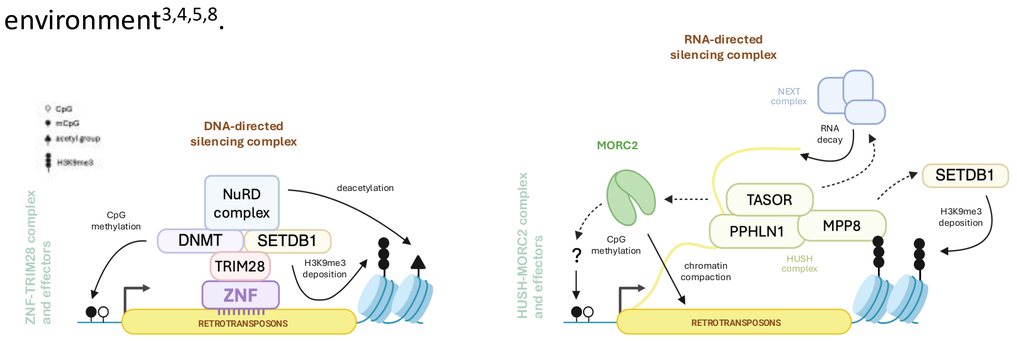
Using DiMeLo-seq to probe the epigenetic control of polymorphic transposable elements

CATERINA FRANCESCONI[1,2,*], ANGELOS KATSIKAS[1,2], FERESHTEH DORAZEHI[1,2], JACOB ALEXANDER QUIROS[3], VIVIEN HORVÁTH[4], CHRISTOPHER DOUSE[1,2]

Background

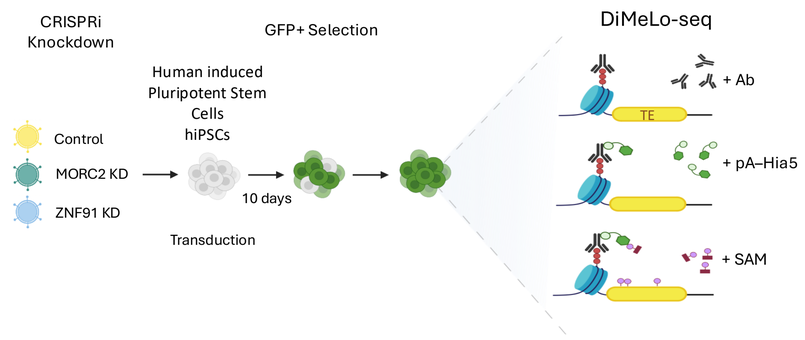
Around 50% of the human genome is occupied by transposable elements (TEs) [1], and because of their potential to cause transcriptome and genome instability they are tightly regulated[2]. In this context, the **ZNF-TRIM28 complex**[3,4] and the transcript-dependent **HUSH-MORC2 complex**[5,6,7] specifically target some of the evolutionarily youngest TEs, such as L1HS and SVAs, respectively. The silencing of the TEs is achieved through recruitment of effectors that modify the chromatin environment[3,4,5,8].

Evolutionary young TEs like L1HS and SVAs have yet to accumulate mutations[9] and are therefore often unmappable with short-reads. Additionally, the sequence of SVAs is highly repetitive[1], presenting an additional hurdle for mapping. The epigenetic characterisation of these young elements is therefore challenging.



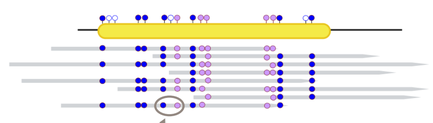
Here we exploit MORC2 KD and ZNF91 KD systems to characterise the epigenetic phenotype over HUSH-MORC2 and ZNF91-TRIM28 targets and test the transformative potential of **DiMeLo-seq**[10] (a recently developed long-read protein-protein/DNA profiling technique) on the study of epigenetic features of otherwise unmappable TEs.

Workflow for DiMeLo-seq



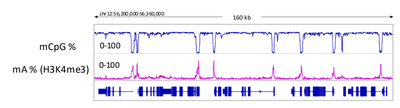
Antibody-targeted adenine methylation (mA) (normally absent in the human genome)

ONT Nanopore long-read sequencing



Detection of mCpG and H3K4me3 signal (mA) on a single long-read (locus level resolution)

Allows epigenetic characterisation of:
- Polymorphic (non-reference) insertions
- Length polymorphisms
- Evolutionarily young transposable elements

