Universal and distinct features of intra-population heterogeneity between differentiated cells and pluripotent stem cells

Zhisheng Jiang*[3], Serena Francesca Generoso*[1], Marta Badia*[2], Bernhard Payer[†1,2], Lucas B. Carey[†2,3]

[1] Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr. Aiguader 88, 08003, Barcelona, Spain.
[2] Universitat Pompeu Fabra (UPF), Barcelona, Spain.
[3] Center for Quantitative Biology and Peking-Tsinghua Center for the Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

[†]Correspondence: bernhard.payer@crg.eu and lucas.carey@pku.edu.cn
*These authors contributed equally to the work

Abstract:

Isogenic cells cultured in the same nutrient-rich environment show heterogeneity in their proliferation rate. To understand the differences between fast and slow-proliferating cells and to identify markers for proliferation rate that can be used at the single-cell level, we developed a method to sort cells by their proliferation rate, and performed RNA sequencing (RNA-Seq) on slow, medium and fast proliferating subpopulations of pluripotent mouse embryonic stem cells (mESCs) and immortalized mouse fibroblasts. We identified a core proliferation-correlated transcriptome that is common to both cell types, to yeast, and to cancer cells: fast proliferating cells have higher expression of genes involved in both protein synthesis and protein degradation. In contrast to cells sorted by proliferation rate, RNA-seq on cells sorted by mitochondria membrane potential revealed a highly cell-type specific mitochondria-state related transcriptome. mESCs with hyperpolarized mitochondria are fast proliferating, while the opposite is true for fibroblasts. In addition, cell-to-cell variation in proliferation rate is highly predictive of pluripotency state in mESCs, with cells of more naïve pluripotent character having a slower proliferation rate. Finally, we show that the proliferation signature learned from sorted cells can predict proliferation from scRNAseq data in both mESCs and in the developing nematode. While the majority of the transcriptional-signature associated with cell-to-cell heterogeneity in proliferation rate is conserved from yeast to embryos to differentiated cells to cancer, the metabolic and energetic details of cell growth are highly cell-type specific.

Introduction

Rates of cell growth and division vary greatly, even among isogenic cells of a single

cell-type, cultured in the same optimal environment [1]. Cell-to-cell heterogeneity in proliferation rate has important consequences for population survival in bacterial antibiotic resistance, stress resistance in budding yeast, and chemo-resistance in cancer [2-10]. Time-lapse fluorescence microscopy has shown that cell-to-cell variability in the expression of some genes, such as *p53* and *p21*, is associated with cell-to-cell variability in proliferation and survival [1, 11]. However, these methods can detect dynamic relationships between gene expression and cell fate, but are limited to measurements of one or two genes. Single-cell RNA sequencing measures transcriptome-level heterogeneity but does not directly link this to cell-biological heterogeneity in organelle state, or to dynamic heterogeneity in proliferation or drug resistance. Transcriptome-level approaches for understanding within-population cell-to-cell heterogeneity in proliferation and other dynamic processes are lacking. While the presence of intrapopulation variation in proliferation, transcriptome, and organelle-state in both steady-state and in differentiation populations is well established, the relationship among the three remains unclear.

One possibility is that the proliferation-correlated gene expression program is the same, regardless of if one looks at interpopulation variation due to genetic or environmental differences, or intrapopulation heterogeneity due to epigenetic differences. However, in the budding yeast *Saccharomyces cerevisiae,* the expression program of intrapopulation heterogeneity in proliferation rate only partially resembles that of cells growing at different rates due to genetic or environmental perturbations [8]. The relation between gene expression and proliferation rate is much less well studied in mammalian cells.

In yeast, in tumors, and in organs, genetic, environmental and developmental changes cause changes in proliferation rate, and changes in the expression of hundreds or possibly thousands of genes [12-16]. Unsurprisingly, many of the genes for which changes in expression are associated with changes in proliferation rate are associated with adverse clinical outcomes in cancer and with antibiotic and antifungal resistance [17, 18]. Within a population of microbes, and within a single multicellular organism, the correct balance of proliferation states and rates is essential. Yet measuring this heterogeneity is difficult, without which, understanding the consequences of this heterogeneity is impossible.

Expression is associated with phenotype, but mRNAs themselves do not necessarily always cause phenotypes. Instead, they can often serve as markers for cell-biological differences between cells. Phenotypes are mostly driven by larger cell-biological differences between cells, such as differences in metabolic state. Cell-to-cell heterogeneity in mitochondria state has been linked to differences in transcription rates, growth rates, proliferation and developmental trajectories [19-21]. Both cancer cells and pluripotent stem cells have atypical metabolisms and use glycolysis to produce much of their ATP, instead of

the mitochondria-based oxidative phosphorylation, which is the predominant form of ATP-generation in differentiated cells [22]. It is unknown if this inter-population variation in proliferation, transcriptome, and mitochondria extents to intra-population variation among single cells within a single isogenic population.

