SAMMY-seq reveals early alteration of heterochromatin and deregulation of bivalent genes in Hutchinson-Gilford Progeria Syndrome

Endre Sebestyén, Fabrizia Marullo, Federica Lucini, Cristiano Petrini, Andrea Bianchi, Sara Valsoni, Ilaria Olivieri, Laura Antonelli, Francesco Gregoretti, Gennaro Oliva, Francesco Ferrari & Chiara Lanzuolo

Hutchinson-Gilford progeria syndrome is a genetic disease caused by an aberrant form of Lamin A resulting in chromatin structure disruption, in particular by interfering with lamina associated domains. Early molecular alterations involved in chromatin remodeling have not been identified thus far. Here, we present SAMMY-seq, a high-throughput sequencing-based method for genome-wide characterization of heterochromatin dynamics. Using SAMMY-seq, we detect early stage alterations of heterochromatin structure in progeria primary fibroblasts. These structural changes do not disrupt the distribution of H3K9me3 in early passage cells, thus suggesting that chromatin rearrangements precede H3K9me3 alterations described at later passages. On the other hand, we observe an interplay between changes in chromatin accessibility and Polycomb regulation, with site-specific H3K27me3 variations and transcriptional dysregulation of bivalent genes. We conclude that the correct assembly of lamina associated domains is functionally connected to the Polycomb repression and rapidly lost in early molecular events of progeria pathogenesis.
Electron microscopy imaging of eukaryotic nuclei shows chromatin compartments with different levels of compaction, known as euchromatin and heterochromatin (reviewed in refs. 1–2). The more accessible and less condensed euchromatin is generally enriched in expressed genes. Instead, heterochromatin contains highly condensed DNA, including pericentric and telomeric regions1–5, and genomic regions with unique packaging properties maintained by the Polycomb-group proteins (PcG)6. These proteins are developmentally regulated factors acting as parts of multimeric complexes named Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2)7.

Another key player of chromatin organization is the nuclear lamina (NL). NL is mainly composed of A- and B-type lamins, intermediate filaments of 3.5 nm thickness which form a dense network at the nuclear periphery8,9. NL preferentially interacts with the genome at specific regions called Lamina-Associated Domains (LADs), with sizes ranging from 100 kb to 10 Mb10. LADs are enriched in H3K9me2 and H3K9me3 histone modifications that are typical of inactive heterochromatic regions. LAD borders are marked by the PRC2-dependent H3K27me3 histone mark11–13, which is characteristic of inactive PcG-regulated chromatin regions. The ensemble of lamins, chromatin marks, and PcG factors around LADs creates a repressive environment14,15, with heterochromatin and PcG target regions adjacent to each other16,17. In line with these observations, we previously demonstrated that Lamin A functionally interacts with PcG18–20, as later also reported by others17,21–23.

In physiological conditions, the association or detachment of LADs from the NL are related to the spatiotemporal regulation of gene expression in cell differentiation processes13,24,25. The crucial function of such dynamics is attested by an entire class of proteins (PcG)6. These proteins are developmentally regulated factors acting as parts of multimeric complexes named Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2)7.

In recent years, the widespread adoption of experimental techniques based on high-throughput sequencing (NGS) has been instrumental in advancing the knowledge of chromatin structure and function33,34. Several of these techniques were originally designed to map active chromatin regions, including ChIP-seq35,36, ATAC-seq36, DNase-seq37, and MNase-seq38. On the other hand, a limited set of options is available for genome-wide mapping of heterochromatin and LADs. The most commonly used ones include ChIP-seq or DamID-seq39 targeting NL components or ChIP-seq for the heterochromatin-associated histone marks (H3K9 methylation). Such techniques suffer major limitations as DamID-seq relies on the exogenous expression of a transgene and it cannot be applied to primary cells or tissues. Instead, ChIP-seq relies on crosslinking and antibodies, which are sources of technical biases and potential issues of sensitivity or cross-reactivity, respectively. More recently proposed high-throughput sequencing methods for genome-wide mapping of LADs and heterochromatin, such as gradient-seq40,41 and protect-seq42, also rely on crosslinking.

Here, we present a high-throughput sequencing-based method to map lamina-associated heterochromatin regions. We named our technology SAMMY-seq (Sequential Analysis of Macromolecules accessibility), as a tribute to Sammy Basso, the founder of the Italian Progeria Association, a passionate and inspiring advocate of research on laminopathies. SAMMY-seq relies on the sequential isolation and sequencing of multiple chromatin fractions, enriched for differences in accessibility. Here we show that our method is robust, fast, easy-to-perform, it does not rely on crosslinking or antibodies, and it can be applied to primary cells.

In healthy human fibroblasts, we can reliably identify genomic regions with heterochromatin marks (H3K9me3) and LADs. Of note, other high-throughput sequencing methods relying on high salt buffers to extract specific chromatin fractions have been proposed in the literature, including “salt fractions profiling”43 and HRS-seq44. However, they adopt different conditions, enzymes, experimental protocol steps (Supplementary Table 1), and cellular models. Moreover, they have been primarily adopted to analyze accessible euchromatin regions, whereas SAMMY-seq allows the study of the condensed heterochromatin domains, thus they are not directly comparable to our technology.

Using SAMMY-seq in early-passage HGPS fibroblasts, we found patient-specific alterations of heterochromatin architecture, but this is not accompanied by H3K9me3 changes or transcriptional activation of silent LAD. Interestingly, we also identified a global decrease of the PRC2-dependent H3K27me3 mark deposition, suggesting an interplay between differentially compacted domains and PcG regulation. These Polycomb-dependent alterations affect the transcriptional regulation of bivalent genes, more susceptible to H3K27me3 variations18,45. Globally, our findings indicate that chromatin structural changes are an early event in HGPS nuclear remodeling and interfere with proper PcG control.

Results

SAMMY-seq maps lamina-associated heterochromatic regions.

We developed the SAMMY-seq method for genome-wide mapping of chromatin fractions separated by accessibility. The method is based on the sequential extraction of distinct nuclear fractions19,46,47 containing: soluble proteins (S1 fraction); DNase-treated chromatin, i.e., the supernatant obtained after DNase treatment (S2 fraction); DNase-resistant chromatin extracted with high salt buffer (S3 fraction); and the most condensed and insoluble portion of chromatin, extracted with urea buffer that solubilizes the remaining proteins and membranes (S4 fraction) (Fig. 1a and Supplementary Fig. 1a). We further adapted the method to leverage high-throughput DNA sequencing for genome-wide mapping of the distinct chromatin fractions (see “Methods” section).

We first applied SAMMY-seq to three independent skin primary fibroblast cell lines, originating from three different healthy individuals (control samples). The average number of sequenced reads was 72 million for S2 fraction, 62 million for S3 fraction, and 65 million for S4 fraction (Fig. 1b and Supplementary Fig. 1b, Supplementary Table 2). On average 58% (S2), 54% (S3), and 52% (S4) of the genome were covered by at least one read (Supplementary Fig. 1c). The pairwise Spearman correlation between control samples (Supplementary Fig. 1d) generally increases when summarizing read counts over larger genomic
bins and reaches a plateau for S3 and S4 fractions (Supplementary Fig. 1e). Around 1 Mb resolution, the mean Spearman correlation between control samples is 0.27 for the S2 fraction, 0.92 for S3, and 0.79 for S4. Overall, these results show that S3 and S4 fractions show high similarity whereas the S2 fraction is more variable across biological replicates of control samples (Supplementary Fig. 1d).

In line with this analysis, the visual inspection of read coverage profiles shows megabase-scale “bumps” in S3 and, even more prominently, in S4 fractions (Fig. 1b). It is worth remarking that the S2 profile is not comparable to the standard DNase-seq profile due to substantial differences in the digestion conditions applied here (see “Methods” section). To detect relative enrichment of specific genomic regions between chromatin fractions, we applied...
the Enriched Domain Detector (EDD)\textsuperscript{48}, an algorithm designed to identify broad regions of enrichment in high-throughput sequencing data. We identified differentially enriched genomic regions by pairwise comparisons of less accessible vs more accessible fractions, the latter used as a baseline in the comparison: i.e., S3 vs S2, S4 vs S3, and S4 vs S2 comparisons (Fig. 1c). We found a well-defined set of regions (SAMMY-seq domains) consistently enriched in the less accessible fractions. The S4 vs S2 comparison is the most consistent with 70.18\% conservation between sample pairs on average (Supplementary Fig. 1f). These results are robust with respect to sequencing depth, as we still identify on average 82.09\% of the original SAMMY-seq domains after down-sampling to 50\% of the sequencing reads (Supplementary Fig. 1g). Further down-sampling to 25\% of reads still led to an average of 61.92\% of conserved domains, with S4 vs S2 still showing the most consistent results (Supplementary Fig. 1g). SAMMY-seq domains have an average size of 2.26 Mb and cover on average 527 Mb (Supplementary Fig. 1h, i and Supplementary Table 3). Based on the observations above, we mostly focused on S4 vs S2 comparison in the subsequent analyses, although similar conclusions can be obtained with other comparisons as well.

