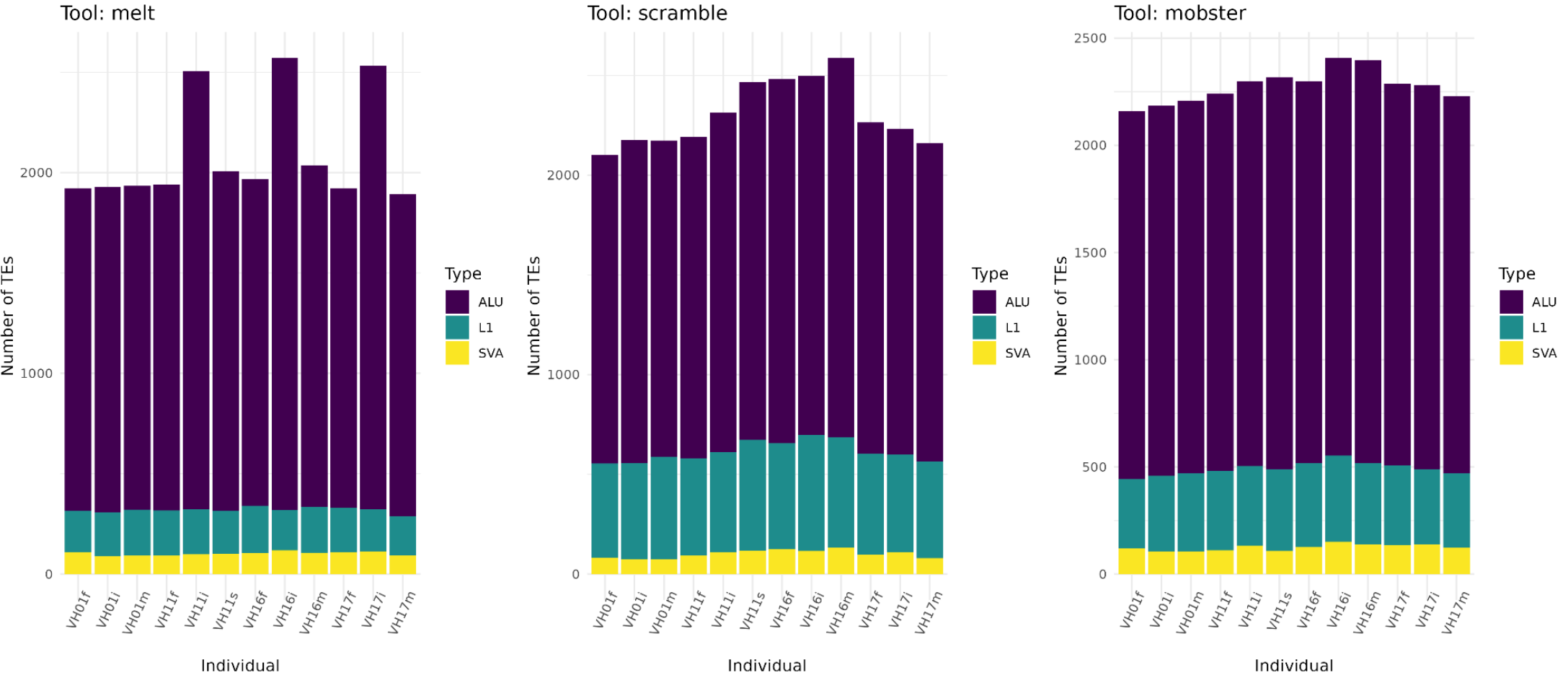


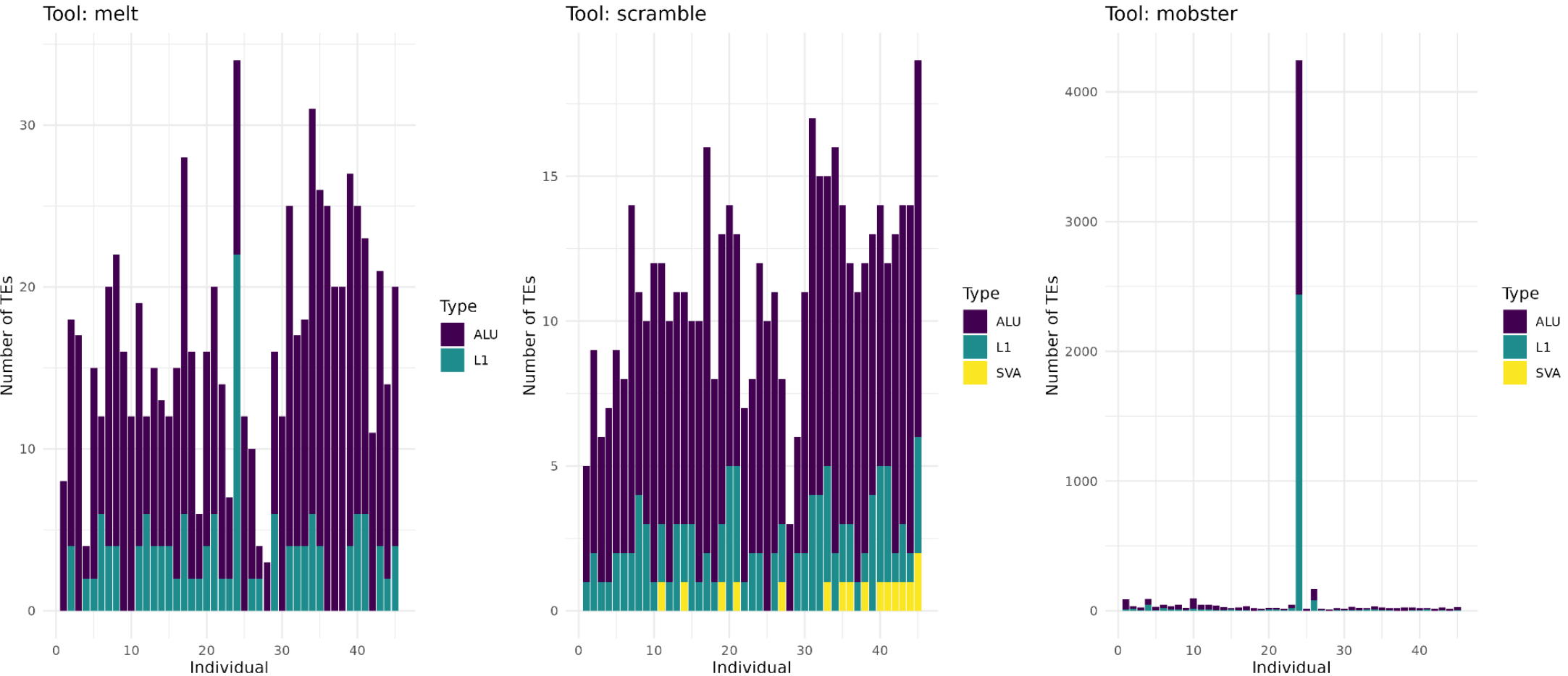
Figure S5: Number of TEs per Individual detected by each tool