Pluripotent stem cells exist in different pluripotency states called naïve or primed based on culture conditions and embryonic origin [23]. Mouse ESCs reflect the naïve pluripotency state of the blastocyst epiblast and can be cultured in either serum+LIF or 2i+LIF conditions, the latter involving inhibitors of FGF/ERK and GSK3 pathways. Culture in 2i+LIF conditions promotes a ground state more closely mirroring the *in vivo* situation with reduced heterogeneity in pluripotency gene expression and different cell cycle profile when compared to cells grown in serum+LIF [24-26]. Nevertheless, even in 2i+LIF conditions, mESCs display a certain amount of cell-to-cell heterogeneity [27, 28] and it is unclear, how this relates to heterogeneity in differentiated cell types when it comes to gene expression and its link to proliferation rate.

To understand the relation between intra-population transcriptome heterogeneity and heterogeneity in proliferation, we developed a FACS-based method to sort cells by proliferation rate. We applied this method to mouse immortalized fibroblasts and to mESCs and performed RNA-seq on fast, medium and slow proliferating cell sub-populations. We find that ribosome-biogenesis (protein synthesis) and proteasome-related (protein degradation) genes are highly expressed in fast proliferating fibroblasts and ESCs. Moreover, the proliferation signature is conserved across cell-type and species, from yeast to cancer cells, allowing us to predict the relative proliferation rate from the transcriptome; we use the gene expression signature to correctly predict proliferation from scRNA-seq data not only in mESCs, but also during *C. elegans* development, in spite of no nematode data going into the initial model In contrast to the generality of this main transcriptional signature, many mitochondria-related genes were upregulated in fast proliferating fibroblasts, yet down-regulated in fast-proliferating mESCs. Consistent with this, the high mitochondria membrane potential is indicative of slow proliferating fibroblasts, while in mESCs it is characteristic of fast proliferating cells. Taken together, these results show the existence of a core protein-synthesis and protein-degradation expression program that is conserved across cell types and species, from yeast to mice, and a metabolic and energy-production program that is highly cell-type specific.

Results

A method to sort single mammalian cells by cell-to-cell heterogeneity in proliferation rate

To understand the causes and consequences of intrapopulation cell-to-cell heterogeneity

118 in proliferation rate in mammalian cells we developed a method for sorting single
119 mammalian cells by their proliferation rate **(Figure 1)**. The cell-permeable dye
120 carboxyfluorescein succinimidyl ester (CFSE) covalently binds to free amines within cells,
121 thus staining most intracellular proteins at lysine residues. We reasoned that in cell types
122 that divide symmetrically, such as embryonic stem cells and immortalized fibroblasts [29],
123 the equal dilution of CFSE into the two daughter cells would enable us to count the number
124 of divisions that each cell had undergone. To eliminate confounding effects due to
125 differences in initial staining we used fluorescence-activated cell sorting (FACS) to obtain an
126 initially homogeneous cell population of cells with identical CFSE signals **(Figure 1A)**. Thus,
127 the CFSE signal should be independent of cell-to-cell variation protein synthesis rates, as the
128 initial signal in each cell is determined by the FACS gate and not by dye update or protein
129 synthesis during staining. In addition, $CFSE_{CFR2}$ conjugates are stable and unable to exit the
130 cell [30]; the dye signal is stable for over eight weeks in non-dividing lymphocytes [31]. The
131 measured CFSE signal should be relatively insensitive to cell-to-cell variation in protein
132 degradation. We cultured this sorted starting cell population for several generations, during
133 which time the CFSE signal decreases with each cell division **(Figure 1B)**. Consistent with the
134 decrease in CFSE being mostly due to cell division, the population-level doubling time of
135 each cell type can be calculated based on the decrease in CFSE signal over time **(Figure 1C,**
136 **D)**, and these doubling times are consistent with those reported by other methods [32, 33].
137 After five days for fibroblasts growing in MEF (mouse embryonic fibroblast) medium, and
138 three for ESCs grown in pluripotent ground-state promoting 2i+LIF conditions [34], we used
139 FACS to isolate cells with high, medium, and low CFSE signal, and performed RNA-seq on
140 each sub-population. This allowed us to identify genes whose expression is positively or
141 negatively correlated with proliferation rate within a single population **(Figure 1E)**.
142

143 *Slow-proliferating ESCs are of more naïve pluripotent character than fast-proliferating ESCs*

