Single-cell dynamics of genome-nucleolus interactions captured by nucleolar laser microdissection (NoLMseq)

Kaivalya Walavalkar[1], Shivani Gupta[1], Jelena Kresoja-Rakic[1], Mathieu Raingeval[1,2], Chiara Mungo[1,2], Raffaella Santoro[1]

1. Department of Molecular Mechanisms of Disease, DMMD, University of Zurich, Zurich, CH-8057, Switzerland
2. Molecular Life Science Program, Life Science Zurich Graduate School, University of Zurich, Zurich, CH-8057, Switzerland

*Corresponding author:
Raffaella Santoro (raffaella.santoro@dmmd.uzh.ch)

Abstract

Gene position in the nuclear space plays an important role in gene regulation. This is exemplified by repressive chromatin domains frequently contacting nuclear lamina or nucleoli. The nucleolus undergoes structural changes in response to various cellular states, potentially impacting genome organization. However, how the 3D-genome organization responds to nucleolar states has remained underinvestigated due to the lack of methods able to identify nucleolar associated domains (NADs) in single cells and under nucleolar stress. To address this, we developed NoLMseq, a method combining laser-capture microdissection and DNA sequencing to map NADs in single cells. NoLMseq identified many unexplored features of chromosome organization around single nucleoli such as NAD heterogeneity among ESCs, culminating in two major populations with distinct chromatin states. The data also revealed that NADs prevalently contact nucleoli in a monoallelic manner and allelic nucleolar contact frequency sets gene expression and chromatin states. NoLMseq also revealed how chromosomes reorganise around nucleoli under nucleolar stress conditions, highlighting the importance of nucleolus integrity in genome organization and 3D-genome response to nucleolar stress. The results demonstrated that NoLMseq accurately measures chromosome contacts around single healthy and stressed nucleoli and it will be a critical tool to study NADs within biological populations and determine how the 3D-genome responds to nucleolar stress in healthy and disease states.

38 **Introduction**

39 The three-dimensional (3D) organization of chromosomes in the cell's nucleus plays an important role
40 in the regulation of gene expression programs and cell fate. One important aspect of this regulation is
41 the position of genes in the nuclear space. This is exemplified by the frequent location of repressive
42 chromatin domains at the nuclear periphery (i.e., lamina associated domains, LADs) or around nucleoli
43 (nucleolar associated domains, NADs) [1-3]. However, while LADs have extensively been studied in many
44 cell types and in single cells, NADs have remained underinvestigated mainly due to technical limitations
45 in identifying DNA sequences associating with a compartment lacking a membrane such as the
46 nucleolus.

47 The nucleolus undergoes structural changes in response to various cellular states, with a potential
48 impact on genome organization. The reason of these changes is that the nucleolus is the compartment
49 where ribosome biogenesis takes place, a process responsible for the assembly of the translational
50 machinery, the ribosome, and tightly regulated according to cell state [4]. Ribosome biogenesis is initiated
51 in the nucleolus by the RNA polymerase I (Pol I)-driven transcription of hundreds of ribosomal RNA
52 (rRNA) genes (rDNA) that generate 45S/47S pre-rRNA. These transcripts are then modified,
53 processed, and assembled with ribosomal proteins (RPs) in the nucleolus, forming pre-ribosomal
54 particles [5]. rRNA genes are distributed among different chromosomes; acrocentric chromosomes in
55 human cells and, in mouse cells, and depending on strain, they can generally be found at chromosomes
56 12, 16, 18, and 19 (rDNA-chromosomes) and close to centromeric regions [3]. The nucleolus structure is
57 very dynamic and promptly responds to external stimuli that affect ribosome biogenesis (i.e., size,
58 number, or physico-chemical properties), such that highly proliferative cells have larger nucleoli
59 compared to quiescent cells [6-9]. Accordingly, alterations in nucleolar activities and structure have been
60 well documented under many cellular stress conditions and in several diseases, such as cancer,
61 neurodegenerative disorders, and premature ageing [10]. For example, nucleolar stress is defined as a
62 condition in which abnormalities in nucleolar structure and function driven by diverse cellular insults,
63 such as nutrient starvation or DNA damage, ultimately leads to activation of stress signalling pathways
64 that downregulate ribosome biogenesis, including rRNA gene transcription, and alter nucleolar structure
65 [11,12]. However, whether these structural alterations might have an impact on the genome organization
66 around nucleoli remains unexplored, mainly due to technical limitations of current methods to map
67 NADs.

68 Currently, there are two major methods to identify NADs; the biochemical purification of nucleoli
69 followed by sequencing and, more recently, the Nucleolar-DamID that relies on the expression of an
70 engineered nucleolar histone H2B fused to Dam that serves to mark NADs with adenine methylation
71 (H2B-Dam-NoLS) (reviewed in [13]). Although based on different methodologies, both methods revealed
72 that NADs are composed of repressive chromatin domains [14-18]. However, both methods have also
73 some limitations. First, they cannot be applied to study how genome organization around nucleoli is
74 affected under nucleolar stress [19,20]. Indeed, under these conditions, nucleoli become frail and cannot
75 be efficiently biochemically purified. On the other side, the consequent downregulation of ribosome
76 biogenesis decreases protein translation, thereby affecting the expression of H2B-Dam-NoLS.
77 Moreover, the physico-chemical alterations of nucleoli upon rDNA downregulation causes the expulsion
78 of the H2B-Dam-NoLS from nucleoli similarly to other nucleolar components [18]. Finally, both methods
79 have only been applied in cell populations and consequentially they have provided an average nucleolar
80 contact frequency, which might not accurately reflect the structure of the genome around single nucleoli.
81 Thus, so far there is no methodology able to identify NADs under nucleolar stress conditions and in
82 single cells.
83 To overcome these challenges, we developed NoLMseq, a novel approach that combines laser-capture
84 microdissection (LCM) and DNA sequencing to identify NADs in single cells. LCM is a technique
85 traditionally used for isolating single cells or specific tissue regions[21,22] that we have expanded it for the
86 organelle-level isolation of single nucleoli. Traditional techniques to study genome organization
87 predominantly rely on either microscopy or DNA sequencing. NoLMseq synergistically integrates
88 strengths of both these approaches, enabling precise examination of the genome architecture
89 surrounding the nucleolus. The application of NoLMseq in mouse embryonic stem cells (ESCs) not only
90 confirmed the general repressive state of NADs but also provided novel insights of genome organization
91 around nucleoli that could not be detected in previous NAD studies that were based on methods that
92 could only be applied in cell populations. These findings included the detection of different levels of
93 NAD heterogeneity among ESCs that culminates in two distinct cell populations with distinct chromatin
94 states. NoLMseq also determined the remodelling of chromosomes around nucleoli during ESC
95 differentiation into neural progenitors with the identification of genes linked to neuronal pathways that
96 detach from nucleoli and become reactivated, indicating that the nucleolus is directly involved in the
97 establishment of repressive chromatin states. Intriguingly, although the majority of genes contacting

98 nucleoli are low expressing, we detected high nucleolar contact frequency of the highly expressing
99 ribosomal protein genes, suggesting a potential role of nucleolus proximity to regulate ribosome
100 biogenesis. Moreover, NoLMseq revealed that NADs prevalently contact nucleoli in a monoallelic
101 manner and that allelic nucleolar contact frequency correlates with gene expression and chromatin
102 states. Finally, we were able to determine the genome reorganization around nucleoli under nucleolar
103 stress conditions, a result that highlights the importance of nucleolus integrity in genome organization
104 and how the genome responds to nucleolar stress conditions. Together, these results not only provided
105 novel insights of how chromosomes are organized around nucleoli and their remodelling during cell fate
106 commitment and stress-induced nucleolar structural changes but also demonstrated that NoLMseq
107 accurately measures chromosome contacts around single healthy and "stressed" nucleoli. Thus,
108 NoLMseq will be a critical tool to study NADs within biological populations and determine how genome
109 organization responds to nucleolar stress in healthy and disease states.
110
111 **Results**
112 **NoLMseq identifies NADs in single cells**
113 To map NADs at a single nucleoli resolution, we combined laser-capture microdissection (LCM) of
114 nucleoli and DNA sequencing. We named this method nucleolus laser microdissection sequencing
115 (NoLMseq). We applied NoLMseq to mouse embryonic stem E14 cells (ESCs), thereby enabling
116 comparisons with NADs from ESC populations, which have recently been identified using Nucleolar-
117 DamID [18]. ESCs were grown on a thin polyester membrane that can be cut with a fine laser (**Fig. 1a,b**).
118 We visualized the nucleolus with brightfield microscopy and used LCM to cut it out and collected it in
119 an isolation cap of a microcentrifuge tube using the MMI CapSure technology. The DNA content of each
120 nucleolus was purified and amplified via quasilinear whole-genome amplification (WGA) [23]. To prevent
121 any loss of DNA, WGA was carried out in the same tube in which the microdissected nucleolus was
122 collected. Before proceeding with library preparation and DNA sequencing, the purity of nucleolar DNA
123 was assessed by PCR to determine the presence of rRNA genes, which are located within nucleoli, and
124 the absence of *Tuba1*, an active gene that previous studies have shown to not contact nucleoli[18]
125 (**Extended Data Fig. 1a**).
126 We generated 53 high-quality single nucleolus contacts maps from ESCs, with a median of $5{\cdot}10^5$
127 uniquely mapped reads per nucleolus (**Fig. 1c, Extended Data Figs. 1b and 2**). We binned the genome

128 into 100 kb segments and called NADs using SICER2 [24] at a 100kb resolution using genomic DNA as
129 a control. The identified reads over the reference genome showed a bimodal distribution, indicating that
130 loci are either in a “contact” or “no contact” state with nucleoli (**Fig. 1c, Extended Data Fig. 1b**). Since
131 the microdissection of nucleoli will inevitably include genomic sequences above and below nucleoli, we
132 analysed the data by calculating the nucleolar contact frequency (CF) that we defined as the proportion
133 of microdissected nucleoli containing a defined NAD sequence (**Fig. 1d**). We found that NADs from
134 rDNA-chromosomes have higher CF than chromosomes not containing rRNA genes, underscoring the
135 specificity of NoLMseq in identifying NADs. Moreover, chromosome X, which is active in the analysed
136 male ESCs, showed the lowest nucleolar CF. For all downstream analyses, we selected NADs with
137 CF> 20%, a value that corresponds to FDR 0.0009. The analysis of NAD coverage/chromosome of all
138 53 nucleoli further supported the high propensity of rDNA-chromosomes to contact nucleoli relative to
139 the other chromosomes (**Fig. 1e**). The average NAD coverage in single nucleoli was about 15%, much
140 lower than the 30% NAD coverage obtained from previous NAD studies in ESC population [16,18] (**Figs.**
141 **1e,f**). The data also showed that the average length of NADs of rDNA-chromosomes was 1.85 Mb and
142 significantly higher than the average NAD size from the rest of the chromosomes (1.30 Mb, **Fig. 1g**).
143 The accuracy of NoLMseq was also supported by finding >90% NADs overlapping with previous NAD
144 profiles from ESC populations using Nucleolar-DamID [18] and SPRITE method [25] (**Fig. 1h**). To further
145 validate the data, we used an orthogonal approach by computing the CF of NADs from NoLMseq
146 genomic data and their distance to Fibrillarin marked nucleoli obtained by seqFISH+ microscopy [26] (**Fig.**
147 **1i**). We observed that nucleolar CFs measured by NoLMseq were inversely proportional to the distance
148 of NADs to the nucleolar marker Fibrillarin, indicating that NoLMseq accurately measures genomic
149 distance to nucleoli.
150 To further support the efficacy of NoLMseq, we analysed known NAD features of the identified NADs.
151 It is well known that centromeres tend to be positioned close to the nucleolus in various model systems.
152 Since in mouse cells all chromosomes are acrocentric, we measured the average NAD coverage at
153 centromere-proximal regions at the 5’-quintile portion of all chromosomes and found that it increases
154 with nucleolar CF, further supporting the accuracy of NoLMseq to identify sequences in contact with
155 nucleoli (**Fig. 2a**). Consistent with previous results in ESC populations obtained with Nucleolar-DamID
156 [18], NADs identified by NoLMseq were significantly depleted of active histone marks relative to
157 sequences within the active A compartment, showed a significant enrichment for the repressive

158 H3K9me2 but not for the other repressive marks H3K27me3 and H3K9me3, and displayed low gene
159 expression levels (**Fig. 2b,c**). Although NADs generally appear to have a repressive chromatin and
160 gene expression state, we could also find cases where highly expressed genes have high nucleolar
161 CFs. In particular, we found high nucleolar CFs for ribosomal protein (RP) genes, which encode proteins
162 for ribosome assembly in nucleoli and are transcribed by RNA Pol II, and tRNA genes, which are
163 transcribed by Pol III (**Fig. 2d**). In contrast, snoRNA genes, which express RNAs playing an important
164 role in rRNA modifications in nucleoli, did not show any preferential location to nucleoli. The nucleolar
165 contacts of RP and tRNA genes was not due to their location at rDNA-chromosomes since most of them
166 were located at chromosomes non containing rRNA genes (92% and 90%, respectively; **Extended**
167 **Data Fig. 3**). These results suggest a potential role of nucleolus proximity for the regulation of RP
168 genes, and consequently ribosome biogenesis, as well as tRNA, thereby influencing protein translation
169 in general (**Fig. 5b**). Finally, these data also suggested that not all NADs are necessarily
170 transcriptionally repressed.
171 NADs have been previously characterized in two categories; NAD-only (or Type-1) which correspond
172 to sequences contacting nucleoli but not the nuclear lamina (NL) and NAD/LAD (or Type-2) which are
173 regions that can be located both at nucleoli and NL [13,16-18]. Since this classification of NADs was
174 obtained from ESC population, it has remained elusive whether NAD/LAD regions contact both the
175 nucleoli as well as the NL in the same or different cells. The analysis of single nucleoli data revealed
176 that NAD-only sequences have higher CF values than NAD/LAD, suggesting that NAD/LAD domains
177 can contact the nucleolus in one cell and the NL in another cell, but rarely both nucleolus and NL in the
178 same cell (**Fig. 2e**). We also observed an inverse correlation of NAD-only CF values and H3K27ac
179 content, indicating that strong NADs generally exhibit low levels of active chromatin (**Fig. 2f**). Similarly,
180 NAD-only domains with high nucleolar CF have high H3K9me2 levels (**Fig. 2g**). However, while
181 NAD/LAD regions showed significantly lower gene expression levels relative to genes within the active
182 A compartment, the expression of genes within NAD-only regions was not significantly reduced, albeit
183 there was a tendency to be low expressed when located at regions with the highest nucleolar CF (**Fig.**
184 **2h**). These results are consistent with the data described above showing that some active genes can
185 also contact nucleoli.
186
187

188 **ESCs are composed of two populations with distinct nucleolar contacts**

189 NADs obtained from single nucleoli allowed to study the degree of cell heterogeneity in genome
190 organization. We calculated the NAD coverage for each chromosome in single cells and performed
191 hierarchical clustering (**Fig. 3a**). We found two main ESC populations that differ according to the level
192 of nucleolar contacts among chromosomes. The major differences between these two groups were on
193 the level of rDNA-chromosome contacts with the nucleolus, which showed a mutually exclusive
194 relationship. rDNA-chromosomes 12 and 19 had high levels of nucleolar contacts in population 1
195 whereas they were almost absent in nucleoli of population 2 (**Fig. 3b**). In contrast, rDNA-chromosomes
196 16 and 18 displayed higher NAD coverage in population 2 than in population 1. In the case of
197 chromosomes not bearing rRNA genes, chromosomes 7, 11, 15, and 2 showed more nucleolar contacts
198 in population 2 than in population 1 whereas chromosomes 3 and 5 preferentially contact nucleoli of
199 population 1. Remarkably, the analysis of histone modifications revealed that NADs enriched in
200 population 1 were significantly depleted of active histone marks and enriched in H3K9me2 relative to
201 NADs of population 2, suggesting different chromatin states between these two ESC populations (**Fig.**
202 **3d**). Thus, the data revealed heterogeneity in the ESC genome structure and chromatin configuration
203 that depend on the genome organization around nucleoli.