Source: Vall d'Hebron WGS

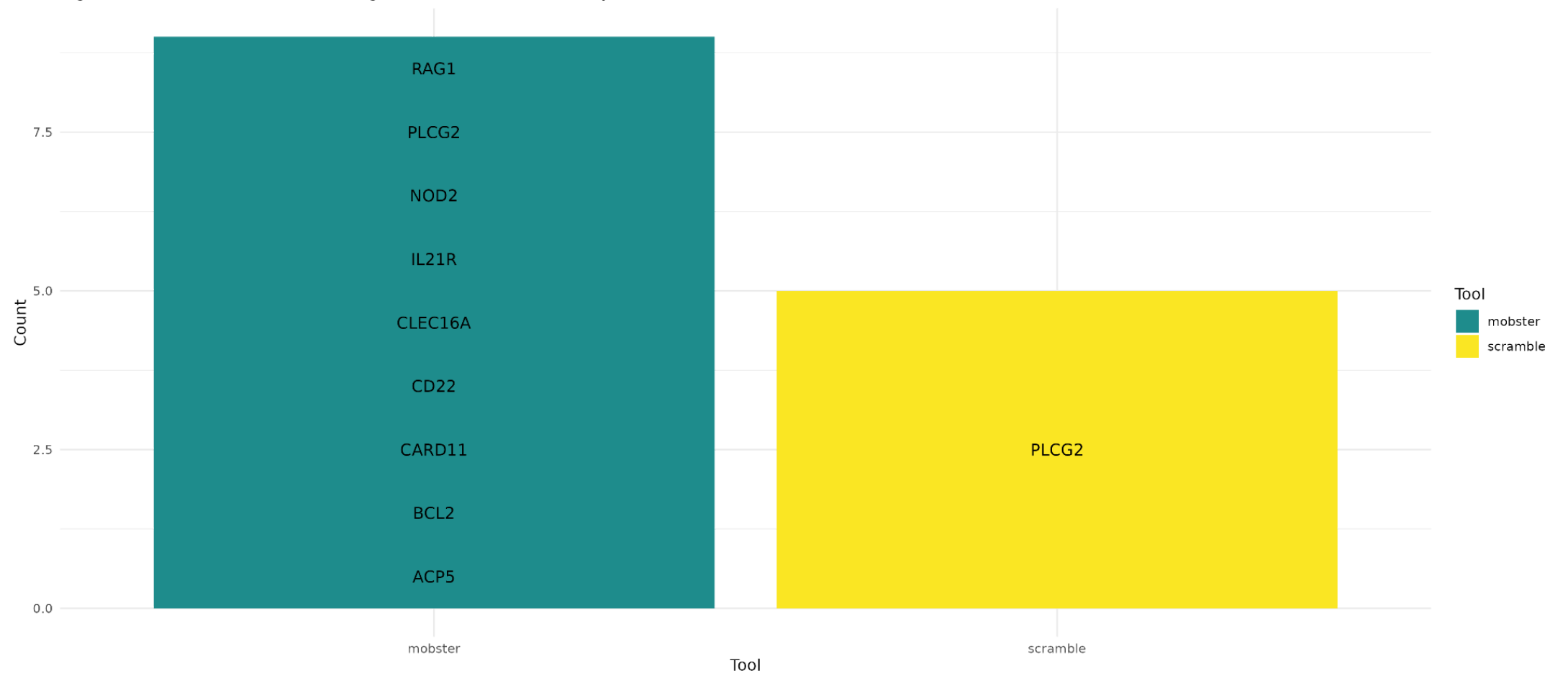


Figure S6: Number of TEs per Individual detected by each tool with S24



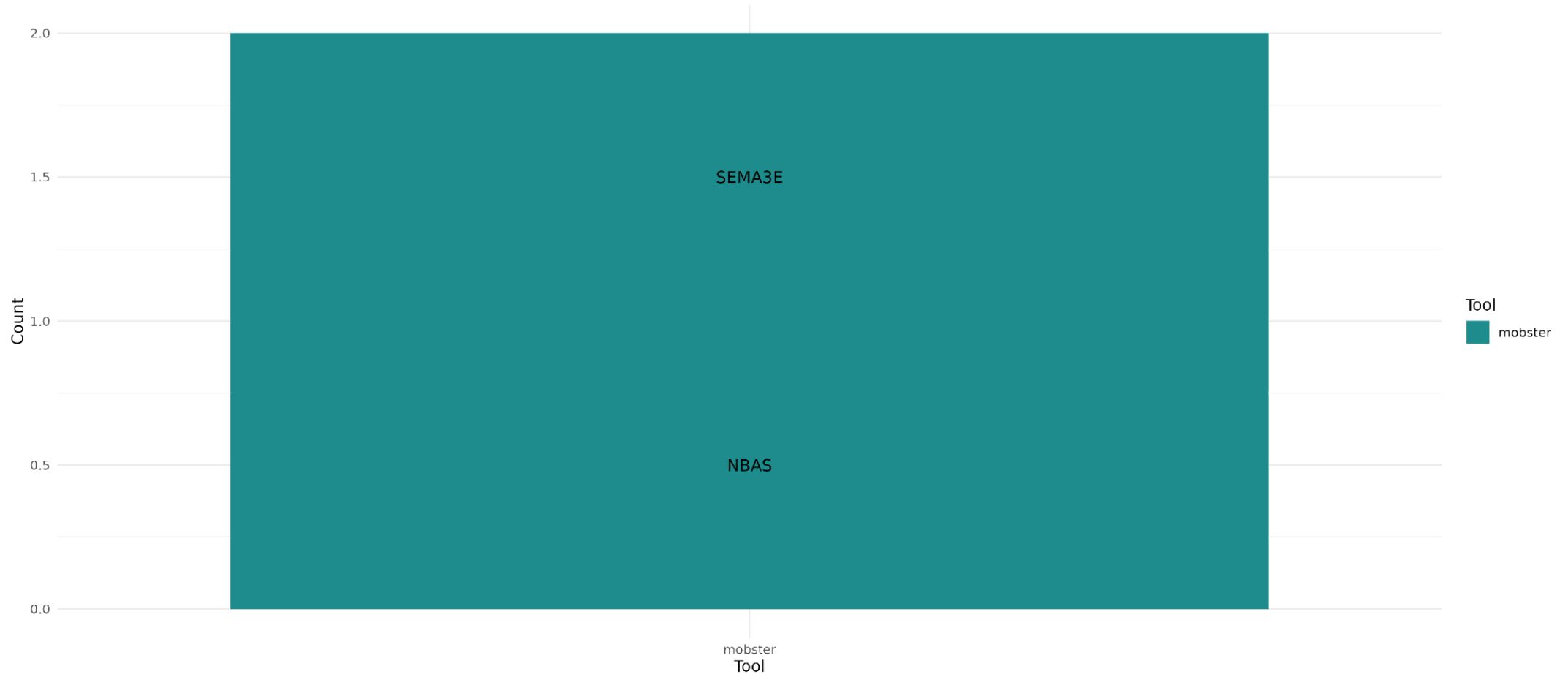
Source: CVID patients

Figure S7. Number of candidate genes (CVID) detected by each method



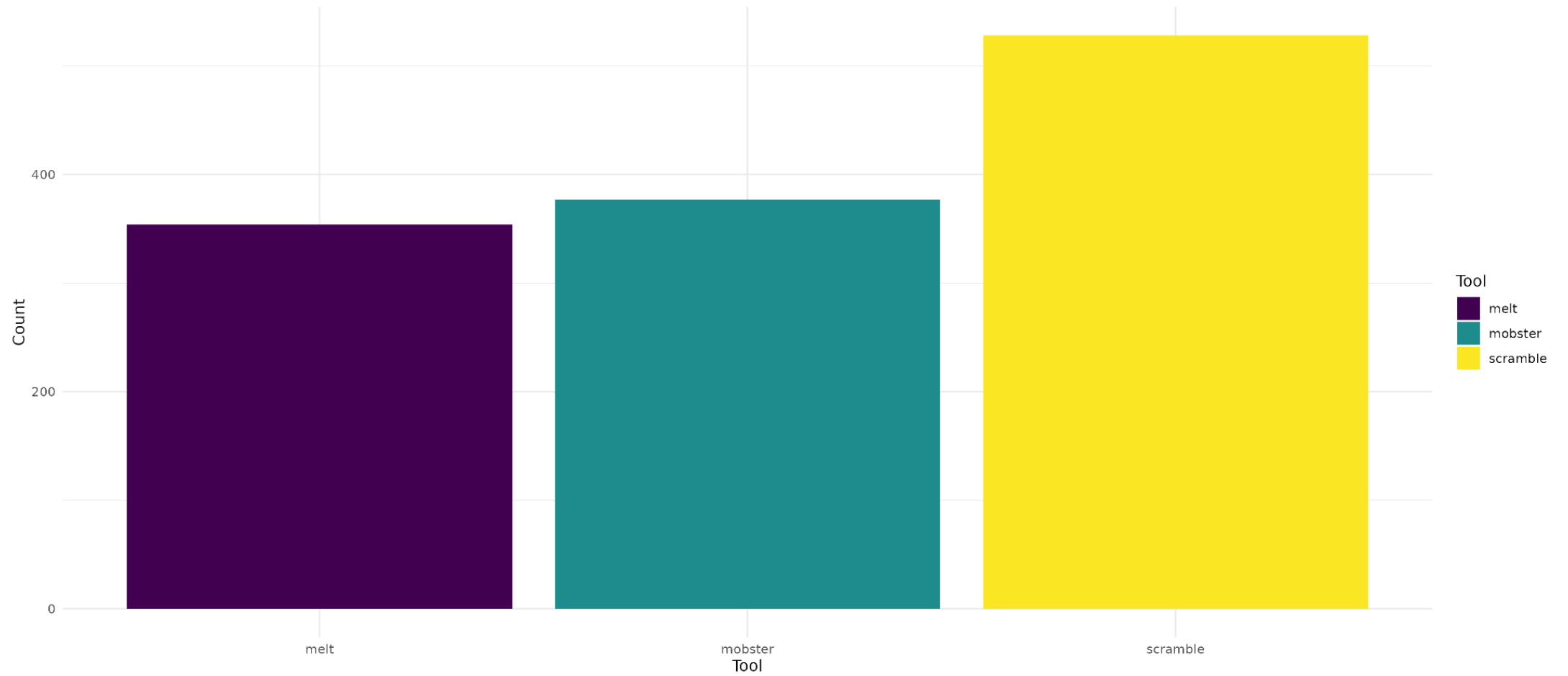
Source: CVID patients

Figure S8. Number of candidate genes (IEI) detected by each method