144 Embryonic stem cells exhibit cell-to-cell heterogeneity in culture based on the
145 expression of naïve pluripotency marker genes such as *Nanog*, *Stella* (*Dppa3*) or *Rex1* (*Zfp42*)
146 [35-37]. Although this heterogeneity is mostly apparent in ESCs cultured in serum+LIF, even
147 when cultured in ground state-pluripotency-promoting 2i+LIF conditions, the sub-population
148 of ESCs with low NANOG-levels displays propensity for lineage-priming and differentiation
149 [28, 38]. To determine if cell-to-cell variation in proliferation rate was caused by a
150 sub-population of mESCs initiating a differentiation program, we determined the fold-change
151 in expression between slow and fast proliferating sub-populations for a set of genes that are
152 upregulated during lineage commitment. We found no consistent enrichment of these
153 differentiation genes in fast versus slow proliferating cells, as they could be found to be
154 expressed in either population **(Figure 2A)**. However, the slow proliferating sub-population
155 did have higher expression of genes that are upregulated in naïve pluripotent cells, and in

156 2-cell stage embryos **(Figure 2B,C)**, suggesting that slow proliferating mESCs are in a more
157 naïve pluripotent cell state than their fast proliferating counterparts.
158
159
160 *Processes correlated with cell-to-cell heterogeneity in proliferation rate that are consistent*
161 *across cell-types and species*
162 To identify functional groups of genes that are differentially expressed between fast and
163 slow proliferating cells within a single population we performed gene set enrichment analysis
164 (GSEA) [39, 40] **(Figure 3A and 3B)** on mRNA-seq data from fast and slow proliferating
165 subpopulations. We found that, in both fibroblasts and ESCs, as well as in the budding yeast
166 *S. cerevisiae*, genes involved in ribosome-biogenesis and the proteasome are more highly
167 expressed in fast proliferating cells **(Figure 3C, 3D and Table S1)**. High expression of
168 ribosomal genes is a common signature for fast proliferating cells [12, 41], and cancer cells
169 often exhibit high proteasome expression [42-44], but it is not clear if this is related to
170 proliferation in-and-of-itself or due to aneuploidy and other genetic alterations [45]. These
171 results suggest that coordinated regulation of the ribosome and proteasome are a signature
172 of fast proliferating cells across both cell-types and species.
173
174 In addition to ribosome-biogenesis and the proteasome, several other gene sets are
175 differentially expressed between fast and slow proliferating cells in both fibroblasts and ESCs
176 **(Figure 3C)**. mTORC1 (mammalian Target Of Rapamycin Complex 1) functions as a nutrient
177 sensor and regulator of protein synthesis, and is regulated by nutrient and cytokine
178 conditions that cause differences in proliferation [46, 47]. We find that, even in the absence
179 of genetic and environmental differences, mTORC1 is more active in fast proliferating cells.
180 Activation of mTORC1 can promote ribosome-biogenesis [46, 48], however, there is still
181 controversy about the regulation of proteasome activity by mTORC1 [47, 49-53]. We
182 observed in both fibroblasts and ESCs, that fast proliferating subpopulations exhibit a
183 transcriptional signature of increased protein synthesis, protein degradation.
184
185 Furthermore, we identified target genes of MYC to be more highly expressed in fast
186 proliferating cells. MYC, a transcription factor frequently amplified in cancer, is estimated to
187 regulate the transcription of at least 15% of all genes [54] and is a master regulator of cell
188 growth [55]. Overexpression of MYC promotes ribosome-biogenesis and cell growth [56, 57],
189 and active mTORC1 can promote MYC activation [58, 59]. Our data suggest that increased
190 expression of MYC and increased mTORC1 activity are general properties of fast-proliferating
191 cells, and those genetic or environmental perturbations are not necessary to cause
192 differential expression of these pathways.
193

