Supplementary Fig. S1

SNP identification

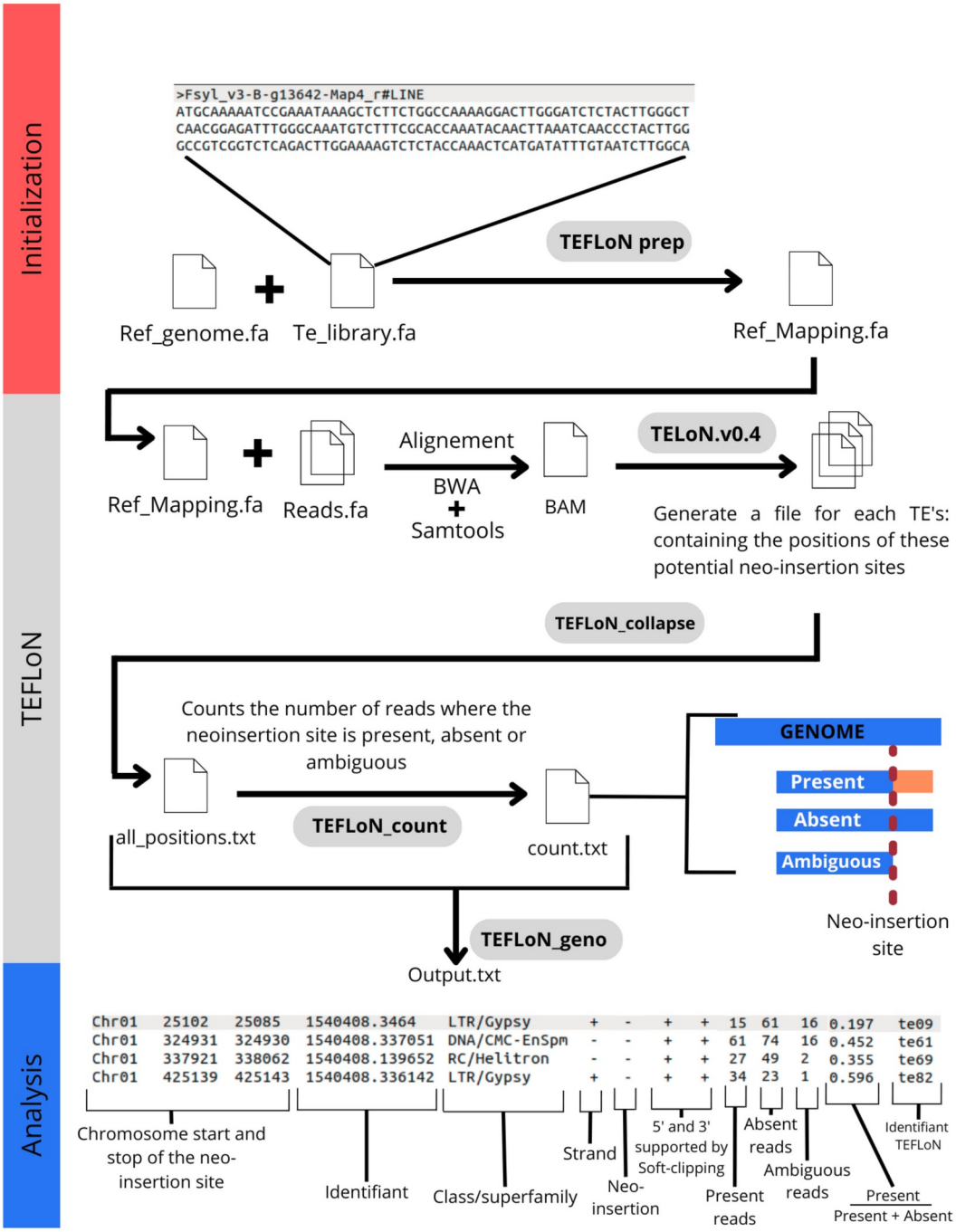
In **poplar**, on average, 11.56% of reads were duplicated and removed and 73.22% of trimmed reads were properly mapped to the reference genome. Following the SNP filtering steps, a total of 16,703,467 SNPs, 14,990,578 SNPs and 6,739,425 SNPs were identified using Bcftools, GATK and FreeBayes, respectively. Considering SNPs identified by at least two out of the three methods, we ended up with a set of **12,432,187** SNPs for poplar.

Following sequencing, **oak genome** coverage ranged from 84.31X to 103.69X (average 94X). Trimming led on average to 215,412,011 reads per sample and 86.05% of trimmed reads were properly mapped on the reference genome. For oak, on average, 7,235,525 SNPs per sample were identified with bcftools. Following the digital normalization step required for FreeeBayes and GATK, on average, 95,984,679 reads per sample were kept (*i.e.* 44.56% of the initial reads) to meet the required 30X sequencing depth. Mapping of those reads using the previously used parameters, resulted on average in 83.68% of properly mapped reads. A total of 5,903,813 SNPs and 8,646,615 SNPs were identified with GATK and FreeBayes, respectively. Considering SNPs identified by at least two out of the three methods, we ended up with a set of reliable **15,727,742** SNPs for oak.

Results of the SNP identification in P. nigra and Q. petraea.