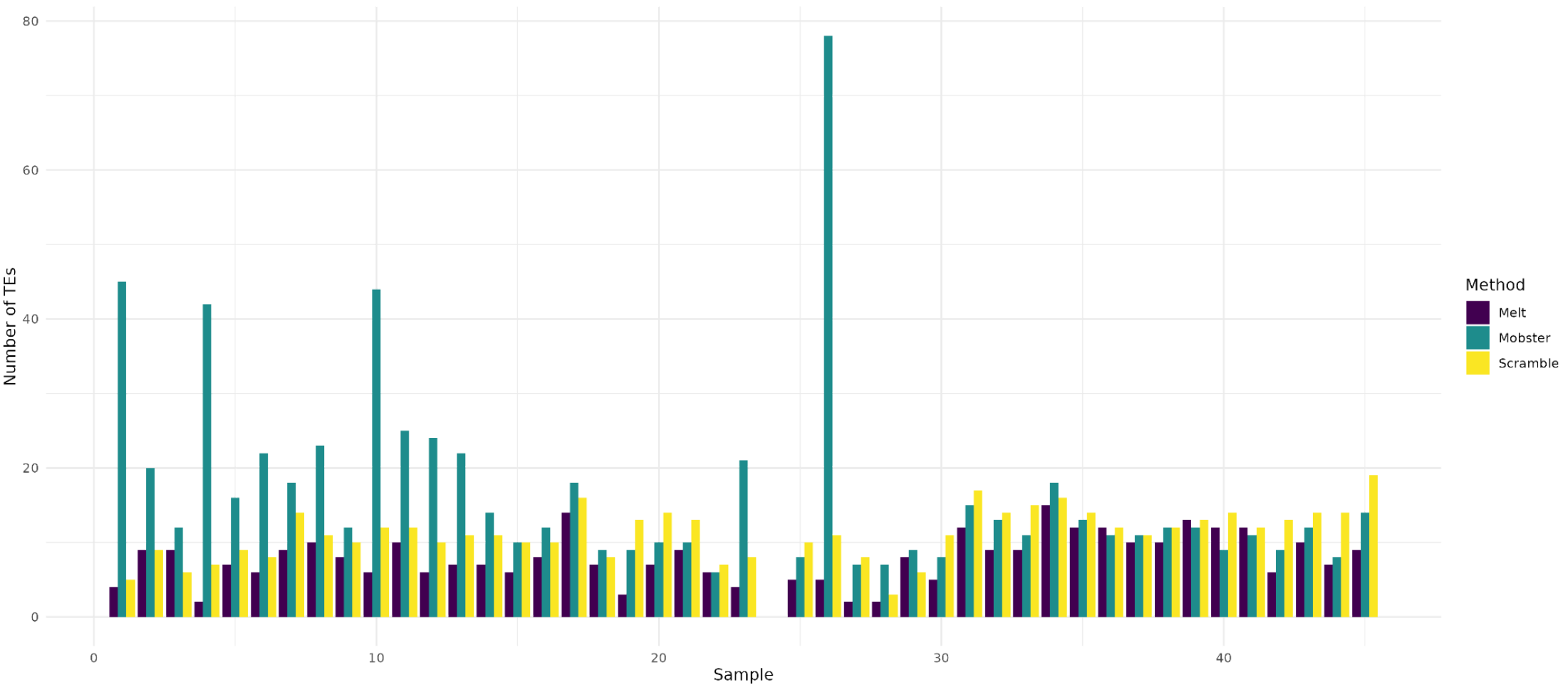
Source: Vall d'Hebron WES

Figure S9. Number of candidate genes (IEI) detected by each method



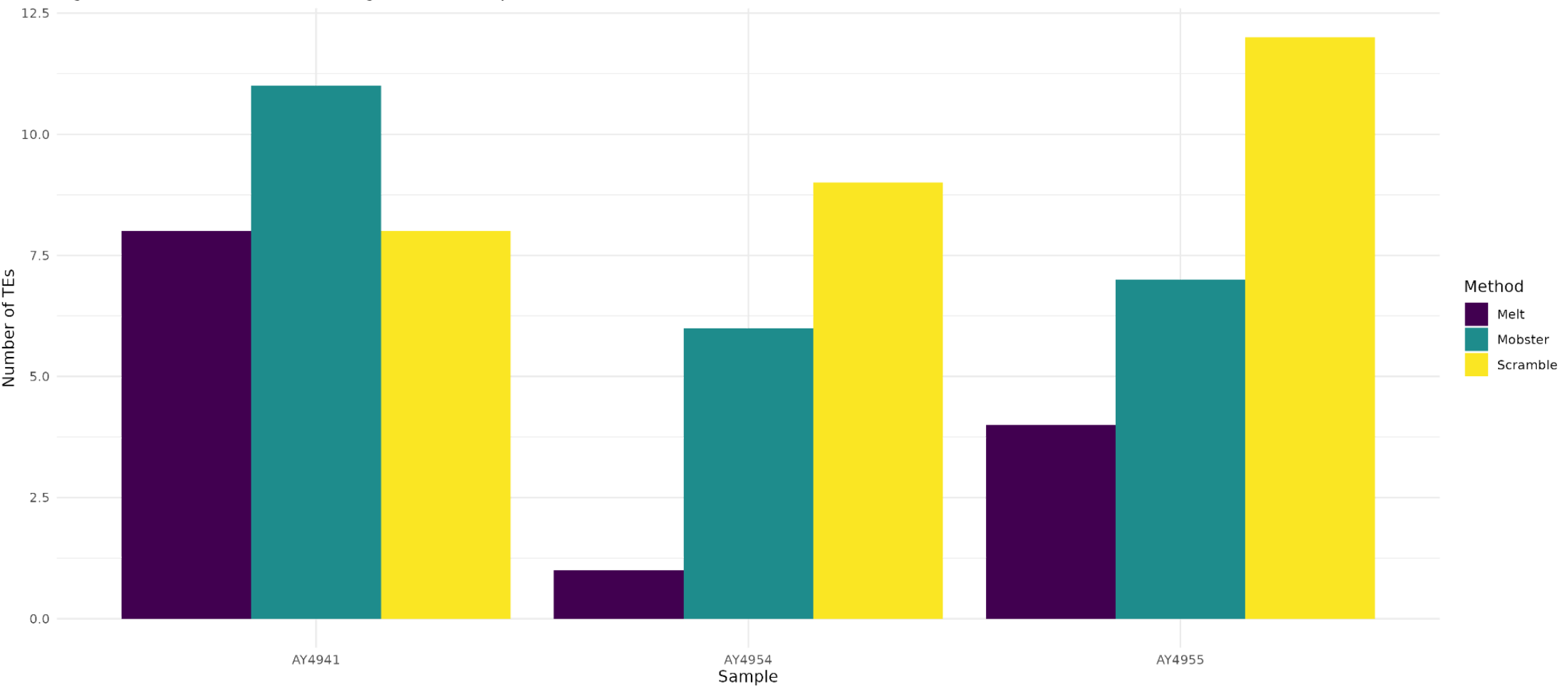
Source: Vall d'Hebron WGS

Figure S10. Number of TEs in CVID patients per tool



Source: CVID patients

Figure S11. Number of TEs in Argentinian trio per tool



Source: Argentinian trio

Figure S12. Number of TEs in Mediterranean Fever patients per tool