To functionally characterize the SAMMY-seq enriched regions, we compared the genome-wide read coverage with reference chromatin mark profiles for the same cell type. Using data from the Roadmap Epigenomics consortium\textsuperscript{49} and other public data sets\textsuperscript{30,48,50–52}, we noticed that the SAMMY-seq signal (S4 vs S2 enrichment) is inversely correlated with open chromatin marks (ATAC-seq and DNase-seq) and is positively correlated with Lamin A/C and B1 ChIP-seq signal (Fig. 2a). We confirmed this is a consistent genome-wide pattern using StereoGene kernel correlation\textsuperscript{53}, a method for the unbiased comparison of different types of chromatin marks (Fig. 2b and Supplementary Fig. 2a). Despite some inter-dataset variability across a comprehensive set of independent Lamin ChIP-seq profiles (Lamin A/C and Lamin B1) for human fibroblasts\textsuperscript{30,48,50–52}, both Lamin A/C and Lamin B1 show similar correlation values (Supplementary Fig. 2a), thus only their average is summarized in (Fig. 2b). Similarly, when considering histone modifications, we observed a positive association with heterochromatin (H3K9me3) and an inverse correlation with active chromatin (H3K4me1) as well as Polycomb-regulated chromatin (H3K27me3) (Fig. 2a, b). The latter observation is compatible with previous reports that H3K27me3 is located at the border of heterochromatic LAD regions\textsuperscript{11–13,17}. However, this should not be interpreted as a co-localization of H3K27me3 and H3K4me1. They are simply concordant in being absent from H3K9me3 enriched SAMMY-seq domains on a large scale. Indeed, a closer inspection of histone mark profiles shows that H3K4me1 peaks are located at positions of relatively lower H3K27me3 enrichment, and vice versa (Fig. 2a). Overall, these observations suggest a similarity between SAMMY-seq and epigenomic profiles for heterochromatin or lamina association.

To further quantify this association, we compared SAMMY-seq domains with Lamin A/C LADs, thus observing an average 0.49 Jaccard Index overlap, when considering all samples and fractions comparisons (Supplementary Table 4). Each of these pairwise comparisons has a significantly high overlap, as assessed by randomizing either LADs or SAMMY-seq domains (Fig. 2c and Supplementary Fig. 2b). On average 73\% of SAMMY-seq domains (S4 vs S2 fractions comparison) overlap with Lamin A/C LADs (Supplementary Table 4). Instead, in the complementary comparison, we observed that 62\% of Lamin A/C LADs overlap with SAMMY-seq domains (S4 vs S2) (Supplementary Fig. 2c). This observation confirms that LADs cover a larger fraction of the genome compared to SAMMY-seq domains. The first percentage would be the SAMMY-seq precision in identifying LADs, whereas the latter would be the sensitivity in calling LADs. In a further analysis, we found that the sensitivity in calling regions that are both heterochromatic (H3K9me3) and associated with the lamina (LADs) is actually higher (75\% on average for S4 vs S2 comparison) than the sensitivity observed when considering only one of the two markers independently (Supplementary Fig. 2c). Overall, these results confirm SAMMY-seq preferential enrichment for the most compact heterochromatin, such as the lamina-associated heterochromatic regions.

It must be noted that SAMMY-seq overcomes specific limitations of other high-throughput sequencing-based methods for genome-wide mapping of LADs and heterochromatin, thus allowing a wide range of applications (Supplementary Table 5). Most notably, SAMMY-seq does not rely on crosslinking, antibodies, or the overexpression of a transgene, thus reducing the sources of biases and extending its applicability. A key feature of SAMMY-seq, that will allow its adoption to a wide range of applications, is the possibility to analyze a small number of cells. To this concern, as a proof of concept, we performed a scale-down experiment by applying SAMMY-seq on a control fibroblast cell line (CTRL004) starting from 250k, 50k, or 10k cells. To further prove the robustness of the technique, the experiment was performed on two replicates by two independent experimentalists. The SAMMY-seq profiles obtained in the scale-down experiment confirm the association of S4 fraction with heterochromatin and lamina (Supplementary Fig. 3a) as well as the good reproducibility of the profile across replicates and scale-down tests. These patterns are confirmed by the genome-wide pairwise correlation between SAMMY-seq samples, as well as their correlation with H3K9me3, and negative correlation with H3K4me1 and H3K27me3 (Supplementary Fig. 3b). The pairwise overlap of scale-down SAMMY-seq domains (Jaccard Index) also confirms the robustness of results (Supplementary Fig. 3c). Globally, these results indicate that SAMMY-seq is a versatile technique that can reliably identify chromatin regions separated on their accessibility.

**Chromatin changes in early-passage progeria fibroblasts.** We applied SAMMY-seq on early-passage HGPS primary fibroblasts, generally used as a cellular model to study early molecular alterations in progeria\textsuperscript{30,54}. Indeed, progerin is known to be resistant to protein degradation within cells\textsuperscript{55} and to accumulate during cell passaging (Supplementary Fig. 4a). According to the literature, primary HGPS fibroblasts entering senescence show alterations in DNA interactions with Lamin A/C\textsuperscript{30}. However, Hi-C experiments for genome-wide mapping of 3D chromatin architecture could not detect alterations in the pairwise contact frequency between genomic loci in early-passage cells. These were evident only at later-passage cells, representing more advanced stages of senescence when cells exhibit severe nuclear blebs and inhibited proliferation\textsuperscript{30}.

We confirmed that early-passage HGPS fibroblasts do not show senescence-associated features such as beta-galactosidase positivity (Supplementary Fig. 4b), reduction in the proliferation rate (Supplementary Fig. 4c), or changes in nuclear area and morphology (Supplementary Fig. 4d–f). We applied SAMMY-seq to investigate chromatin changes in early-passage skin fibroblasts from 3 independent HGPS patients (passage number from 10 to 12) (Supplementary Fig. 1b, c). SAMMY-seq domains (S4 vs S2) appeared more variable in number and dimension across samples with respect to controls (Supplementary Fig. 1h–i), and scattered and with a generally lower signal-to-noise ratio (Fig. 3a) in HGPS, as opposed to the more consistent overlap across control replicates. This would be compatible with a general loss of
Fig. 2 SAMMY-seq preferentially isolates heterochromatic domains. a Visualization of SAMMY-seq enrichment signal along with multiple chromatin marks on a representative region (12 Mb region in chr5:158000000-170000000). From top to bottom: tracks for open chromatin (ATAC-seq and DNase-seq – orange and brown); association to the lamina (Lamin A/C and Lamin B1 ChIP-seq – dark blue and blue); ChIP-seq for histone marks associated to PcG regulation (H3K27me3 – light green), active chromatin (H3K4me1 – dark green) or heterochromatin (H3K9me3 – light blue); SAMMY-seq enrichment signal (S4 vs S2) in 3 control samples at indicated passages (p19-20) (blue or yellow colored track for enrichment or depletion, respectively, over the S2 reference baseline). Gray boxes under each SAMMY-seq track show the SAMMY-seq domains. For ChIP-seq data, the (ChIP – input) reads distribution was computed with SPP package and we are limiting the y axis range to zero as a minimum value. See the “Methods” section for details of each type of signal processing. b Genome-wide kernel correlation (StereoGene) for SAMMY-seq of all control samples at indicated passages (p19-20) (S4 vs S2) compared to ATAC-seq, DNase-seq, and ChIP-seq (Lamin A/C, Lamin B1, H3K27me3, H3K4me1, and H3K9me3). Here we are presenting a subset of data reported in Supplementary Fig. 2a containing a more extended and detailed comparison. For Lamin A/C and Lamin B1, we are reporting the average correlation across multiple data sets (see Supplementary Fig. 2a for individual data sets). c Overlap (Jaccard Index - JI) of three biologically independent (labels on top) control samples SAMMY-seq domains (S4 vs S2) with Lamin-Associated Domains (LADs) (Lamin A/C or Lamin B1 ChIP-seq enrichment domains). For Lamin A/C and Lamin B1, we used a consensus set of regions for each (see “Methods” section). The observed JI (blue bar) is compared to mean JI across 10,000 randomizations of SAMMY-seq domains (pink bar) or LADs (yellow bar) positions along the genome to compute empirical p-values for the probability of the observed JI being larger than the random expectation. All one-tail empirical p-values were significant (p < 0.0001). The whiskers over the randomized values show the mean ± 2 SEM interval.
chromatin organization following the lamina structural alteration, thus resulting in a more random distribution of genomic regions among the different chromatin fractions. Indeed, SAMMY-seq domains are largely overlapping among control samples (average Jaccard Index 0.67) whereas SAMMY-seq domains are more variable in progeria samples, most notably in the S4 vs S2 comparison (average Jaccard Index 0.11) (Fig. 3b and Supplementary Table 6). Moreover, SAMMY-seq domains in control samples fall within regions that are normally enriched for lamins and H3K9me3, whereas this pattern is lost in HGPS SAMMY-seq samples (Fig. 3a). We confirmed this is a genome-wide trend by analyzing chromatin marks at SAMMY-seq domain borders (Fig. 3c). SAMMY-seq domains from control samples (S4 vs S2, S4 vs S3, S3 vs S2) correspond to regions normally enriched in Lamin A/C, Lamin B1 and H3K9me3 ChIP-seq reads, as opposed to flanking regions outside of domain borders. Progeria SAMMY-seq domains (S3 vs S2) are similar to controls. However, when considering a more compacted S4 fraction, the association to reference controls for LADs and heterochromatin marks is lost in 2 out of 3 patients for the S4 vs S2, and 3 out of 3 patients for the
accumulation dependent inhibition of the cell cycle and DNA damage in late-passage HGPS cells can be explained by the progerin-chromatin analysis. Nevertheless, we noticed a partial recovery of the control pattern in the S4 vs S2 comparison, as confirmed by genome-wide correlation (Supplementary Fig. 5c) and association with heterochromatin marks (Supplementary Fig. 5d). Of note, the S4 vs S3 comparison still highlighted consistent differences between the control and HGPS profiles (Supplementary Fig. 5e–h), suggesting that SAMMY fractions give rise to complementary information in chromatin analysis.