194 *scRNA-seq of developing nematodes revleals that, compared to terminally differentiated cells,*
195 *proliferating cells have higher expression of ribosome biogenesis and proteasome genes*
196 Single-cell RNA sequencing is a powerful method for understanding cell-to-cell
197 heterogeneity, but it suffers from high levels of technical noise at the single-gene level. In
198 addition, most commonly used markers (PCNA, Ki67) for measuring proliferation rates in bulk
199 populations are cell-cycle regulated genes; what is really being measured is the fraction of
200 the population that is proliferating, usually the fraction that is in S phase. Thus, these
201 markers cannot be used to predict proliferation rates from scRNAseq data. We reasoned that,
202 as the average expression of large sets of genes, many of which are highly expressed and
203 therefore have lower levels of technical noise, the ribosome biogenesis and proteasome
204 gene sets would be ideal for differentiating proliferating vs non-proliferating cells in
205 scRNA-sequencing data, indpenedent of the cell-cycle position of individual cells. To test this
206 we used a new scRNA-seq dataset of 86,024 cells from C. elegans in which cells have been
207 classified into terminally differentiated and preterminal cell-types[60]. We find that
208 terminally differentiated cells have lower expression of ribosome biogenesis and proteasome
209 genes **(Figure 3E)**, consistent with terminally differentiated cells having proliferation rates of
210 zero.
211
212 *Coordination of protein-synthesis and protein-degradation across cell types, organs and*
213 *species.*
214
215 Significant enrichment results of proteasome and ribosome-biogenesis in fast
216 proliferating fibroblasts, ESCs and yeast suggested that expression of the proteasome and
217 ribosome-biogenesis may serve as cell-type independent reporters of growth rate. To test
218 this hypothesis we analyzed RNA-seq data from 528 cancer cell lines in the Cancer Cell Line
219 Encyclopedia [61] for which the doubling time is roughly known. As GSEA is a measure of
220 differential expression, we created a single common control sample as the median
221 expression of each gene across all 528 cancer cell lines, and used GSEA to calculate the NES
222 (Normalized Enrichment Score) for all gene sets between the single control and each cancer
223 cell line **(Figure 4A)**.
224
225 Ribosome and proteasome-related gene sets were among the gene sets most highly
226 correlated with growth rate across all cancers **(Figure 4B)**. The absolute values of the
227 correlation of all gene sets with the reported doubling time were low, possibly because the
228 doubling times of the cancer cell lines were not measured using exactly the same
229 experimental conditions as were used for the RNA-seq experiments. We also calculated the
230 correlation between measured doubling time and meta-PCNA [17, 62], an RNA-seq-based
231 method for estimating growth rate at the population level that is independent of ribosomal
232 or proteasomal gene expression, but found similar levels of correlation **(Figure 4C)**.
233 Interestingly, proteasome and ribosome gene sets were far more strongly correlated with

234 each other than with proliferation rate **(Figure 4D)**, suggesting a strong mechanistic coupling
235 between increased protein production and a need for increased protein degradation.
236
237 Most functional groups enriched in both fast fibroblast and ESCs are positively
238 correlated with each other across the 528 cancer cell lines **(Figure 4D and Table S1)**. This is
239 not the case for gene sets whose expression is negatively correlated with intra-population
240 variation in proliferation rate. Several *p53* related gene sets are strongly negatively correlated
241 with proliferation within both fibroblasts and ESCs **(Figure 4D and Table S1)**, but the results
242 are much more heterogeneous across cancer cell lines, possibly reflecting the cancer-specific
243 mutation status of genes in this pathway.
244
245 To test if the coupling between ribosome biogenesis and proteasome expression holds
246 across species, we analyzed the bulk RNA-seq data across developmental stages, covering
247 multiple organs in seven species [16]. A high correlation between ribosome biogenesis genes
248 expression and proteasome genes expression was found across all seven species **(Figure 4E)**.
249 The coordinated expression change with developmental stages between ribosome
250 biogenesis genes and proteasome genes across different organs in seven species suggests
251 that the coordination between protein synthesis and degradation is common across all
252 species and cell-types **(Figure S1)**.
253
254 *The major cell-type specific proliferation-correlated expression is in mitochondria and*
255 *metabolism related genes.*
256
257 While the pattern of within-population proliferation-correlated expression in yeast,
258 fibroblasts and ESCs was broadly similar with regard to genes involved in protein synthesis
259 and degradation, the behavior of metabolic and mitochondria-related genes in fast and slow
260 proliferating subpopulations was highly cell-type specific. Mitochondria membrane and
261 respiratory chain-related gene sets were more highly expressed in fast proliferating
262 fibroblasts, but not in fast proliferating ESCs **(Table 1)**. These results are consistent with
263 differential mitochondrial states in ESCs when compared to differentiated cells like
264 fibroblasts [22], which suggest the existence of different types of metabolism and
265 proliferation-related heterogeneity between pluripotent and differentiated cell-types. Fast
266 proliferating sub-populations of different cell-types display differential importance and
267 metabolic states related to mitochondria. We also observed cell-type specific differences in
268 glycolysis, fatty acid metabolism, and other metabolic processes, suggesting fundamental
269 differences in the metabolic pathways required for fast proliferation between pluripotent
270 ESCs and differentiated cells like fibroblasts **(Table 1)**.
271
272

Cell-to-cell heterogeneity in mitochondria state predicts variation in proliferation both in ESCs and fibroblasts, but in opposite directions.

Mitochondrial membrane potential is a major predictor of cell-to-cell heterogeneity in proliferation rate in budding yeast [9]. Mitochondria-related genes are more highly expressed in the fast proliferating subpopulation of fibroblasts. In contrast, these genes are slightly more highly expressed in the slow proliferating subpopulation of ESCs. This suggests that the relation between cell-to-cell heterogeneity in mitochondria state and proliferation may be different in these two cell types. To test the ability of mitochondrial membrane potential to predict in proliferation rate in mammalian cells we used the mitochondria membrane potential-specific dye TMRE to stain fibroblasts and ESCs, and performed both RNA-seq and proliferation-rate assays on high and low TMRE sub-populations **(Figure 5A)**.