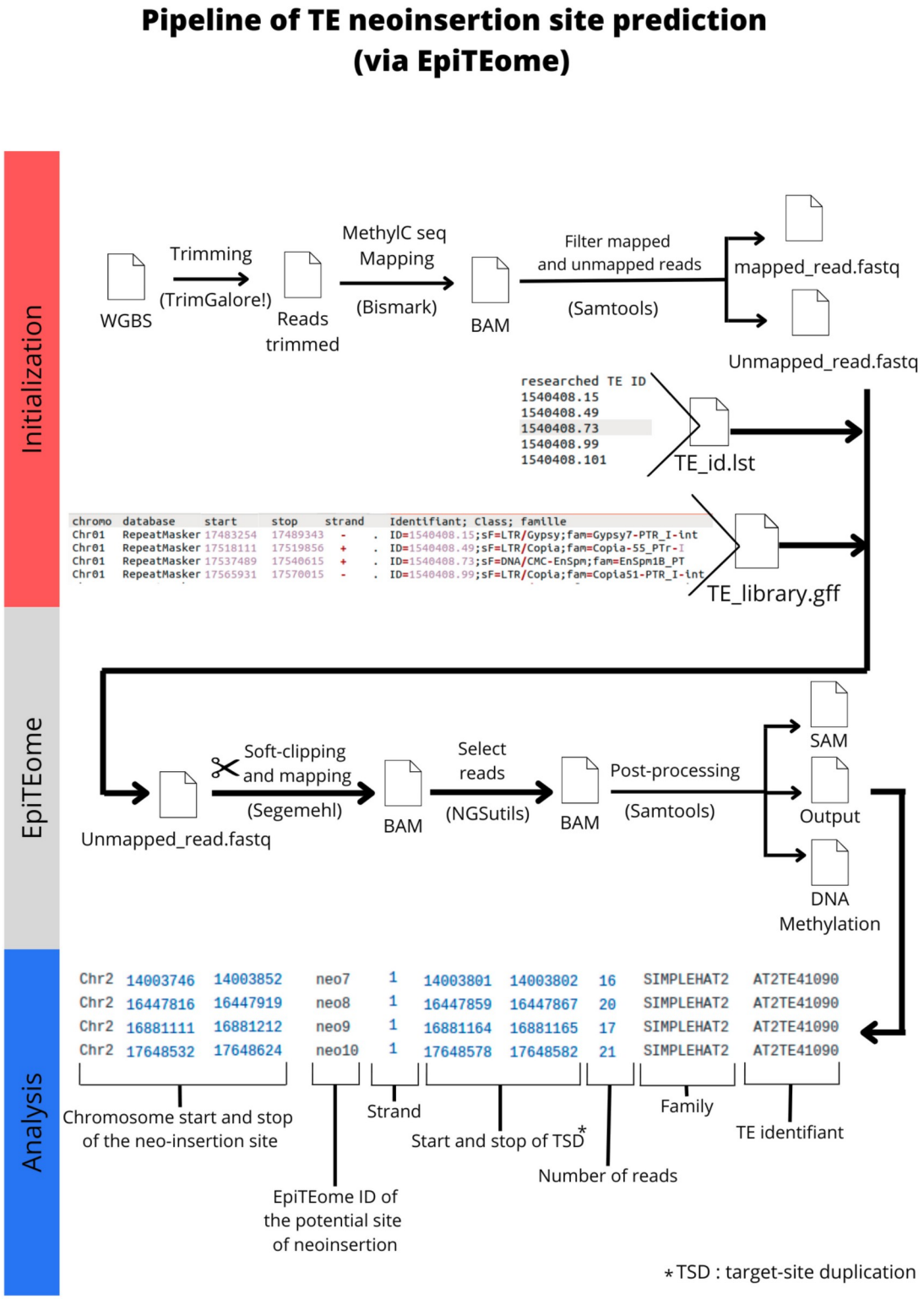
Supplementary Fig. S2

Pipeline of TE neoinsertion site prediction (via TEFLoN)



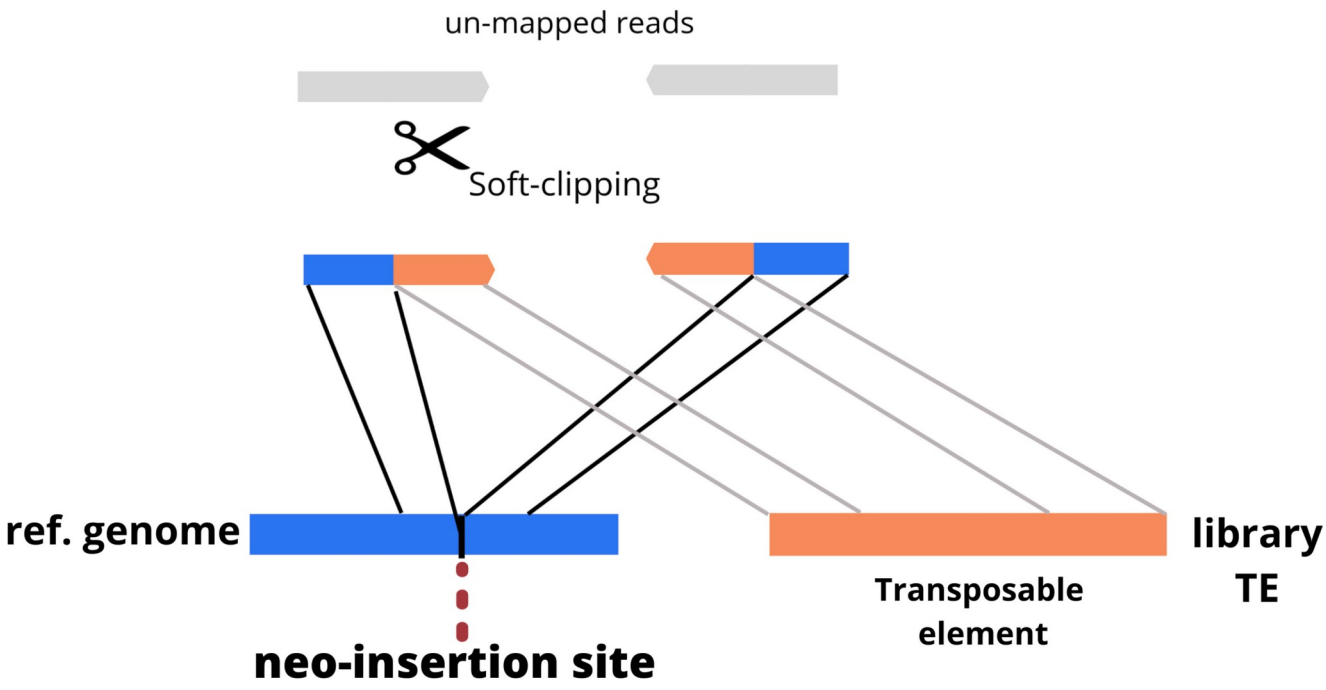
Conceptual diagram of the TEFLoN analysis workflows.
This figure shows the required files and tool modules used to test TEFLoN. This tool generates one tabular output file with, for each TIP, the chromosome number, the start and stop positions of the detected TIPs, the “new insertion” status, the TE identifier, its family and superfamily, its strand (+ or -) and the number of reads in favour (present), ambiguous or against (absent) the breakpoint used to detect the TIPs. A filtration step was applied as previously described (Adrion et al., 2017): a minimum of five "favourable" and five "unfavourable" reads from a total of at least 20 reads for the prediction of the presence or absence of an insertion site. Finally, based on the proportion of reads in favour of detection (reads "present"/reads "present" + reads "absent"), results in the 10 th and 90 th percentiles were excluded. For each population, we analysed two individuals and retained only insertion sites validated in both individuals for a given TE.

Supplementary Fig. S3



Conceptual diagram of the EpiTEome analysis workflow.
This figure lists the files and tools required for the testing of EpiTEome. This tool generates four output files: 1) a tabular file with the coordinates of the TIPs, the coordinates of the target site duplicates (TSD) and the read numbers used to predict this site, 2) a SAM file with 'soft-clipped' reads for visualisation of the alignment and potential target site duplicates (TSD) and 3-4) two tabular files providing an overview of the methylome data around the detected insertion sites (methylation contexts, methylation levels (%), number of methylated cytosines, number of mapped reads).

Supplementary Fig. S4



Representative diagram of the concept of "soft-clipped" reads.
Unmapped reads were isolated and soft-clipped into two fragments of variable size (k-mers). One of these fragments was mapped onto the TE and the second was mapped onto the reference genome, to detect a potential insertion polymorphism for the TE considered.