The seemingly partial recovery of the normal fibroblasts pattern in late-passage HGPS cells can be explained by the progerin-dependent inhibition of the cell cycle and DNA damage accumulation generally resulting in more apoptosis in late-passage HGPS cells. Thus, prolonged culture may actually counter-select the cells with more abundant progerin. Moreover, the permeabilization of cell and nuclear membranes in the initial step of SAMMY-seq protocol would result in disrupting apoptotic bodies and discard their content, thus further under-representing the more damaged cells. It is worth noting that other chromatin analysis methods starting with a fixation step (e.g., crosslinking) would not equally clear apoptotic bodies, thus confounding the results.

H3K9me3 patterns do not change in early-passage HGPS cells. To further investigate the epigenetic regulation in early-passage HGPS fibroblasts, we examined H3K9me3 distribution by ChIP-seq on the same set of control and HGPS fibroblasts analyzed by SAMMY-seq. H3K9me3 profiles were mostly unchanged in early-passage HGPS, despite the chromatin remodeling detected by SAMMY-seq (Fig. 4a). Quantitative analyses confirmed that H3K9me3 enriched regions are largely conserved between controls and HGPS cells, with overlaps ranging between 69 and 80% (Fig. 4b). Only controls exhibit a correlation between SAMMY-seq and H3K9me3, as indicated by the analysis of H3K9me3 pattern at SAMMY-seq domain borders (Supplementary Fig. 6a) as well as by the Jaccard index analysis (Fig. 4c and Supplementary Fig. 6b).

In line with genomic data, at the cellular level, early-passage HGPS cells and controls showed comparable H3K9me3 global levels measured by western blot (Supplementary Fig. 6c, d). Heterochromatin foci were detected in both controls and early-passage HGPS fibroblasts by immunofluorescence analysis with H3K9me3 antibody, revealing high variability across samples (Fig. 4d) but without significant differences in their number per cell (Supplementary Fig. 6e). Instead, their intranuclear localization showed a shift towards the nuclear center in HGPS (Fig. 4d, e), supporting the hypothesis of heterochromatin detachment from the nuclear periphery (Fig. 3). Notably, as previously reported, later passages (p20) HGPS fibroblasts showed an aberrant pattern of H3K9me3 foci (Fig. 4d), further indicating that early-passage chromatin remodeling in HGPS precedes H3K9me3 loss.

Chromatin changes do not affect the expression of LAD genes. Using RNA-seq on the same set of control and HGPS fibroblasts, we assessed the expression at the gene or transcript-level detecting differential expression of 257 genes or 256 transcripts, respectively (Supplementary Fig. 7a) (see “Methods” section). Thus, early stages of senescence in HGPS cells do not cause large-scale changes in the transcriptome. To clarify whether chromatin structural changes detected with SAMMY-seq have an effect on transcription, we examined the expression of protein-coding genes located in SAMMY-seq domains. As expected, genes within SAMMY-seq domains (S4 vs S2) have lower expression levels than those outside in all control samples (Supplementary Fig. 7b). More specifically, in the genomic regions around SAMMY-seq domain borders, we noticed a clear pattern with a transition from lower expression within the domain to higher expression outside in all of the control samples (Supplementary Fig. 7c). A similar pattern was observed in HGPS samples when considering the same genomic regions (Supplementary Fig. 7b, c). This suggests that heterochromatin detachment from NL seen in HGPS (Figs. 3 and 4e) was not sufficient to trigger gene expression deregulation inside SAMMY-seq domains, where H3K9me3 is still present and largely unchanged as shown above (Fig. 4a, b).

The transcriptional activity within H3K9me3 domains of each sample is repressed in HGPS as well as in controls, thus further confirming that epigenetic silencing is still preserved in these regions (Supplementary Fig. 7d).

HGPS chromatin remodeling affects PcG bivalent genes. As we noticed that each HGPS sample has unique SAMMY-seq domains (Fig. 3), we compared individual HGPS expression profiles against the group of three controls to account for sample-specific patterns (see “Methods” section). To verify which biological functions are affected, we performed a pathway enrichment analysis using the MSigDB database and selected the pathways commonly deregulated in all HGPS samples (Fig. 5a). Many of them are related to stem cells and cancer, in line with the hypothesis that accelerated aging affects the stem cell niche. Interestingly, we also found several pathways related to PcG regulation (H3K27me3 and PRC2) (Fig. 5a), in line with previous studies suggesting crosstalk between lamin and PcG17-19,21-23. To further investigate PcG role in early-passage HGPS fibroblasts, we examined the total amounts of PRC1 (Bmi1 subunit) and PRC2 (Ezh2 subunit) at the protein level (Supplementary Fig. 8a, b), as well as H3K27me3 (Supplementary Fig. 8c, d). None of them showed significant differences between control and HGPS samples, despite some degree of inter-individual variability. We also checked the intranuclear architecture of Ezh2 (Supplementary Fig. 8d, e) and H3K27me3 foci (Supplementary Fig. 8g, h) and we found that their number does not change at early-passage HGPS cells.

Finally, we performed ChIP-seq to dissect H3K27me3 distribution at the genome-wide level. By visual inspection, we noticed that the main enrichment peaks seem to be present in both controls and HGPS samples, but the H3K27me3 signal was slightly spreading over flanking regions in HGPS (Fig. 5b). To quantify this pattern, we considered the annotated transcription start site (TSS) for protein-coding genes where H3K27me3 was detected in controls. We examined the H3K27me3 average profile around these TSS regions, where we detected generally lower enrichment in HGPS samples, as well as a less marked relative decrease in TSS distal vs proximal positions, yet not statistically significant due to variability across patients (Supplementary Fig. 9a, b). This observation resembles the findings of H3K27me3 spreading described in another laminopathy model. The progeria-associated differences in the H3K27me3 profile were also confirmed by multi-dimensional scaling of ChIP-seq data (Fig. 5c), showing that control samples are close to each
other and distant from the progeria samples, the latter also showing larger distances between them.