Unlike the proliferation-based sort **(Figure 1)**, sorting ESCs and fibroblasts by mitochondria-state (Figure 5) resulted in highly divergent expression profiles. ESCs with high TMRE signal have high expression of ribosome-biogenesis, proteasome, MYC-targets and mitochondrial-related genes, while in fibroblasts these gene sets are more highly expressed in the low TMRE sub-population **(Figure 5B, 5C and Table S2)**. This is consistent with the opposite behavior of mitochondria-related gene sets in proliferation-rate sorted cells from the two cell types.

The relation between mitochondria and proliferation is highly cell-type specific.

To understand the relationship between heterogeneity in proliferation and mitochondria state across cell types and species we performed principal component analysis (PCA) on RNA-seq data from all our experiments plus data from three additional studies including data from yeast sorted by both proliferation rate and mitochondria membrane potential (TMRE), and mouse CD8+ T-lymphocytes sorted by mitochondria membrane potential (TMRE) [8, 9, 21] **(Figure 6A)**. The first component is correlated with proliferation, and sorting yeast, ESCs and fibroblasts all results in sorting cells along the first PC, with fast cells from each cell type having positive values. The second component is correlated with mitochondria state; high TMRE cells from all four cell types have positive values. However, cells sorted by proliferation, while they behave similarly in PC1 (proliferation), exhibit opposite behaviors in PC2 (mitochondria state), with fast fibroblasts and yeast cells having negative values, similar to low TMRE cells, while fast ESCs have positive values, similar to high TMRE cells **(black boxes in Figure 6B)**. Thus, unlike the relationship between protein synthesis and degradation and proliferation, the relation between mitochondria and proliferation is highly cell-type specific.

These expression data make the following prediction: ESCs with high TMRE should have a shorter doubling time, while fibroblasts with high TMRE should have a longer doubling time.

311 To test this we sorted fibroblasts and ESCs by TMRE, and measured the doubling time. In
312 addition, we tested the effect of ascorbic acid (vitamin C, an antioxidant) and O_2 levels
313 (ambient 21% atmospheric vs. low 5% physiological levels). While there was no significant
314 effect of either ascorbic acid or O_2 in either cell type **(Table S3)**, the transcriptome-data
315 correctly predicted the results of the experiment, with high TMRE fibroblasts proliferating
316 more slowly, while high TMRE ESCs proliferated more rapidly **(Figure 6C)**. Thus, across yeast,
317 ESCs and fibroblasts, while mitochondria state and proliferation rate co-vary within a single
318 popopulation, the direction of this correlation is different between pluripotent ESCs and
319 other cell types.
320

321 *Single mESCs proliferating at different rates due to media fall along the axis of mESCs sorted*
322 *by proliferation rate.*

323 The above suggests that we should be able to use PCA space to predict proliferation
324 rates of single cells from scRNA-seq data. This is in contrast to current proliferation markers,
325 such as PCNA and Ki67 which are cell-cycle regulated, and whose expression will not
326 correlate with proliferation at the single-cell level. In addition, expression measurements for
327 single genes are noisy; we reasoned that the position of a single cell in PCA space should be
328 more robust, as it takes into account the expression of most genes in the cell. As a control we
329 projected expression data from mESCs grown in either serum+LIF or 2i+LIF conditions [63]
330 into the same PCA space from Figure 6. The slower-proliferating 2i+LIF grown cells are
331 perfectly separated from faster-proliferating serum+LIF grown cells by PC1, and, indeed, fall
332 exactly along the fast-slow sorted mESC expression axis **(Figure 7A)**, consistent with the
333 combination of PC1+PC2 representing cell-type specific cell-to-cell variation in
334 proliferation-correlated gene expression. We observed an inter-population relation between
335 PC1 (proliferation) and the expression of pluripotency markers, but no intra-population
336 relation **(Figure 7B)**, which is inconsistent with our results of sorting by proliferation rate
337 **(Figure 2B)**. This may be due to possiblity that technical noise in single cell sequencing
338 drowned out the heterogeneity in proliferation rate, incomparability across experiments and
339 labs, or that PCA is not sufficient to separate by both proliferation and pluripoency. Cells
340 grown in serum+LIF have higher cell-to-cell heterogeneity in gene sets associated with
341 proliferation rate compared with cells grown in 2i+LIF **(Figure 7C,D)**, reflecting the higher
342 homogeneity associated with ground state pluripotency of 2i+LIF grown cells[24].
343

