## **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	Mediterranean	Vall d'Hebron	Vall d'Hebron	Average
		trio	fever	WES	WGS	detection
MELT	18.6%	23%	23.2%	16.3%	9.2%	18.06%
SCRAMb le	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	Mediterranea	Vall d'Hebron	Vall d'Hebron	Average
		trio	n fever	WES	WGS	detection
MELT	26.6%	46.1%	50.7%	20%	20.9%	32.86%
SCRAMb le	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	Mediterranean	Vall d'Hebron	Vall d'Hebron	Average
		trio	fever	WES	WGS	detection
MELT	54.8%	30.9%	26.1%	63.7%	69.9%	49.08%
SCRAMb le	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

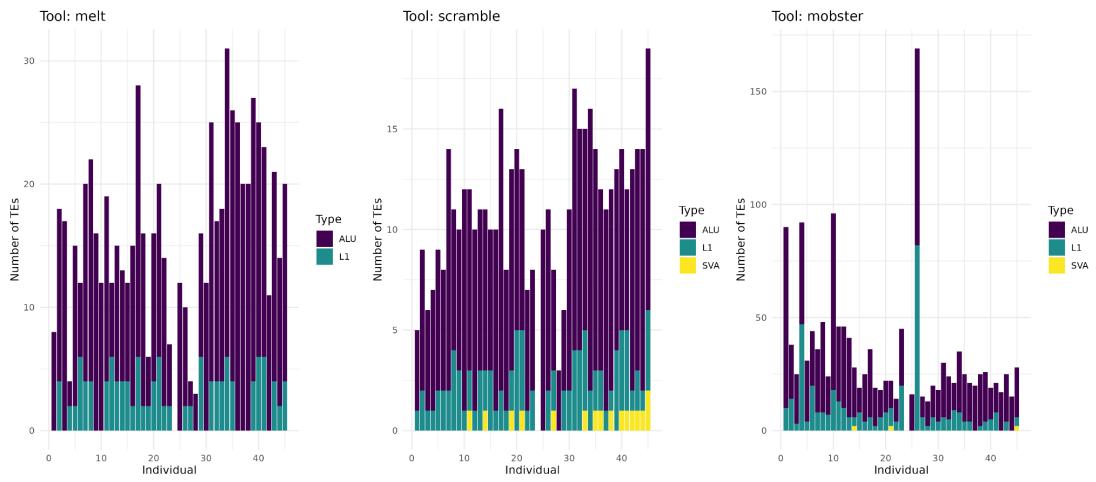
Tool	MELT	SCRAMble	mobster	Average detection
S58/VH17i	68.7%	57.1%	62.9%	62.9%
S59/VH17f	69.2%	63.1%	60%	64.1%
S60/VH17m	72.7%	58.8%	51.8%	61.1%
VH60/VH11s	47.3%	46.1%	40.3%	44.56%
VH61/VH11i	42.8%	55%	28.6%	42.13%
VH63/VH11f	50%	80%	47.8%	59.26%
VH71/VH16i	66.6%	60%	66.6%	64.4%
VH72/VH16m	100%	50%	50%	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%
Average	62.98%	54.83%	56.8%	
detection (tool)				

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

Tool: melt Tool: scramble Tool: mobster 12.5 25 15 10.0 20 Number of TEs Number of TEs Number of TEs Туре Туре Туре L1 SVA 5.0 5 2.5 5 0.0 0 414955

Individual

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

Individual

Source: Argentinian Trio

Individual

Tool: melt Tool: scramble Tool: mobster 10.0 30 30 Number of TEs Number of TEs Number of TEs Туре 10 10 2.5