To further dissect the role of chromatin regulators in the disrupted expression patterns, we examined the overlap of SAMMY-seq domains with the chromatin states of healthy skin fibroblasts. SAMMY-seq domains of control samples are mostly overlapping "Quiescent" or "Heterochromatin" states (Fig. 5d), whereas SAMMY-seq domains of HGPS cells comprise regions normally transcribed or regulated by PcG proteins in healthy skin fibroblasts. To analyze the correlation between gene expression, chromatin states, and altered genome structure, we grouped SAMMY-seq domain genes on the basis of their chromatin state in healthy skin fibroblasts and examined their expression difference between control and HGPS cells (Fig. 6a and Supplementary Fig. 9c). This analysis revealed that PcG-regulated and bivalent genes are deregulated at the transcriptional
level in progeria with respect to controls. To evaluate whether bivalent genes are especially affected in progeria, we performed a Fisher test: we found 76 bivalent out of 257 differentially expressed genes in HGPS (Fisher test p-value < 2.2e−16); or 39 bivalent out of 144 upregulated genes in HGPS (Fisher test p-value 4.32e−12). We thus analyzed the H3K27me3 profile of the subset of bivalent genes upregulated in HGPS, where we observed a clear drop in the H3K27me3 signal (Fig. 6b). These findings overall suggest that early chromatin remodeling has an impact on a subset of PcG targets, i.e., the bivalent genes, more susceptible to variations of PcG occupancy45,61,62, as observed also at the level of H3K27me3 distribution by ChIP-seq (Fig. 6c).

To verify the presence of PcG-progerin crosstalk, we used the proximity ligation assay (PLA), which detects interactions between proteins in close proximity (<30 nm)63. We found that Lamin A/C interacts with PcG proteins in control or HGPS fibroblasts (Fig. 6d, e), and that progerin interacts with both Ezh2 and Bmi1 in late-passage HGPS (Fig. 6d–e). To further investigate the aberrant PcG-progerin interaction in HGPS progression, we tested PcG protein compartmentalization in both early and late-passage HGPS (Fig. 6f, g). When comparing controls and early-passage HGPS fibroblasts, we did not observe major differences in PcG nuclear distribution across chromatin fractions. On the other hand, PcG proteins were predominantly localized in the S4 fraction at late-passage HGPS. Altogether, our results indicate that an increased amount of progerin may alter PcG protein’s nuclear compartmentalization and function.

Discussion

Chromatin within the cell nucleus has a complex structure, which is fundamental for gene expression regulation64. In particular, heterochromatin is highly compacted and shows specific nuclear compartmentalization. LADs, along with pericentricromeric and telomeric regions, are crucial for heterochromatin organization and for preserving chromosome structure. Alterations in heterochromatin are associated with developmental defects and cancer, while its proper conformation is a hallmark of healthy cells. As such, reliable methods to characterize heterochromatin and its alterations are essential for the biomedical scientific community.

Here we present a technology called SAMMY-seq, for genome-wide characterization of lamina-associated heterochromatic regions (Fig. 1). The protocol is based on the sequential extraction of multiple chromatin fractions, corresponding to increasingly compacted and less accessible chromatin regions, which are then mapped by high-throughput sequencing. All of the three SAMMY fractions are sequenced as they convey complementary information about chromatin structure, which may have fundamental differences across distinct cell types67. SAMMY-seq can reproducibly and reliably identify lamina-associated heterochromatic regions as the less accessible portions of chromatin in control wild-type fibroblasts (Fig. 2). Moreover, SAMMY-seq represents a significant improvement in the field of chromatin characterization as it provides information complementary and not overlapping with other high-throughput sequencing-based methods commonly used to study chromatin structure and function (Supplementary Tables 1 and 5). Of note, our protocol overcomes several major limitations of other methods for mapping lamina-associated heterochromatic regions. First of all, the procedure can be applied to primary cells, as it does not require exogenous gene expression as in DamID-seq. Then, SAMMY-seq does not involve chemical modifications of chromatin, which might cause artifacts and biases in sequencing. Additionally, it does not rely on antibodies for enriching specific chromatin fractions, thus avoiding issues related to antibody specificity, production, lot-variability, and cross-reactivity. This is particularly important when studying chromatin changes in cells where protein levels of chromatin-associated factors could be altered, thus allowing more flexibility in terms of experimental design compared to antibody-based techniques. Finally, SAMMY-seq is robust, as it yields reproducible results even at lower sequencing depth (Supplementary Fig. 1g) or with a small number of starting cells (Supplementary Fig. 3) and requires only about 3 h of bench work, excluding DNA extraction and library preparation. Given its characteristics, SAMMY-seq is a versatile technique that can be applied to different experimental systems, including primary cells and rare cell populations.

In early-passage HGPS fibroblasts, SAMMY-seq allowed us to detect changes in chromatin distribution across nuclear compartments characterized by distinct accessibility (Figs. 3 and 6h), despite the recent failure to do so by Hi-C. Indeed, Hi-C is designed to measure pairwise interactions between genomic loci, but it is not able to distinguish the intranuclear compartments where such interactions occur. Remarkably, changes in chromatin accessibility observed in early-passage HGPS are not accompanied by H3K9me3 alterations or transcriptional deregulation in the same genomic regions (Fig. 4 and Supplementary Figs. 6 and 7). Previous studies revealed that chromatin tethering at nuclear lamina and histone marks deposition are independent processes and that nuclear relocation is not sufficient to trigger a transcriptional switch. Overall, our findings support this scenario also in the HGPS model and further indicate that chromatin remodeling detected by SAMMY-seq precedes H3K9me3 decrease observed at later passages44 (Fig. 4d).

On the other hand, we found that chromatin accessibility changes in HGPS functionally affect PcG-regulated genes, by
interfering with H3K27me3 genomic distribution (Fig. 5). The crosstalk between Lamin A and Polycomb, first reported by our- selves19, is now supported by several independent studies17,18,21 – 23. The alteration of Polycomb regulation may be expected to first affect the so-called “bivalent” genes, which are already in an intermediate state, mostly silenced but primed for activation by the presence of both repressive (H3K27me3) and activating (H3K4me3) chromatin marks45,61,62. In fact, the bivalent genes, involved in stemness maintenance and cell identity specification, are mostly affected in early-passage HGPS fibroblasts (Fig. 6).
Finally, it is worth remarking that this model would provide a mechanism for the previously postulated idea that the reduced potential for tissue homeostasis, commonly associated with progeria, is due to a misregulation of stem cell compartments.

Methods

Cell cultures. Primary fibroblast cell lines were cultured in DMEM High glucose with 1× Hyclone-supplemented with 15% FBS and 1% Pen/Strep. HGADFN167 (HGSPS164)-4 years old, HGADFN167 (HGSPS167)-8 years old, HGADFN169-8 years old, HGADFN188 (HGSPS188)-2 years old, HGADFN271 (HGSPS271) -1-year-old human dermal fibroblasts derived from HGSP patients were provided by the Progeria Research Foundation (PRF). AG04989 (CTRL001) -1-year-old and AG07095 (CTRL002)-2-years-old human dermal fibroblasts were obtained from the Correll Institute. Preputial fibroblast strain #2294 (CTRL004)-4 years old was a generous gift from the Laboratory of Molecular and Cell Biology, Istituto Dermopatico dell’Immunoa di (IDI)-IRCCS, Rome, Italy, while control dermal fibroblast CTRL013-13 years old was kindly provided by the Italian Laminopathies Network.

Histochemistry, immunofluorescence assay, and PLA analysis. Beta-galactosidase assay: Cells were fixed at room temperature for 10 min in paraformaldehyde at 4% and were washed twice with 1×PBS, and incubated with fresh staining solution at 37 °C for 16 h in dark room. Then cells were washed twice with 1×PBS, overlaid with 70% glycerol/PBS, and photographed.