204

205 **Detachment of NADs upon differentiation into NPCs results in gene activation**

206 Next, we asked whether the composition of genomic contacts in single nucleoli could change upon
207 differentiation of ESCs into neural progenitors (NPCs). As in the case of ESCs, in NPCs NADs from
208 rDNA- chromosomes consistently exhibited higher nucleolar CF than regions from other chromosomes
209 (**Fig. 4a**). However, NPCs displayed fewer NADs than ESCs, with a large proportion (75%) being the
210 same NADs as found in ESCs (**Fig. 4b-d**). Moreover, we observed more variability in nucleolus
211 association in ESCs than in NPCs as indicated by the Yule's Q coefficient, which provides a measure
212 of similarity between single-cell NADs (**Fig. 4e**). The NAD detection in single NPC nucleoli obtained
213 with NoLMseq is consistent with previous studies performed in NPC populations showing that the
214 architecture of chromosomes surrounding the nucleolus is in a more compact and rigid form whereas
215 in ESCs it appears more flexible and variable [18]. Moreover, these data are consistent with previous
216 reports indicating that ESCs harbour a more open and dynamic chromatin than differentiated cells [27,28].

217 Next, we investigated whether genes contacting nucleoli in ESCs but detaching from nucleoli in NPCs
218 (ESC specific NADs, ESCsp-NADs) could change their expression state in NPCs. We defined ESCsp-
219 NADs those NADs with nucleolar CF > 25% in ESCs and < 10% in NPCs. RNAseq data analysis
220 revealed that ESCsp-NADs are less expressed in ESCs than in NPCs, suggesting that the detachment
221 from nucleoli facilitates gene activation (**Fig. 4f,g**). Moreover, the absoluted fold change in gene
222 expression in NPCs compared to ESCs was higher for the upregulated genes, with a large fraction
223 (60%) not expressed in ESCs (**Fig. 4h,i**). Accordingly, GO terms for these upregulated genes were
224 significantly linked to developmental processes, in particularly nervous system development, whereas
225 downregulated genes were mainly implicated in metabolic processes (**Extended Data Fig. 4**). These
226 results indicate that the detachment of genes from nucleoli during ESC differentiation unlocks them for
227 activation, further supporting a regulatory role of the nucleolus for the establishment or maintenance of
228 repressive chromatin states.
229
230 **Nucleolar integrity is required for genomic contacts with nucleoli**
231 Nucleolus structure is highly dynamic and promptly responds to external stimuli that affect ribosome
232 biogenesis and nucleolar structure. For example, downregulation of rDNA transcription by treatment
233 with low doses of Actinomycin D (ActD) induces a spatial reorganization of the nucleolar structure,
234 including the migration of rRNA genes to the nucleolar periphery, forming the so-called nucleolar caps,
235 and the release of the granular component (GC) protein NPM1 from nucleoli (**Fig. 5a**) [29]. However, until
236 now, it could not be investigated whether these structural changes can impact the organization of the
237 genome around nucleoli mainly due to the technical limitations of the current technologies to map NADs
238 under conditions in which nucleolar integrity is affected. Since nucleoli in ESCs treated with ActD
239 (ESCs+ActD) are still detectable with brightfield microscopy, we reasoned that NoLMseq could be
240 applied to determine how nucleolus integrity affects chromosome contacts with nucleoli. Compared to
241 untreated cells, the average NAD coverage in ESC+ActD significantly decreased from 15% to 10%
242 (**Fig. 5b,c, Extended Data Fig. 5**). Although NoLMseq revealed a substantial loss of nucleolar contacts
243 across all chromosomes, the majority of NADs retained upon ActD treatment were from the rDNA-
244 chromosomes (**Fig. 5d**). We validated this loss of contacts with nucleoli in ESCs+ActD by quantitative
245 DNA-FISH, using probes targeting four NADs at chromosomes 1, 2, 5, and 19 that were previously
246 identified and validated by Nucleolar-DamID [18] (**Fig. 5e-g**). These are strong NADs since they contact

247 nucleoli in 70-95% of ESCs (**Fig. 5f**). The DNA-FISH probe for the rDNA-chromosome 19 targeted a
248 NAD sequence proximal to the rDNA locus. Accordingly, this NAD remained anchored to nucleoli upon
249 ActD treatment, an expected result since rRNA genes remain close to nucleoli by forming nucleolar
250 caps (**Fig. 5f**). In contrast, DNA-FISH probes targeting NADs at chromosomes 1, 2, and 5 showed a
251 significant detachment of NADs from the nucleoli upon ActD treatment, indicating that nucleolar integrity
252 is crucial for NAD association with nucleoli and supporting the results obtained with NoLMseq. Because
253 of the retainment of rDNA-chromosomes to nucleoli in ESC+ActD, we reasoned that regions from rDNA-
254 chromosomes still able to contact nucleoli in ESC+ActD should correspond to centromere-proximal
255 sequences, which are linearly located close to rRNA genes. Accordingly, NAD coverage of centromere-
256 proximal sequences of rDNA-chromosomes was higher than the corresponding NAD coverage in
257 untreated ESCs, suggesting a reorganization of rDNA-chromosomes structure around nucleoli (**Fig.
258 5e**). In contrast, NAD coverage of centromere-proximal sequences of chromosomes not bearing rDNA
259 repeats showed a significant reduction compared to NADs of untreated cells, further indicating that the
260 anchoring of these sequences to nucleoli depends on nucleoli integrity. All these results indicate that
261 nucleolar integrity is an important determinant for chromosome organization around nucleoli. Moreover,
262 they demonstrate that NoLMseq is a unique and only genomic method for studying the role of nucleolus
263 in genome organization under nucleolar stress conditions that affect nucleolar integrity.
264
265 **Genomic contacts with nucleoli are predominantly mono-allelic**
266 Recent data have started to detect allelic asymmetry in gene expression and 3D-genome organization
267 [30,31]. So far, however, all studies on NADs were performed using cells whose maternal and paternal
268 chromosomes could not be distinguished from each other, as in the case of the E14 ESC line that
269 derives from a fully inbred homozygous mouse strain. Consequently, the information on NAD coverage
270 has been based on the assumption that both alleles contact nucleoli. To determine the allelic distribution
271 of NADs in single nucleoli, we applied NoLMseq to the male mouse F123-ESC line that was derived
272 from the F1 generation of two fully inbred homozygous mouse strains: *Mus musculus castaneus* (CAST,
273 paternal chromosomes) and *Mus musculus domesticus* S129S4/SvJae (S129, maternal chromosomes)
274 [32]. The F123 DNA sequence has been well characterized and contains high density of single nucleotide
275 variants (SNVs), with an average of 1 SNV every 124 nucleotides across autosomes [33]. We analysed
276 NAD profiles from 50 microdissected nucleoli of F123-ESCs by aligning them against the F123 genome

277 to identify allele-specific NADs (i.e., NAD_{S129} and NAD_{CAST}) (**Fig. 6a**). The vast majority (93%) of total
278 NADs identified in F123-ESCs corresponded to NADs identified in E14-ESCs, indicating that genomic
279 contacts with nucleoli are generally conserved between cell lines with the same developmental stage
280 but from different strains (**Extended Data Fig. 6a**). NADs of F123-ESCs showed a slightly lower
281 genome coverage (11.3%, calculated to haploid genome) relative to 15% NAD coverage measured in
282 E14-ESCs (**Fig. 6B**). Surprisingly, we found that a large fraction of NADs (86%) are mono-allelic (**Fig.**
283 **6a-c**). Monoallelic NADs were equally distributed among S129 and CAST chromosomes (**Extended**
284 **Data Fig. 6b**). We validated these results by DNA-FISH for two strong NADs, one located at
285 chromosome 5 and another located at the rDNA-chromosome 19 and close to the rDNA locus (**Fig.**
286 **6d**). The NAD of chromosome 5 showed monoallelic contacts in almost all the cells. Remarkably, even
287 the NAD of chromosome 19 located close to the rDNA locus showed mono-allelic contacts with nucleoli
288 in 50% of the cells, supporting the results of NoLMseq. This prevalent mono-allelic NAD distribution
289 resulted in a much lower NAD coverage per cell (5.6%, 297 Mbp) compared to the previously assumed
290 bi-allelic distribution (**Fig. 6e**). To determine whether mono- and bi-allelic NADs reflect differences in
291 chromatin modifications, we performed ChIPseq analysis for H3K9me2, H3K27ac, and H3K4me3 in
292 F123-ESCs (**Fig. 6f**). We found that bi-allelic NADs are in a more repressed chromatin states than
293 mono-allelic NADs since they display reduced levels of H3K27ac and H3K4me3 and higher levels of
294 H3K9me2, the latter the repressive histone mark characterizing NADs in ESCs [18]. Similarly, RNAseq
295 analyses showed a significantly lower expression of genes located at bi-allelic NADs (**Fig. 6g**). Among
296 genes located at bi-allelic NADs, we found cell type specific transcription factors, such as *Gata6* and
297 *TAF4b*, as well as genes linked to metabolic pathways and neuroactive ligand-receptor interactions
298 (**Extended Data Fig. 6c**). These results further indicate that high-frequency contact with nucleoli, even
299 at the allelic levels, results in a more repressive state.
300 Next, we asked whether the mono-allelic distribution of NADs might be parent specific. We applied
301 stringent criteria to assign parent specific NADs, considering, for example, that a specific NAD_{S129}
302 should have >25% nucleolar CF with the S129 genome and <10% nucleolar CF with the CAST genome
303 (**Fig. 6h**). However, very few NADs (1% on both alleles) showed a parent specific distribution. This
304 indicates that, generally, in one cell it is the paternal allele contacting nucleoli while in another cell it is
305 the maternal allele (**Fig. 6h**).
306

307 **Discussion**

308 We described NoLMseq, a method able to measure genome organization around nucleoli in single
309 cells. NoLMseq synergistically integrates the strengths of both microscopy and DNA sequence
310 technologies, enabling precise examination of the genome architecture surrounding the nucleolus.
311 Relative to previous methods, NoLMseq not only detects NADs in single nucleoli but it can also be
312 applied to study genome organization around nucleoli under conditions of nucleolar stress, which
313 profoundly alter nucleolar structure [6-9] and that could not be examined with current methods, making
314 elusive how genome structure around nucleoli responds to nucleolar stress. The data revealed a
315 profound reorganization of the genome around "stressed" nucleoli, such an increase in nucleolar
316 contacts of centromere-proximal sequences of rDNA-chromosomes while the rest of the chromosomes
317 lose their contact. The genomic data obtained by NoLMseq are also consistent with recent image
318 analyses showing that perturbation of nucleolus structure upon downregulation of the RNA polymerase
319 I subunit RPA194 reduced centromere-nucleolus interactions [34]. These results indicate that nucleolar
320 structure and integrity plays an important role in genome organization and future studies will address
321 alterations in chromosomal interactions with nucleoli under pathological stress conditions impacting
322 ribosome biogenesis and how these perturbations are caused and their functional significance. The
323 results also showed that NADs are very dynamic and NoLMseq represents an unique technique able
324 to capture NADs at the present state rather than averages of the past and present states as in the case
325 of DamID as recently discussed [35].
326 The detection of NADs in single nucleoli using NoLMseq has proven to be highly precise compared to
327 previous methods and capable of providing unexplored insights in genome organization. For example,
328 NoLMseq determined that active genes, such as ribosomal protein genes and tRNA genes, also have
329 high frequency contacts with nucleoli. Interestingly, previous imaging analyses in living yeast cells
330 reported the proximity of some RP genes to nucleoli [36], suggesting that the nucleolus could act as a
331 compartment for the regulation of the expression of ribosomal components. The presence of some
332 active genes at nucleoli is also consistent with a recent study showing that the localization of genes at
333 the repressive B compartment or proximal to nucleoli does not necessarily preclude transcription [37].
334 The accuracy of NoLMseq in detecting NADs relative to cell population studies is also shown by the
335 identification of ESCsp-NADs that contain genes activated in NPCs, which could not be detected in
336 previous cell population studies [18], further indicating that the nucleolus serves in general as a repressive