344 Discussion

345 In summary, we have developed a method to sort cells by their proliferation rate and
346 have examined the whole picture of gene expression patterns related to cell-to-cell
347 heterogeneity in proliferation **(Table S4)**. We found with that genes involved in protein
348 synthesis (ribosome-biogenesis translation intitiation), and in protein degradation (the

349 proteasome and proteasome-related protein degradation) are highly expressed in fast
350 proliferating mammalian cells and yeast cells. Previous studies have reported that high
351 expression of the proteasome in fast-growing cells can degrade misfolded protein because
352 the fast protein synthesis in fast-growing cells will produce more incorrectly folded proteins
353 [47, 64, 65], which is consistent with our enrichment of proteasome-related gene sets in fast
354 proliferating cells.
355
356 In all non-cancer mammalian cells, we also found the mTORC1 signaling pathway
357 enriched in fast proliferating cells and P53-targets enriched in slow proliferating cells. Our
358 results show both upregulations of the mTORC1 signaling pathway and proteasome activity
359 in fast proliferating cells, which is consistant with several previous studies [9, 12-15].
360
361 Our analysis of fast versus slow proliferating ESCs cultured in 2i+LIF conditions indicated
362 at several levels that slow proliferating cells were of more naïve ground state pluripotent
363 character than fast proliferating cells. First, this was supported by the fact that they displayed
364 higher expression of naïve pluripotency marker genes and markers of 2C-like cells (**Figure**
365 **2B,C**). Second, we observed enrichment of E2F targets and genes involved in G1 S cell cycle
366 phase transition (**Table 1**) in our fast cycling ESC population, indicative of a shortened G1
367 phase as described normally for ESCs cultured in serum+LIF conditions [26]. This is also
368 consistant with the observation, that our ESC line proliferates much faster when cultured in
369 serum+LIF, when compared to the 2i+LIF conditions used in this study (S.F.G., unpublished).
370 Finally, although we could find differentiation genes to be expressed both in fast and slow
371 proliferating cells (**Figure 2A**), we saw a number of differentiation pathways to be enriched
372 specifically in fast dividing ESCs (**Table 1**). In summary, even when ESCs are cultured in
373 ground-state pluripotency promoting 2i+LIF conditions, they display heterogeneity in
374 proliferation rate, with the slow proliferating being of more naïve pluripotent character when
375 compared to fast dividing cells.
376
377 While we observed ESCs to behave similar to other cell types like fibroblasts or yeast
378 when it comes to gene expression signatures characteristic of fast proliferating cells related
379 to protein synthesis and turnover (**Figure 3C**), we found a very different behavior when it
380 comes to regulation of metabolism. Although the growth rate can be predicted by
381 mitochondrial membrane potential in *Saccharomyces cerevisiae* [23], where it is negatively
382 correlated with proliferation rate like in fibroblasts as we show in this study, our results show
383 mitochondrial membrane potential to be positively correlated with proliferation rate in ESCs
384 (**Figure 6**), which suggests mitochondrial membrane potential has different functions in
385 pluripotent cells when compared to differentiated cell types or yeast. This is corroborated by
386 our gene expression analysis of cells with high vs. low mitochondrial membrane potential
387 (**Figure 5B-C**), where we found pathways linked with fast proliferating cells to be enriched in

388 fibroblasts with low mitochondrial membrane potential but on the contrary enriched in ESCs
389 with high mitochondrial membrane potential. Surprisingly, primed pluripotent stem cells
390 have been described to rely more non-oxidative, glycolysis-based metabolism than naïve
391 pluripotent stem cells [66-68], which appears in contradiction with our result that our slow
392 proliferating, mitochondria activity low ESCs being more naïve-like. However, TMRE is not a
393 direct measure of ATP generation by mitochondria; yeast cells that are repsiring and
394 producing all of their ATP using their mitochondria, and yeast cells unable to respire and
395 unable to produce any ATP using their mitochondria both have high TMRE signals[9].
396 Differentiated cells in general rely more on oxidative metabolism than pluripotent cells,
397 therefore our fast proliferating ESCs could potentially reflect a more differentiated state. In
398 conclusion, we were able to show that pluripotent ESCs behave similarly to other cell types
399 in their relation between proliferation rate and aspects like protein turnover, but in the
400 opposite direction when it comes to their metabolic state. For our full understanding of the
401 pluripotent state it will be important to reveal why and how metabolism and proliferation
402 rate are regulated so differently when compared to differentiated cells.
403
404
405 ## MATERIALS AND METHODS
406
407 **Cell culture growth conditions**
408 Tail tip fibroblasts (TTFs) were isolated from a female newborn mouse from a *Mus musculus*
409 *x Mus Castaneus* cross and immortalized with SV40 large T antigen [69]. The clonal line
410 68-5-11 [70] was established and maintained in DMEM supplemented with 10% serum
411 (LifeTech), HEPES (30mM, Life Tech), Sodium Pyruvate (1mM, Life Tech), non-essential
412 amminoacids (NEAA) (Life Tech), penicillin-streptomycin (Ibian Tech), 2-mercaptoethanol
413 (0.1mM, Life Tech).
414 The mouse embryonic stem cell (ESC) line EL16.7 (40XX, Mus musculus/*M.castaneus* hybrid
415 background) [71] was maintained on gelatin coated tissue culture dishes in 2i+LIF medium.
416 This contains a 1:1 mixture of DMEM/F12 supplemented with N2 (LifeTech) and neurobasal
417 media (LifeTech) supplemented with glutamine (LifeTech), B27 (LifeTech), insulin (Sigma),
418 penicillin-streptomycin (Ibian Tech), 2-mercaptoethanol (LifeTech), LIF (Orfgenetics),
419 PD0325901 (Sigma) and CHIR9021 (Sigma).
420
421 **Proliferation and doubling time analysis**
422 ESCs and fibroblasts were plated on 10 cm plates at $5.3x10^6$ and $7.3x10^5$ concentration,
423 respectively. Cells were expanded and counted for 7 days. To monitor distinct generations of
424 proliferating cells, carboxyfluorescein succinimidyl ester (CFSE, Thermo Fisher Scientific) was
425 used to stain the cells and the dilution of the dye was detected by flow cytometry every day.