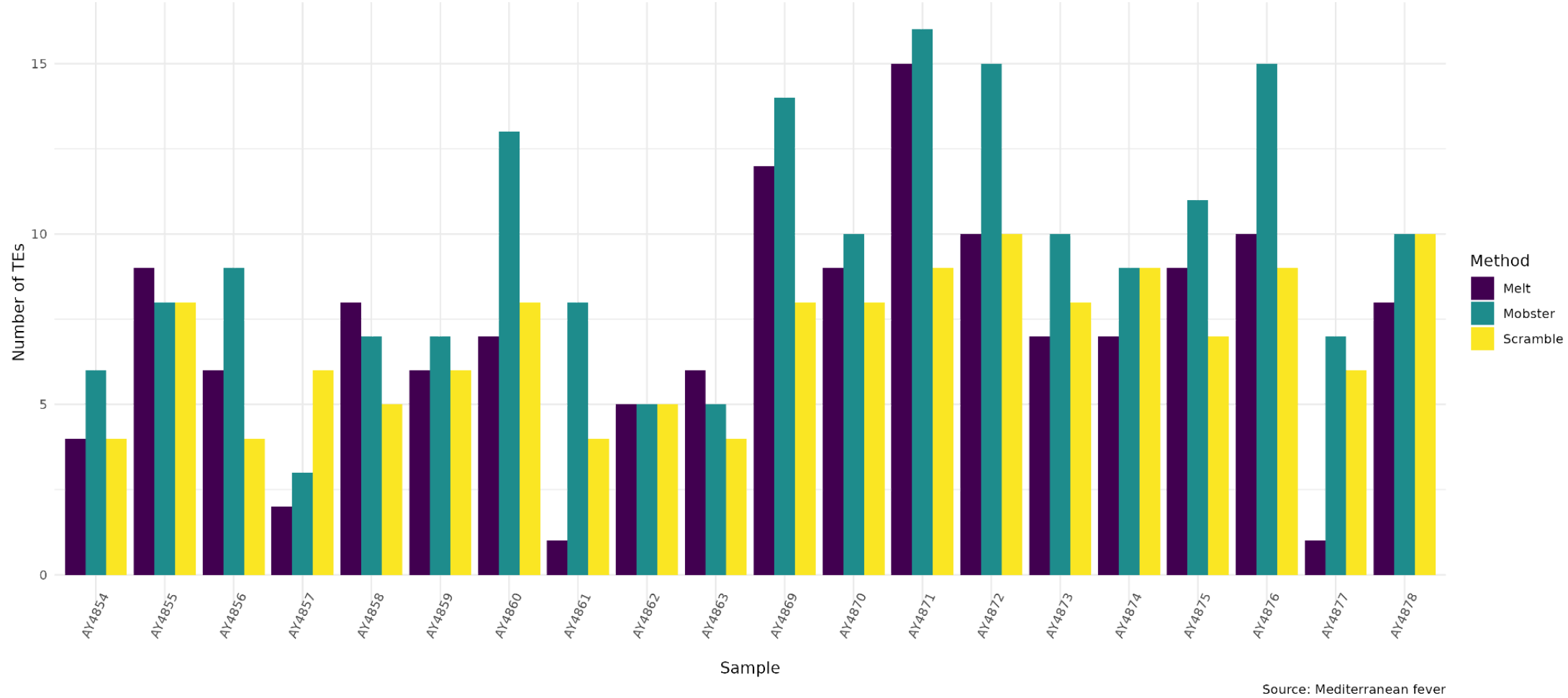


Figure S13. Number of TEs in Vall d'Hebron WES data per tool

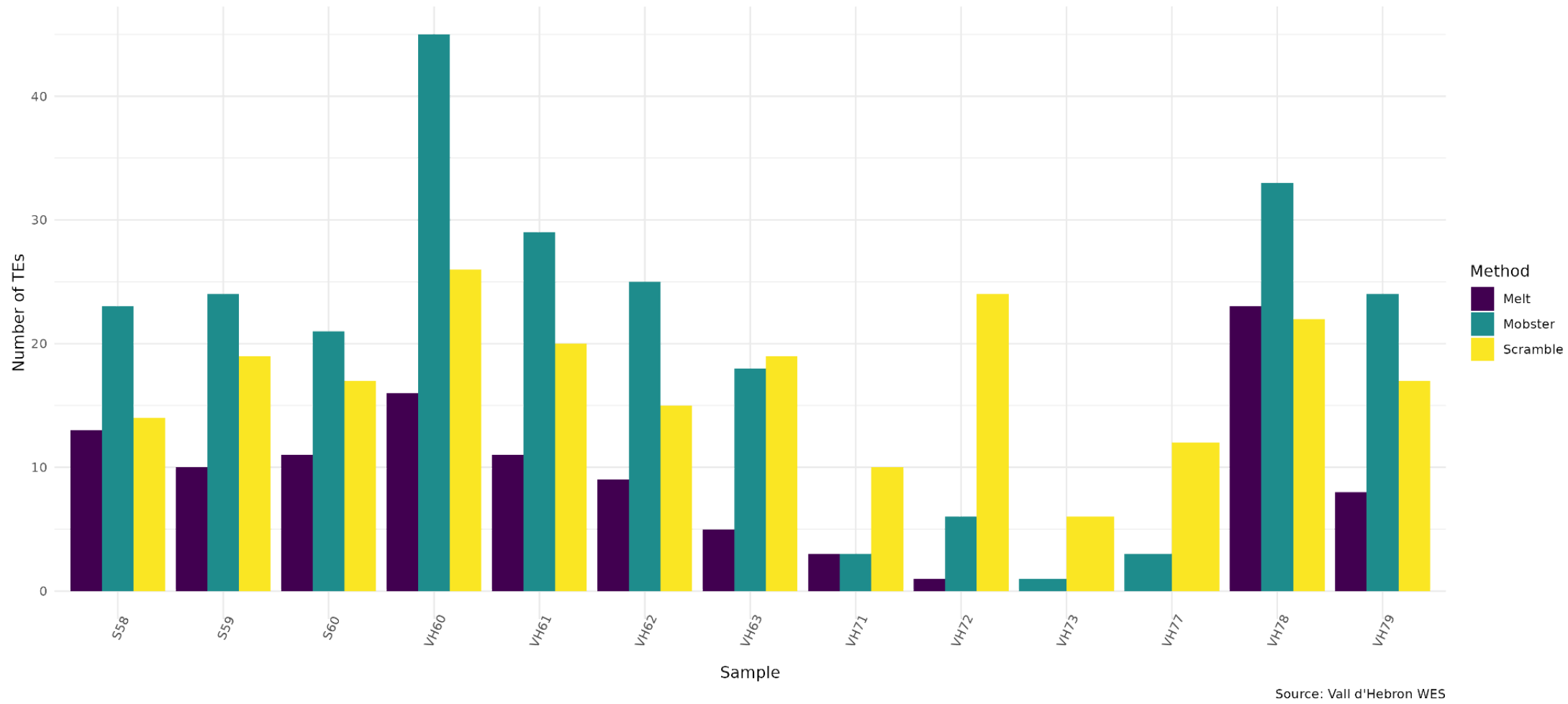
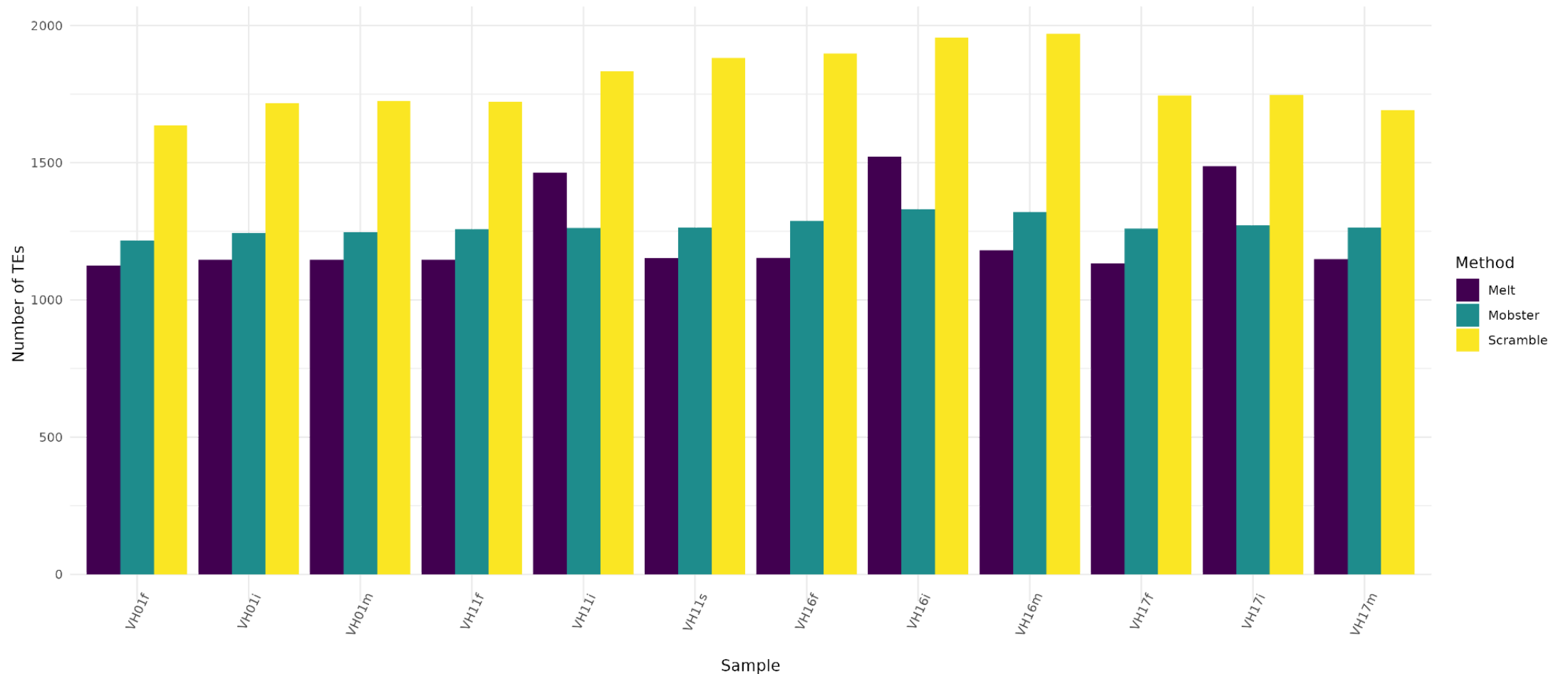


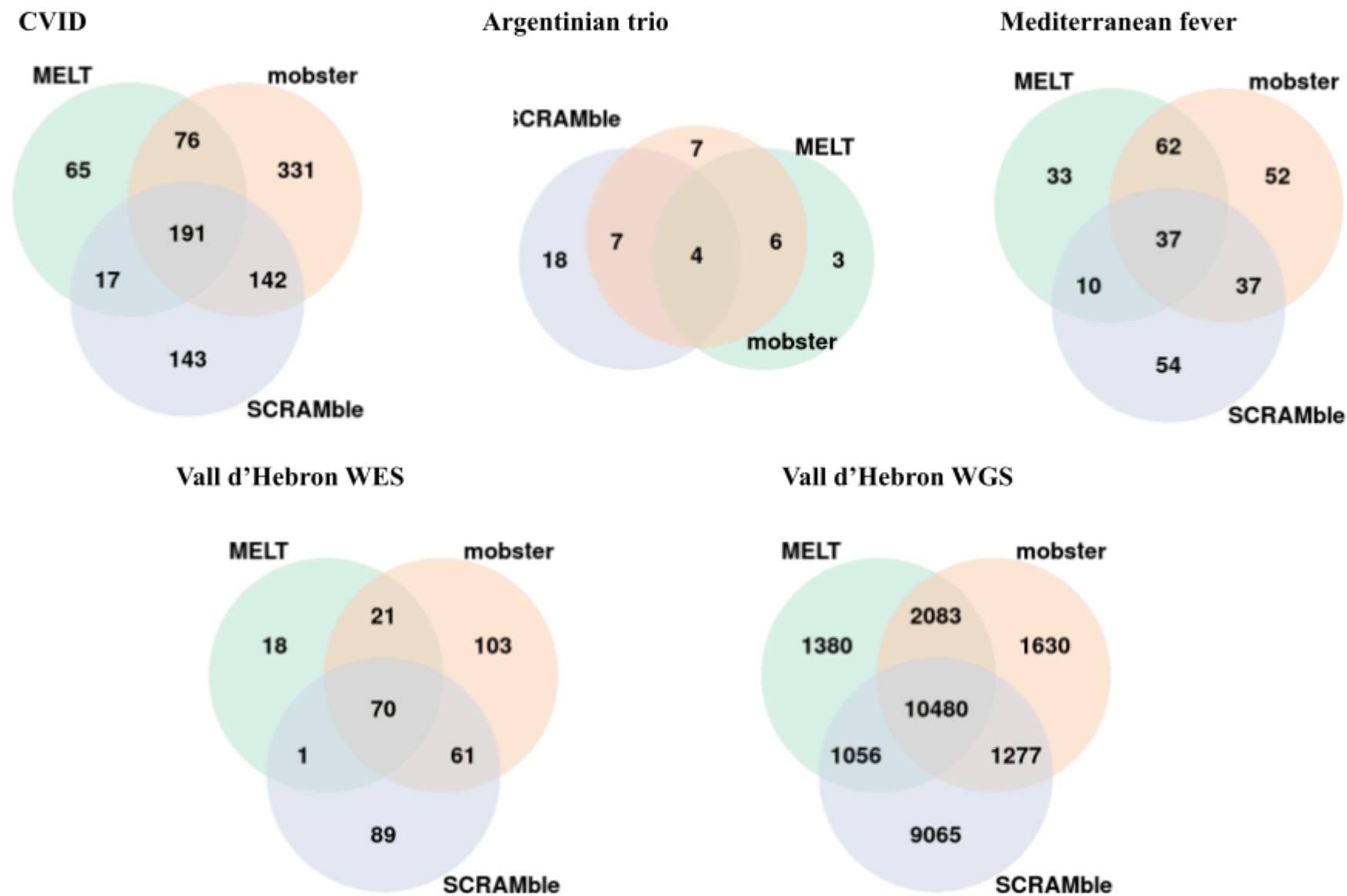


Figure S14. Number of TEs in Vall d'Hebron WGS data per tool



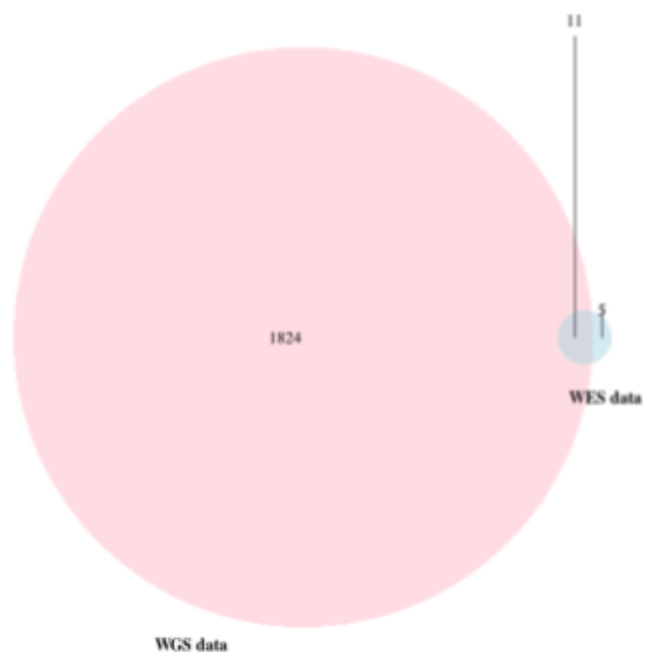
Source: Vall d'Hebron WGS

Figure S15. Venn diagrams of detected insertions per tool in each dataset.