337 compartment. Because of the single nucleolus detection, NoLMseq revealed that ESCs display a certain
338 heterogeneity in genomic contacts, with nucleoli showing two major cell populations with distinct
339 chromatin features. Whether the presence of such populations, one more repressed than the other one,
340 might reflect differences in gene expression, pluripotency states, or other functions remains to be
341 investigated. These results are also consistent with recent data showing heterogeneity of nuclear
342 organization between individual single ESCs [38]. Finally, the application of NoLMseq in the hybrid F123-
343 ESC line showed that genomic contacts with nucleoli are mainly monoallelic. This result clearly provided
344 a more accurate value of NAD coverage, which was previously calculated based on either a haploid
345 genome or by assuming that all NADs are bi-allelic in the diploid genome. The average NAD coverage
346 5.6% (i.e., ca. 297 Mbp) per cell obtained in the F123-ESC line is much lower than the 30% NAD
347 coverage measured in ESC population studies using both purified nucleoli and Nucleolar-DamID
348 methods [16,18]. Although this calculated NAD coverage value does not include repetitive sequences such
349 as major and minor satellites, which are known to contact nucleoli but are not present in the reference
350 genome, it clearly represents a more realistic scenario than to have one third of chromosomes
351 contacting nucleoli. Moreover, the data showed that bi-allelic NADs are in a more repressive state than
352 monoallelic NADs, further supporting the efficacy of NoLMseq in identifying NADs and the nucleolus
353 acting as a repressive compartment.
354 Together, the results demonstrate that NoLMseq not only accurately measures chromosome contacts
355 around single nucleoli but also provides novel insights into genome organization around nucleoli that
356 could not be detected by previous studies. We predict that NoLMseq will not only be a critical tool to
357 study genome organization around the nucleolus in biological populations but also to study how the 3D-
358 genome organization responds to nucleolar stress in disease states and understand the functional
359 consequence of these alterations.
360
361
362
363
364
365
366

367 **Materials and Methods:**

368 **Cell lines**

369 129 mouse embryonic stem cells (E14 line) were cultured in serum medium containing Dulbecco's
370 modified Eagle's medium + Glutamax (Life Technologies), 15% FCS (Life Technologies; Cat no.
371 10270106 FBS South American), 1× MEM NEAA (Life Technologies), 100 μM β-mercaptoethanol,
372 recombinant leukemia inhibitory factor, LIF (Polygene, 1000 U/ml), 1× penicillin/streptomycin (Life
373 Technologies). ESCs were seeded at a density of 50,000 cells/cm2 in culture dishes (Corning CellBIND
374 surface) coated with 0.1% gelatin without a feeder layer. Propagation of cells was carried out every 2
375 days using enzymatic cell dissociation. ESCs were treated with Actinomycin D (10ng/ml) for 24h before
376 downstream analyses.

377 Neural progenitor cells were generated from ESCs, according to a previously established protocol [39].
378 In brief, differentiation used a suspension-based embryoid bodies formation (Bacteriological Petri
379 Dishes, Bio-one with vents, Greiner). The neural differentiation media (DMEM, 10% fetal calf serum, 1×
380 MEM NEAA, 2 mM Pen/Strep, β-mercaptoethanol, and sodium pyruvate) was filtered through 0.22 μm
381 filters and stored at 4 °C. During the 8-day differentiation procedure, media was exchanged every 2
382 days. In the last 4 days of differentiation, the media was supplemented with 2 μM retinoic acid to
383 generate neural precursors that are Pax-6- positive radial glial cells.

384 F123-ESCs (mouse male, hybrid cell line, F1 S129/Jae and Cast mouse cross) [32] were cultured on a
385 layer of mitotically inactivated feeder murine embryonic fibroblasts under standard conditions (DMEM,
386 supplemented with 15 % KSR, 1x Glutamax, 10 mM non-essential amino acids, 50 μM beta-
387 mercaptoethanol, 1,000 U/ml LIF). Media was changed every 24h. Before harvesting, mESCs were
388 passaged onto feeder-free ESGRO-gelatin (EMD Millipore SF008). coated plates for at least 2
389 passages to remove feeder cells. Cells were harvested after approximately 48 h at 70 – 80 %
390 confluency. F123-ESCs were a gift from Bing Ren, University of California, San Diego.

391

392 **NoLMseq**

393 For laser-microdissection, cells were grown on ESGRO-gelatin coated and UV-irradiated PET
394 membrane slides (1.4um thick, RNase-free, MMI, 50102) for 24h. Cells were washed twice with 1X PBS
395 and fixed with EM-grade freshly-depolymerised 3% PFA in 1X PBS for 30mins. The slides were then
396 placed in 70% ethanol for 1min and air dried. The slides were then washed with filtered ultra-pure H_2O

397 (3x 1 min each) and air dried. Single nucleoli were laser microdissected with the MMI Laser Capture
398 Microdissection Nikon microscope (Molecular Machines & Industries) into single 0.2 ml tubes
399 (transparent MMI caps, 50208). Nucleoli were identified by bright-field imaging. A 0.2ml cap filled with
400 transparent adhesive material was lowered onto the cells. A laser was used to cut the PET membrane
401 surrounding each nucleolus and the cut PET membrane section along with the single nucleolus was
402 lifted off by the sticky cap. Each single nucleolus was collected in a separate cap and stored for further
403 processing.
404 Whole-genome amplification (WGA) was carried out using the MALBAC single cell WGA kit (YK001A)
405 following manufacturer's recommendations with minor modifications. The entire WGA reaction was
406 carried out in the same 0.2ml tube with a single nucleolus. Another tube, with an empty section was
407 used as a negative control whereas 30 pg of whole cell DNA were used as a positive control for the
408 WGA reaction. The cell lysis reaction was first carried out by using a heated lid at 50 °C for 50 mins
409 and then samples were collected by briefly spinning and proceeding with the cell lysis reaction again at
410 50 °C for 50 mins, followed by inactivation at 80 °C for 10mins. The reaction is then cooled at 4 °C. The
411 lysis step is followed by a pre-amplification step which involves the quasilinear amplification using
412 random primers at 20 °C progressively with 10 °C steps till 70 °C. After cooling, the samples were
413 subjected to a final amplification step by 20 cycles of PCR.
414 WGA-amplified DNA was purified using a Qiagen MinElute PCR Purification Kit and eluted in 140 µl
415 ultrapure library-grade H_2O. The amplified nucleolar DNA was checked on a 2% agarose gel to select
416 positive samples showing smears corresponding to amplified DNA similar to the positive control. The
417 positive samples were further checked for the presence of rDNA and the absence of Tubulin gene by
418 PCR on an agarose gel. The samples which passed this quality control check were then used for library
419 preparation and sequencing. 120 µl of the reaction was first sonicated in split-cap tubes by Covaris
420 E220 to obtain fragments around 300-500nt in length (Peak power- 75, Duty factor- 20, cycle/bursts-
421 1000, duration- 35s). The samples were then purified and concentrated with AMPure XP beads
422 (Beckman Coulter, A63881) and used for Illumina library using the NEBNext ultra II kit (NEB E7600).
423 Briefly, NoLMseq samples (10-100 ng) were end-repaired and polyadenylated before the ligation of
424 Illumina compatible adapters. The adapters contain the index for multiplexing. Library concentrations
425 were estimated using a Qubit® 4 fluorometer (Thermo Fisher Scientific) and the 4200 TapeStation
426 system and libraries were pooled together in batches averaging 40 samples. Each library pool was

427 sequenced in single end 100bp in one lane of Novaseq6000 to get an average 10 million reads per
428 library.
429 For data analysis, NoLM-seq reads were aligned to the mouse mm10 reference genome using Bowtie2
430 (version 2.5.0; [40]) with default parameters. PCR duplicates were removed using samtools [41]. Positive
431 NAD bins were called using SICER2 [42], using the sequenced gDNA file as the control file and the
432 parameters w=100000, g=100000, fdr=0.0001. Bigwig files were created from the bam files using
433 bamCoverage from deeptools [43]. For F123-ESCs, N-masked genome was created using the mm10
434 reference genome and the Variant Call *Format* (VCF) file from the mouse genome project containing
435 the SNP information for all strains. The SNPsplit package [44] was used to read the VCF file and extract
436 the SNP information for the S129/SvJae and CAST strains and N-mask for the mm10 genome. The
437 reads were aligned using Bowtie2 to this N-masked reference genome using default parameters.
438 SNPsplit was then used to tag the SNPs for S129/SvJae and CAST alleles in the bam file and sort the
439 haplotype-specific bam files for both the alleles. bamCoverage and Sicer2 positive NAD calling was
440 performed independently on the allele-specific bam files. Integrative Genome Viewer (IGV, version
441 2.15.4) [45] was used to visualize and extract representative single nucleolus NoLMseq bigwig and bed
442 tracks.
443 Nucleolar contact frequency was calculated at a 100kb resolution. We defined as NADs sequences with
444 a > 20% nucleolar contact frequency that correspond to a false discovery rate (FDR) 0.0009 or 0.09%,
445 meaning that more than 99% of called NADs are truly NADs. This was calculated using the cumulative
446 distribution function (CDF) $\mathrm{P}(\mathrm{X} \leq \mathrm{k}) = \sum i = 0k(in)pi(1-p)n-1$ where "n" is the number of single
447 nucleoli samples, "p" is the probability that a false-positive region is present in the sample, "k" is the
448 contact frequency threshold. For the hierarchical clustering of NADs, NAD chromosome coverages in
449 single nucleoli were used for clustering single nucleoli based on divisive clustering (Diana) using the
450 "cluster" package (version 2.1.6) in R.
451
452 **RNAseq**
453 Total RNA was purified with TRIzol reagent (Life Technologies). The quality of the isolated RNA was
454 determined with a Qubit® 4 fluorometer (Life Technologies, California, USA) and 4200 TapeStation
455 system (Agilent, Santa Clara, California, USA). Only those samples with a 260 nm/280 nm ratio between
456 1.8–2.1 and a 28S/18S ratio within 1.5-2 were further processed. The TruSeq Stranded mRNA (Illumina,

457 Inc, California, USA) was used in the succeeding steps. Briefly, total RNA samples (100–1000 ng) were
458 polyA enriched and then reverse- transcribed into double-stranded cDNA. The cDNA samples were
459 fragmented, end-repaired and adenylated before ligation of TruSeq adapters containing unique dual
460 indices (UDI) for multiplexing. Fragments containing TruSeq adapters on both ends were selectively
461 enriched with PCR. The quality and quantity of the enriched libraries were validated using Qubit® 4
462 fluorometer. The product is a smear with an average fragment size of approximately 260 bp. Libraries
463 were normalized to 10 nM in Tris-Cl 10 mM, pH8.5 with 0.1% Tween 20. The NovaSeq X (Illumina, Inc,
464 California, USA) was used for cluster generation and sequencing according to standard protocol.
465 Sequencing was paired end at 2 × 150 bp.
466 The quality of the reads was checked by FastQC, a quality control tool for high throughput sequence
467 data [46]. The quality of the reads was increased by applying: a) SortMeRNA [47] (version 2.1) tool to filter
468 ribosomal RNA; b) Trimmomatic [48] (version 0.40) software package to trim the sorted (a) reads. N-
469 masked genome was created using the mm10 reference genome and the VCF file from the mouse
470 genome project containing the SNP information for all strains. The SNPsplit package [44] was used to
471 read in the VCF file and extract the SNP information for the S129/SvJae and CAST strains and N-mask
472 the mm10 genome. The sorted (a), trimmed (b) reads were mapped against the Nmasked mouse
473 genome (mm10) using the default parameters of the STAR (Spliced Transcripts Alignment to a
474 Reference, version 2.6, [49]). SNPsplit was then used to tag the SNPs for S129/SvJae and CAST alleles
475 in the bam file and sort the haplotype-specific bam files for both the alleles. For each gene, exon
476 coverage was calculated using a custom pipeline and then normalized in reads per kilobase per million
477 (RPKM) [50].
478
479 **ChIPseq**
480 ChIP analysis was performed as previously described [51]. Briefly, 1% formaldehyde was added to
481 cultured cells to cross-link proteins to DNA. Isolated nuclei were then lysed with lysis buffer (50 mM
482 Tris-HCl, pH 8.1, 10 mM EDTA, pH 8, 1% SDS, 1X protease inhibitor complete EDTA-free cocktail,
483 Roche). Nuclei were sonicated using a Bioruptor ultrasonic cell disruptor (Diagenode) to shear genomic
484 DNA to an average fragment size of 200 bp. 20 µg of chromatin was diluted to a total volume of 500 µl
485 with ChIP buffer (16.7 mM Tris-HCl, pH 8.1, 167 mM NaCl, 1.2 mM EDTA, 0.01% SDS, 1.1% Triton X-
486 100) and incubated overnight with the ChIP-grade antibodies against H3K9me2, H3K9me3,

487 H3K27me3, H3K4me1, H3K4me3 and H3K27ac. After washing, bound chromatin was eluted with the
488 elution buffer (1% SDS, 100 mM $NaHCO_3$). Upon proteinase K digestion (50 °C for 3 h) and reversion
489 of cross- linking (65 °C, overnight), DNA was purified with phenol/chloroform, and ethanol precipitated.
490 The quality and quantity of the isolated DNA were determined with a Qubit® 4 Fluorometer (Life
491 Technologies, California, USA) and the 4200 TapeStation system. Briefly, ChIP samples (1 ng) were
492 end-repaired and polyadenylated before the ligation of Illumina compatible adapters. The adapters
493 contain the index for multiplexing. The quality and quantity of the enriched libraries were validated using
494 Qubit® 4 Fluorometer and the 4200 TapeStation system. The libraries were pooled to 10 nM in Tris-Cl
495 10 mM, pH8.5 with 0.1% Tween 20. Sequencing was performed on the NovaSeq 6000 for 100bp single
496 end reads (Illumina, Inc, California, USA).
497 For F123-ESCs, N-masked genome was created using the mm10 reference genome and the Variant
498 Call Format (VCF) file from the mouse genome project containing the SNP information for all strains.
499 The SNPsplit package [44] was used to read the VCF file and extract the SNP information for the
500 S129/SvJae and CAST strains and N-mask for the mm10 genome. ChIPseq reads were aligned to the
501 N-masked mm10 reference genome with Bowtie2 (version 2.5.0; (version 2.5.0; [40]) using default
502 parameters. PCR duplicates were removed using samtools [41]. SNPsplit was then used to tag the SNPs
503 for S129/SvJae and CAST alleles in the bam file and sort the haplotype-specific bam files for both the
504 alleles. Bigwig files were created from the bam files using bamCoverage from deeptools [43] using default
505 parameters. Integrative Genome Viewer (IGV, version 2.15.4) [45] was used to visualize and extract
506 representative ChIPseq tracks.
507
508 **Gene Ontology (GO) enrichment:**
509 GO enrichment analysis of genes was performed with DAVID 6.8 [52]. All genes in *Mus musculus* were
510 used as the background.
511
512 **DNA-FISH**
513 DNA-FISH probes for chromosomes 1, 2, 5, and 19 were generated as previously described [18]. The
514 probes were obtained with oligopaint libraries that were constructed the PaintSHOP interface created
515 by the Beliveau lab and were ordered from CustomArray/Genscript in the 12 K Oligopool format. Each
516 library contains a universal primer pair used to amplify all the probes in the library, followed by a specific