Supplementary Fig. S5

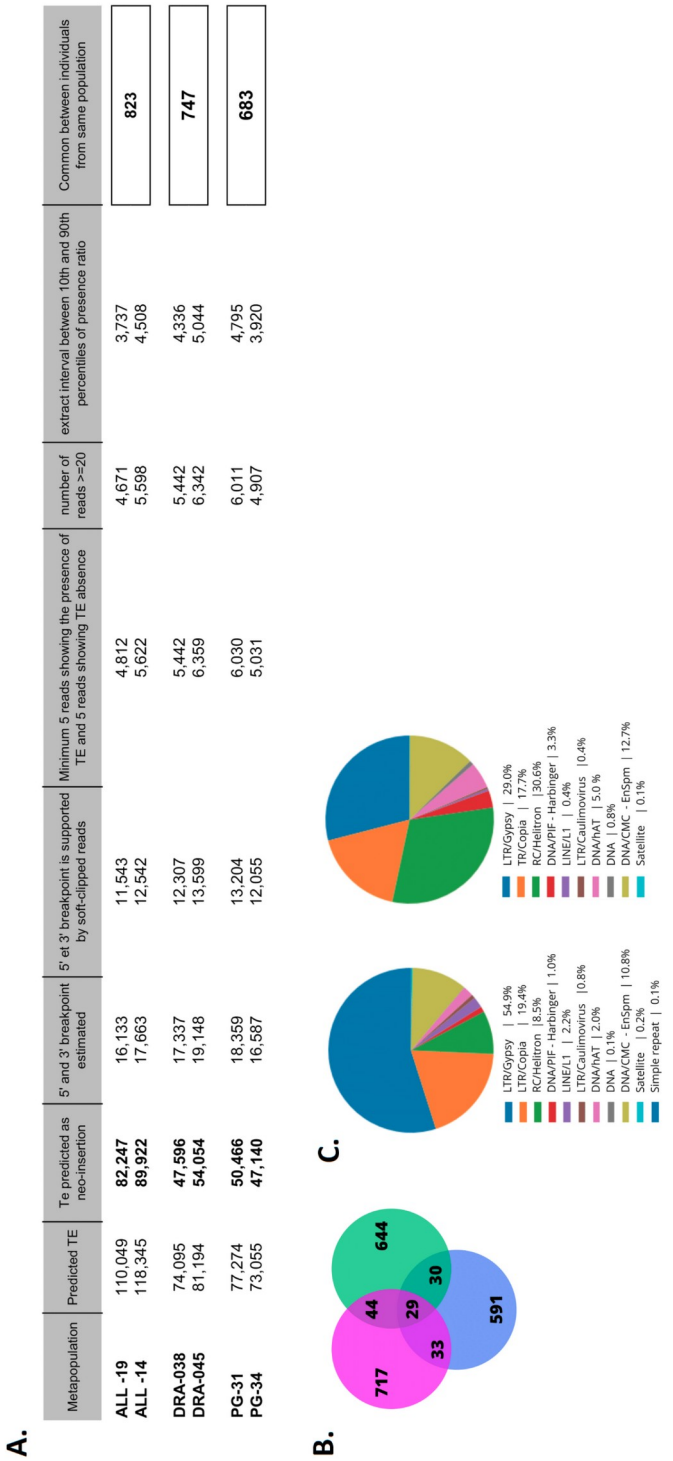
Paglia (PG)		
TE Families	TE Super-families	Nbr of neo-insertion site
Helitron-N2_PTr	RC/Helitron	149
Gypsy-71_PTr-LTR	LTR/Gypsy	133
Copia-91_PTr-LTR	LTR/Copia	62
Copia-3_PTri-I	LTR/Copia	45
EnSpm1B_PT	DNA/CMC-EnSpm	44
ENSPM1_PT	DNA/CMC-EnSpm	38
Helitron-N3_PTr	RC/Helitron	36
Gypsy-72_PTr-LTR	LTR/Gypsy	28
hAT-6_PTr	DNA/hAT-Tag1	21
DNA-3-1_PTr	DNA/PIF-Harbinger?	19
Helitron-N1_PTr	RC/Helitron	14
DNA-3-2B_PTr	DNA	10
Copia-93_PTr-I	LTR/Copia	10
POPGY2_LTR	LTR/Gypsy	7
Gypsy-35_PT-I	LTR/Gypsy	5
Gypsy-25_PTr-LTR	LTR/Gypsy	4
Ogre-PT1_I-int	LTR/Gypsy	4
Gypsy-79_PTr-LTR	LTR/Gypsy	4
Gypsy-73_PTr-LTR	LTR/Gypsy	3
Copia-56_PTr-I	LTR/Copia	3
Caulimovirus-1_PTr	LTR/Caulimovirus	3
hAT-1N_PTr	DNA/hAT-Ac	3
Helitron-1_PTr	RC/Helitron	3
hAT-1_PTr	DNA/hAT-Tip100	3
Copia43-PTR_I-int	LTR/Copia	2
hAT-4_PTr	DNA/hAT-Tip100	2
Copia-66_PT-I	LTR/Copia	2
EnSpm2_PTr	DNA/CMC-EnSpm	2
Gypsy-78_PTr-LTR	LTR/Gypsy	2
Gypsy13-PTR_LTR	LTR/Gypsy	1
Gypsy-78_PTr-I	LTR/Gypsy	1
POPGY1_LTR	LTR/Gypsy	1
hAT-3_PTr	DNA/hAT-Ac	1
Copia-54_PTr-I	LTR/Copia	1
Gypsy-26_PTr-LTR	LTR/Gypsy	1
Helitron-N4_PTr	RC/Helitron	1
GYPOT1_LTR	LTR/Gypsy	1
Gypsy18-PTR_I-int	LTR/Gypsy	1
Ogre-PT3_LTR	LTR/Gypsy	1
Copia-92_PTr-I	LTR/Copia	1
hAT-5_PTr	DNA/hAT-Tag1	1
Copia-55_PTr-I	LTR/Copia	1
Copia40-PTR_I-int	LTR/Copia	1
Copia-81_PT-I	LTR/Copia	1
L1-3_PTr	LINE/L1	1
Gypsy-27_PTr-I	LTR/Gypsy	1
Gypsy13-PTR_I-int	LTR/Gypsy	1
L1-5_PTr	LINE/L1	1
GYPOT1_I-int	LTR/Gypsy	1
Copia42-PTR_I-int	LTR/Copia	1
EnSpm3_PT	DNA/CMC-EnSpm	1

Dranse (DRA)		
TE Families	TE Super-families	Nbr of neo-insertion site
Helitron-N2_PTr	RC/Helitron	146
Gypsy-71_PTr-LTR	LTR/Gypsy	140
Copia-3_PTri-I	LTR/Copia	60
Copia-91_PTr-LTR	LTR/Copia	51
EnSpm1B_PT	DNA/CMC-EnSpm	49
ENSPM1_PT	DNA/CMC-EnSpm	43
Helitron-N3_PTr	RC/Helitron	38
DNA-3-1_PTr	DNA/PIF-Harbinger?	33
Gypsy-72_PTr-LTR	LTR/Gypsy	29
hAT-6_PTr	DNA/hAT-Tag1	28
Helitron-N1_PTr	RC/Helitron	24
Copia-93_PTr-I	LTR/Copia	11
POPGY2_LTR	LTR/Gypsy	9
Helitron-1_PTr	RC/Helitron	8
DNA-3-2B_PTr	DNA	5
Gypsy-79_PTr-LTR	LTR/Gypsy	5
EnSpm3_PT	DNA/CMC-EnSpm	4
Gypsy18-PTR_LTR	LTR/Gypsy	4
Gypsy-35_PT-I	LTR/Gypsy	4
EnSpm2_PTr	DNA/CMC-EnSpm	3
hAT-1_PTr	DNA/hAT-Tip100	3
Harbinger1_PTr	DNA/PIF-Harbinger	3
hAT-4_PTr	DNA/hAT-Tip100	3
Gypsy-73_PTr-LTR	LTR/Gypsy	2
Gypsy-27_PTr-I	LTR/Gypsy	2
Helitron-N4_PTr	RC/Helitron	2
hAT-1N_PTr	DNA/hAT-Ac	2
Copia-56_PTr-I	LTR/Copia	2
Gypsy-30_PTr-LTR	LTR/Gypsy	2
Ogre-PT2_LTR	LTR/Gypsy	2
Copia1-PTR_I-int	LTR/Copia	2
Caulimovirus-2_PTr	LTR/Caulimovirus	2
Copia-55_PTr-I	LTR/Copia	1
Gypsy-26_PTr-LTR	LTR/Gypsy	1
GYPOT1_I-int	LTR/Gypsy	1
hAT-7_PTr	DNA/hAT-Ac	1
hAT-3_PTr	DNA/hAT-Ac	1
Copia22-PTR_I-int	LTR/Copia	1
Caulimovirus-1_PTr	LTR/Caulimovirus	1
Gypsy13-PTR_I-int	LTR/Gypsy	1
Gypsy-25_PTr-LTR	LTR/Gypsy	1
Gypsy-39_PT-I	LTR/Gypsy	1
Ogre-PT3_LTR	LTR/Gypsy	1
Copia11-PTR_I-int	LTR/Copia	1
L1-1_PTr	LINE/L1	1
Gypsy13-PTR_LTR	LTR/Gypsy	1
GYPOT1_LTR	LTR/Gypsy	1
Copia26-PTR_I-int	LTR/Copia	1
Copia30-PTR_I-int	LTR/Copia	1
L1-4_PTr	LINE/L1	1
Ogre-PT1_I-int	LTR/Gypsy	1
L1-3_PTr	LINE/L1	1
Copia-55_PTr-LTR	LTR/Copia	1
Ogre-PT1_LTR	LTR/Gypsy	1
L1-2_PTr	LINE/L1	1
Ogre-PT2_I-int	LTR/Gypsy	1
Helitron-2_PTr	RC/Helitron	1
POPGY1_LTR	LTR/Gypsy	1

