screen. Several factors could have limited our ability to identify all genes controlling L1 retrotransposition to saturation, such as: (i) a subset of regulators may function in a cell-type specific manner not captured by either K562 or HeLa screens, (ii) essential genes with strong negative effects on cell growth may have dropped out, (iii) regulators that strictly require native L1 UTR sequences may have been missed due to our reporter design. Nonetheless, our combined screens identify many novel candidates for L1 retrotransposition control in human cells and provide a rich resource for mechanistic studies of TEs.

Select screen hits were further validated in K562 cells using a well-characterized L1-GFP reporter[18] (Extended Data Fig. 1a), confirming 13 suppressors and 1 activator (SLTM) out of 16 examined genes (Fig. 1e). Interestingly, chromatin regulators (TASOR, MORC2, MPP8, SAFB and SETDB1) suppress the retrotransposition of L1-GFP reporter, but not that of a previously described codon-optimized L1-GFP reporter (hereinafter referred to as (opt)-L1-GFP)[19,20], indicating that these factors regulate L1 retrotransposition in a manner dependent upon the native L1 ORF nucleotide sequence (Extended Data Fig. 3f,g). An additional secondary screen against the codon-optimized (opt)-L1-G418^R reporter in K562 cells confirmed the sequence-dependent feature of these L1 regulators, and systematically partitioned our top screen hits into native L1 sequence-dependent and –independent candidates (Extended Data Fig. 3h, see Table S2 for full list).

We next examined whether the identified regulators influence the expression of endogenous L1Hs, the youngest and only retrotransposition-competent L1 subfamily in humans. CRISPR-deletion of some genes (TASOR, MPP8, SAFB and MORC2) significantly increased expression of endogenous L1Hs, whereas deletion of other genes, such as SETX, RAD51 or FA complex components, had little effect (Fig. 1f). Since all interrogated genes restrict L1-GFP retrotransposition into the genome (Fig. 1e and Extended Data Fig. 4a), our results suggest that identified suppressors can function at either transcriptional or posttranscriptional level.

We further investigated three candidate transcriptional regulators of L1: MORC2, TASOR and MPP8. TASOR and MPP8 (along with PPHLN1), comprise the HUSH complex and recruit the H3K9me3 methyltransferase SETDB1 to repress genes[8]. Notably, PPHLN1 and SETDB1 also came up as L1 suppressors in our screen (Fig. 1d and Extended Data Fig. 3b). MORC2, which has recently been shown to biochemically and functionally interact with HUSH[21], is a member of the microrchidia (MORC) protein family that has been implicated in transposon silencing in plants and mice[22,23]. While MORC2/HUSH have been previously implicated in heterochromatin formation, most heterochromatin factors had no impact on L1 retrotransposition, suggesting a selective effect (Fig. 2a and Extended Data Fig. 4b).

Several independent experiments in clonal knockout (KO) K562 lines confirmed that HUSH and MORC2 suppress the retrotransposition of the L1-GFP reporter by silencing its transcription (Fig. 2b,c and Extended Data Fig. 4c-f). Additionally, HUSH/MORC2 repressed endogenous (non-reporter) L1Hs RNA and protein expression in both K562 and human embryonic stem cells[24] (hESC, H9) (Fig. 2d and Extended Data Fig. 4g-k). PolyA-selected RNA sequencing (RNA-seq) experiments revealed up-regulated expression of evolutionarily younger L1PA families (including L1Hs) upon HUSH or MORC2 KO in K562 cells (Fig. 2e). Taken together, these data demonstrate that HUSH/MORC2 silence both the reporter transgene as well as endogenous evolutionarily young L1s.