Individual

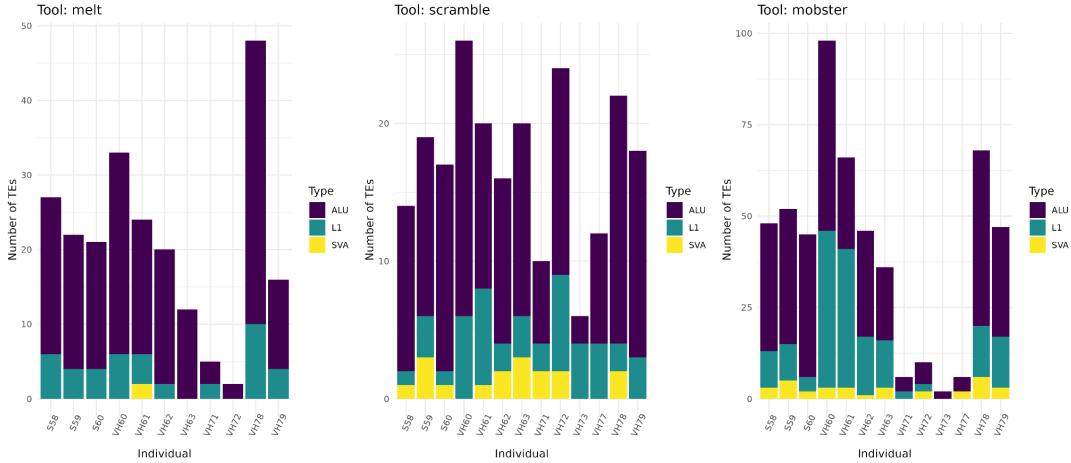
Figure S3: Number of TEs per Individual detected by each tool

Individual

Source: Mediterranean Fever

Individual

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



Source: Vall d'Hebron WES

Tool: melt Tool: scramble Tool: mobster 2500 2000 2000 Number of TEs Number of TEs Number of TEs Type Type Type SVA SVA 500 14015 14011 14011 14115 14115 14165 14015 14015 14017 14115 14115 14115 WO11 WO1m WH115 VH11, VH175 VH115 VH165 VH16; VH17F W165 W161 W161 W1275 W1277 W1277 VH16m VH16; VH16h VH17; VH17h VH17;

Individual

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

Individual

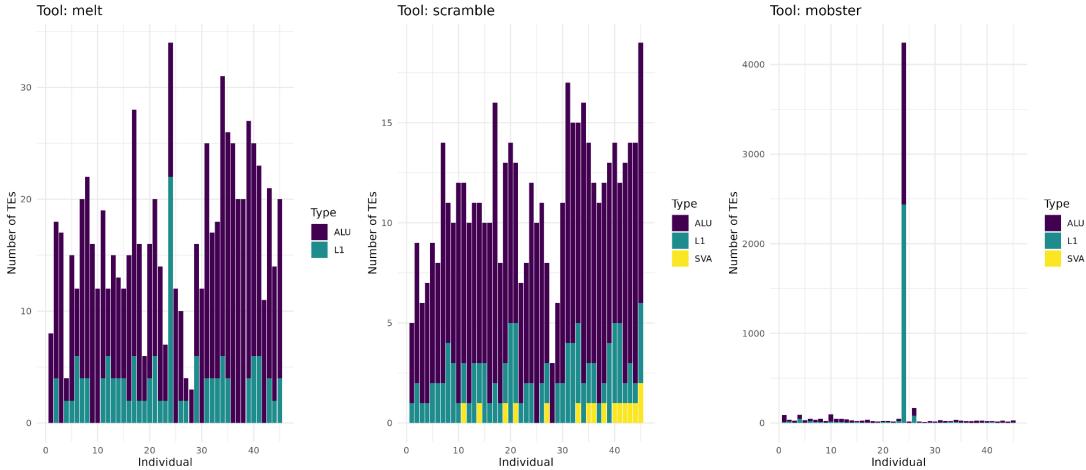
Source: Vall d'Hebron WGS

Individual

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

Tool: melt

Tool: scramble



Source: CVID patients

RAG1 PLCG2 7.5 NOD2 IL21R Tool CLEC16A scramble CD22 CARD11 PLCG2 2.5 BCL2 ACP5 0.0