β-strand (5-Bromo-2-deoxy-uridine, Sigma B9285) labeling: Cells were grown for 8 h in the presence of 50 μM of BrdU, and then fixed in 4% PFA. After Triton X-100 treatment, cells were incubated for 2 min at RT in 0.07 N NaOH, briefly rinsed twice in PBS, and blocked in PBS/1% BSA. Reaction with BrdU antibody (1:10, Becton Dickinson 347580) was performed at room temperature for 1 h in a PBS solution containing 1% BSA. Immunofluorescence assay: Coverslips were fixed with paraformaldehyde at 4% in PBS for 10 min. Then, cells were permeabilized with 0.5% Triton X-100 in PBS and non-specific signals were blocked with 1% BSA in PBS for 30 min at room temperature. The following antibodies were used: Bmi1 (Millipore 05-637, mouse) diluted 1:100; Lamin A/C (Santa Cruz sc-6215, goat) diluted 1:200; Eh2 (Cell signaling AC22 3147 S, mouse) diluted 1:100; H3K9me3 (Abcam ab8898, rabbit) diluted 1:500; H3K27me3 (Millipore 07-449, rabbit) diluted 1:100. Incubation was performed at 12–16 h at 4 °C or at room temperature for 2 h for Lamin A/C. Primary antibodies were diluted in a PBS solution containing BSA 1%. Cells were stained with appropriate secondary antibodies, diluted 1:200 for 1 h at room temperature. Washes were done in PBT. As secondary antibodies, we used Alexa Fluor 488 Donkey anti-mouse IgG (715-545-150), Alexa Fluor 594 Donkey anti-goat IgG (705-585-03) or Alexa Fluor 488 Donkey anti-rabbit IgG (711-545-152) from Jackson Immunoresearch Laboratories, and Alexa Fluor 647 Chicken anti-goat IgG (Invitrogen, A21469). Finally, DNA was counterstained with DAPI, and glasses were mounted in Vectashield Antifade (Vector Laboratories) or Prolong Gold Antifade Reagent (Invitrogen). For PLA experiments, coverslips were fixed first with paraformaldehyde at 4% in PBS for 10 min. Then, cells were permeabilized with 0.5% Triton X-100 in PBS and blocked with 1% BSA in PBS for 1 h at room temperature. Incubation with primary (Alexis human mAb, 13A4, ALX-804-662-R200, mouse) diluted 1:20 or Lamin A/C (Santa Cruz sc-6215, goat) diluted 1:200 was performed for 12–16h at 4 °C. After PBT washes, cells were incubated with Eh2 (Cell signaling 49055, rabbit) 1:100 or Bmi1 (Abcam ab56858 rabbit) 1:100 for 12–16 h at 4 °C. Primary antibodies were diluted in a PBS solution containing BSA 1%. Finally, detection of protein interactions was performed using the Duolink system (Systagna) following the manufacturer’s instructions.

Image processing and analysis. Fluorescent images were taken with a Nikon ECLIPSE 90i microscope (1006 objective), equipped with a digital camera (Nikon Coolpix 990) and NIS Element Digital software (Supplementary Fig. 4b and data for determination of nuclear distances among nuclei of different ages were collected for Supplementary Fig. 4c) or with confocal Leica SP5 supported by LAS-AF software (version 2.6.0) (Figs. 4d, 6d, e and Supplementary Figs. 4d, 8e, g). Confocal image size was 1024×1024 with the following laser wavelengths: 405-Diode; Argon; HeNe-633. Argon laser emission power was set to a range of 30–50% of total power amplitude. For acquisition settings, excitation power was set at 30% of Argon; 52% of 405-Diode, and 25% of HeNe-633. PLA blob quantification was performed using Cell Profiler 2.0. Nuclei were detected using the Otsu method with a global two-class threshold strategy, and foci were detected using the Otsu method with a per-object three-class threshold strategy. In order to automatically detect and quantify PCG bodies in fluorescence cell images, we improved a method derived from a previous work.

The algorithm has been implemented in MATLAB77 following the pseudocode reported in Supplementary Methods.

The algorithm computes the number of PCG bodies, the area of any PCG body, and the minimum Euclidean distance of any PCG body from the nuclear periphery. In order to compare these minimum distances among nuclei of different ages, we compute for each PCG body a measure of its closeness to the nuclear periphery (proximity). This measure is size-independent, being computed by dividing the minimum distance by the distance of the nuclear centroid from the point in the nuclear periphery closest to the PCG body.

Protein extraction and western blot analyses. Total proteins were prepared by resuspending 1 × 106 cells in extraction buffer (50 mM Tris-HCl pH 7.6; 0.15 M NaCl; 5 mM EDTA; 16 Protease Inhibitors; 1% Triton X-100). One 40 s pulse of sonication (UP100H manual sonicator, Hielsser) at 40% amplitude was performed to allow dissociation of protein from chromatin and solubilization. Extracts were analyzed by SDS-PAGE using an 8% gel (37:1 Acryl/Bis Acrylamide). The following primary antibodies were used: Beta-Actin (Santa-Cruz sc1616, rabbit 1:4000), H3 total (Abcam ab1791, rabbit 1:6000), Lamin A/C (Santa-Cruz sc-6215, goat 1:4000), Lamin B (Santa-Cruz sc6216, goat 1:2000), progerin (13A4 mouse, Abcam ab22927, mouse 1:1000), Lamin A/C (Santa-Cruz sc-222, mouse 1:1000), Bmi1 (D4B23 Cell signaling, rabbit 1:1000), H3K9me3 (Abcam ab8898, rabbit 1:1000), H3K27me3 (Millipore 07-449 rabbit 1:1000). HRP-conjugated secondary antibodies were revealed with the ECL chemiluminescence kit (ThermoFisher Scientific). The following secondary antibodies were used: Anti-Mouse IgG-Peroxidase (Sigma, A9044), Anti-Rabbit IgG-Peroxidase (Sigma, A9169), Anti-Goat IgG-Peroxidase (Sigma, A5420).

Chromatin fractionation. Chromatin fractionation was carried out as described in ref. 78 with minor adaptations. Briefly, 4 million fibroblasts were washed in PBS 1×, and were lysed in cytoskeleton buffer (CSK: 10 mM PIPES pH 6.8; 190 mM NaCl; 1 mM EGTA; 300 mM Sucrose; 3 mM MgCl2; 1× protease Inhibitors by Roche Diagnostics; 1 mM PMSF) supplemented with 1 mM DTT and 0.5% Triton X-100. After 5 min at 4 °C the cytoskeletal structure was separated from soluble proteins by centrifugation at 900 g for 3 min, and the supernatant was labeled as S1 fraction. The pellets were washed with an additional wash of cytoskeleton. Chromatin was solubilized by DNA digestion with 10U of RNase–free DNase (Turbo DNADase; Invitrogen AM2238) in CSK buffer for 60 min at 37 °C. To stop digestion, ammonium sulfate was added in CSK buffer to a final concentration of 250 mM and, after 5 min at RT samples were pelleted at 2350 g for 3 min at 4 °C and the supernatant was labeled as S2 fraction. After a wash in CSK buffer, the pellet was further extracted with 2 M NaCl in CSK buffer for 5 min at 4 °C, centrifuged at 2350 g 3 min at 4 °C and the supernatant was labeled as S3 fraction. This treatment removed the majority of histones from chromatin. After 2 washing in NaCl 2 M CSK, the pellets were solubilized in 8 M urea buffer to remove any remaining protein component by applying highly denaturing conditions. This remaining fraction was labeled as S4. For the scaled-down experiment, samples of 25,000, 50,000, or 10,000 cells were treated analogously, except for a reduction of
buffers volumes to half of those used for 10 million cells and a decrease of DNase to 8U.

For protein analysis, 70% of the volume of each fraction was analyzed by SDS-PAGE and immunoblotting. Anti-tubulin alpha (Sigma T5168, mouse 1:10,000), H3 (Abcam ab1791, rabbit 1:4000), Beta-Actin (Santa-Cruz sc1616, rabbit 1:4000), Lamin A/C (Santa Cruz sc-6215, goat 1:4000), Lamin B (Santa Cruz sc6216, goat 1:2000), progerin (13A4 mouse, Abcam 66587, mouse 1:1000), Ezh2 (AC22 Cell Signaling 3147S, mouse 1:1000), Bmi1 (D42B3 Cell signaling, rabbit 1:1000) were used as primary antibodies. HRP-conjugated secondary antibodies were revealed with the ECL chemiluminescence kit (Thermo Fisher Scientific).