517 primer pair hooked to the 40-46-mer genomic sequences, for a total probe of around 124-130-mers.
518 Oligopaint libraries were produced by emulsion PCR from the pool, followed by a “two-step PCR” and
519 lambda exonuclease as described before [18,53]. Specifically, the emulsion PCR with the universal primers
520 allowed amplifying all pool of probes in the library. The “two-step PCR” led to the addition of a tail to the
521 specific probes bound to the Alexa Fluor 488 fluorochrome. All oligonucleotides were purchased from
522 Integrated DNA Technologies (IDT, Leuven, Belgium).
523 DNA-FISH was adapted from the protocol of Cavalli’s lab [54]. Cells fixed with 4% PFA for 10 min at room
524 temperature 24h after seeding. Cells were washed three times with 1X PBS and then permeabilized
525 with 0.5%Triton-X100/PBS for 10 min on ice, followed by 10 min at room temperature. After three
526 additional washes with 1X PBS, cells were incubated with 0.1 M HCl for 10 min at room temperature,
527 washed twice with 2X SCCT, and twice with 50% Formamide-2X SCCT for 20min, firstly at room
528 temperature and then at 60°C. Probe mixture contained around 45pmol of Oligopaint probe, 0,8µL of
529 ribonuclease A (ThermoScientific, 10 mg/mL), and FISH hybridization buffer for a total mixture volume
530 of 20uL, which was directly added on the slide. Cell DNA was denaturated at 83 °C for 8 min, and
531 hybridization was performed in a humid dark chamber overnight at 42 °C. Cells were washed with 2X
532 SCCT for 15 min once at 60°C and once at room temepature, once with 0.2X SCC for 10 min, twice
533 with 1X PBS for 2 min, and three times with 1X PBT for 2 min. Slides were then incubated for 1 h at
534 room temperature with blocking solution in a dark humid chamber, and overnight at 4 °C with primary
535 antibody against Nucleolin. Cells were washed four times with 1X PBT at increasing incubation times
536 (1 × 2 min, 1 × 3 min, 2 × 5 min) and then incubated for 2 h at room temperature with goat anti-mouse
537 IgG (H + L) Alexa Fluor 546 in a dark humid chamber. Slides were washed four times with 1X PBT at
538 increasing incubation times (1 × 2 min, 1 × 3 min, 2 × 5 min), three times with 1X PBS for 2 min. Then,
539 slides were mounted with Vectashield DAPI mounting media (Vector, H-1200) and, after drying, stored
540 at 4 °C.
541 DNA FISH/IF samples were imaged using a Zeiss LSM 980 Airyscan, with a z-stack collected for each
542 channel (step size, 0.3 um), using the oil objective Plan-Apochromat/1.40. Images were processed by
543 ImageJ (version 2.14.0/1.54F). The individual cells were identified by Hoechst/DAPI staining and cells
544 containing signal for DNA-FISH channel were identified manually on the corresponding fluorescent
545 channel. Distance between the DNA-FISH signal and the nucleolar marker immunofluorescence signal

546 was calculated using ImageJ (version 2.14.0/1.54F) and used to count the number of foci contacting
547 nucleolus and the number of cells with at least one contact with the nucleolus.
548
549 **Published datasets**
550 Published datasets used in the study were from: H3K27ac-, H3K4me1-, H3K27me3-ChIP-seq [39]
551 (GSE72164); H3K4me3- and H3K9me3-ChIPseq [55] (GSE23943); H3K9me2-ChIPseq [56] (GSE77420);
552 A/B compartments [57]; Nucleolar DamID and RNAseq in ESCs and NPCs [18] (GSE150822); Lamin-
553 DamID [58] (GSE17051); SPRITE [25] (GSE114242); seqFISH+ [26].
554
555 **Acknowledgements**
556 This work was supported by the ERC grant (ERC-AdG-787074-NucleolusChromatin), Swiss National
557 Science Foundation (31003B-201268), and EMBO Postdoctoral Fellowship (ALTF 209-2021 to K.W).
558 We thank Catherine Aquino and the Functional Genomic Center Zurich for the assistance in
559 sequencing. We thank Ana Pombo for her technical advises on laser-microdissection and library
560 preparation and Bing Ren for the F123-ESC line.
561
562 **Author contributions**
563 K.W. and J.K-R. established the NoLMseq. K.W. performed NoLMseq in all the described conditions,
564 ChIPseq and RNAseq in F123-ESCs and all data analyses of NADs. S.G. and M.R. performed DNA-
565 FISH analyses. C.M. performed IF analyses. K.W. and R.S. wrote the manuscript. R.S conceived and
566 supervised the project.
567
568 **Declaration of interests**
569 The authors declare no competing interests.
570

571 **Figure legends**

572 **Figure 1**

573 **NoLMseq identifies NADs in single cells**

574 **a.** Scheme representing the nucleolar laser capture microdissection (NoLMseq).

575 **b.** Representative images of ESCs before and after nucleolar laser capture microdissection.

576 **c.** Nucleolar contacts maps from 53 single nucleoli for chromosomes 6, 8, and 19. Black bars indicates

577 NADs.

578 **d.** Cumulative histogram of genome-wide nucleolar contact frequency (CF) values with respect to their

579 location at chromosomes containing rRNA genes (rDNA-chromosomes), without (w/o rDNA), and X and

580 Y chromosomes.

581 **e.** Coverage of NADs with CF > 20% on each chromosome in single nucleoli. Dotted grey line indicates

582 the mean NAD coverage in single cells. Black lines indicate the mean on each chromosome.

583 **f.** Violin plot showing NAD coverage in single cells. Red dashed line indicates the median.

584 **g.** Size of NADs located at rDNA-chromosomes and chromosomes not bearing rRNA genes (w/o rDNA).

585 Red and blue lines indicate medians, and the dotted lines indicate quartiles. Statistical significance (*P*-

586 values) was calculated using unpaired t-test (**** < 0.0001).

587 **h.** Coverage of NADs identified by NoLMseq relative to NADs identified in ESC populations with the

588 Nucleolar-DamID [18] and SPRITE [25] method.

589 **i.** Genome-wide comparison of DNA seqFISH+ distance to exterior of nucleoli marked with Fibrillarin

590 (µm) [26] and NoLMseq NAD contact frequency for 3713 paired genomic regions. Tukey boxplot where

591 box limits represent the 25th and 75th percentiles and the black line represents the median. Statistical

592 significance (*P*-values) was calculated using unpaired t-test (** < 0.01; **** < 0.0001).

593

594 **Figure 2**

595 **NADs identified by NoLMseq correspond to repressive chromatin domains**

596 **a.** NAD coverage on the centromere-proximal regions (5' -quintile) as a function of nucleolar contact

597 frequency.

598 **b.** Levels of active histone marks (H3K27ac, H3K4me1, H3K4me3) and repressive histone marks

599 (H3K9me2, H3K9me3, H3K27me3) at NADs relative to genomic regions located at the active A

600 compartment (A comp.). Values are shown as average RPKM. Tukey boxplot where box limits represent

601 the 25th and 75th percentiles. The horizontal line within the boxes represents the median. Statistical
602 significance (*P*-values) was calculated using the unpaired two-tailed t test (**<0.01, ****<0.0001, ns:
603 nonsignificant).
604 **c.** Expression values (RPKM) of genes within the active A compartment (A Comp.) and NADs. Tukey
605 boxplot where box limits represent the 25th and 75th percentiles. The horizontal line within the boxes
606 represents the median. Statistical significance (*P*-values) was calculated using the unpaired two-tailed
607 t test(****<0.0001).
608 **d.** Gene densities of ribosomal protein genes (red), tRNA genes (black), snoRNA genes (pink) and all
609 genes (grey) as a function of nuclear CF.
610 **e.** Left panel shows nucleolar contact frequency for NAD-only and NAD/LAD sequences. Right panel
611 depicts how NAD-only and NAD/LAD contact nucleoli in single cells.
612 **f,g,h.** Levels of H3K27ac (**f**), H3K9me2 (**g**), and gene expression (**h**) at NAD-only and NAD/LAD as a
613 function of nucleolar CF relative to the active A compartment (A comp.). at the active A compartment
614 (A Comp.) and at NAD-only and NAD/LAD relative to nucleolar CF. Values are shown as average
615 RPKM. Tukey boxplot where box limits represent the 25th and 75th percentiles. The horizontal line
616 within the boxes represents the median. Statistical significance (*P*-values) was calculated using the
617 unpaired two-tailed t test (**<0.01, ***< 0.001, ****<0.0001, ns: nonsignificant).
618
619 **Figure 3**
620 **ESCs showed two populations with distinct nucleolar contacts**
621 **a.** Hierarchical clustering of nucleoli according to their NAD coverage/chromosome
622 **b.** NAD coverage at the indicated chromosomes in the top quartile NADs of the two ESC populations.
623 **c.** Scheme representing the contact with nucleoli of the rDNA-chromosomes with the nucleoli in the two
624 ESC populations.
625 **d.** Levels of active histone marks (H3K4me1, H3K4me3, H3K27ac) and repressive histone marks
626 (H3K9me2, H3K9me3, H3K27me3) at NADs enriched in ESC population (Pop) 1 or 2. Values are shown
627 as average RPKM. Tukey boxplot where box limits represent the 25th and 75th percentiles. The
628 horizontal line within the boxes represents the median. Statistical significance (*P*-values) was calculated
629 using the unpaired two-tailed t test (*<0.05, ****<0.0001, ns: nonsignificant).
630

631 **Figure 4**

632 **NADs from single nucleoli of neural progenitors**

633 **a.** Cumulative histogram of genome-wide of nucleolar contact frequency measured from 23
634 microdissected nucleoli with respect to their location at chromosomes containing rRNA genes (rDNA-
635 chromosomes), without (w/o rDNA), and X and Y chromosomes in NPCs.

636 **b.** Coverage of NADs with CF > 20% on each chromosome in single nucleoli of NPCs. Dotted grey line
637 indicates the average NAD coverage in single cells of NPCs. Black lines indicate the mean on each
638 chromosome.

639 **c.** Violin plots showing NAD coverage in single ESCs and NPCs. Data from ESCs were also shown in
640 **Figure 1f**. Red dashed lines indicate the median and black dotted lines indicate the quartiles.

641 **d.** Venn diagram showing the distribution of NADs in ESCs and NPCs.

642 **e.** Violin plots showing the distribution of Yule's Q values as a measure of NAD cell-cell similarity in
643 ESCs and NPCs. Red dashed lines indicate the median and black dotted lines indicate the quartiles.

644 **f.** Expression levels of genes located at ESC_{sp}-NADs in ESCs and NPCs. Tukey boxplot where box
645 limits represent the 25th and 75th percentiles. The horizontal line within the boxes represents the
646 median. Statistical significance (*P*-values) was calculated using the paired two-tailed t test
647 (****<0.0001).

648 **g.** Scatter plot showing the expression levels of genes at ESC_{sp}-NADs in ESCs and NPCs. Expression
649 of genes significantly (P < 0.05) upregulated (blue) and downregulated (red) are shown.

650 **h.** Plot showing the absolute log_2 fold changes of upregulated and downregulated genes at ESC_{sp}-
651 NADs in NPCs. Black lines indicate the median. Statistical significance (*P*-values) was calculated using
652 the unpaired two-tailed t test (****<0.0001).

653 **i.** Proportion of low expressing genes (< 0.1 RPKM) in ESCs that are located at ESC_{sp}-NADs and are
654 up- or downregulated in NPCs.

655

656 **Figure 5**

657 **Nucleolar integrity is required for genomic contacts with nucleoli**

658 **a.** Left panel represents alterations in nucleolus structure upon treatment with Actinomycin D (ActD),
659 showing the re-organization of rRNA genes, which form nucleolar caps. Right panel shows
660 immunofluorescence images of the nucleolar markers UBF and NPM1. Scale bar is 9 μm.

661 **b.** Histogram showing the average NAD coverage/chromosome from single nucleoli of untreated ESCs
662 (blue) and 34 microdissected nucleoli from ESCs treated with ActD (grey).
663 **c.** Violin plots showing NAD coverage in single untreated ESCs (blue) and ESCs treated with ActD
664 (grey). Data from ESCs were also shown in **Figure 1F**. Red dashed lines indicate the median and black
665 dotted lines indicate the quartiles. Statistical significance (P-value) was calculated using the unpaired
666 two-tailed t test (****<0.0001).
667 **d.** Distribution of NADs in ESCs treated with ActD as a function of nucleolar CF across rDNA-
668 chromosomes, chromosomes without (w/o) rDNA, and NADs lost relative to untreated ESCs.
669 **e.** Representation of the DNA-FISH probes targeting NADs at chromosomes 1, 2, 5, and 19, the latter
670 containing rRNA genes at the 5' end, close to the centromeric region. The corresponding NADs profiles
671 from Nucleolar-DamID [18] are shown. iNAD: genomic domains not interacting with nucleoli.
672 **f.** Representative immuno-FISH images for NAD at chromosome 2. Nucleolin serves as nucleolar
673 marker. Scale bar is 5 µm.
674 **g.** Box plots represent the quantifications of immuno-DNA-FISH analyses showing the percentage of
675 cells displaying NADs contacting nucleoli. Error bars represent s.d. Statistical significance (*P*-values)
676 from three independent experiments was calculated using Mann-Whitney test (** < 0.01; *** < 0.001;
677 ns: non-significant).
678 **h.** NAD coverage at the centromere-proximal regions (the first 5'-quintile) between rDNA-chromosomes
679 and chromosomes without (w/o) rDNA. rDNA on untreated ESCs and ESCs treated with ActD. Red
680 dashed lines indicate the median and black dotted lines indicate the quartiles. Statistical significance
681 (P-value) was calculated using the unpaired two-tailed t test (**<0.01).
682
683 **Figure 6**
684 **Genomic contacts with nucleoli are mono-allelic**
685 **a.** NoLMseq tracks displaying NAD sequences from four microdissected single nucleoli from hybrid
686 F123-ESCs (S129/CAST) on the total mm10 genome (black), maternal S129 genome (pink) and
687 paternal CAST genome (blue). Bars below the tracks indicate NADs called on respective genomes.
688 **b.** NAD coverage from single nucleoli calculated over the haploid genome. Median and quartiles are
689 shown. Statistical significance (P-value) was calculated using the unpaired two-tailed t test
690 (****<0.0001, ns: non-significant).