Tool

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

mobster

Source: CVID patients

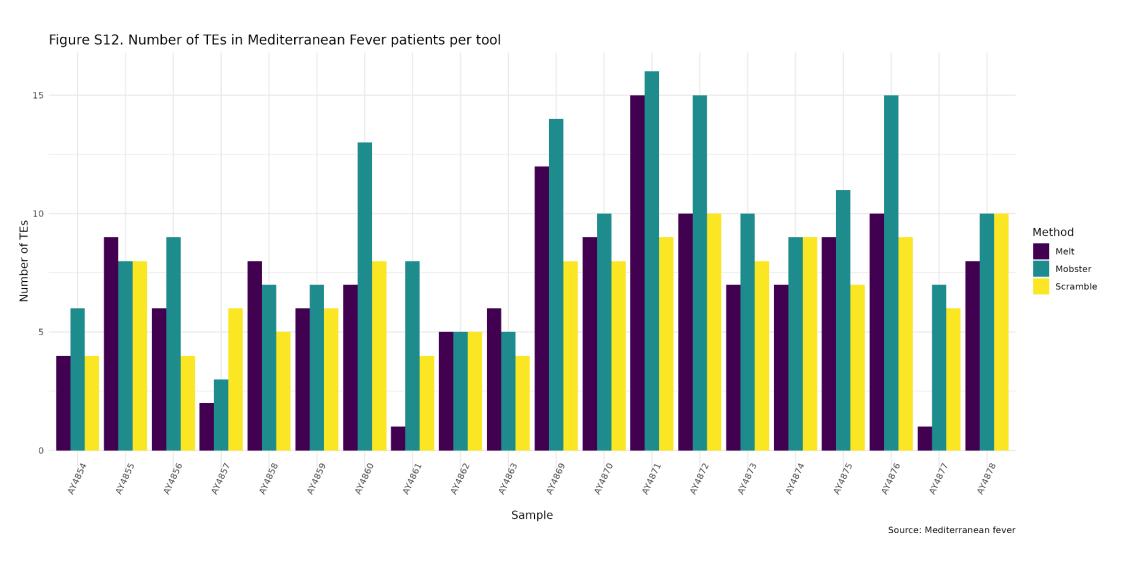
scramble

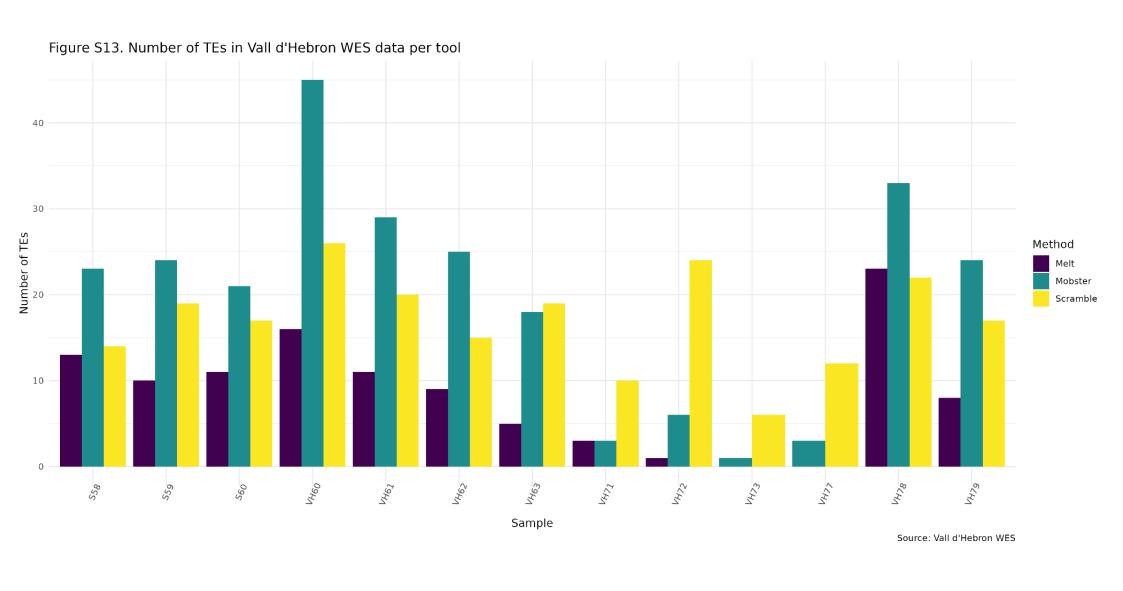
Figure S8. Number of candidate genes (IEI) detected by each method SEMA3E 1.5 Tool mobster NBAS 0.5 0.0 mobster Tool Source: Vall d'Hebron WES

Figure S9. Number of candidate genes (IEI) detected by each method 400 Tool melt mobster scramble 200 mobster melt scramble Tool Source: Vall d'Hebron WGS

Figure S10. Number of TEs in CVID patients per tool 60 Number of TEs Method Mobster Scramble 20 20 10 Sample Source: CVID patients

Figure S11. Number of TEs in Argentinian trio per tool 10.0 7.5 Number of TEs Method Melt Mobster Scramble 2.5 0.0 AY4954 Sample AY4941 AY4955 Source: Argentinian trio