**DNA sonication and sequencing for chromatin fractionation.** For DNA analysis, the remaining 30% of the volume of each fraction was diluted to 300 μl with TE,
incubated with RNase A (Roche) (90 min at 37 °C) and Proteinase K (Sigma) (150 min at 55 °C). DNA was extracted by standard phenol/chloroform extraction, precipitated and resuspended in 25 μl Mill-Q H2O. After Nanodrop (260/280 = 1.7–1.8), 30 μl of each HSP DNA quantification, we added H2O to a final volume of 105 μl. Then samples were transferred to 96 well plates and sonicated 4 times with Bioruptor sonicator (10 min 30 s ON - 30 s OFF, High Power). The DNA profiles were finally checked by capillary electrophoresis (Agilent 2100 Bioanalyzer with 2100 Expert Software). Finally, DNA libraries were prepared by using NuGEN Ovation Ultralow Library Prep System kit and then sequenced using an Illumina HiSeq 2500 instrument according to manufacturer’s instructions (Illumina). High passage samples were sonicated using Bioruptor or Covaris sonicator to reach a fragment dimension range of 100–500 bps. Samples were then quantified using Qubit HS DNA kit (Life Technologies) and technology prepared from each one of them using NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB E7645L) with Unique Dual Index NEBNext Multiplex Oligos for Illumina (NEB, E6440S). The library was quantified using Qubit HS DNA kit (Life Technologies) and profiles were checked by capillary electrophoresis on an Agilent 2100 Bioanalyzer. Finally, DNA libraries were sequenced using an Illumina NovaSeq 6000 instrument at the IEO Genomic Unit in Milan. In the scaled-down experiment, due to the low amount of starting material, the entire volume of each fraction was sampled the raw fastq files were converted, and quality checked using the same tools as for the SAMMY-seq reads. ChIP-seq sequencing read analysis. Sequencing reads were aligned to the hg38-noalt human genome available in the bcbio-nextgen pipeline, using bwa aln (version 0.7.12) with options -n -k 2 and saved the results in sam format with bwa sam. The sam files were converted to bam and named sorted with samtools (version 1.3.1). We marked PCR duplicates using the biobambam2 toolset (version 2.0.54). We discarded reads mapping to non-autosomal chromosomes, PCR duplicates, qcfail, unmapped, or mapping quality 0 reads with samtools. We converted the bam files to bedgraph using bedtools (version 2.25.0) and bedgraph to bigwig using UCSC’s bedgraphToBigWig tool (version 4) for further distribution visualization. Read coverage was normalized for sequencing library size, before converting bedgraph files to bigWig. Sample quality was assessed using fastqc (version 0.11.5) (http://www.bioinformatics.ubahram.ac.uk/projects/fastqc). For the down-sampling analysis (Supplementary Fig. 1g), we sampled the raw fastq files to 50 and 25% using the seqtk sample command (https://github.com/lh3/seqtk) and ran the same pipeline, as above. For the SAMMY-seq scale-down (Supplementary Fig. 3) and late passage (Supplementary Fig. 5) data sets analysis, we used the same pipeline as above with newer software sub-versions for bwa (version 0.7.17), samtools (version 1.9) and bedtools (version 2.29.0).

RNA sequencing. For high-throughput sequencing, cDNA libraries were prepared from total RNA, extracted with Trizol, using Illumina TruSeq Stranded Total RNA Kit with Ribo-Zero GOLD. cDNA fragments of ~300 bp were purified from each library and were sequenced for 125 bp, using an Illumina HiSeq 2500 instrument according to manufacturer’s instructions (Illumina).
Gaussian smoothed sequencing reads distribution, and we refer to this as "ChIP-seq signal" in the genomic tracks and metaplotfiles.

We de defines H3K27m containing gene TSS regions. As we called ChIP-seq peaks with the get.broad.enrichment.clusters function of SPP on the preprocessed data with $\text{window.size} = 2000$ and $\text{zthr} = 3$, tag.shif t = ((cross correlation peak position)/2) parameters. We defined a set of peaks region present in all control samples using the bedtools intersect command and selected those protein coding genes, that had a common H3K27m3 peak region $\pm 500$ m around the TSS. We defined a list of 4101 genes this way, which were used in downstream analyses.

MDS analysis was done using the plotMDS function of edgeR (version 3.2.4) after importing the 10k gene level data on reads for all samples into the R statistical environment, and selected informative reads with the get.billing.characteristics and select.informative.tags functions, removed anomalous positions with an extremely high number of reads using the remove.local.tag.anomalies function, and calculated the differential signal, smoothed by a Gaussian kernel, using the get.smoothed.tag.density function with the default bandwidth and tag.shift parameter.

We used deepTools (version 3.2.1) to visualize the ChIP-seq signal around TSS regions. We used the following parameters for the computeMatrix reference-point subcommand: -b 5000 -a 5000 -binSize 100, to calculate a matrix of SPP signal values per bin and gene. For visualization, we used the plotProfile command.

**Literature data processing**

We collected publicly available data sets from the following sources: ATAC-seq from (GEO: GSE80639), Lamin A/C ChIP-seq from (GEO: GSE471757 and GSE43332), Lamin B1 ChIP-seq from (GEO: GSE43341 and GSE36440). Sequencing reads were aligned, filtered, and converted using the same tools as for the SAMMY-seq reads. We calculated the genome-wide ChIP-seq signal for Lamin A/C and Lamin B1 data, using the SP package (version 1.16.0). We imported bam files into the R (version 3.5.1) statistical environment, and selected informative reads with the get.billing.characteristics and select.informative.tags functions, removed anomalous positions with an extremely high number of reads using the remove.local.tag.anomalies function, and calculated the differential signal, smoothed by a Gaussian kernel, using the get.smoothed.tag.density function with the default bandwidth and tag.shift parameter. We used the filtered read count as the genome-wide signal for the ATAC-seq analysis.

We downloaded from Roadmap Epigenomics the aligned reads of the consolodate bigWig file (together with the bigWigToBedGraph tool, liftOver over desired features for the human genome, and the bedGraphToBigWig tool with the UCSC toolkit (version 4)). We calculated genome-wide correlations between SAMMY-seq samples and all of the public data sets listed and processed above using StereoGene (version 2.20). StereoGene uses kernel correlation to calculate genome-wide correlations for spatially related but not completely overlapping features irrespective of their discrete or continuous nature.

We called laminator-associated domains with the EDD tool (version 1.1.19). For the Lamin A/C data set, we used the following options: gap-penalty 25 – bin-size 200 – write-bin-scores, and for Lamin B1 data set, we used: gap-penalty 5 – bin-size 100 – write-log-ratios – write-bin-scores. We also changed the fraction of informative_bins parameter to 0.98.

For Lamin A/C and Lamin B1, we defined a consensus set of regions for each of these lads for specific downstream analyses with bedops (version 2.4.37), by selecting those genomic regions, that were called lads by the EDD tool in at least two different ChIP-seq experiments (for either Lamin A/C or Lamin B1 data sets). We used a total of 6 ChIP-seq experiments for Lamin A/C and 5 for Lamin B1. We used these consensus regions for downstream analysis in Fig. 2c and Supplementary Fig. 2b, c.

**SAMMY-seq domain analyses and visualization.** We performed relative comparisons of SAMMY-seq fractions within each sample using EDD (version 1.1.15), optimizing 10k resolution enrichment dataset. EDD was originally designed for lamin ChIP-seq data, comparing IP and input samples. As SAMMY-seq data also shows broad enrichment regions, we used EDD to select significantly enriched SAMMY-seq domains by comparing less accessible fractions to more accessible ones (S3 vs S2, S4 vs S3, and S4 vs S2 comparisons) in each sample, with the following parameters: gap-penalty 25–bin-size 200–write-log-ratios–write-bin-scores and also excluding blacklisted genomic regions containing telomeric, centromeric, and certain heterochromatic regions. We also changed the required_fraction of informative_bins parameter to 0.98. We used the same set of parameters for the downsampled data (Supplementary Fig. 1g), except changing _bin-size to 100 for 50% down sampling or 200 for 25% down-sampling. To account for the lower sequencing depth, in the SAMMY-seq scale-down experiment (results reported in Supplementary Fig. 3), we run EDD (version 1.1.19) with –gap-penalty 25 and –bin-size 250 parameters, and in late-passage cells (results reported in Supplementary Fig. 5), we used: gap-penalty 45 and –bin-size 250.

Additionally, we calculated the genome-wide differential signal for all comparisons, using the SP package (version 1.15.4). We imported bam files into the R (version 3.3.1) statistical environment, and selected informative reads with the get.billing.characteristics and select.informative.tags functions, removed anomalous positions with an extremely high number of reads using the remove.local.tag.anomalies function, and calculated the differential signal, smoothed by a Gaussian kernel, using the get.smoothed.tag.density function with the default bandwidth and tag.shift parameter = 0.

We used bedtools (version 2.25.0) and the bedtools jaccard or bedtools fisher command to calculate jaccard Index or Fisher test p values for the various overlap analyses. We randomized SAMMY-seq domains and LADs for the SAMMY-seq – LAD overlap analysis using the bedtools shuffle with --notOverlapping option and also excluding blacklisted genomic regions containing telomeric, centromeric, and certain heterochromatic regions with the –exclude option.

For the SAMMY-seq scale-down and late-passage data sets analysis, we used newer software sub-versions for R (version 3.5.1), the SP package (version 1.16.0), and bedtools (version 2.29.0).

We used the Gviz (version 1.26.5) and karyoplotter (version 1.2.2) tools. Biocomputer packages to visualize SAMMY-seq read coverages, differential signals, and domains. We used ggplot2 (version 3.3.0) for additional plotting, and the generalized additive model (GAM) method (version 4.3.3) from the mgcv package (version 1.8–12) for smoothing data while plotting SAMMY-seq border region signals.