691 **c.** Proportion of bi-allelic and mono-allelic NADs in single nucleoli. Black dashed lines indicate median
692 and black dotted lines indicate quartiles. Statistical significance (P-value) was calculated using the
693 unpaired two-tailed t test (****<0.0001).
694 **d.** Left panel shows DNA-FISH images of a NAD located at chromosome 19. Right panel is the
695 quantification of bi-allelic and mono-allelic contacts with nucleoli. Scale bar is 5 μm. Error bars indicate
696 s.d.
697 **e.** NAD coverage from single nucleoli over the diploid genome. Black dashed line indicates the median
698 and black dotted lines indicate the quartiles.
699 **f,g.** Levels of H3K9me2, H3K27ac, and H3K4me3 (**F**) and gene expression (**G**) at bi-allelic and mono-
700 allelic NADs. Values are shown as average RPKM. Violin plots where black dashed lines indicate the
701 median and black dotted lines indicate the quartiles. Statistical significance (*P*-values) was calculated
702 using the unpaired two-tailed t test (**<0.01; ***<0.001; ****<0.0001).
703 **h.** Pie chart showing the amounts of paternal (NAD_{CAST}) and maternal (NAD_{S129}) NADs.
704 **i.** Model showing that mono-allelic distribution of NADs at nucleoli of single cells is generally not parent
705 specific.
706
707
708
709
710
711
712
713

714 **Extended Data figure legends**

715 **Extended Data Figure S1**

716 **NoLMseq identifies NADs in single cells**

717 **a.** Gel electrophoresis showing whole-genome amplification (WGA) of isolated nucleoli (left panel) and
718 the PCR amplification for rRNA genes and *Tuba1* sequences.

719 **b.** NoLMseq tracks from 6 microdissected nucleoli at chromosomes 6, 8, at 19. Black bars below the
720 tracks indicate NADs called in respective single nucleoli.

721

722 **Extended Data Figure S2**

723 **NAD contact frequency for each chromosome**

724 Nucleolus contact frequency for the 53 ESC nucleoli along the length of each chromosome. rDNA-
725 chromosomes are depicted with an orange line.

726

727 **Extended Data Figure S3**

728 **Ribosomal protein and tRNA genes show high nucleolar contact frequency**

729 Distribution of ribosomal protein and tRNA genes at rDNA-chromosomes and chromosomes not
730 containing rRNA genes (other chromosomes).

731

732 **Extended Data Figure S4**

733 **NADs from single nucleoli of neural progenitors**

734 Gene ontology (GO) terms for upregulated and downregulated genes at ESC_{sp}-NADs in NPCs.

735

736 **Extended Data Figure S5**

737 **Nucleolar integrity is required for genomic contacts with nucleoli**

738 Coverage of NADs at each chromosome of single nucleoli from ESCs treated without and with
739 Actinomycin D (ActD). Dotted grey line indicates the average NAD coverage in single cells. Black lines
740 indicate the mean on each chromosome.

741

742 **Extended Data Figure S6**

743 **Genomic contacts with nucleoli are mono-allelic, but not parent specific.**

a. Pie chart showing the proportion of NADs identified in F123-ESCs that have also been found in E14-ESCs.

b. Mono-allelic distribution of 129 and CAST genome around nucleoli of single cells. Black dashed lines indicate median and black dotted lines indicate quartiles. Statistical significance (P-value) was calculated using the unpaired two-tailed t test (ns: non-significant).

c. Gene ontology (GO) terms of genes located on bi-allelic NADs.

761 **References**

762 1 Willemin, A., Szabo, D. & Pombo, A. Epigenetic regulatory layers in the 3D nucleus. *Mol Cell* **84**, 415-428
763 (2024). https://doi.org:10.1016/j.molcel.2023.12.032

764 2 Misteli, T. The Self-Organizing Genome: Principles of Genome Architecture and Function. *Cell* **183**, 28-45
765 (2020). https://doi.org:10.1016/j.cell.2020.09.014

766 3 Bersaglieri, C. & Santoro, R. Genome Organization in and around the Nucleolus. *Cells* **8(6)**, 579 (2019).
767 https://doi.org:10.3390/cells8060579

768 4 Bassler, J. & Hurt, E. Eukaryotic Ribosome Assembly. *Annu Rev Biochem* **88**, 281-306 (2019).
769 https://doi.org:10.1146/annurev-biochem-013118-110817

770 5 Pelletier, J., Thomas, G. & Volarevic, S. Ribosome biogenesis in cancer: new players and therapeutic avenues.
771 *Nat Rev Cancer* **18**, 51-63 (2018). https://doi.org:10.1038/nrc.2017.104

772 6 Derenzini, M. *et al.* Nucleolar size indicates the rapidity of cell proliferation in cancer tissues. *The Journal of*
773 *pathology* **191**, 181-186 (2000). https://doi.org:10.1002/(SICI)1096-9896(200006)191:2<181::AID-
774 PATH607>3.0.CO;2-V

775 7 Boisvert, F. M., van Koningsbruggen, S., Navascues, J. & Lamond, A. I. The multifunctional nucleolus. *Nat Rev*
776 *Mol Cell Biol* **8**, 574-585 (2007). https://doi.org:nrm2184 [pii] 10.1038/nrm2184

777 8 Goodfellow, S. J. & Zomerdijk, J. C. Basic mechanisms in RNA polymerase I transcription of the ribosomal RNA
778 genes. *Subcell Biochem* **61**, 211-236 (2013). https://doi.org:10.1007/978-94-007-4525-4_10

779 9 Ruggero, D. Revisiting the nucleolus: from marker to dynamic integrator of cancer signaling. *Sci Signal* **5**, pe38
780 (2012). https://doi.org:10.1126/scisignal.2003477

781 10 Corman, A., Sirozh, O., Lafarga, V. & Fernandez-Capetillo, O. Targeting the nucleolus as a therapeutic strategy
782 in human disease. *Trends Biochem Sci* **48**, 274-287 (2023). https://doi.org:10.1016/j.tibs.2022.09.006

783 11 Yang, K., Yang, J. & Yi, J. Nucleolar Stress: hallmarks, sensing mechanism and diseases. *Cell Stress* **2**, 125-
784 140 (2018). https://doi.org:10.15698/cst2018.06.139

785 12 Boukoura, S. & Larsen, D. H. Nucleolar organization and ribosomal DNA stability in response to DNA damage.
786 *Curr Opin Cell Biol* **89**, 102380 (2024). https://doi.org:10.1016/j.ceb.2024.102380

787 13 Bersaglieri, C. & Santoro, R. Methods for mapping 3D-chromosome architecture around nucleoli. *Curr Opin*
788 *Cell Biol* **81**, 102171 (2023). https://doi.org:10.1016/j.ceb.2023.102171

789 14 Nemeth, A. *et al.* Initial genomics of the human nucleolus. *PLoS Genet* **6**, e1000889 (2010).
790 https://doi.org:10.1371/journal.pgen.1000889

791 15 Koningsbruggen, S. v. *et al.* High-Resolution Whole-Genome Sequencing Reveals That Specific Chromatin
792 Domains from Most Human Chromosomes Associate with Nucleoli. *Molecular Biology of the Cell* **21**, 3735-
793 3748 (2010). https://doi.org:10.1091/mbc.e10-06-0508

794 16 Bizhanova, A., Yan, A., Yu, J., Zhu, L. J. & Kaufman, P. D. Distinct features of nucleolus-associated domains
795 in mouse embryonic stem cells. *Chromosoma* (2020). https://doi.org:10.1007/s00412-020-00734-9

796 17 Vertii, A. *et al.* Two contrasting classes of nucleolus-associated domains in mouse fibroblast heterochromatin.
797 *Genome Res* **29**, 1235-1249 (2019). https://doi.org:10.1101/gr.247072.118

798 18 Bersaglieri, C. *et al.* Genome-wide maps of nucleolus interactions reveal distinct layers of repressive chromatin
799 domains. *Nat Commun* **13**, 1483 (2022). https://doi.org:10.1038/s41467-022-29146-2

800 19 Shore, D. & Albert, B. Ribosome biogenesis and the cellular energy economy. *Curr Biol* **32**, R611-R617 (2022).
801 https://doi.org:10.1016/j.cub.2022.04.083

802 20 Lempiainen, H. & Shore, D. Growth control and ribosome biogenesis. *Curr Opin Cell Biol* **21**, 855-863 (2009).
803 https://doi.org:10.1016/j.ceb.2009.09.002

21 Datta, S. *et al.* Laser capture microdissection: Big data from small samples. *Histol Histopathol* **30**, 1255-1269 (2015). https://doi.org:10.14670/HH-11-622

22 Beagrie, R. A. *et al.* Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**, 519-524 (2017). https://doi.org:10.1038/nature21411

23 Baslan, T. *et al.* Genome-wide copy number analysis of single cells. *Nat Protoc* **7**, 1024-1041 (2012). https://doi.org:10.1038/nprot.2012.039

24 Zang, C. *et al.* A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**, 1952-1958 (2009). https://doi.org:10.1093/bioinformatics/btp340

25 Quinodoz, S. A. *et al.* Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. *Cell* **174**, 744-757 e724 (2018). https://doi.org:10.1016/j.cell.2018.05.024

26 Takei, Y. *et al.* Integrated spatial genomics reveals global architecture of single nuclei. *Nature* **590**, 344-350 (2021). https://doi.org:10.1038/s41586-020-03126-2

27 Meshorer, E. *et al.* Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* **10**, 105-116 (2006). https://doi.org:S1534-5807(05)00421-1 [pii] 10.1016/j.devcel.2005.10.017

28 Gaspar-Maia, A., Alajem, A., Meshorer, E. & Ramalho-Santos, M. Open chromatin in pluripotency and reprogramming. *Nature reviews. Molecular cell biology* **12**, 36-47 (2011). https://doi.org:10.1038/nrm3036

29 Reynolds, R. C., Montgomery, P. O. & Hughes, B. Nucleolar "Caps" Produced by Actinomycin D. *Cancer Res* **24**, 1269-1277 (1964).

30 Sun, Y. *et al.* MSL2 ensures biallelic gene expression in mammals. *Nature* **624**, 173-181 (2023). https://doi.org:10.1038/s41586-023-06781-3

31 Irastorza-Azcarate, I. *et al.* Extensive folding variability between homologous chromosomes in mammalian cells. *bioRxiv* (2024). https://doi.org:10.1101/2024.05.08.591087

32 Gribnau, J., Hochedlinger, K., Hata, K., Li, E. & Jaenisch, R. Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. *Genes Dev* **17**, 759-773 (2003). https://doi.org:10.1101/gad.1059603

33 Markowski, J. *et al.* GAMIBHEAR: whole-genome haplotype reconstruction from Genome Architecture Mapping data. *Bioinformatics* **37**, 3128-3135 (2021). https://doi.org:10.1093/bioinformatics/btab238

34 Wang, C., Ma, H., Baserga, S. J., Pederson, T. & Huang, S. Nucleolar structure connects with global nuclear organization. *Mol Biol Cell* **34**, ar114 (2023). https://doi.org:10.1091/mbc.E23-02-0062

35 Kefalopoulou, S. *et al.* Time-resolved and multifactorial profiling in single cells resolves the order of heterochromatin formation events during X-chromosome inactivation. *bioRxiv*, 2023.2012.2015.571749 (2023). https://doi.org:10.1101/2023.12.15.571749

36 Berger, A. B. *et al.* High-resolution statistical mapping reveals gene territories in live yeast. *Nat Methods* **5**, 1031-1037 (2008). https://doi.org:10.1038/nmeth.1266

37 Goronzy, I. N. *et al.* Simultaneous mapping of 3D structure and nascent RNAs argues against nuclear compartments that preclude transcription. *Cell Rep* **41**, 111730 (2022). https://doi.org:10.1016/j.celrep.2022.111730

38 Arrastia, M. V. *et al.* Single-cell measurement of higher-order 3D genome organization with scSPRITE. *Nat Biotechnol* **40**, 64-73 (2022). https://doi.org:10.1038/s41587-021-00998-1

39 Joshi, O. *et al.* Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. *Cell Stem Cell* **17**, 748-757 (2015). https://doi.org:10.1016/j.stem.2015.11.010

40 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012). https://doi.org:10.1038/nmeth.1923

41 Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10** (2021). https://doi.org:10.1093/gigascience/giab008

42 Zang, C., Wang, Y. & Peng, W. RECOGNICER: A coarse-graining approach for identifying broad domains from ChIP-seq data. *Quant Biol* **8**, 359-368 (2020). https://doi.org:10.1007/s40484-020-0225-2

43 Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187-191 (2014). https://doi.org:10.1093/nar/gku365

44 Krueger, F. & Andrews, S. R. SNPsplit: Allele-specific splitting of alignments between genomes with known SNP genotypes. *F1000Res* **5**, 1479 (2016). https://doi.org:10.12688/f1000research.9037.2

45 Robinson, J. T. *et al.* Integrative genomics viewer. *Nature Biotechnology* (2011). https://doi.org:10.1038/nbt0111-24

46 Andrews, S. FastQC: a quality control tool for high throughput sequence data. (2010).

47 Kopylova, E. N., L.; Touzet, H. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* (2012).

48 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014). https://doi.org:10.1093/bioinformatics/btu170

49 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013). https://doi.org:10.1093/bioinformatics/bts635

50 Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**, 621-628 (2008). https://doi.org:10.1038/nmeth.1226

51 Lerra, L. *et al.* An RNA-dependent and phase-separated active subnuclear compartment safeguards repressive chromatin domains. *Mol Cell* **84**, 1667-1683 e1610 (2024). https://doi.org:10.1016/j.molcel.2024.03.015

52 Sherman, B. T. *et al.* DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res* **50**, W216-W221 (2022). https://doi.org:10.1093/nar/gkac194

53 Beliveau, B. J. *et al.* Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. *Nat Commun* **6**, 7147 (2015). https://doi.org:10.1038/ncomms8147

54 Szabo, Q., Cavalli, G. & Bantignies, F. in *Capturing Chromosome Conformation: Methods and Protocols* (eds Beatrice Bodega & Chiara Lanzuolo) 221-237 (Springer US, 2021).