Val d'allier (ALL)		
TE Families	TE Super-families	Nbr of neo-insertion site
Helitron-N2_PTr	RC/Helitron	195
Gypsy-71_PTr-LTR	LTR/Gypsy	184
Copia-3_PTri-I	LTR/Copia	62
Copia-91_PTr-LTR	LTR/Copia	56
EnSpm1B_PT	DNA/CMC-EnSpm	52
ENSPM1_PT	DNA/CMC-EnSpm	48
Helitron-N3_PTr	RC/Helitron	43
hAT-6_PTr	DNA/hAT-Tag1	37
Gypsy-72_PTr-LTR	LTR/Gypsy	26
Helitron-N1_PTr	RC/Helitron	26
DNA-3-1_PTr	DNA/PIF-Harbinger?	18
Copia-93_PTr-I	LTR/Copia	10
POPGY2_LTR	LTR/Gypsy	7
DNA-3-2B_PTr	DNA	4
Gypsy-35_PT-I	LTR/Gypsy	4
Helitron-1_PTr	RC/Helitron	4
hAT-1_PTr	DNA/hAT-Tip100	3
Gypsy-26_PTr-LTR	LTR/Gypsy	3
Ogre-PT1_I-int	LTR/Gypsy	3
EnSpm2_PTr	DNA/CMC-EnSpm	2
hAT-1N_PTr	DNA/hAT-Ac	2
Gypsy-73_PTr-LTR	LTR/Gypsy	2
hAT-5_PTr	DNA/hAT-Tag1	2
POPGY1_I-int	LTR/Gypsy	2
L1-2_PTr	LINE/L1	2
Gypsy13-PTR_I-int	LTR/Gypsy	1
Copia43-PTR_I-int	LTR/Copia	1
Gypsy13-PTR_LTR	LTR/Gypsy	1
Caulimovirus-1_PTr	LTR/Caulimovirus	1
Gypsy-30_PTr-LTR	LTR/Gypsy	1
POPGY1_LTR	LTR/Gypsy	1
Copia-80_PT-I	LTR/Copia	1
L1-3_PTr	LINE/L1	1
Gypsy-25_PTr-LTR	LTR/Gypsy	1
Ogre-PT3_LTR	LTR/Gypsy	1
Copia-55_PTr-LTR	LTR/Copia	1
Gypsy-79_PTr-LTR	LTR/Gypsy	1
Copia11-PTR_I-int	LTR/Copia	1
Caulimovirus-2_PTr	LTR/Caulimovirus	1
Harbinger1_PTr	DNA/PIF-Harbinger	1
hAT-4_PTr	DNA/hAT-Tip100	1
SAT-1_PTr	Satellite	1
Gypsy-27_PTr-LTR	LTR/Gypsy	1
Copia-56_PTr-I	LTR/Copia	1
Copia52-PTR_I-int	LTR/Copia	1
Ogre-PT3_I-int	LTR/Gypsy	1
Gypsy-28_PTr-I	LTR/Gypsy	1
Copia-55_PTr-I	LTR/Copia	1
POPCOP2_I-int	LTR/Copia	1
Gypsy-25_PTr-I	LTR/Gypsy	1
Ogre-PT2_LTR	LTR/Gypsy	1

Numbers of TIPs in each TE family, for each population of P. nigra.