RNA-seq sequencing read analysis. Transcripts and gene-level quantification were done with Kallisto (version 0.43.0) to estimate transcript-level read counts and TPM (Transcripts Per Million values). The GENCODE v27 annotation was used to build the transcriptome index. Kallisto was run with the--bias option to perform sequence-based bias correction on each sample. We used sleuth (version 2.0.9.20) to calculate gene or transcript-level differential expression comparing all three controls against all three HGPS samples, where the linear model used in sleuth included the sample type (control or HGPS), sex, and library id. Additionally, we analyzed transcript-level differential expression, comparing the three controls with a single progeria sample, where the linear model used in sleuth included the sample type (control or HGPS) and sex. For both cases, we used sleuth’s Wald test to calculate p and q-values. After differential expression, we grouped transcripts according to MSigDB pathways and aggregated transcript p-values using the Lancaster aggregation method from the approximate R package (version 0.1) motivated by a recently described analysis pipeline in ref. 86. Lancaster p-values were corrected for multiple testing using the Benjamini–Hochberg method. Alternatively, we aggregated transcript p-values based on their Roadmap chromatin states and differential expression direction. We used two different criteria: for the Txs, Tlx, ReprPC, ReprPCW, Het, and ZNF/Rpts, we ranked chromatin states, as they contain transcripts overlapping 50% with the state. For the Enhbiv, Tslbiv, and Bvflnk Roadmap chromatin state regions, we required that the 200 nt region around the TSS region overlaps at least 50% with the state, and these were considered bivalent genes for specific downstream analyses. We aggregated p-values of upregulated (sleuth b-value > 0) or downregulated (sleuth b-value < 0) transcripts separately, according to their chromatin state overlap and additionally filtering for being in a specific (S3 vs S3 or S4 vs S2) enrichment region.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability.** The high-throughput sequencing data generated for this study are available in the GEO repository with accession number "GSE116633". Other previously published genomes data used in this article were released in public repositories by the original publication authors as indicated in the article or Methods section above. These include GEO data sets for ATAC-seq (GEO:GSE80639), Lamin A/C ChIP-seq (GSE471757 and GSE53342), Lamin B1 ChIP-seq (GSE49341 and GSE63440), in addition to Roadmap Epigenomics consolodated ChIP-seq data sets for E055 (human foreskin fibroblasts) for histone marks (H3K9me3, H3K27me3, H3K4me1, H3K36me3, H3K27ac, and H3K4me3) retrieved from the Roadmap Epigenomics on line data repository at URL (https://regrid2.wustl.edu/roadmap/data/byFileType/alignments/consolidated). We also changed the fraction of informative_bins parameter to 0.98.

We refer to this as "ChIP-seq signal" in the genomic tracks and metaplotfiles. We defined a set of peaks region present in all control samples using the bedtools intersect command and selected those protein coding genes, that had a common H3K27m3 peak region $\pm 500$ m around the TSS. We defined a list of 4101 genes this way, which were used in downstream analyses.
References

1. Gilbert, N. et al. Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. Cell 118, 555–566 (2004).
2. Huisinga, K. L., Brower-Toland, B. & Elgin, S. C. The contradictory definitions of heterochromatin: transcription and silencing. Chromosoma 115, 110–122 (2006).
3. Blasco, M. A. The epigenetic regulation of mammalian telomeres. Nat. Rev. Genet. 8, 299–309 (2007).
4. Gilbert, N. & Allan, J. Distinctive higher-order chromatin structure at mammalian centromeres. Proc. Natl Acad. Sci. USA 98, 11949–11954 (2001).
5. Slotkin, R. K. & Matreisenn, R. Transposable elements and the epigenetic regulation of the genome. Nat. Rev. Genet. 8, 272–285 (2007).
6. Boettiger, A. N. et al. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. Nature 529, 418–422 (2016).
7. Schaetttergruber, B., Bourbon, H. M., Di Croce, L. & Cavalli, G. Genome folding for different epigenetic states.
8. Turgay, Y. et al. The molecular architecture of lamins in somatic cells. Nat. Rev. Genet. 19, 226 (2018).
9. Shimi, T. et al. Structural organization of nuclear lamins A, C, B1, and B2 revealed by superresolution microscopy. Mol. Cell Biol. 26, 4075–4086 (2015).
10. van Steensel, B. & Belmont, A. S. Lamin-associated domains: links with chromosome architecture, heterochromatin, and gene repression. Cell 169, 780–791 (2017).
11. Kind, J. et al. Single-cell dynamics of genome-nuclear lamina interactions. Cell 153, 178–192 (2013).
12. Kind, J. & van Steensel, B. Genome-nuclear lamina interactions and gene repression. Curr. Opin. Cell Biol. 22, 320–325 (2010).
13. Lund, E. et al. Lamin A/C-promoter interactions specify chromatin state-dependent transcription outcomes. Genome Res. 23, 1580–1589 (2013).
14. Collas, P., Lund, E. G. & Oldenburg, A. R. Closing the (nuclear) envelope on the genome: how nuclear lamins interact with promoters and modulate gene expression. Bioessays 36, 75–83 (2014).
15. Meuleman, W. et al. Constitutive nuclear-laminae interactions are highly conserved and associated with A/T-rich sequence. Genome Res. 23, 270–280 (2013).
16. Guelen, E. et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 453, 948–951 (2008).
17. Zheng, X. et al. Lamin Acomparative genome from the nuclear periphery. Mol. Cell 71, 802–815.e7 (2018).
18. Bianchi, A. et al. Dysfunctional polycystic transcriptional repression contributes to Lamin A/C-dependent muscular dystrophy. J. Clin. Invest. 130, 2408–2421 (2020).
19. Cesarini, E. et al. Lamin A/C sustains PcG protein architecture, maintaining transcription of alternate fate genes. Mol. Cell 68, 1023–1037.e15 (2017).
20. Nicetto, D. et al. H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification. Science 363, 294–297 (2019).
21. Spracklin, G. & Pradhan, S. Protect-seq: genome-wide profiling of nuclease inaccessible domains reveals physical properties of chromatin. Nucleic Acids Res. 48, e16 (2020).
22. Henikoff, S., Henikoff, J. G., Sakai, A., Loeb, G. B. & Ahmad, K. Genome-wide profiling of salt fractions maps physical properties of chromatin. Genome Res. 19, 460–469 (2009).
23. Baudement, M. O. et al. High-salt-recovered sequences are associated with the active chromosomal compartment and with large ribonucleoprotein complexes including nuclear bodies. Genome Res. 28, 1733–1746 (2018).
24. Jadhav, U. et al. Replicational dilution of H3K27me3 in mammalian cells and the role of poised promoters. Mol. Cell 78, 141–151.e5 (2020).
25. De, H. C., Nickerson, J. A. & Penman, S. Core filaments of the nuclear matrix. J. Cell Biol. 110, 569–580 (1990).
26. Marasca, F., Marullo, F. & Lanzuoco, C. Determination of polycomb group of protein compartmentalization through chromatin fractionation procedure. Methods Mol. Biol. 1480, 167–180 (2016).
27. Lund, E., Oldenburg, A. R. & Collas, P. Enriched domain detector: a program for detection of wide genomic enrichment domains robust against local variations. Nucleic Acids Res. 42, e92 (2014).
28. Roadmap Epigenomics, G. et al. Integrative analysis of 111 reference human epigenomes. Nature 518, 317–330 (2015).
29. Lund, E. G., Duband-Goulet, I., Oldenburg, A., Buendia, B. & Collas, P. Distinct features of lamin A-interacting chromatin domains mapped by ChIP-sequencing from sonicated or micrococcal nuclease-digested chromatin. Nature 56, 30–39 (2015).
30. Dou, Z. et al. Autophagy mediates degradation of nuclear lamina. Nature 527, 105–109 (2015).
31. Sadie, M. et al. Redistribution of the Lamin B1 genomic binding profile affects rearrangement of heterochromatic domains and SAHF formation during senescence. Genes Dev. 27, 1800–1808 (2013).
32. Stavrovskaya, E. D. et al. SteroGene: rapid estimation of genome-wide correlation of continuous or interval feature data. Bioinformatics 33, 3158–3165 (2017).
33. Zhang, H. et al. Loss of H3K9me3 correlates with ATM activation and histone H2AX phosphorylation deficiencies in Hutchinson-Gilford progeria syndrome. Proc. Natl Acad. Sci. USA 110, e0167454 (2016).
34. Wu, D., Yates, P. A., Zhang, H. & Cao, K. Comparing lamin proteins post-translational relative stability using a 2A peptide-based system reveals elevated resistance of progerin to cellular degradation. Nature 57, 585–596 (2016).
35. Hamczyk, M. R. et al. Progerin accelerates atherosclerosis by inducing endoplasmic reticulum stress in vascular smooth muscle cells. EMBO Mol. Med. 11, e07369 (2019).
36. Rebollo, N. & Motels, T. Shared molecular and cellular mechanisms of premature ageing and age–associated diseases. Nat. Rev. Mol. Cell Biol. 18, 595–609 (2017).