55 Marks, H. *et al.* The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**, 590-604 (2012). https://doi.org:10.1016/j.cell.2012.03.026

56 von Meyenn, F. *et al.* Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. *Mol Cell* **62**, 848-861 (2016). https://doi.org:10.1016/j.molcel.2016.04.025

57 Bonev, B. *et al.* Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* **171**, 557-572.e524 (2017). https://doi.org:10.1016/j.cell.2017.09.043

58 Peric-Hupkes, D. *et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell* **38**, 603-613 (2010). https://doi.org:S1097-2765(10)00321-7 [pii]10.1016/j.molcel.2010.03.016

Figure 1

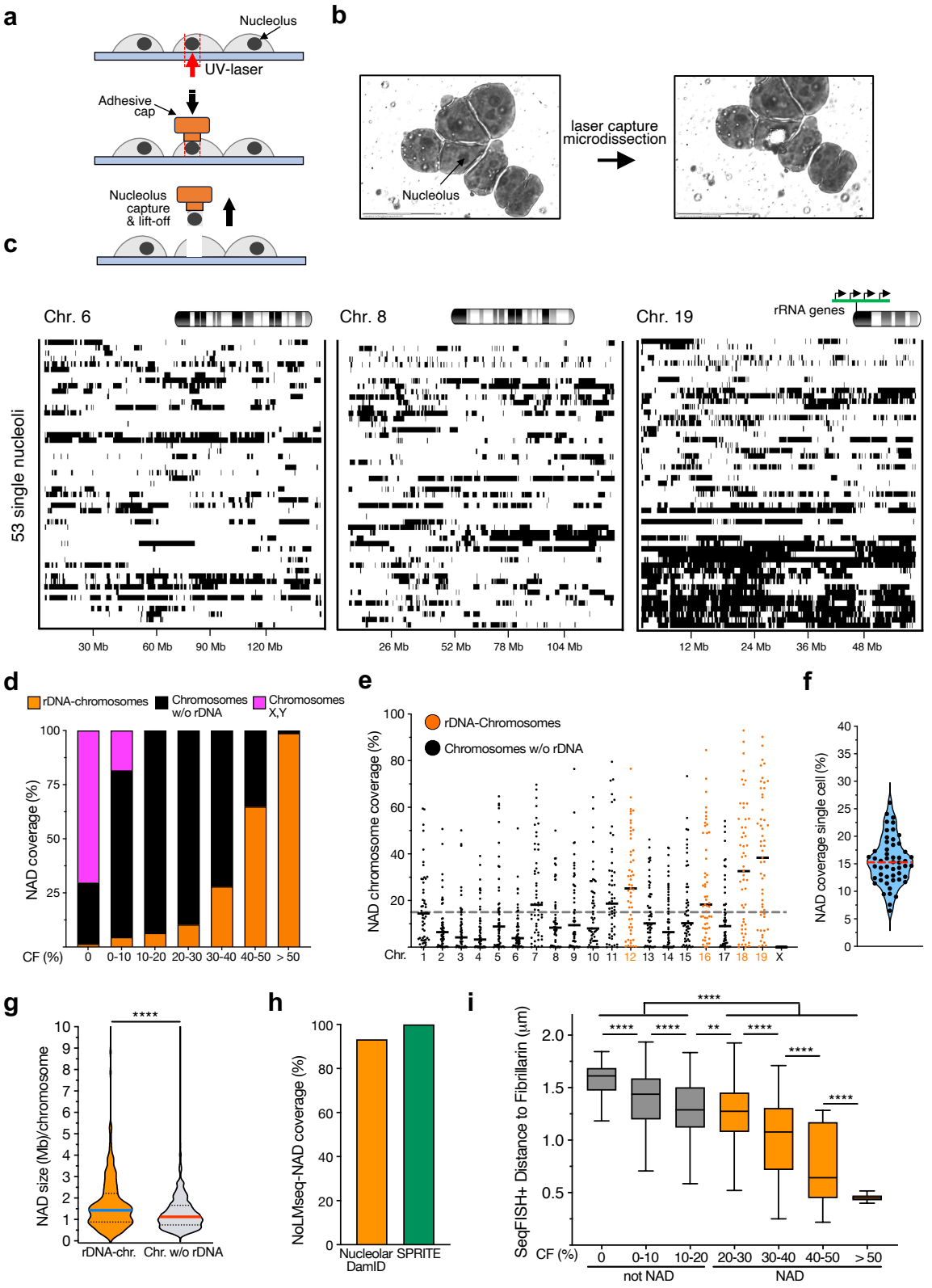

a
Nucleolus
UV-laser
Adhesive cap
Nucleolus capture & lift-off
b
Nucleolus
laser capture microdissection
c
Chr. 6
Chr. 8
Chr. 19
rRNA genes
53 single nucleoli
30 Mb
60 Mb
90 Mb
120 Mb
26 Mb
52 Mb
78 Mb
104 Mb
12 Mb
24 Mb
36 Mb
48 Mb
d
rDNA-chromosomes
Chromosomes w/o rDNA
Chromosomes X,Y
NAD coverage (%)
CF (%)
0
0-10
10-20
20-30
30-40
40-50
> 50
e
rDNA-Chromosomes
Chromosomes w/o rDNA
NAD chromosome coverage (%)
Chr.
f
NAD coverage single cell (%)
g
NAD size (Mb)/chromosome
rDNA-chr.
Chr. w/o rDNA
****
h
NoLMseq-NAD coverage (%)
Nucleolar DamID
SPRITE
i
SeqFISH+ Distance to Fibrillarin (µm)
CF (%)
not NAD
NAD
****
**

Figure 2

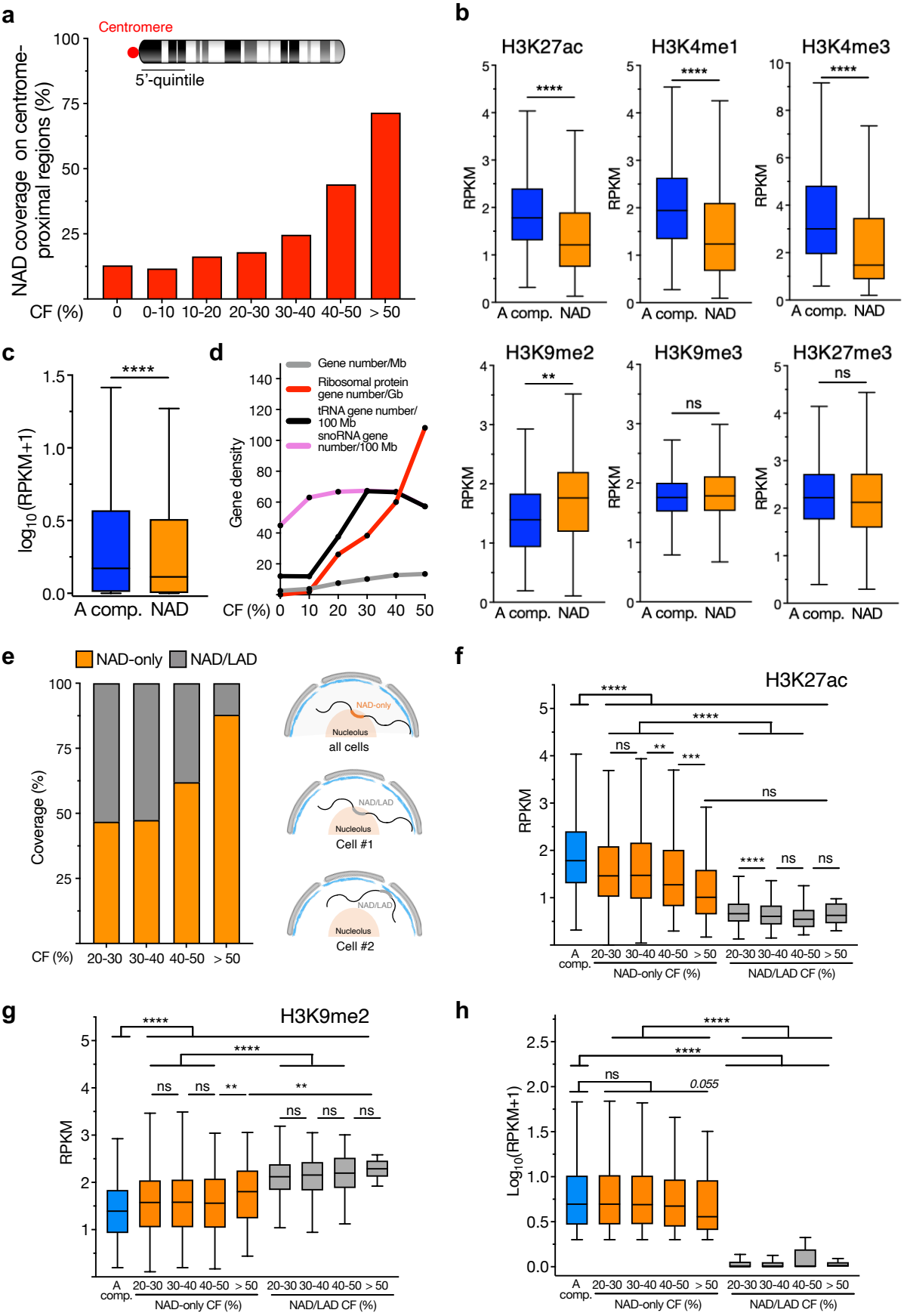

a
Centromere
5'-quintile
NAD coverage on centrome-proximal regions (%)
CF (%)
0
0-10
10-20
20-30
30-40
40-50
> 50
b
H3K27ac
H3K4me1
H3K4me3
RPKM
A comp.
NAD
****
c
log10(RPKM+1)
****
d
Gene number/Mb
Ribosomal protein gene number/Gb
tRNA gene number/ 100 Mb
snoRNA gene number/100 Mb
Gene density
CF (%)
H3K9me2
H3K9me3
H3K27me3
**
ns
e
NAD-only
NAD/LAD
Coverage (%)
CF (%)
20-30
30-40
40-50
> 50
Nucleolus
all cells
Cell #1
Cell #2
f
H3K27ac
A comp.
NAD-only CF (%)
NAD/LAD CF (%)
g
H3K9me2
h
Log10(RPKM+1)
0.055

Figure 3

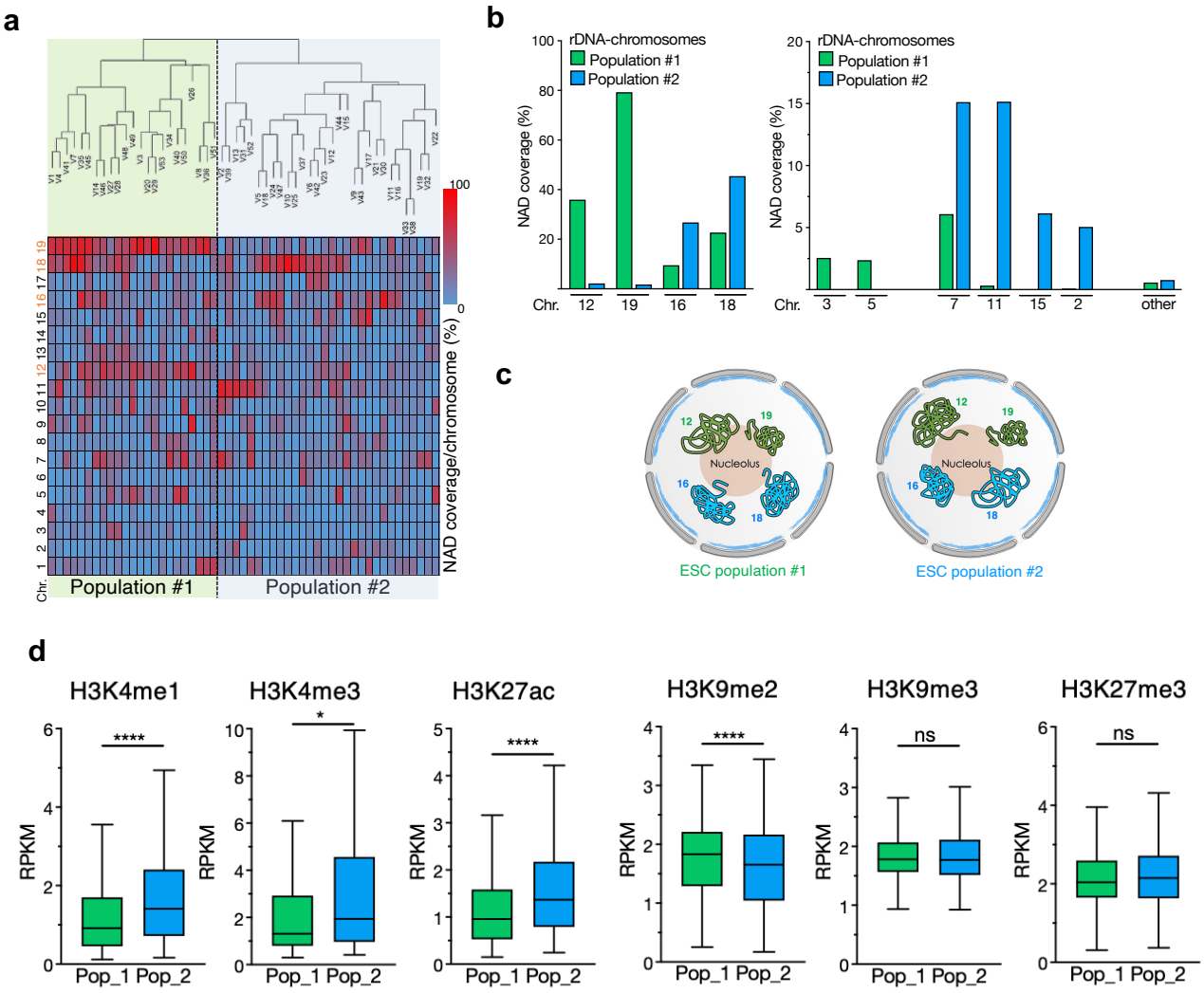

a
Population #1
Population #2
Chr.
NAD coverage/chromosome (%)
b
rDNA-chromosomes
Population #1
Population #2
NAD coverage (%)
Chr.
12
19
16
18
3
5
7
11
15
2
other
c
Nucleolus
ESC population #1
ESC population #2
d
H3K4me1
H3K4me3
H3K27ac
H3K9me2
H3K9me3
H3K27me3
RPKM
Pop_1
Pop_2
****
*
ns