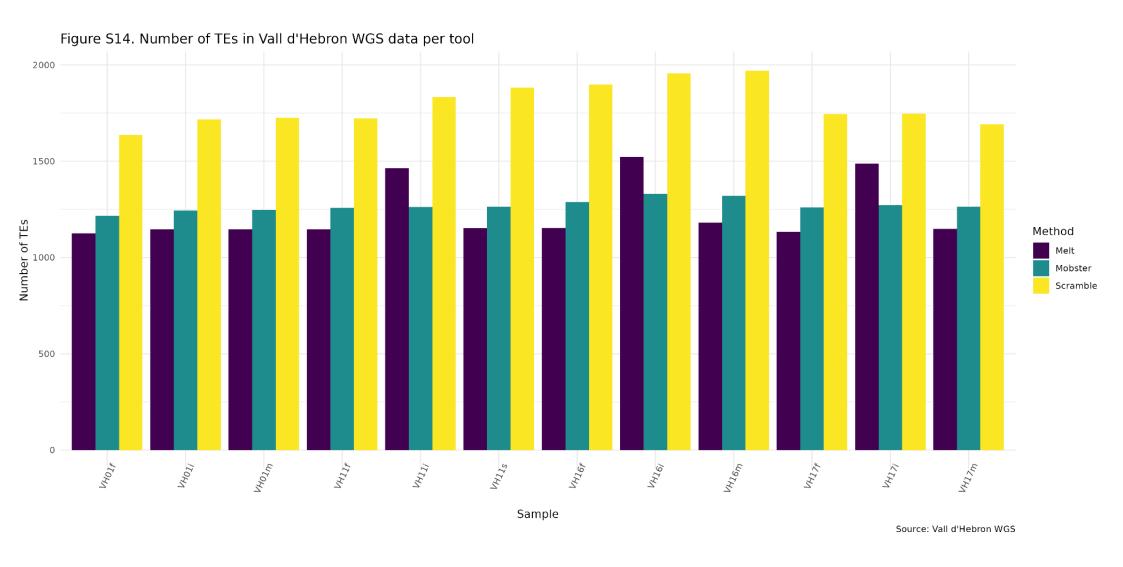


Figure S15. Venn diagrams of detected insertions per tool in each dataset.

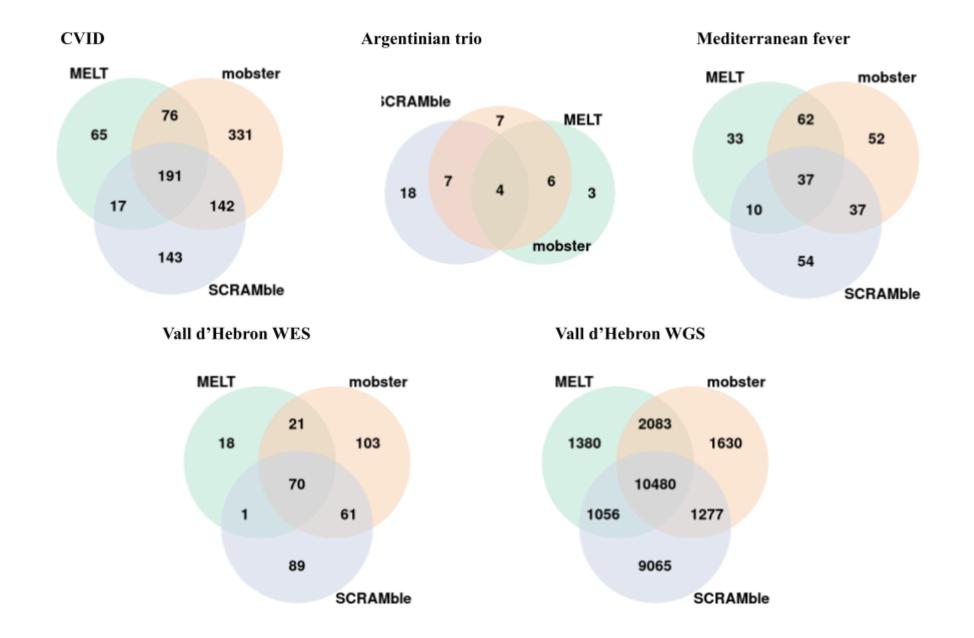


Figure S16. Venn diagrams VH17i/S58 comparison MELT mobster 1824 WES data WES data WGS data WGS data **SCRAMble** 

Figure S17. Venn diagrams VH17f/S59 comparison MELT mobster 1388 WES data WES data WGS data WGS data **SCRAMble** 2253 WES data

Figure S18. Venn diagrams VH17m/S60 comparison MELT mobster 1370 WES data WES data WGS data WGS data **SCRAMble** 2149 WES data