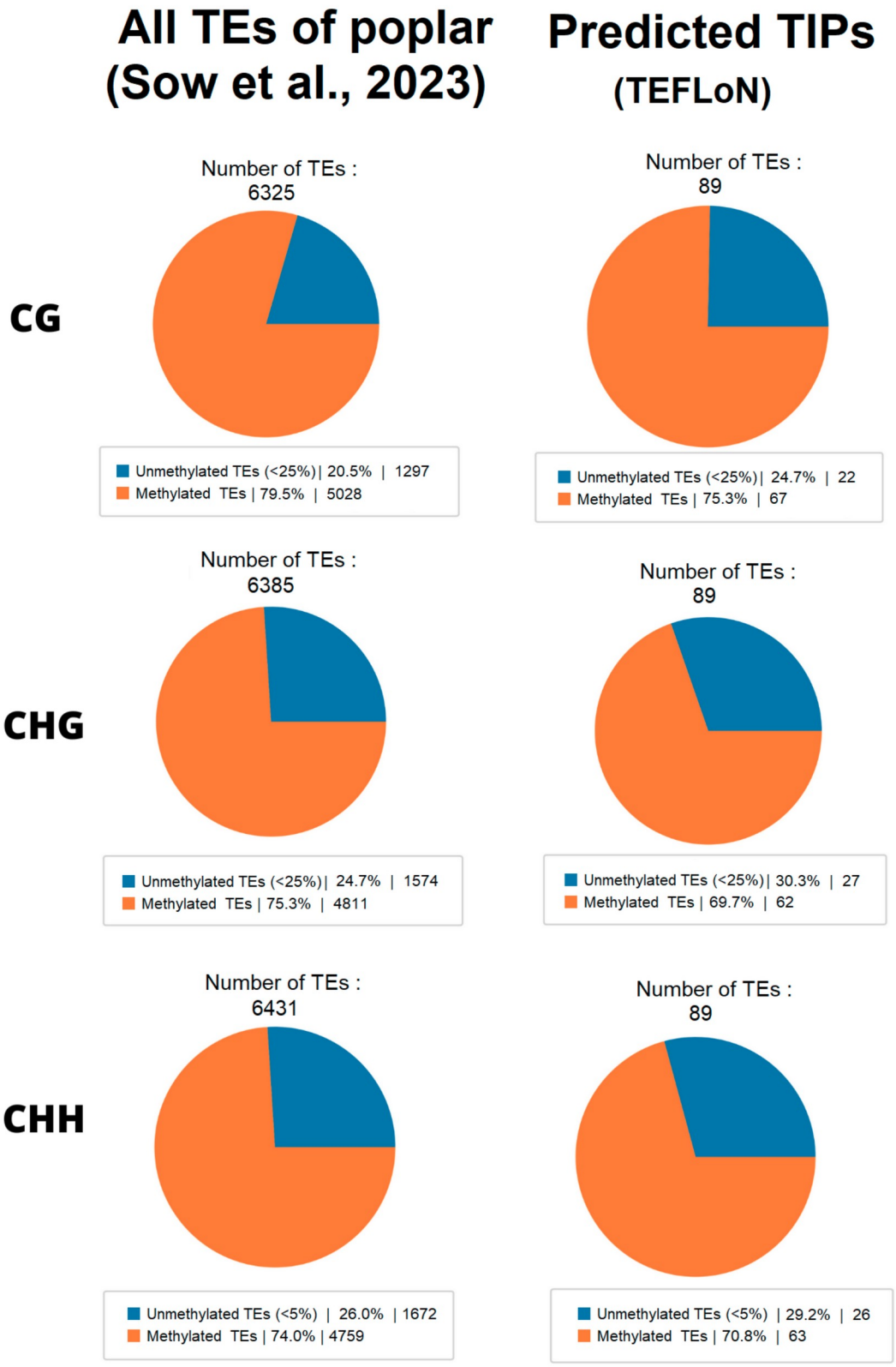
Supplementary Fig. S6



Metapopulation	Predicted TE	Te predicted as neo-insertion	5' and 3' breakpoint estimated	5' et 3' breakpoint is supported by soft-clipped reads	Minimum 5 reads showing the presence of TE and 5 reads showing TE absence	number of reads >=20	extract interval between 10th and 90th percentiles of presence ratio	Common between individuals from same population
ALL -19	110,049	**82,247**	16,133	11,543	4,812	4,671	3,737	**823**
ALL -14	118,345	**89,922**	17,663	12,542	5,622	5,598	4,508	
DRA-038	74,095	**47,596**	17,337	12,307	5,442	5,442	4,336	**747**
DRA-045	81,194	**54,054**	19,148	13,599	6,359	6,342	5,044	
PG-31	77,274	**50,466**	18,359	13,204	6,030	6,011	4,795	**683**
PG-34	73,055	**47,140**	16,587	12,055	5,031	4,907	3,920	

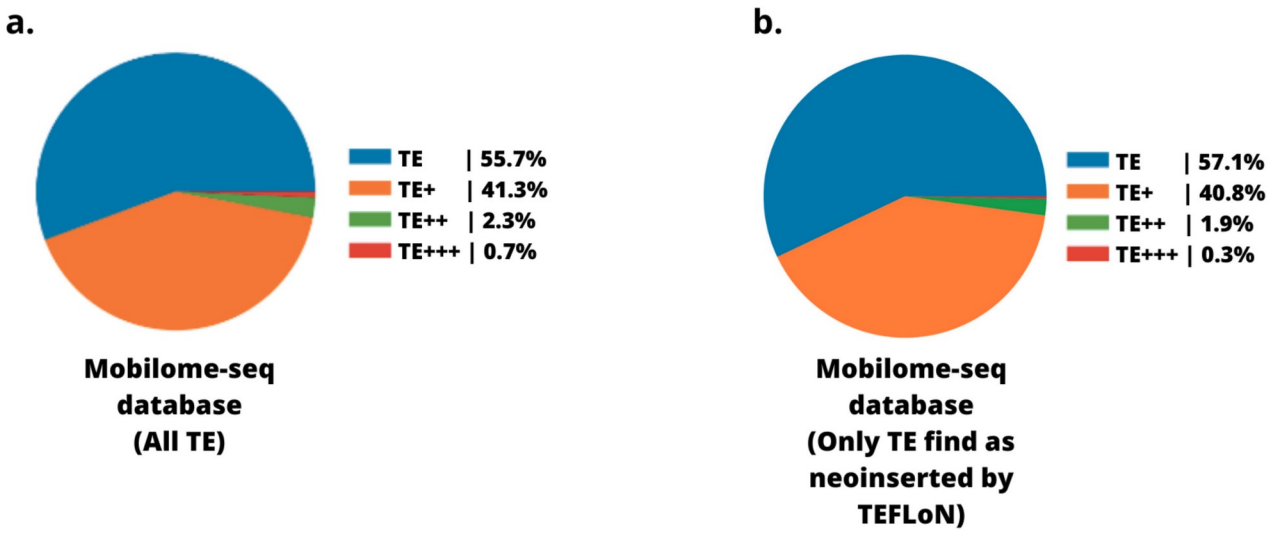
Numbers of TIPs and TE superfamilies predicted by TEFLoN in each poplar population.
A. Numbers of TIPs predicted by TEFLoN after each filtering step on WGS data for the following individuals: ALL-14, ALL-19, DRA-038, DRA-045, PG-31 and PG-34. Each filtering step has been described elsewhere (Adrion et al., 2017): Results are filtered based on start and stop values, 3' and 5' soft-clipping and a minimum of five "favourable" reads or five "unfavourable" reads from a total of at least 20 reads for insertion site detection. Based on the proportion of reads in favour of detection (reads "present"/ (reads "present" + reads "absent")), results in the 10th and 90th percentiles were excluded for each population; we retained only TIPs detected in both individuals of each population. **B.** Venn diagram of the number of TIPs detected in each poplar population by TEFLoN tools. TIPs common to two populations represent a same TIPs at the same insertion site (or overlap) in both populations. **C.** Proportion of TE superfamilies. On the left, the proportion of TE superfamilies of the 23,728 TE as TE library for TEFLoN. On the right, the proportion of cumulated TE superfamilies predicted for the 3 populations.

Supplementary Fig. S7



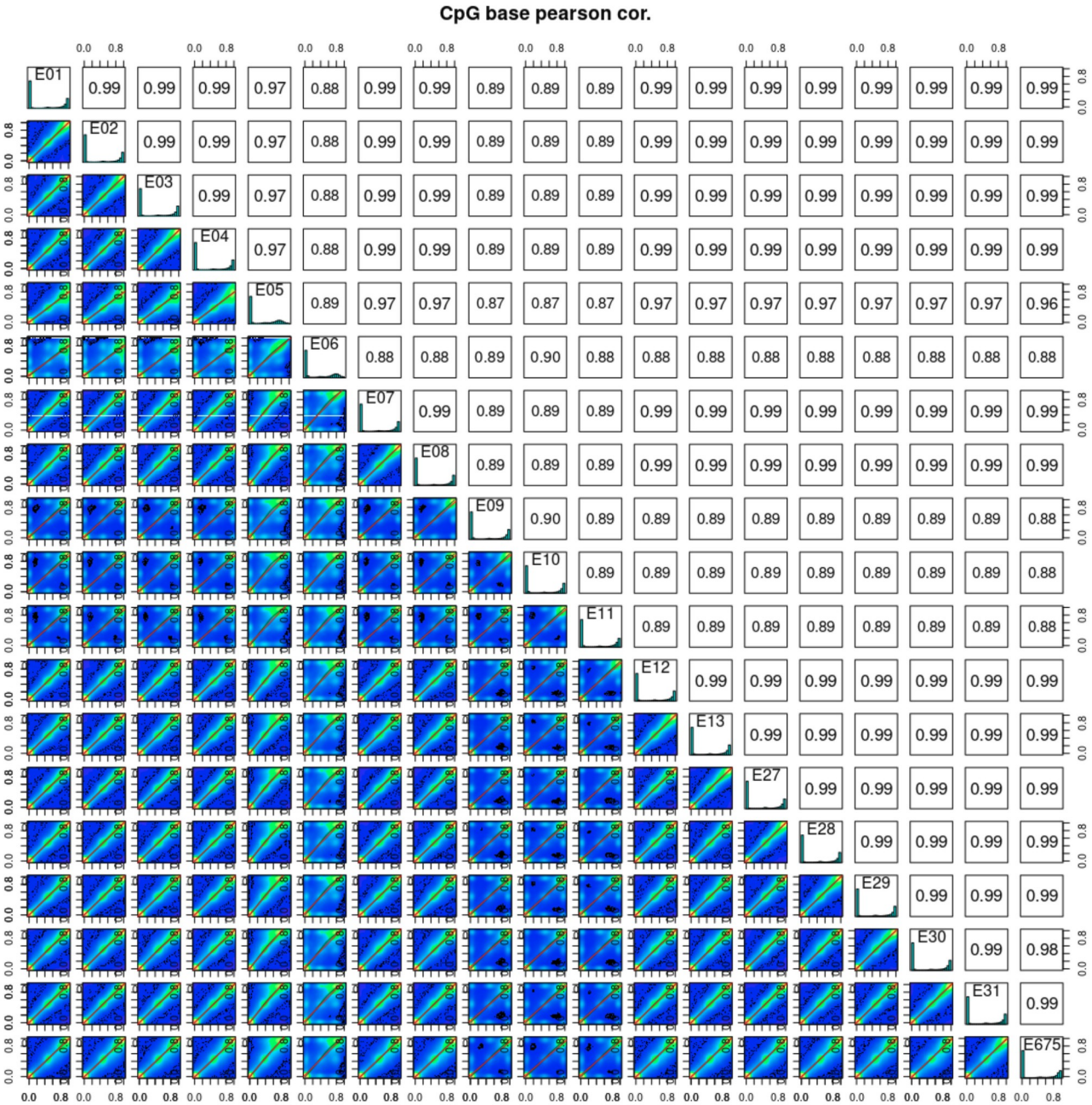
Correlations between TE families detected in the methylome data from Sow et al. 2023 with TEs predicted by TEFLoN to be newly inserted.
Each line corresponds to a methylation context: CG, CHG, CHH. The first column shows the proportion of methylated cytosines if only one population (DRA, PG or ALL) has a methylation value above the threshold of 25% for CG and CHG and 5% for CHH. The second column shows the proportion of TEs methylated, which is correlated with TEFLoN prediction.