Received: 11 November 2019; Accepted: 5 November 2020; Published online: 08 December 2020
58. Petr, M. A., Tulika, T., Carmona-Marin, L. M. & Scheibye-Knudsen, M. Protocating the aging genome. Trends Cell Biol. 30, 117–120 (2020).
59. Liberezon, A. et al. The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst. 1, 417–425 (2015).
60. Scaffidi, P. & Misteli, T. Lamin A-dependent misregulation of adult stem cells associated with accelerated aging. Nat. Cell Biol. 10, 452–459 (2008).
61. Bianchi, A. & Lanzuolo, C. Into the chromatin world: role of nuclear architecture in genome regulation. AIMS Biophysics. 2, 585–612 (2015).
62. Minoux, M. et al. Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. Science 355, eaal2913 (2017).
63. Fredriksson, S. et al. Protein detection using proximity-dependent DNA ligation assays. Nat. Biotechnol. 20, 473–477 (2002).
64. Sewitz, S. A., Fahmi, Z. & Lipkow, K. Higher order assembly: folding the genome. Curr. Opin. Struct. Biol. 42, 162–168 (2017).
65. Criscione, S. W., Teo, Y. V. & Neretti, N. The chromatin landscape of cellular senescence. Trends Genet. 32, 751–761 (2016).
66. Madakshira, B. P. & Sadler, K. C. DNA methylation, nuclear organization, and cancer. Front. Genet. 8, 76 (2017).
67. Tan, L., Xing, D., Chang, C. H., Li, H. & Xie, X. S. Three-dimensional genome architecture in epigenome regulation. Nat. Rev. Genet. 14, 677–697 (2016).
68. Petr, M. A., Tulika, T., Carmona-Marin, L. M. & Scheibye-Knudsen, M. Chromatin architecture associates with accelerated ageing. Trends Cell Biol. 30, 1919–1920 (2020).
69. Scaf, A. The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst. 1, 417–425 (2015).
70. Scaffidi, P. & Misteli, T. Lamin A-dependent misregulation of adult stem cells associated with accelerated aging. Nat. Cell Biol. 10, 452–459 (2008).
71. Bianchi, A. & Lanzuolo, C. Into the chromatin world: role of nuclear architecture in genome regulation. AIMS Biophysics. 2, 585–612 (2015).
72. Minoux, M. et al. Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. Science 355, eaal2913 (2017).
73. Fredriksson, S. et al. Protein detection using proximity-dependent DNA ligation assays. Nat. Biotechnol. 20, 473–477 (2002).
74. Sewitz, S. A., Fahmi, Z. & Lipkow, K. Higher order assembly: folding the genome. Curr. Opin. Struct. Biol. 42, 162–168 (2017).
75. Criscione, S. W., Teo, Y. V. & Neretti, N. The chromatin landscape of cellular senescence. Trends Genet. 32, 751–761 (2016).
76. Madakshira, B. P. & Sadler, K. C. DNA methylation, nuclear organization, and cancer. Front. Genet. 8, 76 (2017).
77. Tan, L., Xing, D., Chang, C. H., Li, H. & Xie, X. S. Three-dimensional genome architecture in epigenome regulation. Nat. Rev. Genet. 14, 677–697 (2016).
78. Petr, M. A., Tulika, T., Carmona-Marin, L. M. & Scheibye-Knudsen, M. Chromatin architecture associates with accelerated ageing. Trends Cell Biol. 30, 1919–1920 (2020).
79. Scaf, A. The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst. 1, 417–425 (2015).
80. Scaffidi, P. & Misteli, T. Lamin A-dependent misregulation of adult stem cells associated with accelerated aging. Nat. Cell Biol. 10, 452–459 (2008).
81. Bianchi, A. & Lanzuolo, C. Into the chromatin world: role of nuclear architecture in genome regulation. AIMS Biophysics. 2, 585–612 (2015).
82. Minoux, M. et al. Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. Science 355, eaal2913 (2017).
83. Fredriksson, S. et al. Protein detection using proximity-dependent DNA ligation assays. Nat. Biotechnol. 20, 473–477 (2002).
84. Sewitz, S. A., Fahmi, Z. & Lipkow, K. Higher order assembly: folding the genome. Curr. Opin. Struct. Biol. 42, 162–168 (2017).
85. Criscione, S. W., Teo, Y. V. & Neretti, N. The chromatin landscape of cellular senescence. Trends Genet. 32, 751–761 (2016).
86. Madakshira, B. P. & Sadler, K. C. DNA methylation, nuclear organization, and cancer. Front. Genet. 8, 76 (2017).
87. Tan, L., Xing, D., Chang, C. H., Li, H. & Xie, X. S. Three-dimensional genome architecture in epigenome regulation. Nat. Rev. Genet. 14, 677–697 (2016).
88. Petr, M. A., Tulika, T., Carmona-Marin, L. M. & Scheibye-Knudsen, M. Chromatin architecture associates with accelerated ageing. Trends Cell Biol. 30, 1919–1920 (2020).
89. Scaf, A. The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst. 1, 417–425 (2015).
90. Scaffidi, P. & Misteli, T. Lamin A-dependent misregulation of adult stem cells associated with accelerated aging. Nat. Cell Biol. 10, 452–459 (2008).
91. Bianchi, A. & Lanzuolo, C. Into the chromatin world: role of nuclear architecture in genome regulation. AIMS Biophysics. 2, 585–612 (2015).
92. Minoux, M. et al. Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. Science 355, eaal2913 (2017).

Acknowledgements
We thank Giovanna Lattanzi, Sammy Basso, the Italian network of Laminopathies, and members of the laboratory for stimulating discussions and constructive criticisms. We thank Beatrice Bodega, Marina Lusci, Maria Vivo, Daniela Palacios, Chiara Mozzetta, Vincenzo Costanzo, Fabrizio d’Adda di Fagagna, Judith Hartiprakash, Koustav Pal, Paolo Mauri, and Mattia Forcato for critical feedback on the manuscript. We are grateful to Mariangela Panetta, Elisa Cesarini, Tom Misteli, Roland Foisner, Giannino del Sal, Francesco Napoletano, and Valentina Sacccone for providing support with cell cultures. We are grateful to Chiara Cordiglieri for the precious help at the confocal microscope and Luca Madaro for providing support in image processing during the revision. We thank Elisa Salvato for advice on statistical analyses and Judith Hariprakash for advice on the definition of active enhancers. We gratefully acknowledge the Progeria Research Foundation for providing primary human fibroblasts of HGPS patients. The following cell lines were obtained from the NIGMS Human Genetic Cell Repository at the Corell Institute for Medical Research: AG08498, AG07095. This work was supported by grants from the Flagship CNR projects, (Epigen and Interomics) to C.L and G.O, My First AIRC Grant (MFAG) n. 18535 to C.L, AFM-Telethon n. 23030 to C.L and F.F.; by AIRC Start-up grant 2015 n.16841 to F.F.; and Carplo 2017-0649 to C.L and F.F. E.S. was supported by the Structural National Post-doc Program of SEIM (SIPOD) and the AFM TELETHON fellowship n. 21835.

Author contributions
In our opinion, F.L and C.P. equally contributed to this work and should be both considered as second authors. C.L designed and F.M., F.L., A.B., and I.O. performed the experiments. E.S., C.P., and F.F. analyzed all sequencing data. S.V. contributed to analyze ChIP-seq data. L.A., F.G. and G.O. performed image processing and analysis. E.S., F.M., C.L. and F.F. interpreted the results and wrote the manuscript. All authors reviewed and approved the manuscript for submission.

Competing interests
E.S., F.M., F.L., F.F., and C.L. declare a patent application has been submitted for the SAMMY-seq technique, thus constituting a potential competing interest. The remaining authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-20048-9.

Correspondence and requests for materials should be addressed to F.F or C.L.

Peer review information Nature Communications thanks Alessandro Bonetti, Aki Minoda, Eric Schirmer and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.  © The Author(s) 2020