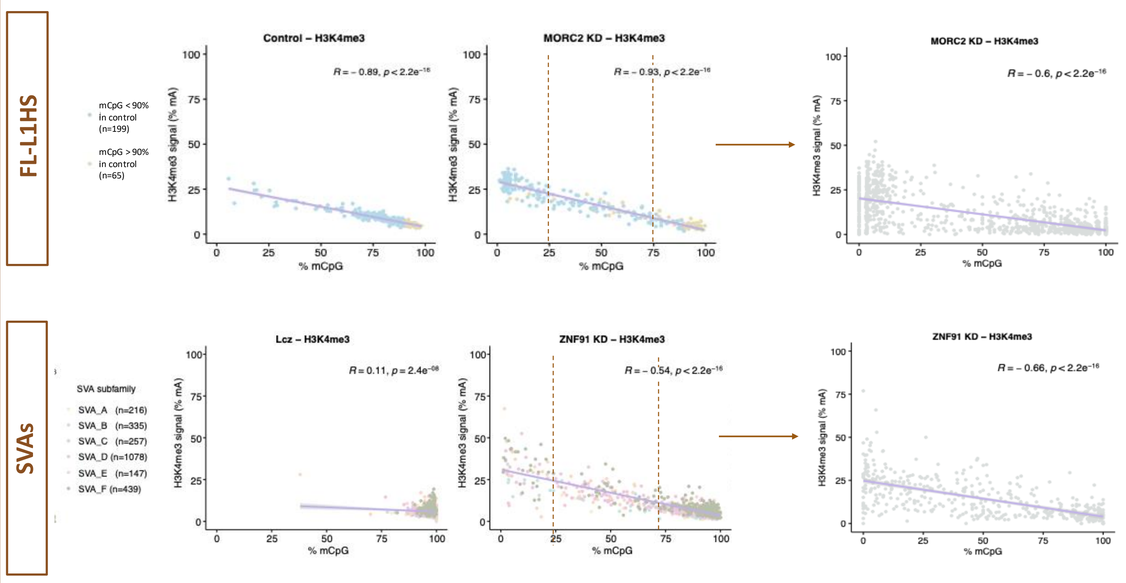


Anticorrelation of mCpG and H3K4me3 over gene promoters

Haplotype-specific mCpG and H3K4me3 information. In the example: an imprinted gene

Results

mCpG and H3K4me3 are perfectly anticorrelated over the promoter of L1HS and the body of SVAs. This anticorrelation persists at a single read level.

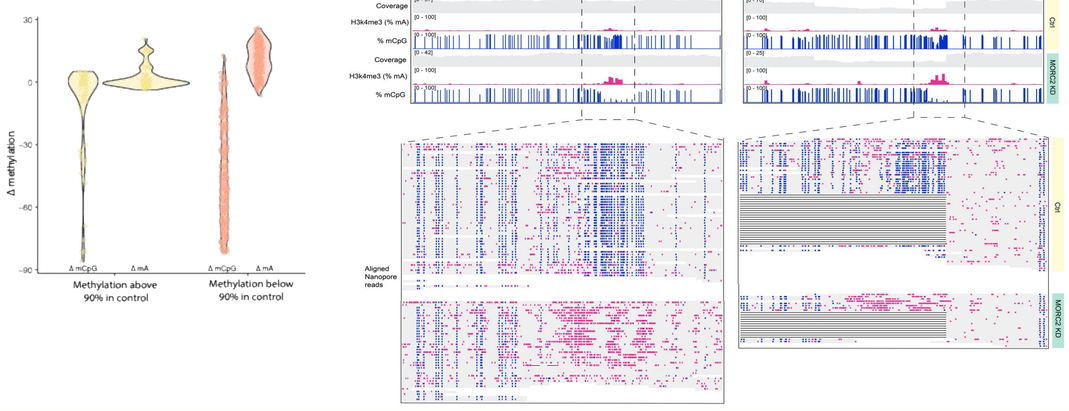


Per-element mCpG and H3K4me3 correlation over the promoter of FL-L1HS (top; n=264) and all body of SVAs (bottom; SVA > 700 bps; n=2467) . Each dot represents an individual promoter (L1HS) or an individual SVA.

Per-read mCpG and H3K4me3 correlation. Each dot represents a read spanning the entire promoter (L1HS) or element body (SVA).

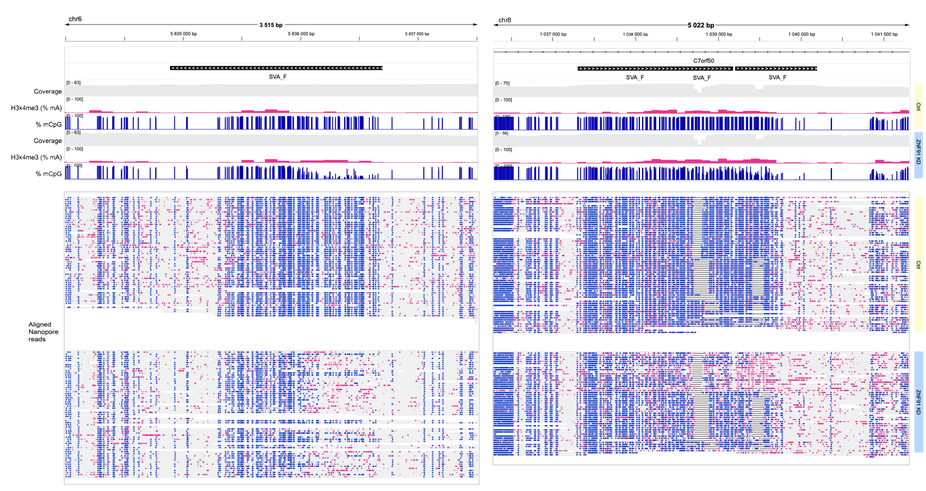
Active transcription of L1HS confers sensitivity to HUSH-MORC2 silencing.