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) from K562 cells and hESCs demonstrated that MORC2, MPP8 and TASOR co-bind genomic regions characterized by specific L1 instances. Elements from the primate-specific L1P family showed higher enrichment than the older L1M family elements (Fig. 3a,b and Extended Data Fig. 5a,b, 7a,b), consistent with the preferential derepression of the former upon HUSH or MORC2 KO (Fig. 2e). Moreover, this enrichment was specific to L1s, as other major repeat classes were not enriched (Fig. 3b and Extended Data Fig. 7b), although all three proteins also targeted expressed KRAB-ZNF genes (Extended Data Fig. 5c,d). HUSH KO in K562 cells almost completely abrogated MORC2 binding at L1s (consistent with recently published observations that HUSH recruits MORC2 for transcriptional repression[21]), whereas MORC2 deletion led to a modest, but appreciable decrease of HUSH subunit binding (Extended Data Fig. 6). In mouse ESCs, MPP8 bound retrotransposition-competent L1Md-A and L1Md-T, as well as IAP elements, a class of murine endogenous retroviruses that remain currently mobile in the mouse genome (Extended Data Fig. 7c,d), suggesting that regulators uncovered by our study in human cells may in other species target additional active transposons beyond L1s.

Interestingly, even within younger human L1Ps only a subset is bound by HUSH/MORC2 in either K562 cells or hESCs, and we sought to identify genomic or epigenomic features that could explain this selectivity. We found that HUSH/MORC2 selectively target young full-length L1s, particularly the L1PA1-5 in human cells (Fig. 3c,d) and L1Md-A/T in mice (Extended Data Fig. 7e). Both MPP8 and MORC2 bind broadly across the L1: while MORC2 binding is skewed towards the 5' end, MPP8 shows higher enrichments within the body and at 3' end of L1PAs, including the L1Hs (L1PA1) elements (Extended Data Fig. 7f,g).

Nonetheless, preference for the full-length, evolutionarily younger L1PAs can only partially explain observed HUSH/MORC2 selectivity, as only a subset of such elements is targeted by the complex (Fig. 3d). We found that the additional layer of selectivity can be explained by the state of surrounding chromatin, with HUSH/MORC2-occupied L1s preferentially immersed within the transcriptionally permissive euchromatic environment marked by modifications such as H3K4me3 and H3K27ac (Fig. 3e). In agreement, HUSH/MORC2-bound L1s are enriched within introns of actively transcribed genes (Extended Data Fig. 8a,b). Furthermore, although most HUSH/MORC2-bound L1s are concordant between K562 and hESCs, those that are bound in a cell type-specific manner tend to be associated with genes that are differentially active between the two cell types (Extended Data Fig. 8c). To understand the role of transcription in HUSH/MORC2 targeting of L1s, we investigated MORC2 and MPP8 occupancy at the inducible L1 transgene. We observed increased binding of these factors upon transcriptional induction (Extended Data Fig. 8d), suggesting that transcription through L1 sequences facilitates HUSH/MORC2 binding. Taken together, HUSH/MORC2 selectively target young, full-length L1s located within transcriptionally permissive euchromatic regions, which are precisely the elements that pose the highest threat to genome integrity, as a subset of them remains mobile and transcription is the first step of L1 mobilization.

Despite their immersion within the euchromatic environment, HUSH/MORC2-bound L1s themselves are heavily decorated with the transcriptionally repressive H3K9me3 (Fig. 3e), consistent with the role of HUSH in facilitating H3K9me3 deposition at target sites[8]. HUSH/MORC2 KO decreased H3K9me3 level preferentially at L1 versus non-L1 HUSH/MORC2 genomic targets, and at bound versus unbound L1s (Fig. 4a and Extended Data Fig. 9a,b). Since HUSH/MORC2-bound L1s are significantly enriched within introns of transcriptionally active genes (Extended Data Fig. 8a-c), we examined whether HUSH/MORC2 recruitment and its associated H3K9me3 deposition can influence chromatin modification and expression of the host genes. Despite the transcriptionally active status (Extended Data Fig. 8a,b), promoters and especially bodies of genes harboring MORC2/HUSH-bound L1s show appreciable levels of H3K9me3. This enrichment is substantially diminished in the KO lines (Extended Data Fig. 9c) with the concomitant upregulation of genes harboring MORC2/HUSH-bound L1s, but not those with unbound intronic L1s (Fig. 4b). Thus, HUSH/MORC2 binding at intronic L1s leads to a modest, but significant down-regulation of the active genes that harbor them (Fig. 4c and Extended Data Fig. 9d-g, 10a).