426 CFSE was dissolved in dimethyl sulfoxide at a concentration of 5 mM as stock solution and
427 CFSE was added to a 1 ml cell suspension, to a final concentration of 5uM or 10uM. After
428 addition of CFSE, cells were incubated at 37°C for 20 min. Then the cells were washed twice
429 with complete medium and maintained on ice until use in a buffer containing PBS, 2% serum
430 and 1% pen-strep. Cell viability was determined by DAPI (Biogen Cientifica) staining. Dye
431 signals were measured on an LSRII flow cytometer.
432
433 **RNA-seq**
434 To collect cells with different growth rates, cells were isolated by sorting at room
435 temperature according to the CFSE signal (median and high CFSE signal). ESCs and fibroblasts
436 were sorted into 1.5 ml Eppendorf tubes containing medium and were cultured for 3 days
437 and 5 days respectively in specific culture conditions as described earlier. For each cell line
438 three bins were sorted: the lowest 10%, the median 10% and the highest 10% CFSE. Cells
439 were sorted into prechilled 1.5-ml Eppendorf tubes containing 200 µl medium each. Cells
440 were then centrifuged at 1000 rpm for 5 min, the media removed and the resulting cell pellet
441 was used for RNA extraction. All bins were treated identically throughout the process.
442 Cellular RNA was extracted using the Maxwell RNA Purification Kit and processed for RNA
443 sequencing.
444
445 **Mitochondrial Membrane Potential Measurements.**
446 The relative mitochondrial transmembrane potential (ΔΨm) was measured using with the
447 membrane-potential-dependent fluorescent dye TMRE (Tetramethylrhodamine, Ethyl Ester,
448 Perchlorate) (Molecular Probes, Thermo Fisher Scientific) [72]. For TMRE staining fibroblasts
449 and ESCs were grown, washed in PBS, trypsinized and resuspended in PBS with 0.1% BSA and
450 TMRE added at a final concentration of 50nM, from a 10uM stock dissolved in DMSO. Cells
451 were incubated for 20min at 37C, washed with PBS and were analyzed by flow cytometry or
452 sorted.
453
454 **Cell sorting**
455 For the CFSE sort (no TMRE), cells were stained with CFSE and DAPI, and we used FACS to
456 obtain a population of viable cells the same CFSE signal. We then grew cells for 3 or 5 days,
457 and every 24 hours measured the CFSE signal using flow cytometry.
458 For the TMRE sort for proliferation rate, cells were stained with CFSE and TO-PRO-3, and we
459 used FACS to obtain a population of G1 cells with the same CFSE signal. We then grow cells
460 for 3 or 5 days, and every 24 hours measured the CFSE signal using flow cytometry.
461 In order to have a homogeneous starting population, both cell types were stained with
462 Hoechst (10 ug/ml, Life Technologies) to pick cells in G0/G1 phase. Within this population,
463 cells were selected according to the proliferation rate on the peak of CFSE signal prior
464 staining them with the dye. Then cells were sorted by TMRE into three bins: low, medium