L1HS insertion polymorphisms are under the same epigenetic control as reference L1HS.

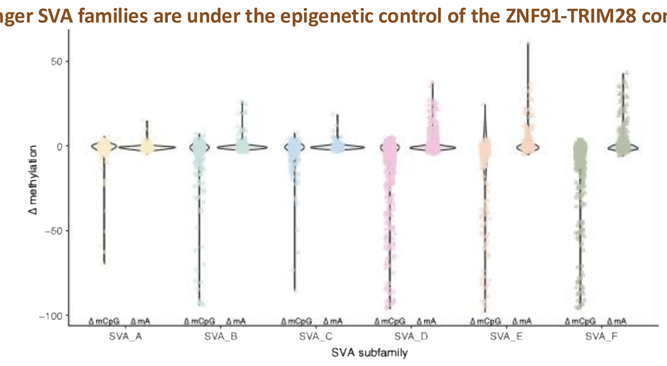


The ZNF91-TRIM28 complex controls only a subset of SVAs.

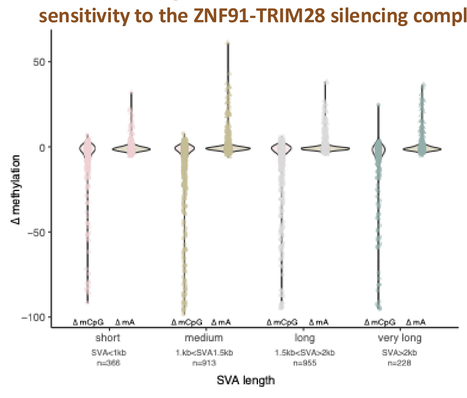
SVA length polymorphisms appear to be subjected to differential epigenetic regulation.



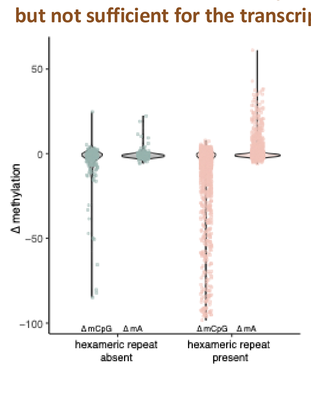
Younger SVA families are under the epigenetic control of the ZNF91-TRIM28 complex.



The overall length of the SVA does not affect sensitivity to the ZNF91-TRIM28 silencing complex.



Presence of the hexameric repeat is necessary but not sufficient for the transcription of SVAs.



Conclusions

DiMeLo-seq is a powerful tool for the study of TEs: it allows epigenetic characterisation of otherwise unmappable TEs while providing information on multiple levels of epigenetic regulation with a single locus resolution.

Future work

- Explore possible further layers of complexity of HUSH-MORC2 and ZNF91-TRIM28 recruitment.
- Use DiMeLo-seq to characterise haplotype-dependent epigenetic regulation of polymorphic TEs, i.e., polymorphic insertion and length polymorphisms.
- Apply DiMeLo-seq to complex samples such as tissue samples.

Affiliations and Funders

1. Laboratory Of Epigenetics And Chromatin Dynamics, Department Of Experimental Medical Science, Lund University, Sweden. 2. Lund Stem Cell Center, Lund University, Sweden. 3. Faculty Of Arts & Science, University Of Toronto, Canada. 4. Laboratory Of Molecular Neurogenetics, Department Of Experimental Medical Science, Lund University, Sweden

* caterina.francesconi@med.lu.se

Swedish Research Council
Chan Zuckerberg Initiative
SVENSKA SÄLLSKAPET FÖR MEDICINSK FORSKNING

References

1. International Human Genome Sequencing Consortium, Nat. 409, 860-921 (2001).
2. Jönsson et al., TiG 36(8), 610-623 (2020).
3. Rowe et al., Nat. 463(7278):237-40 (2010).
4. Turelli et al., Genome Res. 24(8):1260–1270 (2014).
5. Tchasovnikarova et al., Science 348(6242), 1481-1485 (2015).
6. Tchasovnikarova et al., Nat. Genet. 49(7), 1035-1044 (2017).
7. Douse et al., Nat. Commun. 11(1), 4940 (2020).
8. Garland et al., Mol Cell. 82(9):1691–1707.e8 (2022).
9. Khan et al., Genome Res. 6(1):78-87 (2006).
10. Altemose et al., Nat. Methods 19, 711-723 (2022).