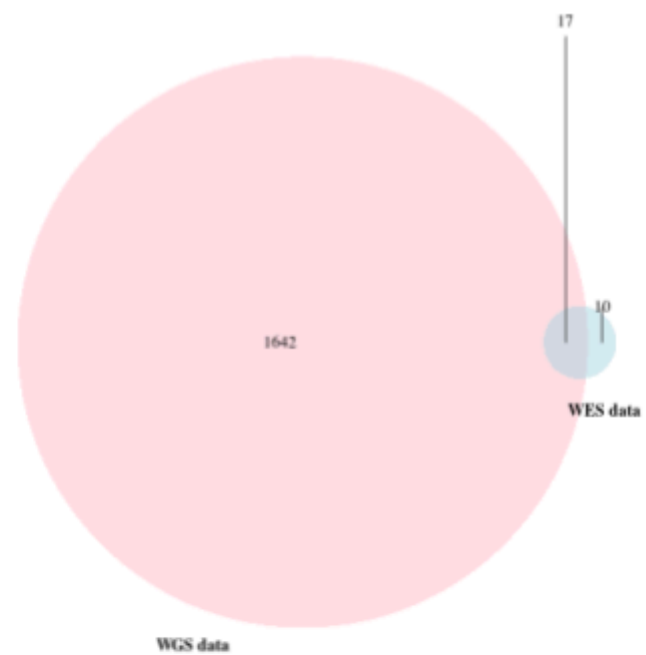


**Figure S16. Venn diagrams VH17i/S58 comparison**

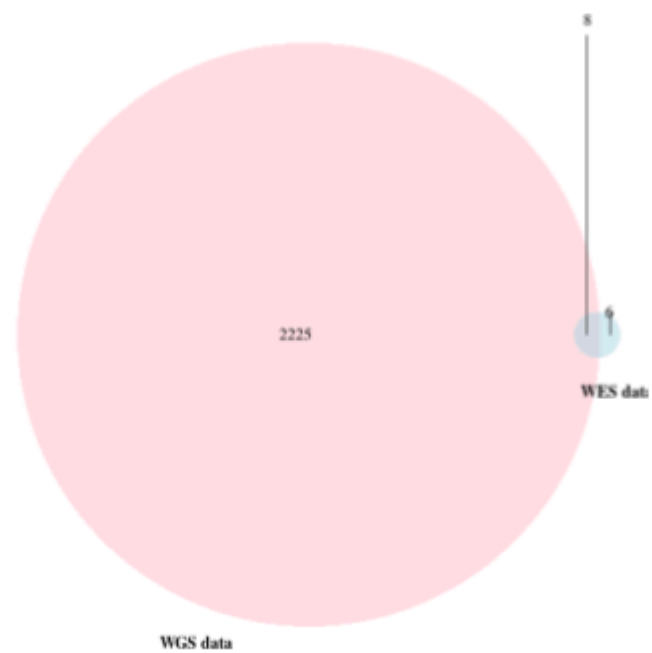
**MELT**



**mobster**

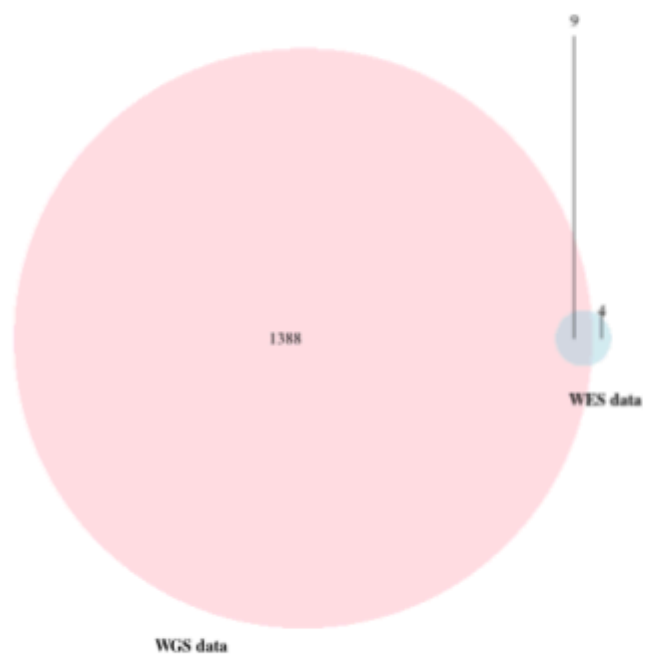


**SCRAMble**

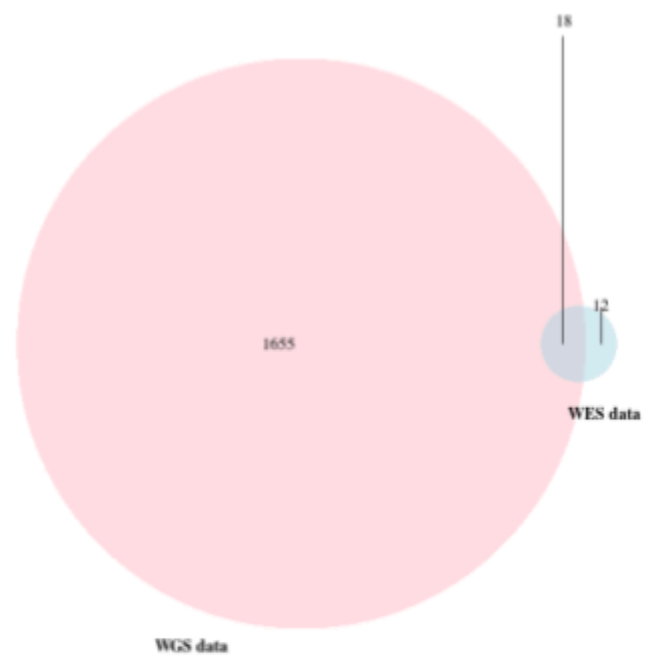


**Figure S17. Venn diagrams VH17f/S59 comparison**

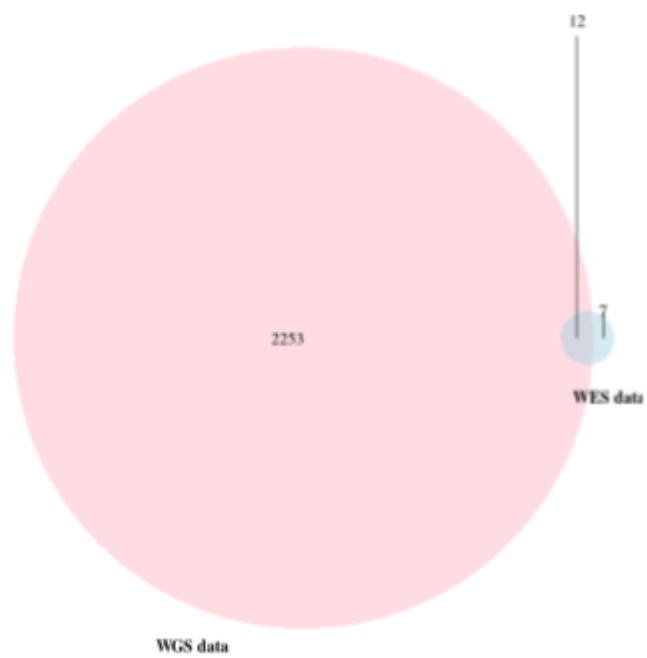
**MELT**



**mobster**

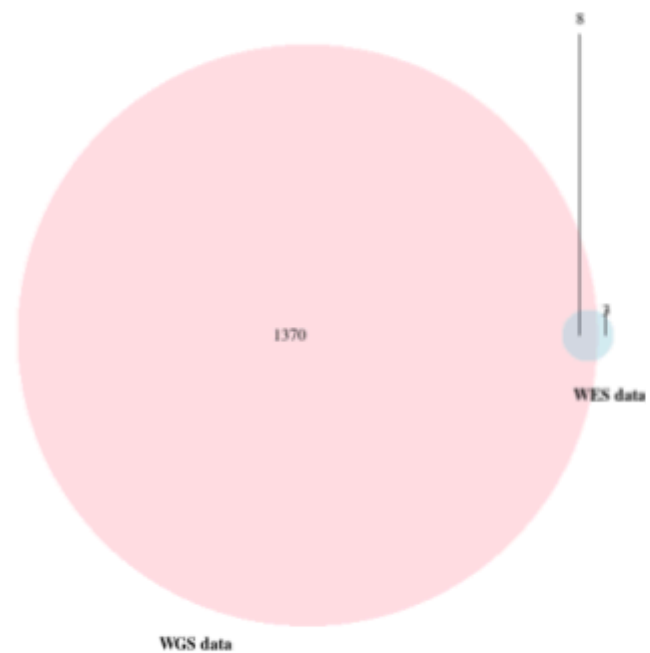


**SCRAMble**

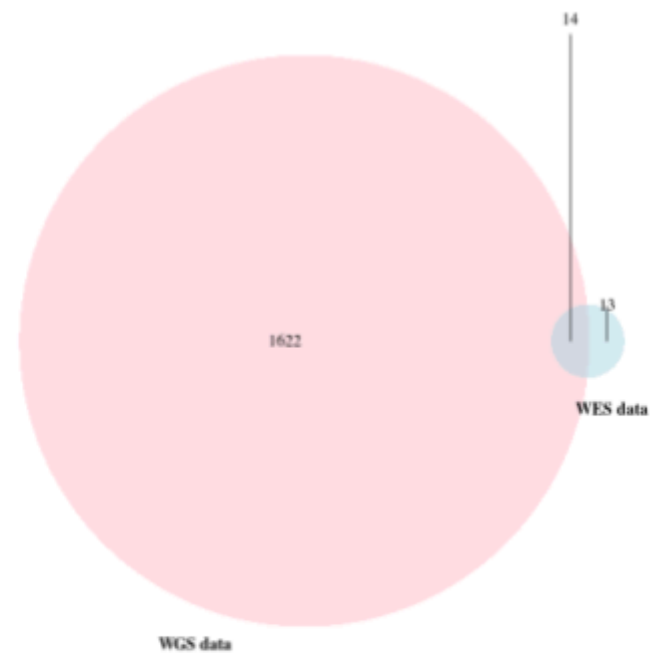


**Figure S18. Venn diagrams VH17m/S60 comparison**

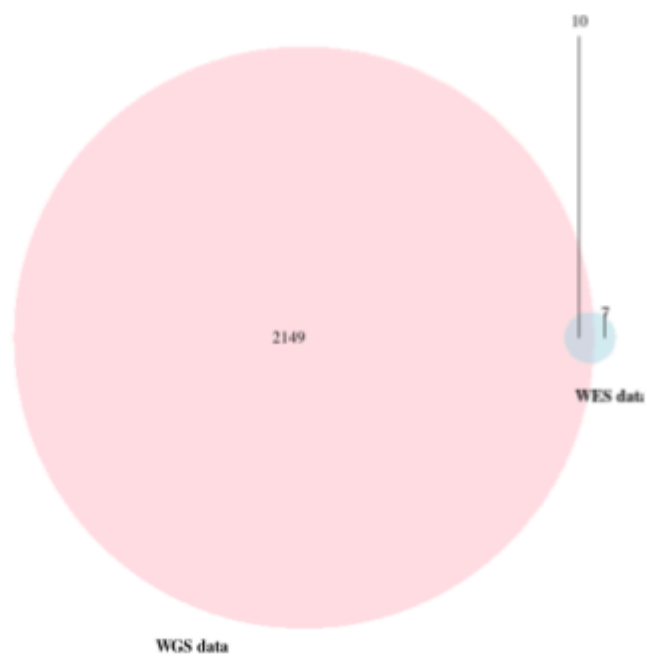
**MELT**



**mobster**

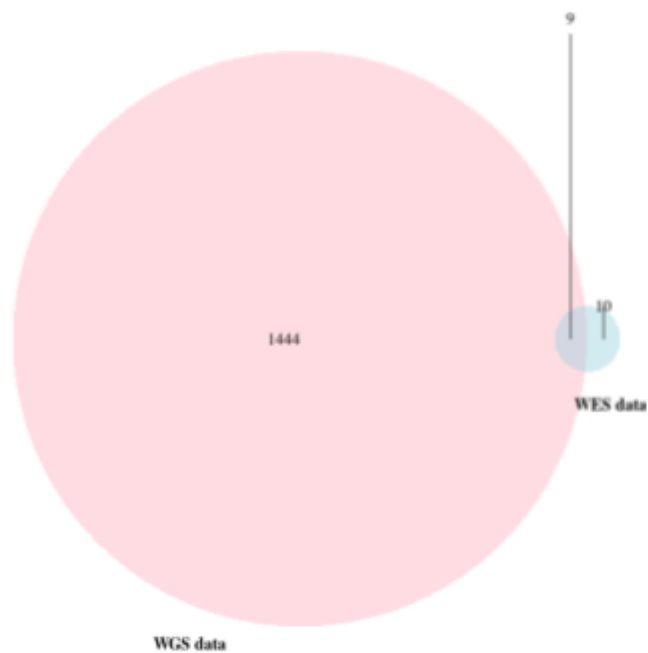


**SCRAMble**



**Figure S19. Venn diagrams VH11s/VH60 comparison**

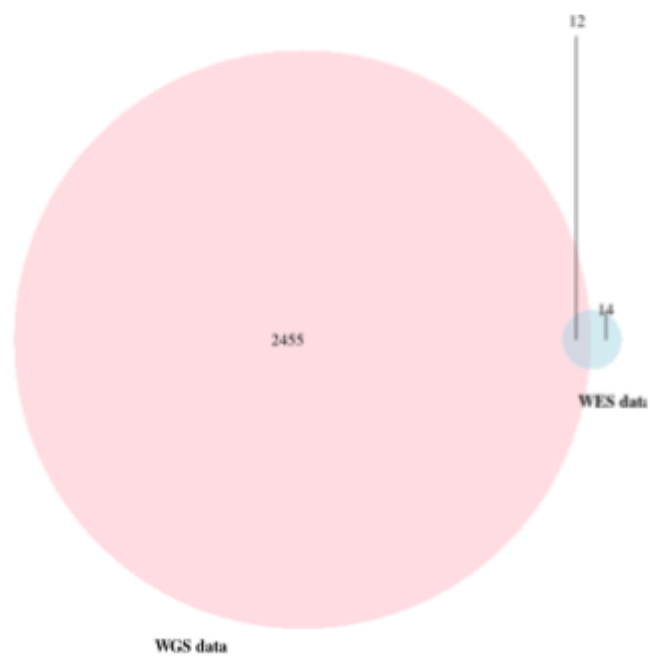
**MELT**



**mobster**

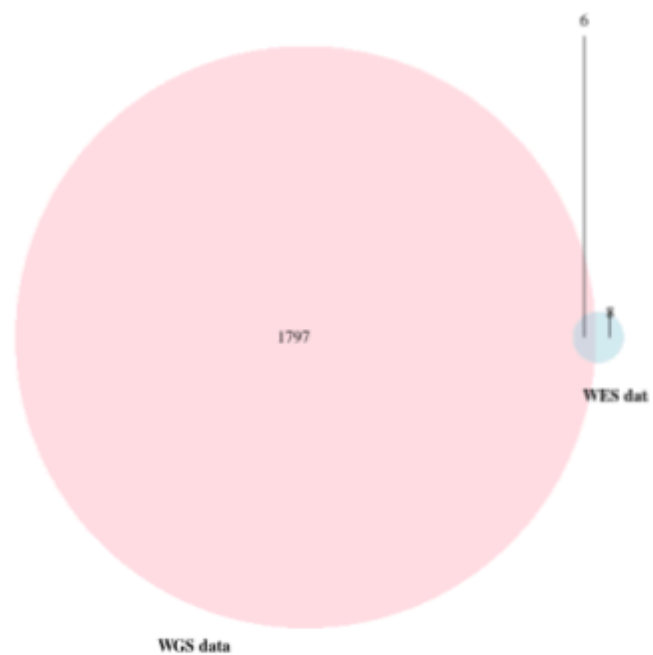


**SCRAMble**

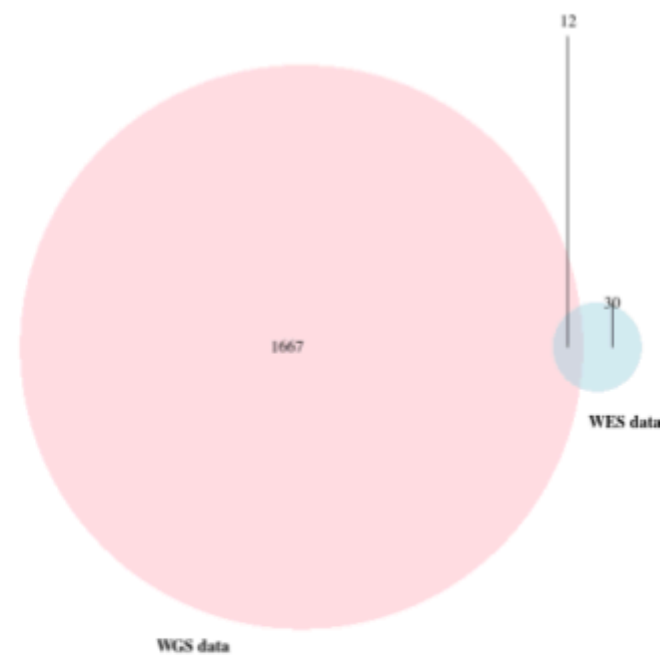