Figure 4

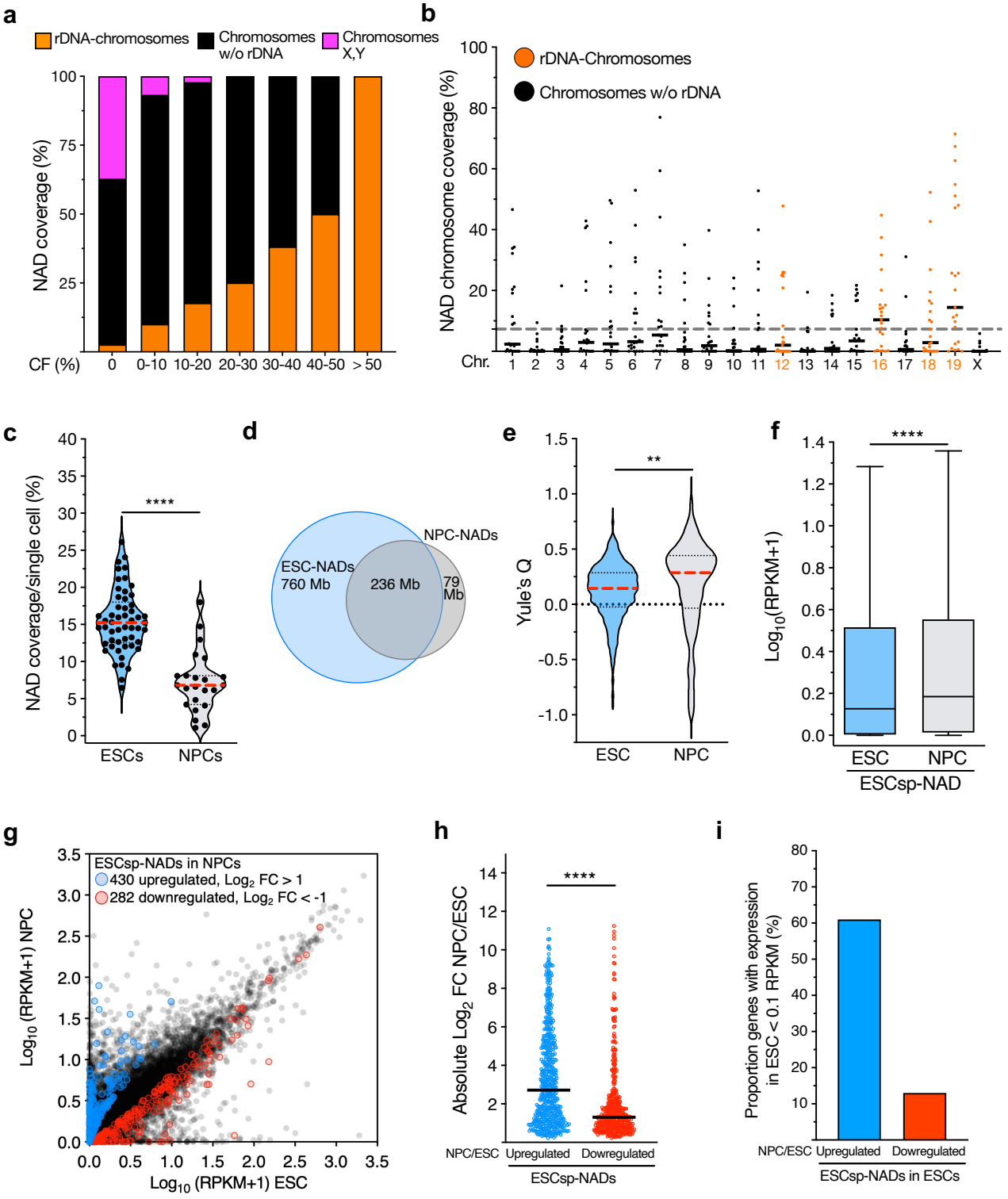

a
rDNA-chromosomes
Chromosomes w/o rDNA
Chromosomes X,Y
NAD coverage (%)
CF (%)
0
0-10
10-20
20-30
30-40
40-50
> 50
b
rDNA-Chromosomes
Chromosomes w/o rDNA
NAD chromosome coverage (%)
Chr.
c
NAD coverage/single cell (%)
****
ESCs
NPCs
d
ESC-NADs
760 Mb
236 Mb
NPC-NADs
79 Mb
e
Yule's Q
**
ESC
NPC
f
Log10(RPKM+1)
****
ESC
NPC
ESCsp-NAD
g
ESCsp-NADs in NPCs
430 upregulated, Log2 FC > 1
282 downregulated, Log2 FC < -1
Log10 (RPKM+1) NPC
Log10 (RPKM+1) ESC
h
Absolute Log2 FC NPC/ESC
****
NPC/ESC
Upregulated
Downregulated
ESCsp-NADs
i
Proportion genes with expression in ESC < 0.1 RPKM (%)
NPC/ESC
Upregulated
Downregulated
ESCsp-NADs in ESCs

Figure 5

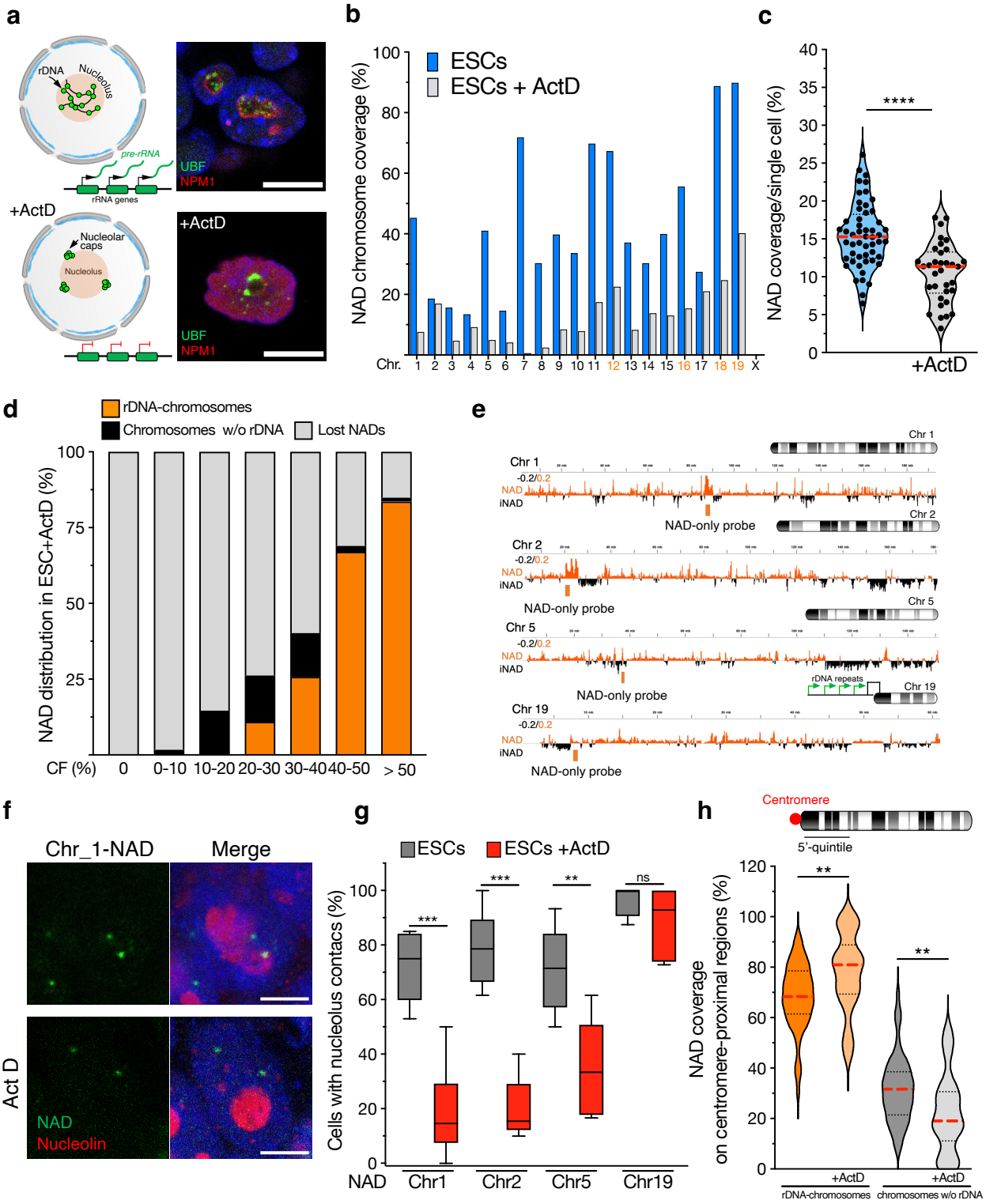
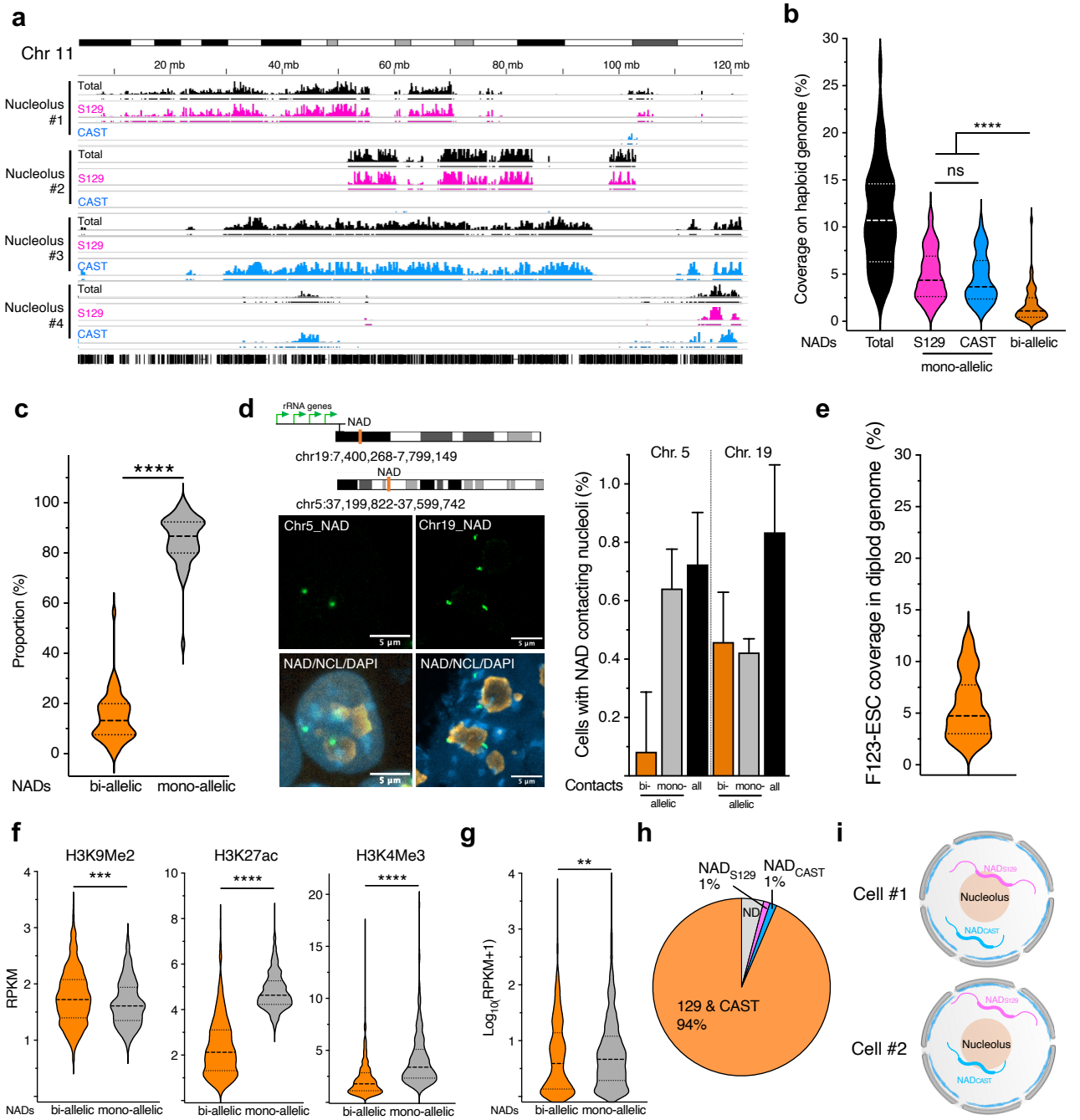
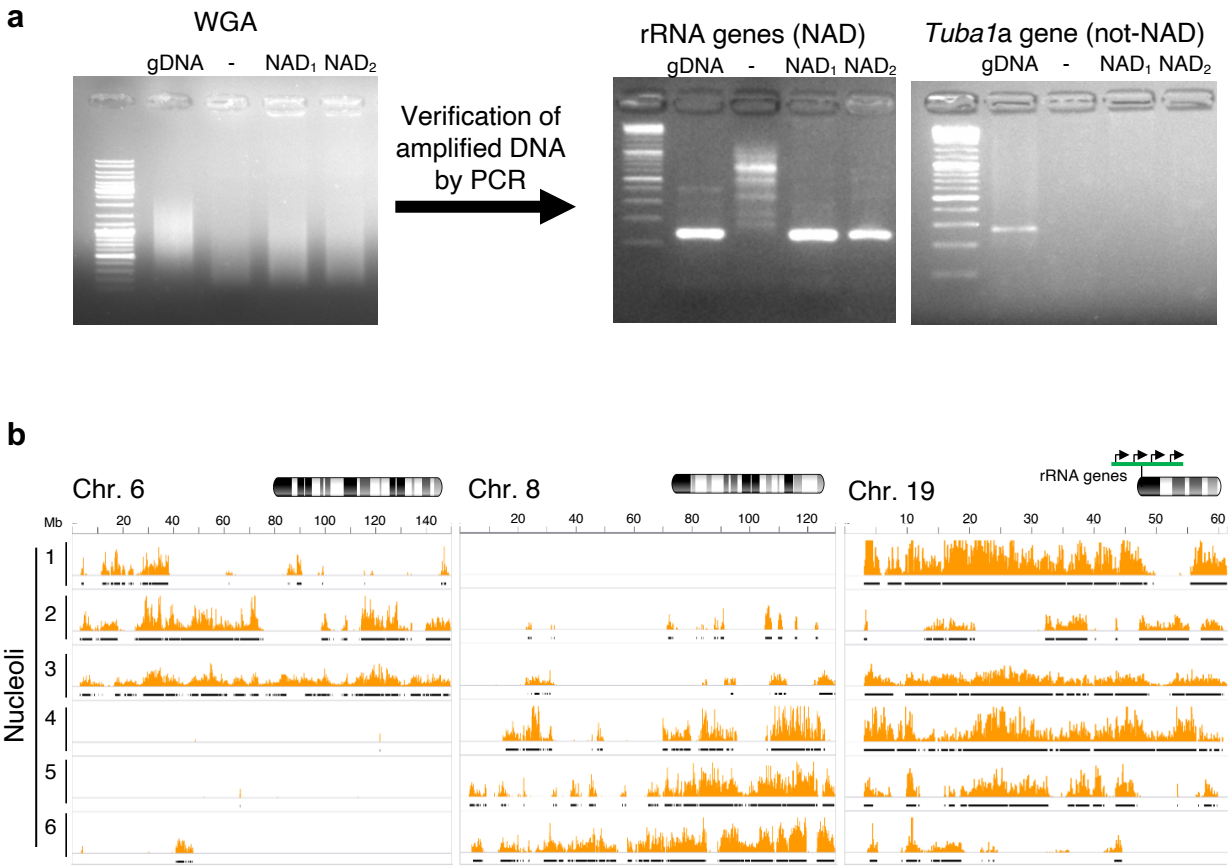