Figure S19. Venn diagrams VH11s/VH60 comparison MELT mobster 1444 WES data WES data WGS data WGS data SCRAMble 2455 WES data

Figure S20. Venn diagrams VH11i/VH61 comparison MELT mobster 1797 WES data WES dat WGS data WGS data SCRAMble WES data

Figure S21. Venn diagrams VH11f/VH63 comparison MELT mobster 1623 WES data 1397 WGS data WGS data **SCRAMble** 2175 WES data

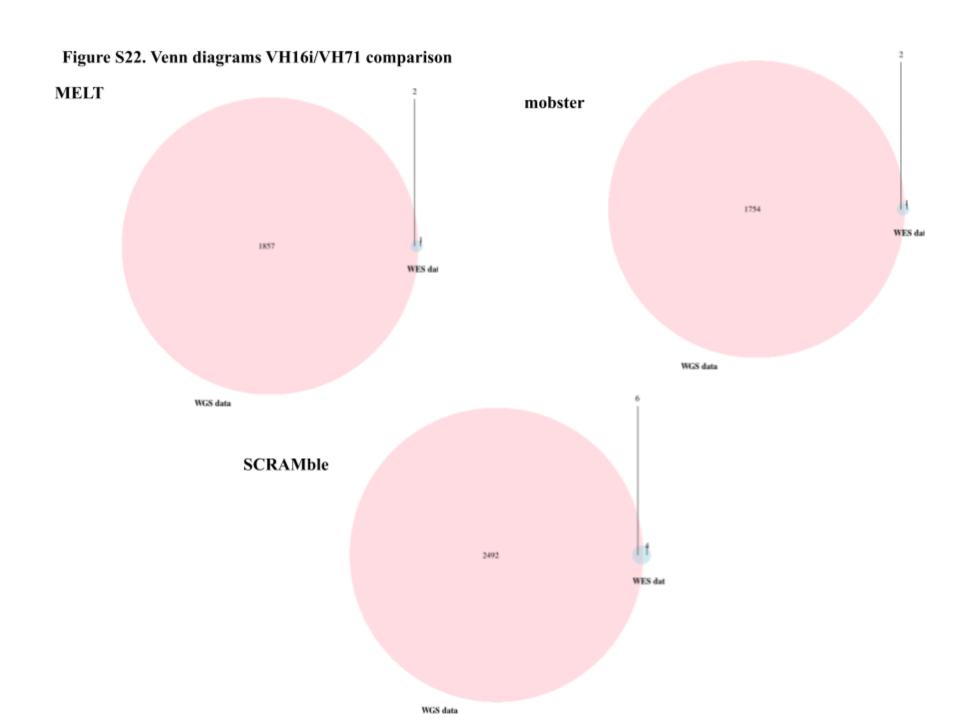


Figure S23. Venn diagrams VH16m/VH72 comparison MELT WGS data mobster 1745 WES data WES dat 1467 WGS data 12 **SCRAMble** WES data

Figure S24. Venn diagrams VH16f/VH73 comparison WGS data MELT mobster WGS data WES data WES data 1417 **SCRAMble** 2480 WES da

Figure S25. Venn diagrams VH01i/VH77 comparison WGS data MELT mobster WGS data WES data WES data 1585 1378 **SCRAMble** 2167 WES data

Figure S26. Venn diagrams VH01m/VH78 comparison MELT mobster 1378 WES data WES data WGS data WGS data SCRAMble 2164 WES data

Figure S27. Venn diagrams VH01f/VH79 comparison MELT mobster WES data WES data WGS data WGS data SCRAMble WES data

400 Number of De novo detected Method Mobster Scramble 100 VH16 Family VH17 VH01 Source: Vall d'Hebron WGS

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

1.00 0.75 Number of De novo detected Genes COL21A1 DQ592463 0.25 0.00 VH16 VH17 Family Source: Vall d'Hebron WGS

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

Figure S30. Number of De novo insertions detected in WES trios per tool Number of De novo detected  $\stackrel{\sim}{\sim}$ Method Mobster Scramble VH16 Family VH17 VH01 Source: Vall d'Hebron WES

Figure S31: Number of type of TEs detected by rMETL 250 200 150 Number of TEs Туре L1 50 1.00 Sample 0.75 1.25 Source: Long-Read Sample

Number of TEs  $^{\sim}$ Туре 1.00 Sample 0.75 1.25 Source: Long-Read Sample

Figure S32: Number of type of TEs detected by SCRAMble

Figure S33. Venn diagram detection comparison rMETL-SCRAMble