**Figure S20. Venn diagrams VH11i/VH61 comparison**

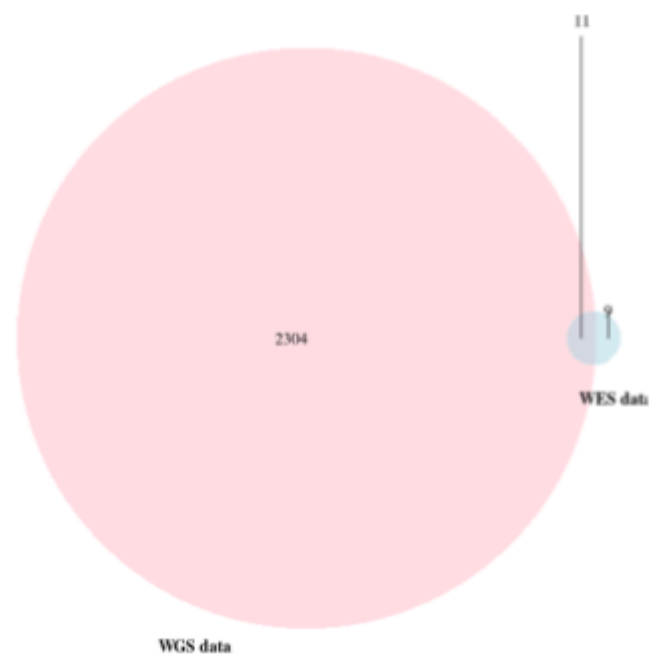
**MELT**



**mobster**

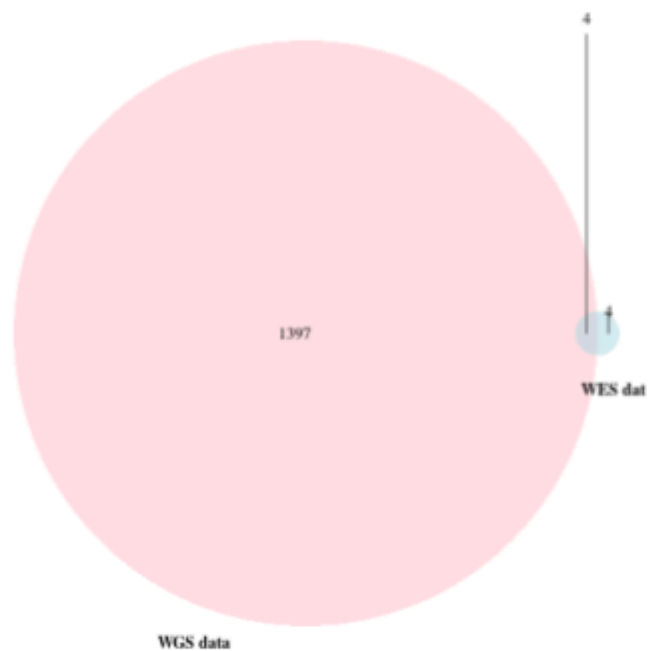


**SCRAMble**

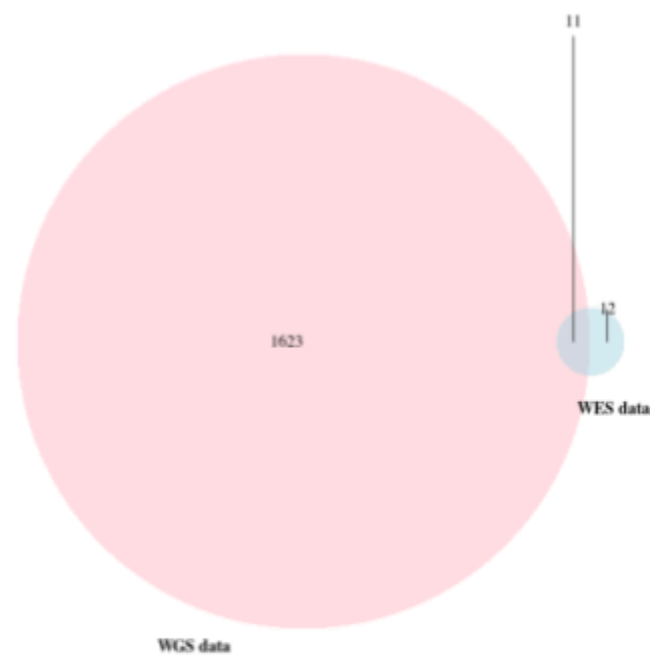


**Figure S21. Venn diagrams VH11f/VH63 comparison**

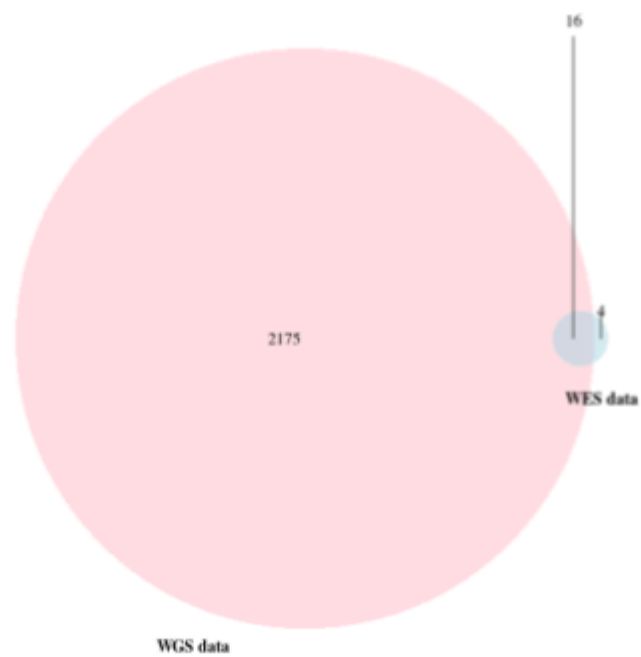
**MELT**



**mobster**



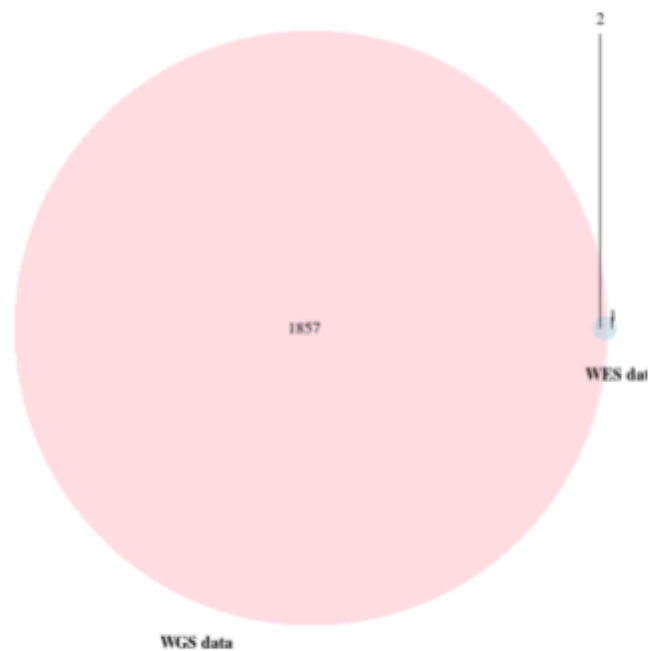
**SCRAMble**



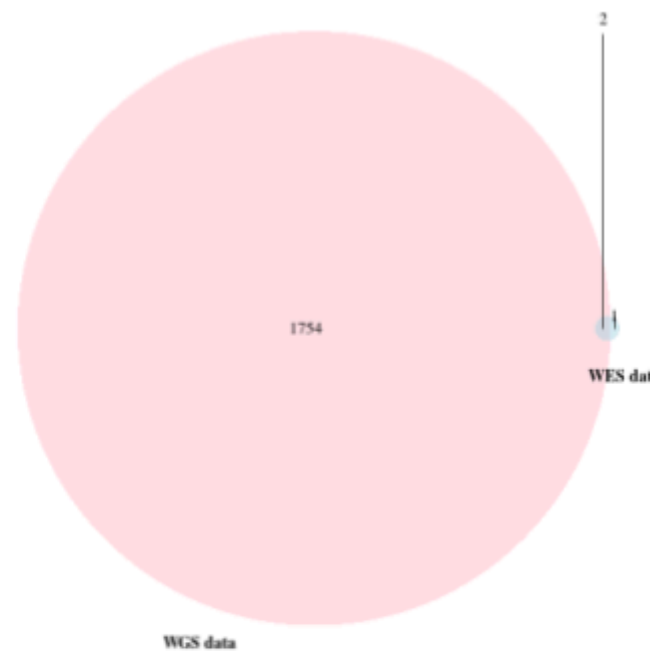


**Figure S22. Venn diagrams VH16i/VH71 comparison**

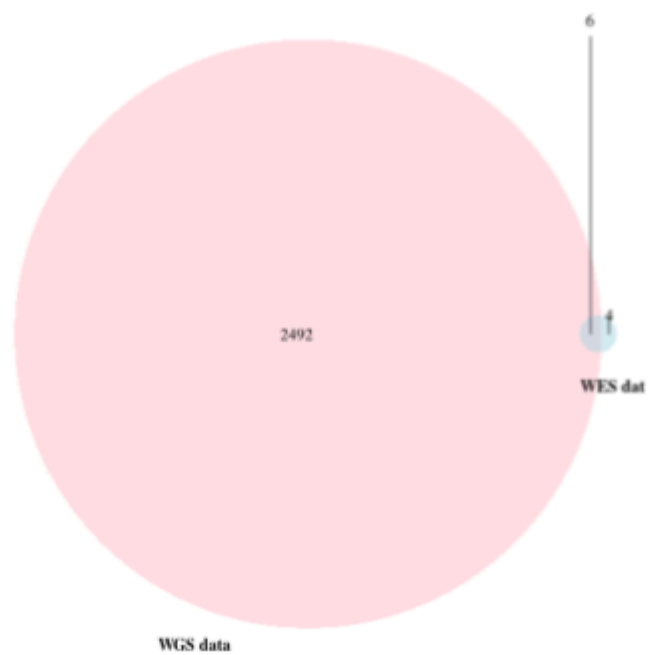
**MELT**



**mobster**

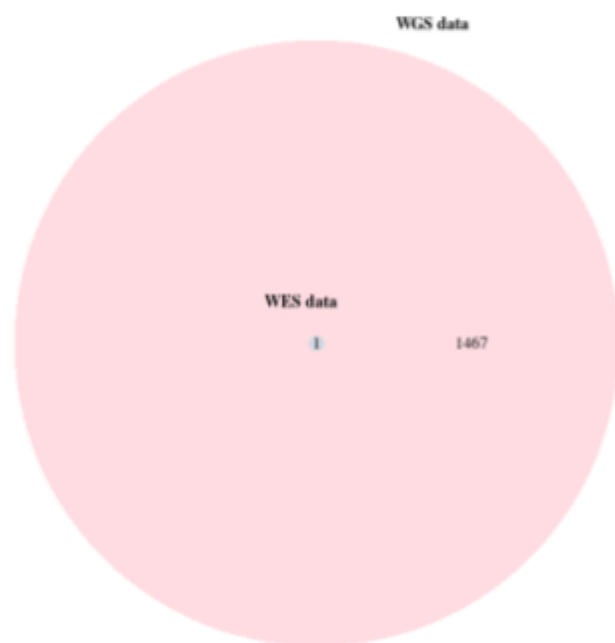


**SCRAMble**

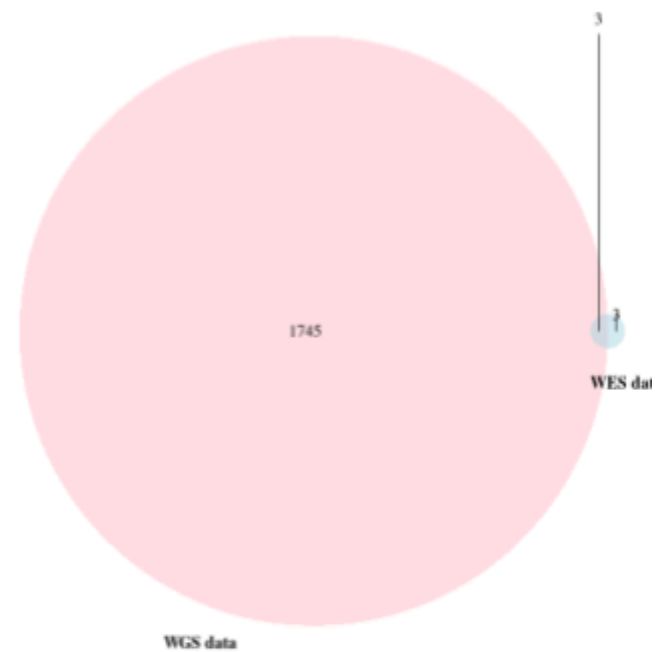


**Figure S23. Venn diagrams VH16m/VH72 comparison**

**MELT**



**mobster**



**SCRAMble**

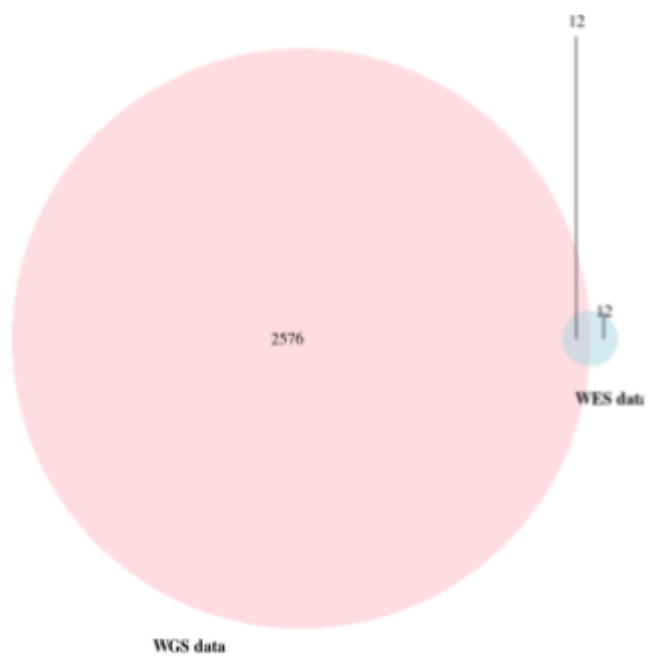
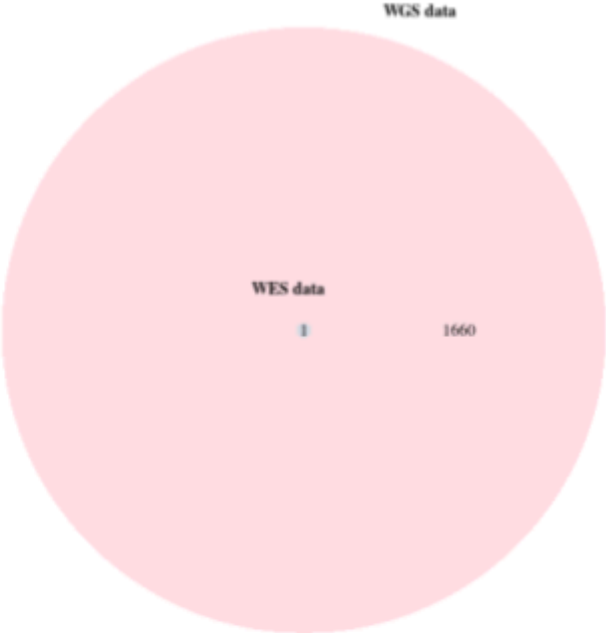