a
rDNA
Nucleolus
pre-rRNA
rRNA genes
UBF
NPM1
+ActD
Nucleolar caps
Nucleolus
+ActD
UBF
NPM1
b
ESCs
ESCs + ActD
NAD chromosome coverage (%)
Chr. 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
c
NAD coverage/single cell (%)
****
+ActD
d
rDNA-chromosomes
Chromosomes w/o rDNA
Lost NADs
NAD distribution in ESC+ActD (%)
CF (%) 0 0-10 10-20 20-30 30-40 40-50 > 50
e
Chr 1
Chr 2
Chr 5
Chr 19
-0.2/0.2
NAD
iNAD
NAD-only probe
rDNA repeats
f
Chr_1-NAD
Merge
Act D
NAD
Nucleolin
g
ESCs
ESCs +ActD
Cells with nucleolus contacs (%)
***
***
**
ns
NAD Chr1 Chr2 Chr5 Chr19
h
Centromere
5'-quintile
NAD coverage on centromere-proximal regions (%)
**
**
+ActD
+ActD
rDNA-chromosomes
chromosomes w/o rDNA

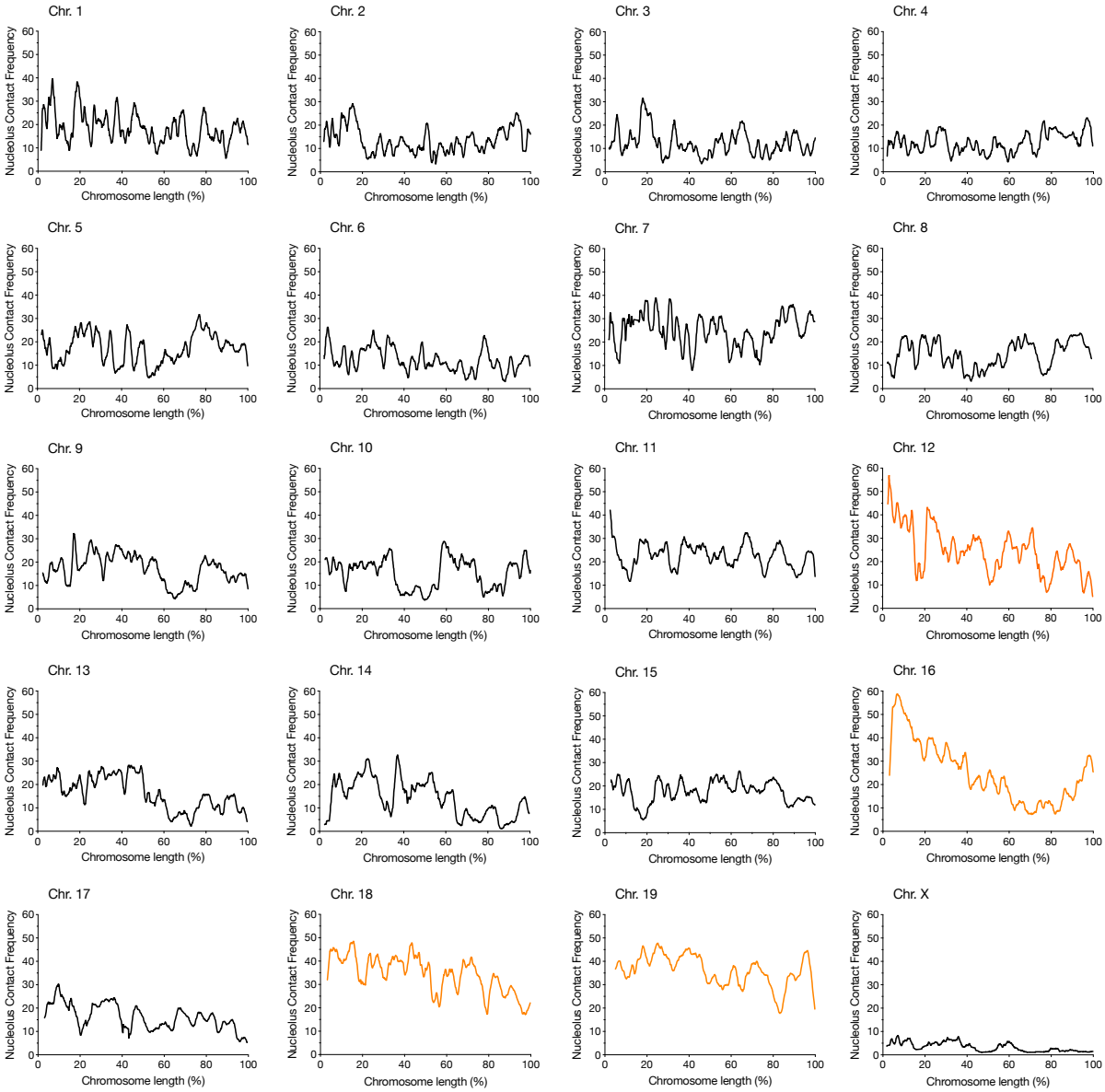
Figure 6


a
Chr 11
20 mb
40 mb
60 mb
80 mb
100 mb
120 mb
Nucleolus #1
Nucleolus #2
Nucleolus #3
Nucleolus #4
Total
S129
CAST
b
Coverage on haploid genome (%)
****
ns
NADs
Total
S129
CAST
bi-allelic
mono-allelic
c
Proportion (%)
****
NADs
bi-allelic
mono-allelic
d
rRNA genes
NAD
chr19:7,400,268-7,799,149
chr5:37,199,822-37,599,742
Chr5_NAD
Chr19_NAD
NAD/NCL/DAPI
NAD/NCL/DAPI
5 µm
Cells with NAD contacting nucleoli (%)
Chr. 5
Chr. 19
Contacts
bi-
mono-
all
allelic
e
F123-ESC coverage in diplod genome (%)
f
H3K9Me2
H3K27ac
H3K4Me3
RPKM
***
****
****
NADs
bi-allelic
mono-allelic
g
Log10(RPKM+1)
**
NADs
bi-allelic
mono-allelic
h
NAD S129 1%
NAD CAST 1%
ND
129 & CAST 94%
i
Cell #1
Cell #2
Nucleolus
NADS129
NADCAST

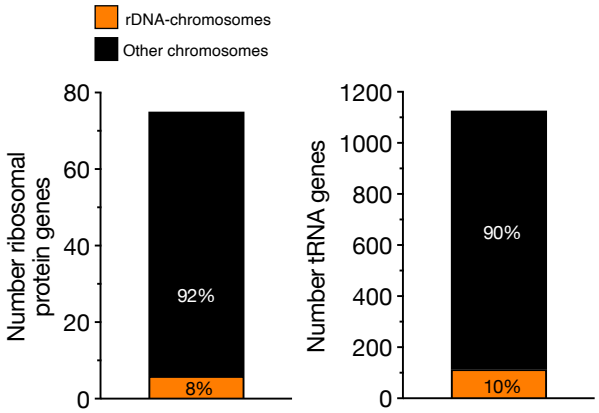
Extended data Figure 1



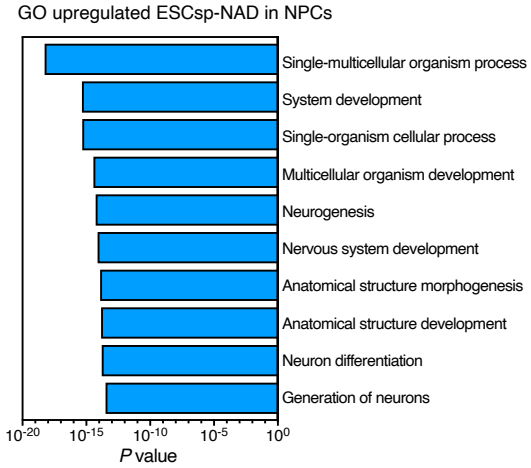
Extended data Figure 2

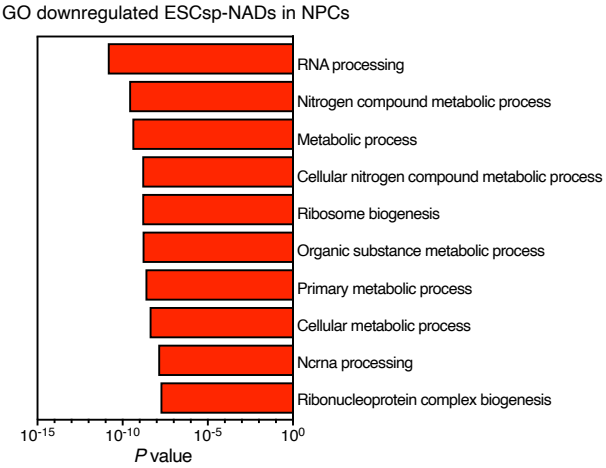


Extended data Figure 3

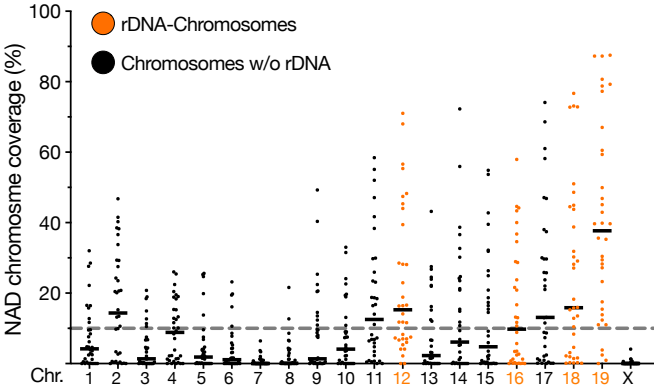
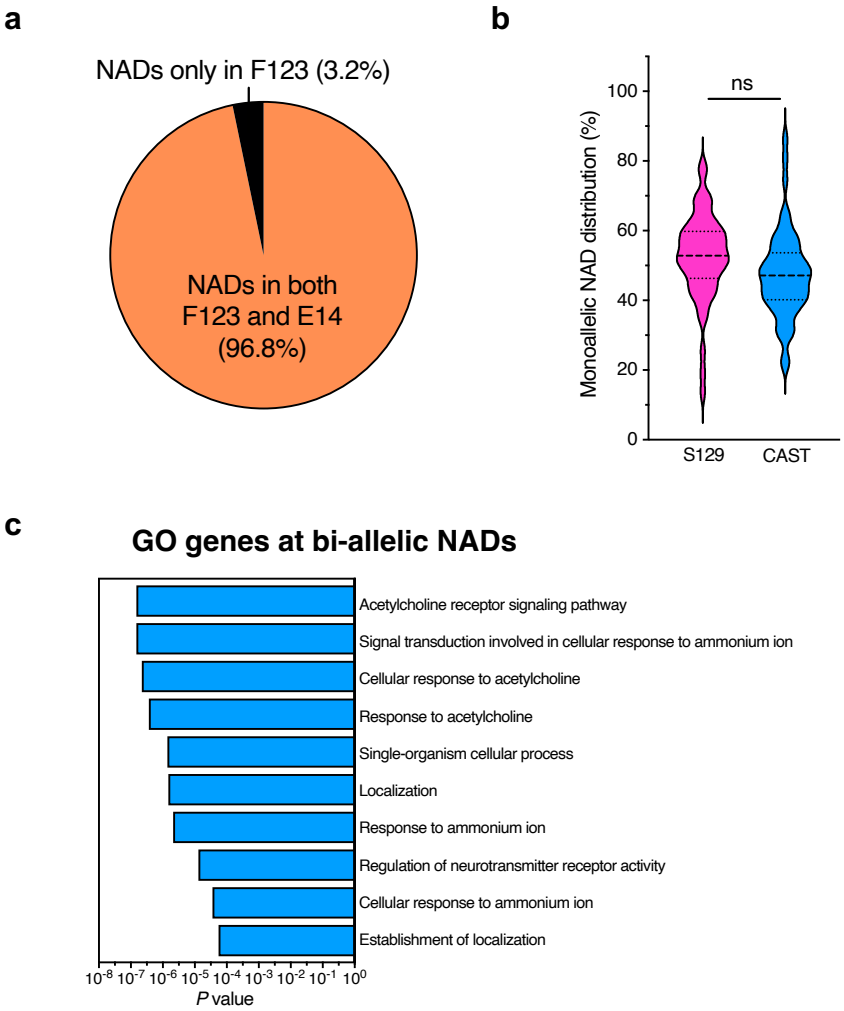


Extended data Figure 4





Extended data Figure 5



Extended data Figure 6