Supplementary Fig. S8



Proportions of active TEs according to Mobilome-seq P. nigra data.
A. Pie chart of all TEs detected in the mobilome-seq database (Sow et al., 2021) for P. tremula x alba. **B.** Pie chart of TIPs detected by TEFLoN in the three poplar populations (P. nigra): ALL-14, ALL-19 (Val d'Allier), DRA-038, DRA-045 (Dranse), PG-31, PG-34 (Paglia). TEs were classified according to their depth of coverage (DOC), as follows: TE for a DOC of 1 to 200 reads (2689 TEs); TE+ (1994 TEs) for a DOC of 201-2,000 reads, TE++ (111 TEs) for a DOC of 2,001-10,000 reads and TE+++ (34 TEs) for a DOC over 10,000.

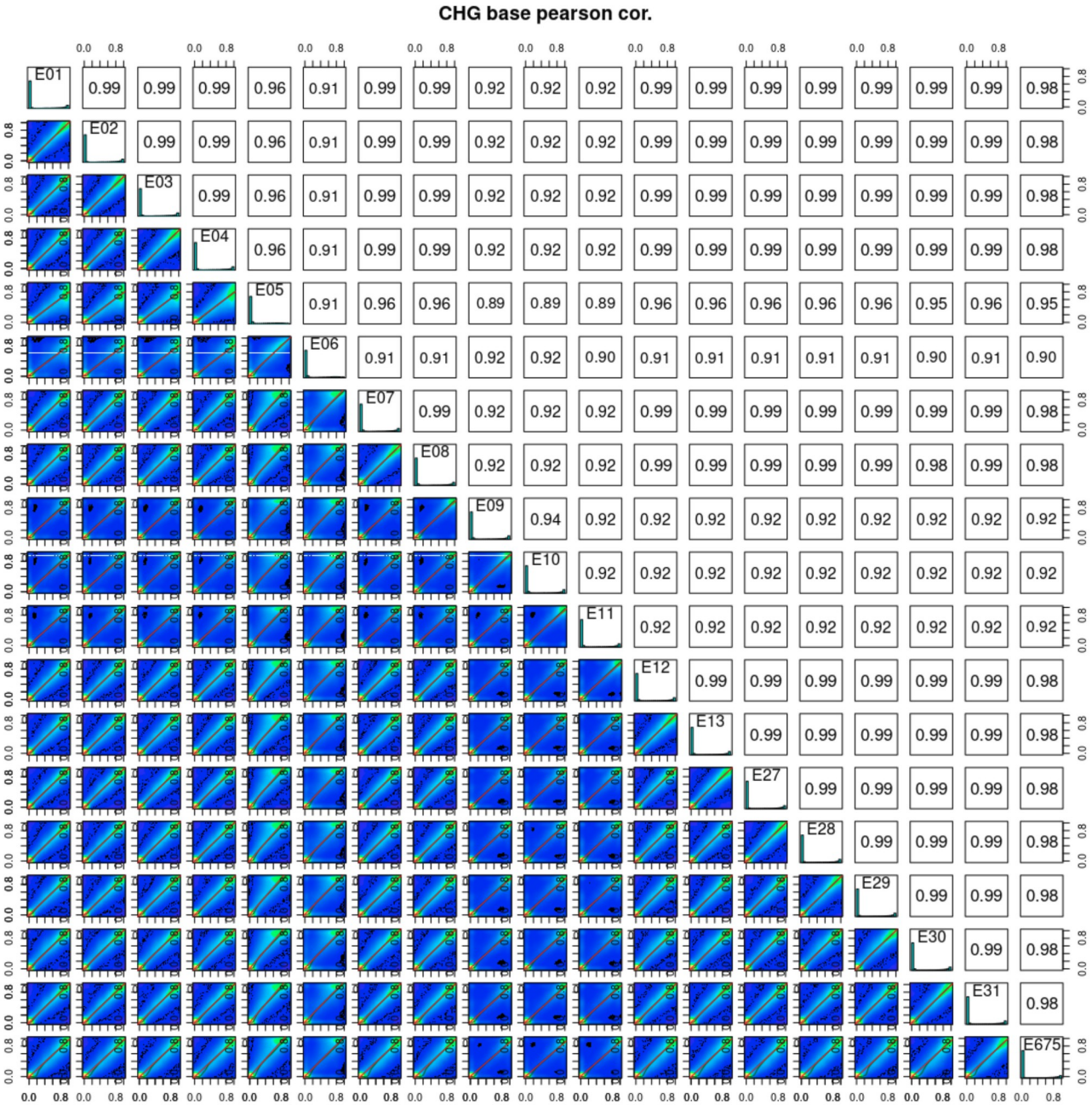
Supplementary Fig. S9



Correlations between SeqCapBis and WGBS results for P. nigra samples in the CG context.

Pearson correlation coefficient for SMPs common to WGBS and SeqCapBis in the CG context.