Figure S24. Venn diagrams VH16f/VH73 comparison

MELT



mobster

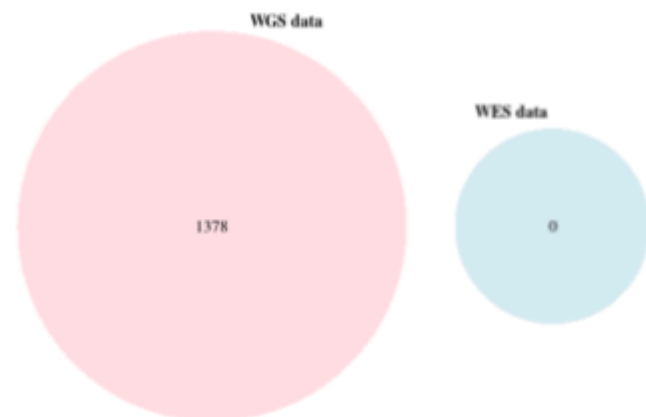


SCRAMble

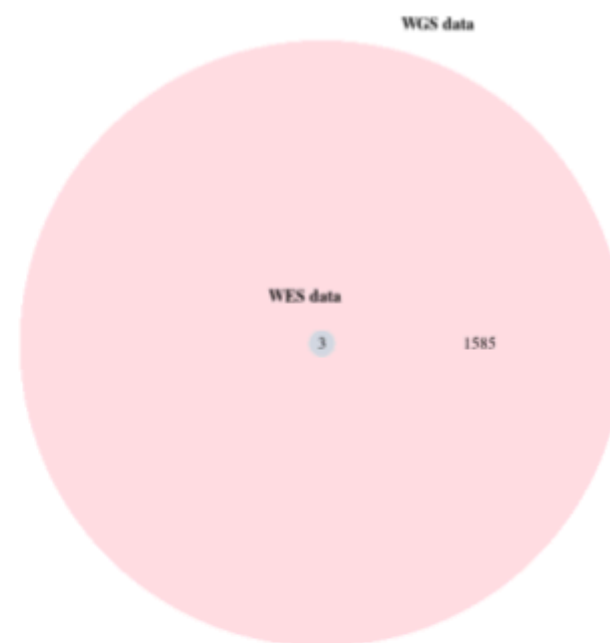


**Figure S25. Venn diagrams VH01i/VH77 comparison**

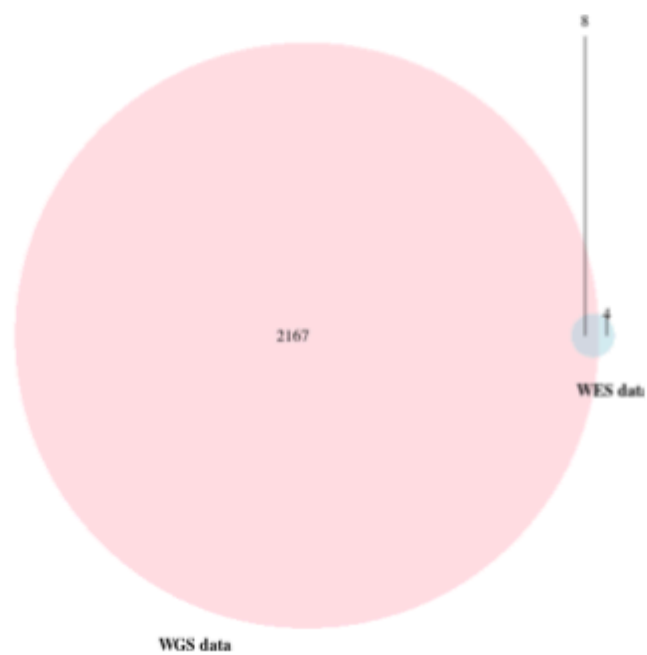
**MELT**



**mobster**

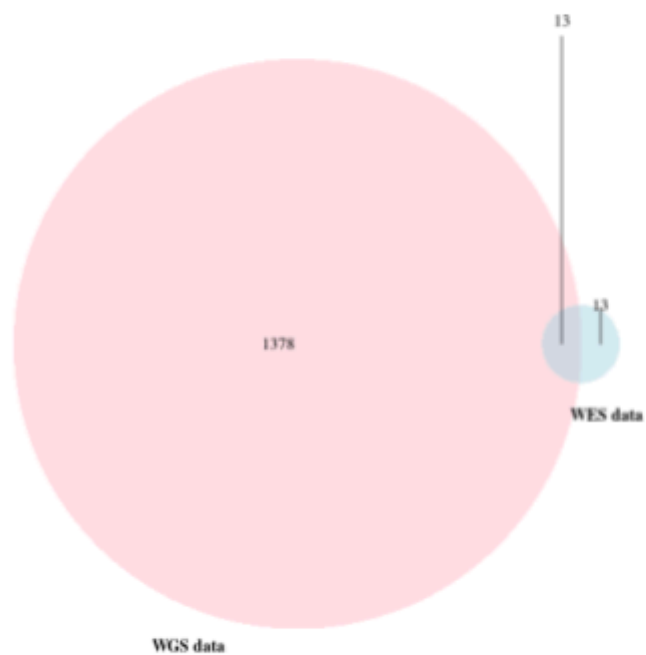


**SCRAMble**

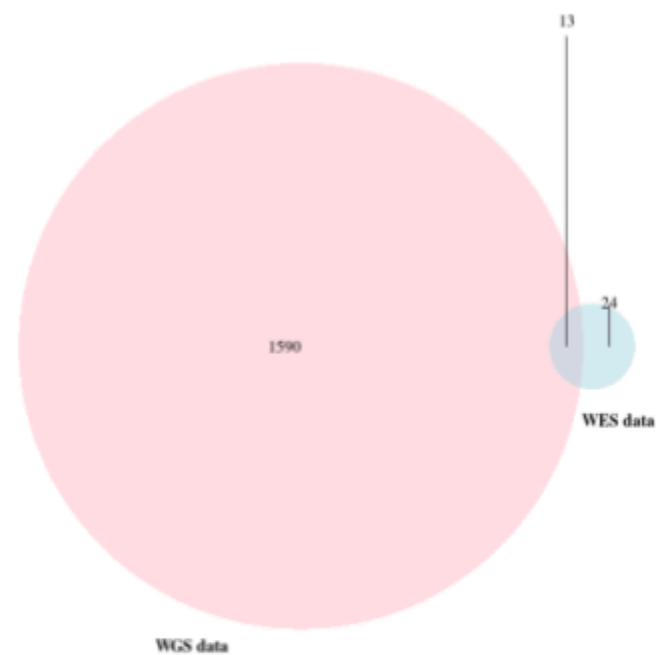


**Figure S26. Venn diagrams VH01m/VH78 comparison**

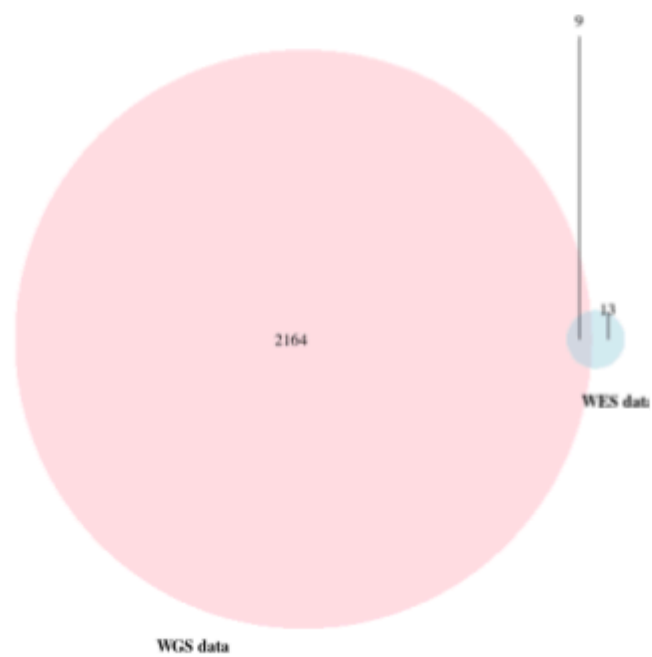
**MELT**



**mobster**

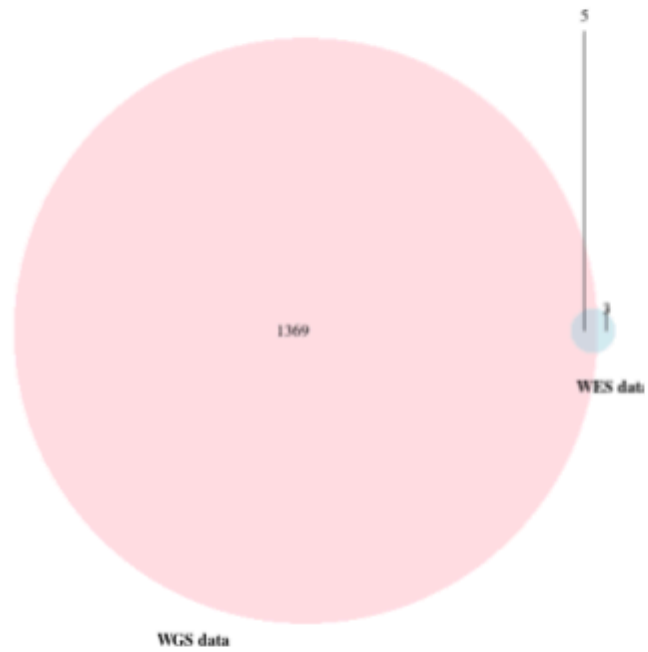


**SCRAMble**

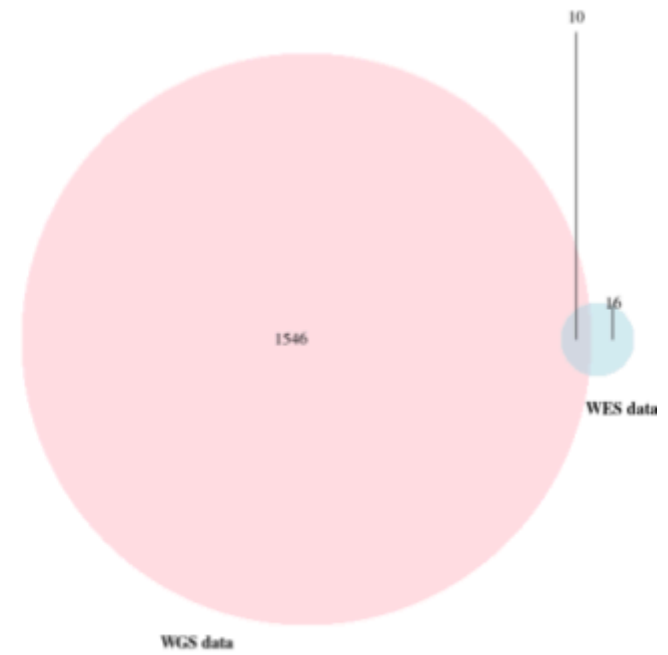


**Figure S27. Venn diagrams VH01f/VH79 comparison**

**MELT**



**mobster**



**SCRAMble**

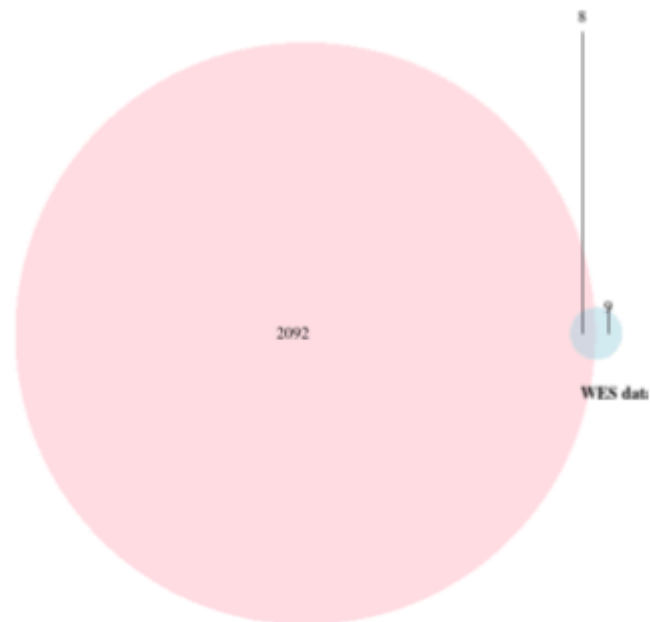
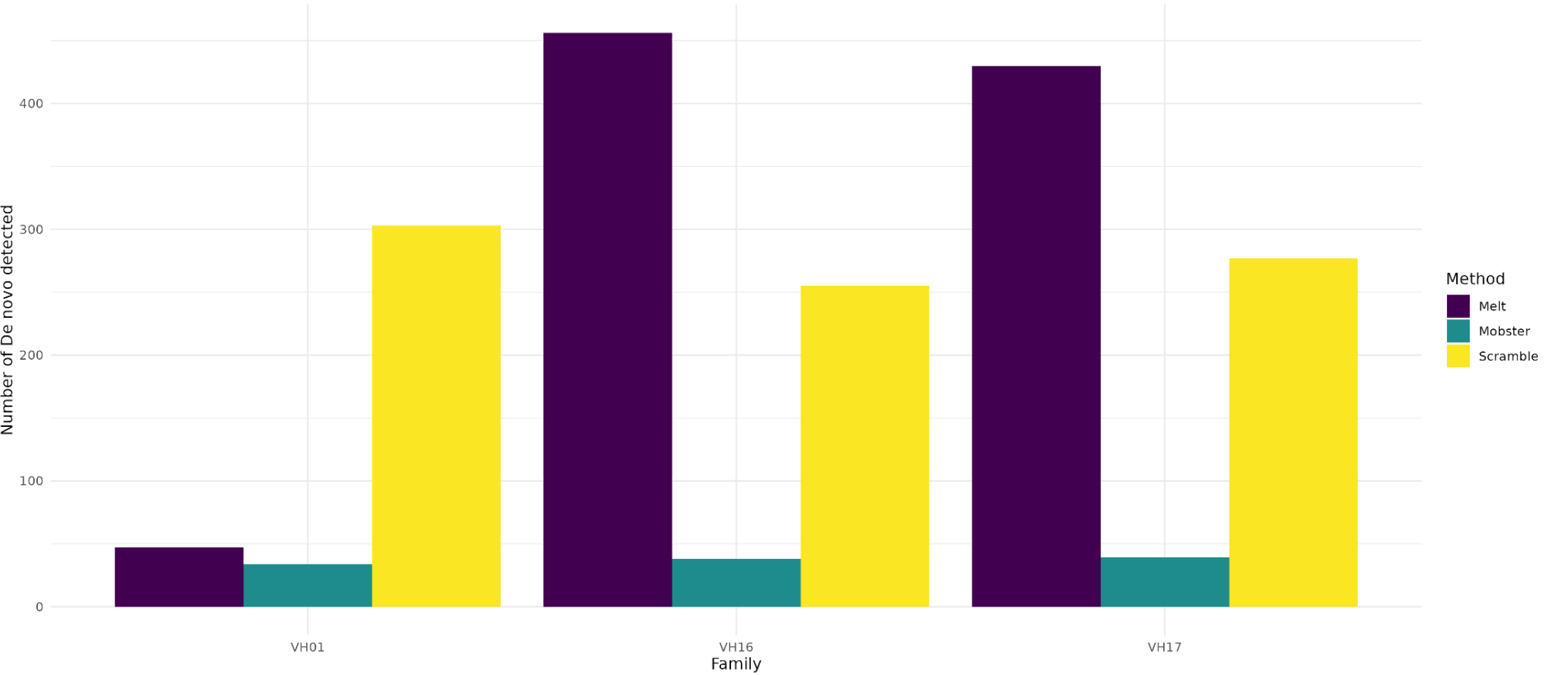
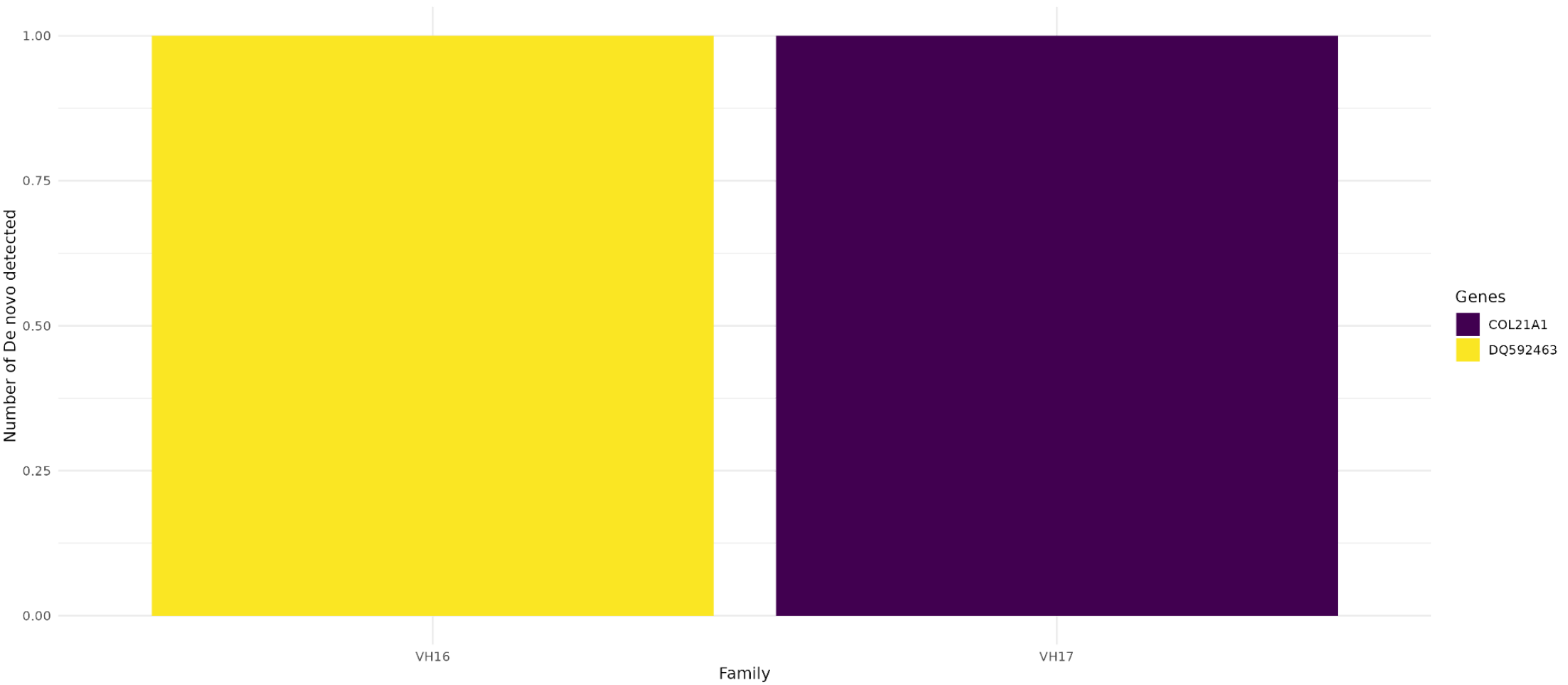