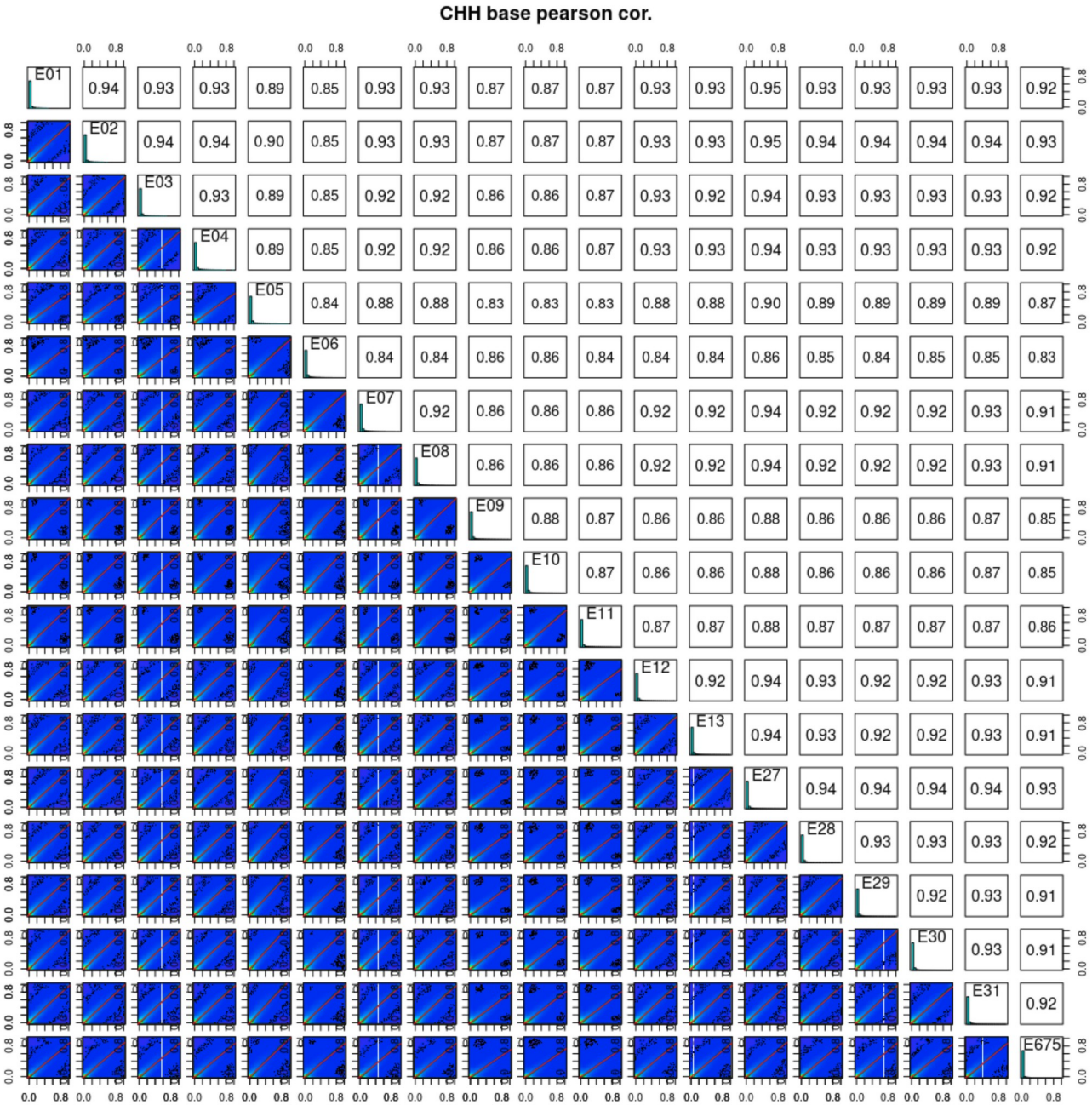
Supplementary Fig. S10



Correlations between SeqCapBis and WGBS results for P. nigra samples in the CHG context.

Pearson correlation coefficient for SMPs common to WGBS and SeqCapBis in the CHG context.

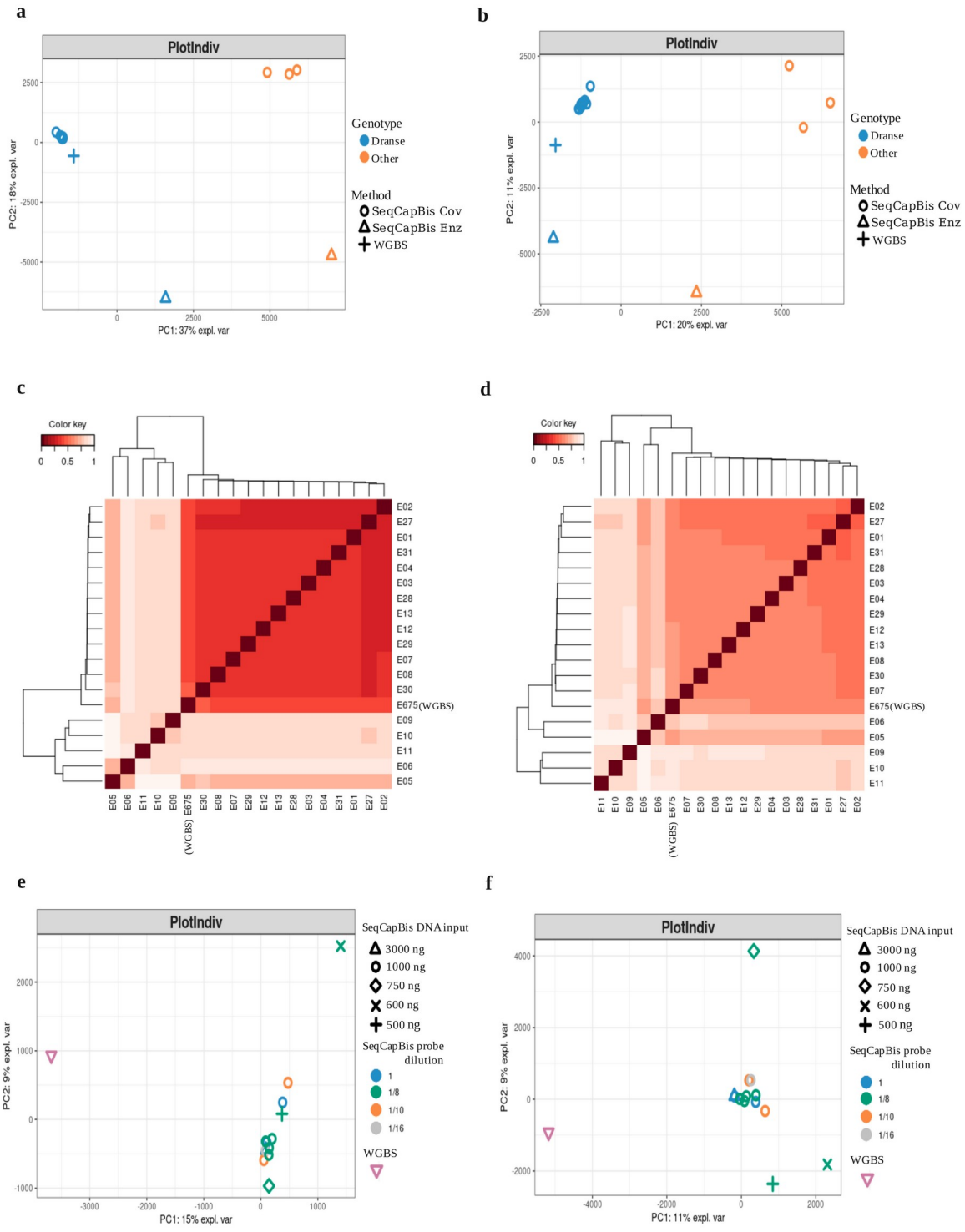
Supplementary Fig. S11



Correlations between SeqCapBis and WGBS results for P. nigra samples in the CHH context.

Pearson correlation coefficient for SMPs common to WGBS and SeqCapBis in the CHH context.