Figure S28. Number of De novo insertions detected in WGS trios per tool



Source: Vall d'Hebron WGS

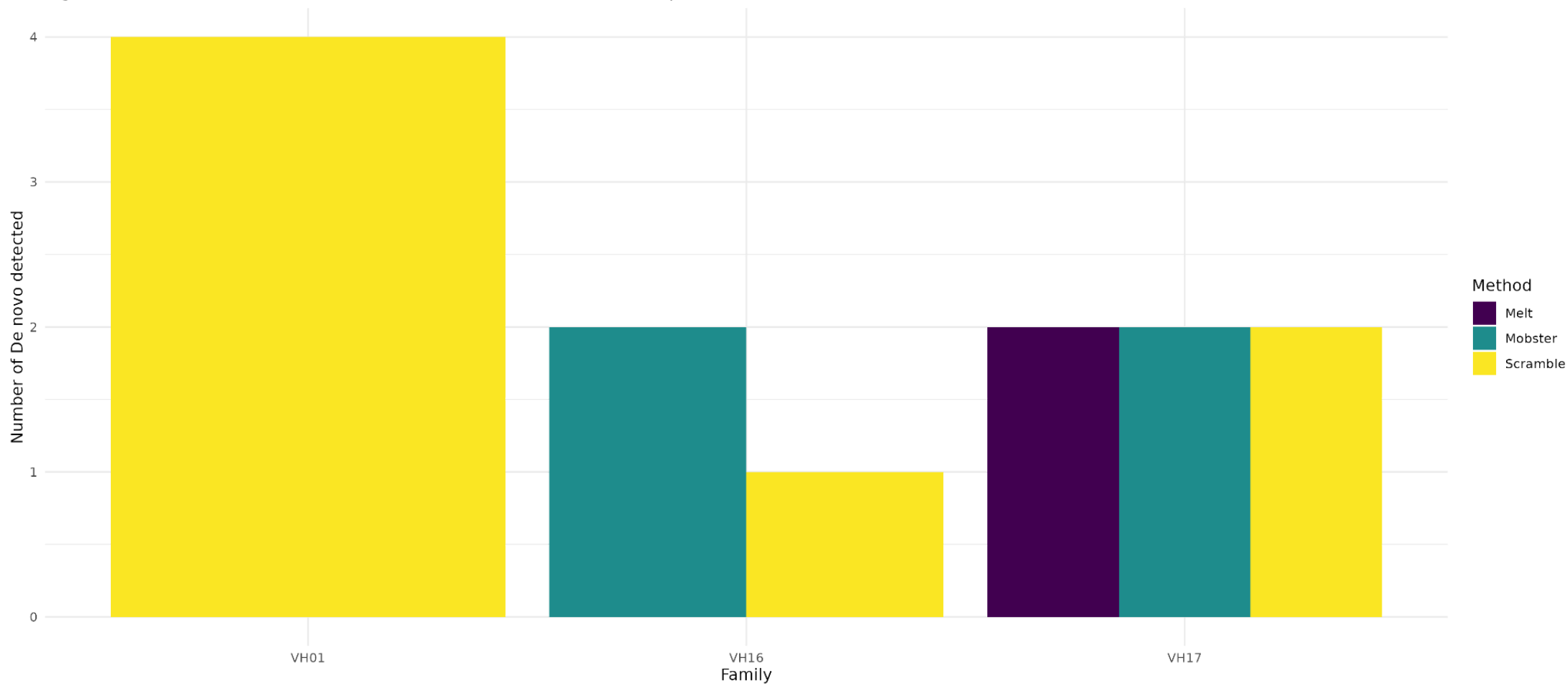
Figure S29. Genes from De novo insertion detected in WGS trios



Source: Vall d'Hebron WGS

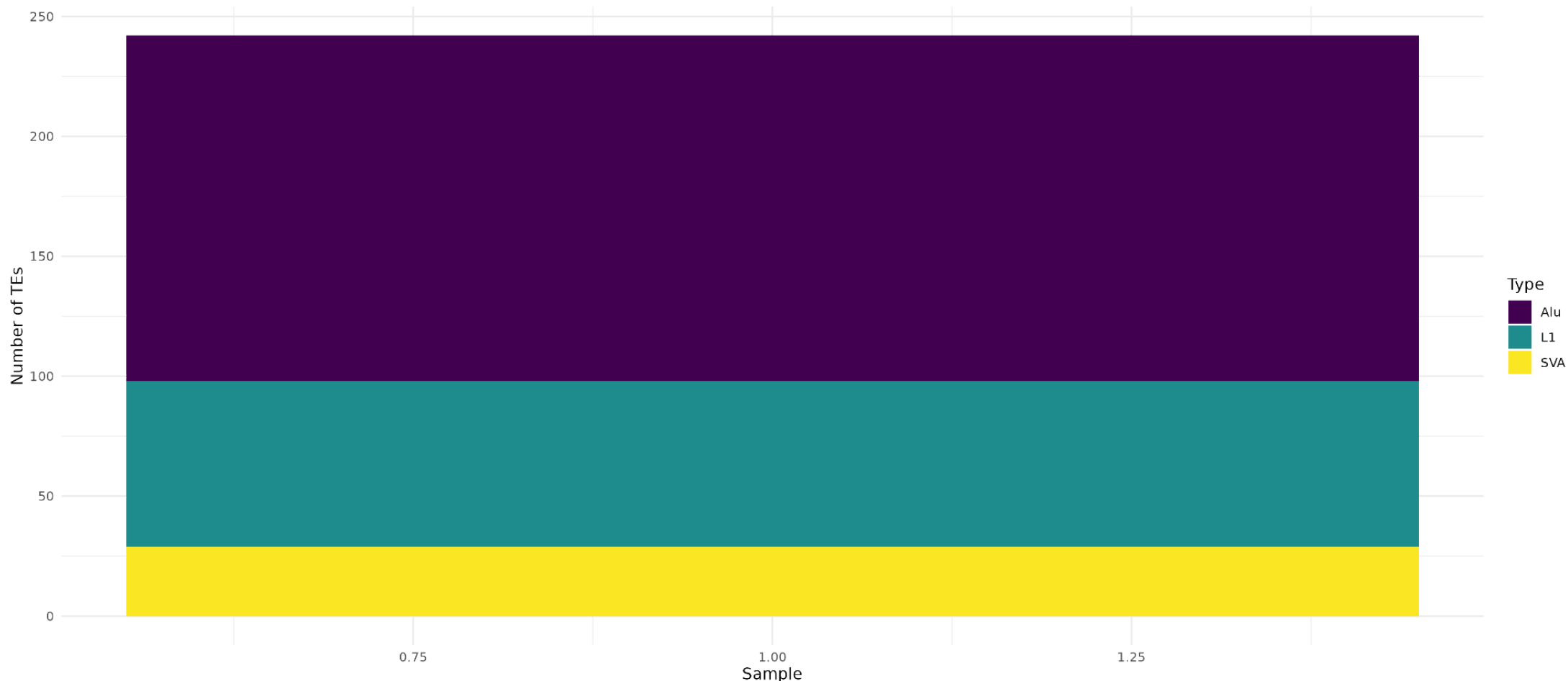


Figure S30. Number of De novo insertions detected in WES trios per tool



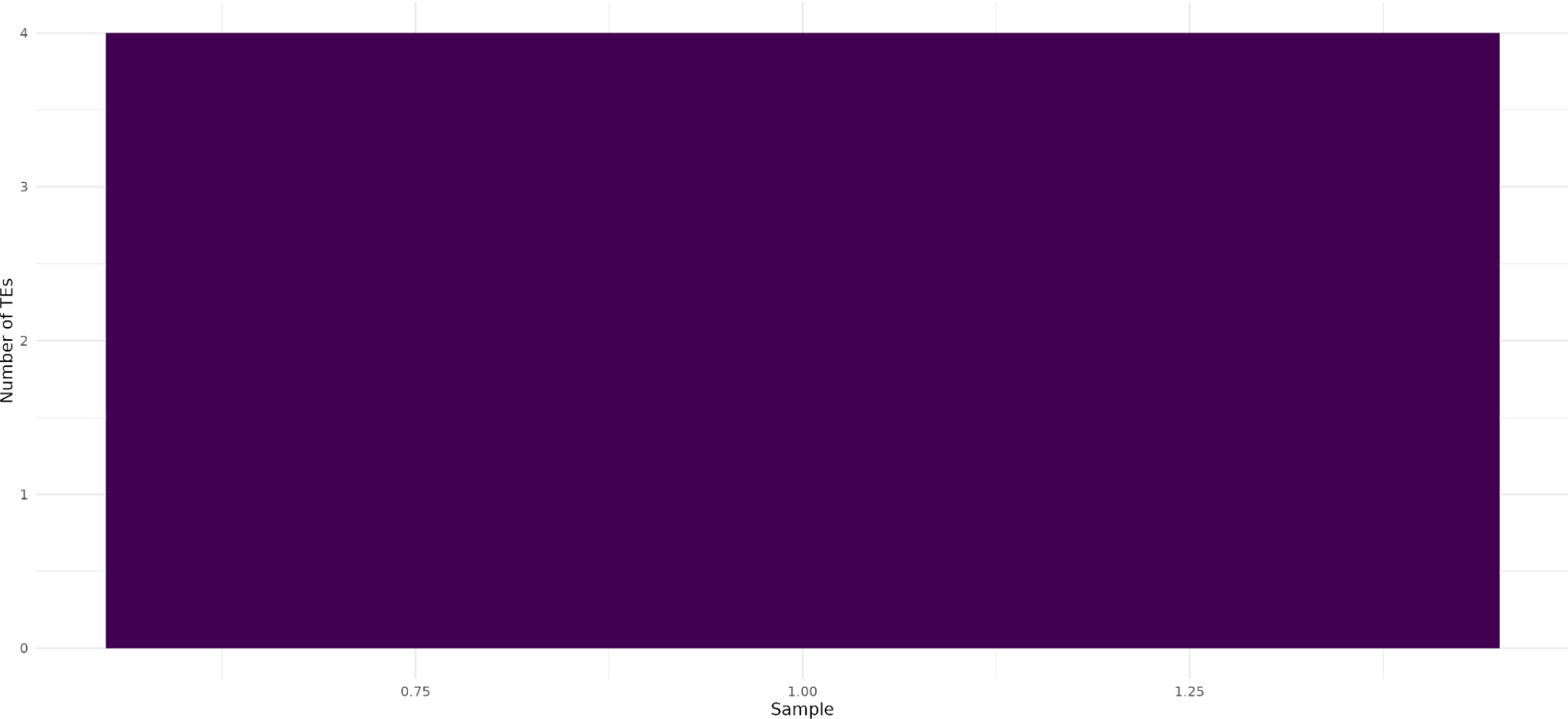
Source: Vall d'Hebron WES

Figure S31: Number of type of TEs detected by rMETL



Source: Long-Read Sample

Figure S32: Number of type of TEs detected by SCRAMble



Source: Long-Read Sample

**Figure S33. Venn diagram detection comparison rMETL-SCRAMble**