Supplementary Fig. S12



Principal component analysis (PCA) and heatmap for SeqCapBis and WGBS in the CHG and CHH contexts.

a. PCA on WGBS and SeqCapBis results for two P. nigra genotypes (Dranse in blue and others in orange, see Table 2) and two fragmentation approaches (acoustic shearing (Covaris) and enzymatic digestion) for the CHG context. **b.** PCA on WGBS and SeqCapBis results for two P. nigra genotypes (Dranse in blue and others in orange, see Table 2) and two fragmentation approaches (acoustic shearing (Covaris) and enzymatic digestion) for the CHH context. **c.** Heatmaps (Euclidean distance for clustering) based on SeqCapBis and WGBS data obtained in different experimental setups in the CHG context. **d.** Heatmaps based on SeqCapBis and WGBS data obtained in different experimental setups in the CHH context. **e.** PCA on CHG methylation data for Dranse samples fragmented by acoustic shearing. **f.** PCA on CHH methylation data for Dranse samples fragmented by acoustic shearing.

Supplementary Table. S1

Sample ID	Exp. Condition	Raw reads	Duplicates (%)	Properly mapped reads (%)	Reads ON-Target (%)	CG	CHG	CHH	Total
Dranse_DRA-038_CC	E01	28,311,331	27.4	45.9	63.9	1,651,768	2,460,998	13,402,495	17,515,261
Dranse_DRA-038_CC	E02	31,397,407	27.7	46.9	61.7	1,762,876	2,609,903	14,119,374	18,492,153
Dranse_DRA-038_CC	E03	26,752,606	34.1	45.3	60.4	1,691,561	2,513,765	13,517,857	17,723,183
Dranse_DRA-038_CC	E04	26,111,601	34.7	46.1	59.1	1,736,287	2,585,860	13,673,900	17,996,047
Dranse_DRA-038_CC	E05	27,345,856	41.4	43.0	56.7	1,731,123	2,544,811	13,331,464	17,607,398
Loire_SPM-034	E06	24,026,040	37	43.4	57.3	1,901,994	2,859,895	15,142,331	19,904,220
Dranse_DRA-038_CC	E07	26,031,231	41.9	46.1	54.7	1,668,747	2,444,243	12,789,973	16902,963
Dranse_DRA-038_CC	E08	22,546,694	33.1	45.8	59.1	1,631,743	2,430,118	12,886,868	16,948,729
Loire_SPM-004	E09	24,895,439	33.6	44.1	59.3	1,757,730	2,622,119	14,200,412	18,580,261
Loire_VDL-052	E10	28,350,316	44.5	45.2	55.8	1,831,152	2,760,750	14,678,900	19,270,802
Ticino_N-30	E11	25,729,274	34.6	45.3	59.8	1,706,486	2,541,907	13,688,207	17,936,600
Dranse_DRA-038_CC	E12	24,846,345	33.9	45.6	60.4	1,696,627	2,535,353	13,567,280	17,799,260
Dranse_DRA-038_CC	E13	29,686,867	42.1	46.1	57.7	1,941,441	2,975,150	15,437,966	20,354,557
Dranse_DRA-038_CC	E27	36,867,205	30.7	47.9	57.8	1,918,427	2,822,291	14,809,802	19,550,520
Dranse_DRA-038_CC	E28	24,751,793	29.4	46.8	61.3	1,774,362	2,676,680	14,225,631	48,676,673
Dranse_DRA-038_CC	E29	23,183,272	28.4	46.7	61.9	1,711,015	2,575,099	13,724,648	18,010,762
Dranse_DRA-038_CC	E30	27,505,283	40	48.8	51	1,708,586	2,461,130	12,013,675	16,183,391
Dranse_DRA-038_CC	E31	25,874,987	33.2	47.2	58.9	1,763,456	2,613,577	13,715,239	18,092,272

Statistics for the 18 experimental conditions tested on five P. nigra samples with the SeqCapBis approach in the CG, CHG and CHH methylation contexts.
A total of 17.84 Mb of sequence corresponding to 25,434 DMRs is considered for each sample.

Supplementary Table. S2

WGS

NCBI ID	Sample ID	Total read pairs	Properly mapped	SNP Samtools	SNP GATK	SNP FreeBayes
SAMN26818645	Bezange_82	167,369,201	78.98	7,051,865	5,134,514	4,431,359
SAMN26818644	Gresigne_37	244,496,293	85.60	7,413,814	5,384,298	4,647,074
SAMN26818643	Lappwald_108	223,042,876	85.60	7,415,315	5,392,681	4,661,098
SAMN26818642	Berce_193	194,147,247	85.77	7,241,681	5,304,389	4,577,253
SAMN26818641	StSauvant_6	185,113,993	86.67	7,191,650	5,252,899	4,558,342
SAMN26818640	Bourran_274	254,070,813	90.17	6,846,400	5,396,314	4,658,437
SAMN26818639	Bourran_214	206,603,533	89.83	6,928,104	5,374,525	4,616,138
SAMN26818638	Gohrde_89	231,229,747	87.11	7,704,570	5,638,236	4,842,614
SAMN26818637	Troncais_189	192,150,012	86.77	7,249,255	5,269,930	4,551,542
SAMN26818636	Longchamps_136	258,680,310	83.97	7,312,595	5,366,521	4,631,983

WGBS

NCBI ID	Sample ID	Total read pairs	Properly mapped	Coverage x	CG	CHG	CHH
SAMN26818645	Bezange_82	295,416,479	48.3	45	21,687,540	31,149,273	206,835,072
SAMN26818644	Gresigne_37	267,854,604	64.0	54	21,720,368	31,211,753	207,458,959
SAMN26818643	Lappwald_108	229,120,755	55.1	40	21,440,880	30,876,124	205,497,023
SAMN26818642	Berce_193	303,969,825	58.4	56	21,821,871	31,331,555	208,097,689
SAMN26818641	StSauvant_6	277,093,047	61.2	53	21,710,249	31,198,800	207,418,420
SAMN26818640	Bourran_274	336,584,932	66.4	70	22,802,963	32,479,615	215,183,473
SAMN26818639	Bourran_214	269456029	64.1	54	22,687,271	32,338,993	214,289,161
SAMN26818638	Gohrde_89	250,519,557	62.7	49	21,718,208	31,197,273	207,268,829
SAMN26818637	Troncais_189	413,725,043	61.5	80	22,046,594	31,614,934	209,923,445
SAMN26818636	Longchamps_136	296,584,323	61.7	58	21,759,083	31,252,613	207,700,986
Before filtering					21,939,503	31,465,093	208,967,306
After filtering					**4,256,014**	**14,429,956**	**96,137,556**

Summary of the WGS and WGBS sequencing data analysis for Q. petraea.
Before and after filtering: mapping statistics and number of SNPs detected with samtools/bcftools, GATK and FreeBayes.

Supplementary Table. S3

		4 samples altogether			Individual samples	
Sample ID	Sequencing type	Methylated positions	Depth >10X	Correlation coefficient	Common pos. WGBS – SeqCapbis	Correlation coefficient
T189	WGBS	15,608,479	15,592,853	0.951	202,534	0.927
T193	WGBS	18,050,016	18,031,956	0.943	195,277	0.925
T6	WGBS	19,613,356	19,593,724	0.944	169,509	0.930
T82	WGBS	16,496,039	16,479,527	0.950	162,971	0.936
T189	SeqCapBis	363,564	363,196	0.951		
T193	SeqCapBis	343,431	343,084	0.944		
T6	SeqCapBis	273,487	273,213	0.945		
T82	SeqCapBis	275,721	275,444	0.950		

Correlations between the SeqCapBis and WGBS results for the four Q. petraea samples. Pearson correlation analysis for positions common to the WGBS and SeqCapBis analyses in the CG context. When the four Q. petraea samples were considered together, 110,957 positions with at least 10X sequencing coverage were common to all datasets. Considering the samples separately increased number of common positions between SeqCapBis and WGBS to 182,573 per sample on average.