465 and high with a BD Influx cell sorter into prechilled 1.5 ml Eppendorf tubes containing 200 µl
466 medium each. Cells were then centrifuged at 1000 rpm for 5 min, the cell pellet was washed
467 with PBS and used for RNA extraction. All bins were treated identically throughout the
468 process. Cellular RNA was extracted using the Maxwell RNA Purification Kit and processed for
469 RNA sequencing.
470 Cell viability was determined by TO-PRO-3 (Thermo Fisher Scientific) staining.
471
472 To test the effect of O_2 levels and ascorbic acid/vitamin C in both cell types, sorted cells from
473 each bin were plated into each of the four different conditions (low O_2 (5%), normal oxygen
474 growing conditions, and with or without ascorbic acid/vitamin C (25 ug/ml, Sigma-Aldrich))
475 in duplicate. After one day of recovery from the sorting, the cells were washed in PBS, were
476 trypsinized, and counted. After seeding the same initial number, the rest of the cells was
477 analyzed on a BD Fortessa analyzer. Every day a sample from each condition and replicate
478 was taken for counting, and stained with 50 nM TMRE, up to 3 days for ESCs and 5 days for
479 fibroblasts, and both TMRE and CFSE were measured by flow cytometry.
480
481 **Gene set enrichment analysis (GSEA)**
482 GSEA was performed using the GSEA software and the MSigDB (Molecular Signature
483 Database v6.2) [39, 40]. We use signal to noise (requires at least three replicates) or log2
484 ratio of classes (for experiments with less than three replicates) to calculate the rank of each
485 gene. The maximum number of genes in each gene set size was set to 500, the minimum to
486 15, and GSEA was run with 1000 permutations.
487
488 ***C.elegans* scRNAseq data analysis**
489 Preterminal cell lineage and terminal cell type scRNAseq data of *C.elegans* were downloaded
490 [60]. For each cell we calculate average log2(TPM+1) for genes in "GO preribosome" gene set
491 and for genes in "GO proteasome complex" gene set, and a t-test was used to compare the
492 mean expression of all cells in each of the two groups.
493
494 **Differential expression of pluripotency and lineage commitment-related genes in mESCs**
495 **sorted by proliferation rate (CFSE)**
496 To see the corresponding pluripotent cell state of fast and slow proliferating mESCs, we
497 calculated mean expression of naïve pluripotent markers in four fast-proliferating and four
498 slow-proliferating replicates and log2(fast/slow) was calculated to compare genes expression
499 in fast proliferating subpopulation and slow proliferating sub-population. The same method
500 was applied to lineage commitment gene markers and 2C-like state gene markers.
501
502 **GSEA of cancer cell lines**
503 RNAseq data for cancer cell lines that have corresponding doubling time were obtained from

CCLE [61]. To perform GSEA on there cancer cell data, we create a “control” sample in which a gene’s expression is the median expression all 528 samples. Then we apply GSEA to each of these cancer cell lines with the “control” sample as control.

Spearman correlation between each gene set NES and growth rate across 528 cancer cell lines was calculated to find the gene sets correlated with growth rate. Average expression level (log2(TPM+1)) of 11 proliferation marker genes (*PCNA, ZWINT, RFC3, LBR, TFDP1, SNRPB, SMC4, NUSAP1, BIRC5, UBE2C,* and *TROAP*) was calculated as meta-PCNA. To see the behavior of gene sets that belongs to the functional groups in Figure 2C in cancer cell data, we calculate spearman correlation of gene sets NES across 528 cancer cell lines.

Principal Component Analysis (PCA)

The GSEA results were first filtered to extract gene sets that at are significant (FDR<0.1) in at least one of the samples. The NES values of the selected gene sets served as the input into PCA without scaling or normalization. FactoMineR [73] was used to perform the PCA using a covariance matrix.

Projection of scRNA-seq data into PCA space and calculation of weighted Euclidean distance

To project scRNA-seq data into the PCA space, we first perform GSEA on the scRNA-seq data from publication [63]. Identical to the method we used for the GSEA of 528 cancer cell lines, we create a “control” sample, which is the average of the (median expression of serum+LIF grown cells and the median expression of 2i+LIF grown cells). GSEA for each single cell from both conditions, vs this single control, was used to get the NES for each gene set and for each single cell. We then used this NES matrix multiplied by the covariance matrix of the PCA to project the scRNA-seq data into the PCA space.

The weighted Euclidean distance was calculated by set the percent of variance that the principal component can explain as the weight of the corresponding dimension of this principal component in PCA space, Euclidean distance was calculated between each two samples after every samples coordinates were multiplied with the corresponding weight.

Coefficient of Variation (CV) of mESCs scRNA-seq data

Gene sets in MSigDB were first filtered to remove gene sets with fewer than 15 genes or more than 500 genes, leaving 13794 gene sets for analysis. The expression level of each gene set was calculated as the mean of log2(TPM+1) for genes in the gene set and the CV for each gene set was calculated seperatly for both serum+LIF and 2i+LIF grown cells.

AUTHOR CONTRIBUTIONS

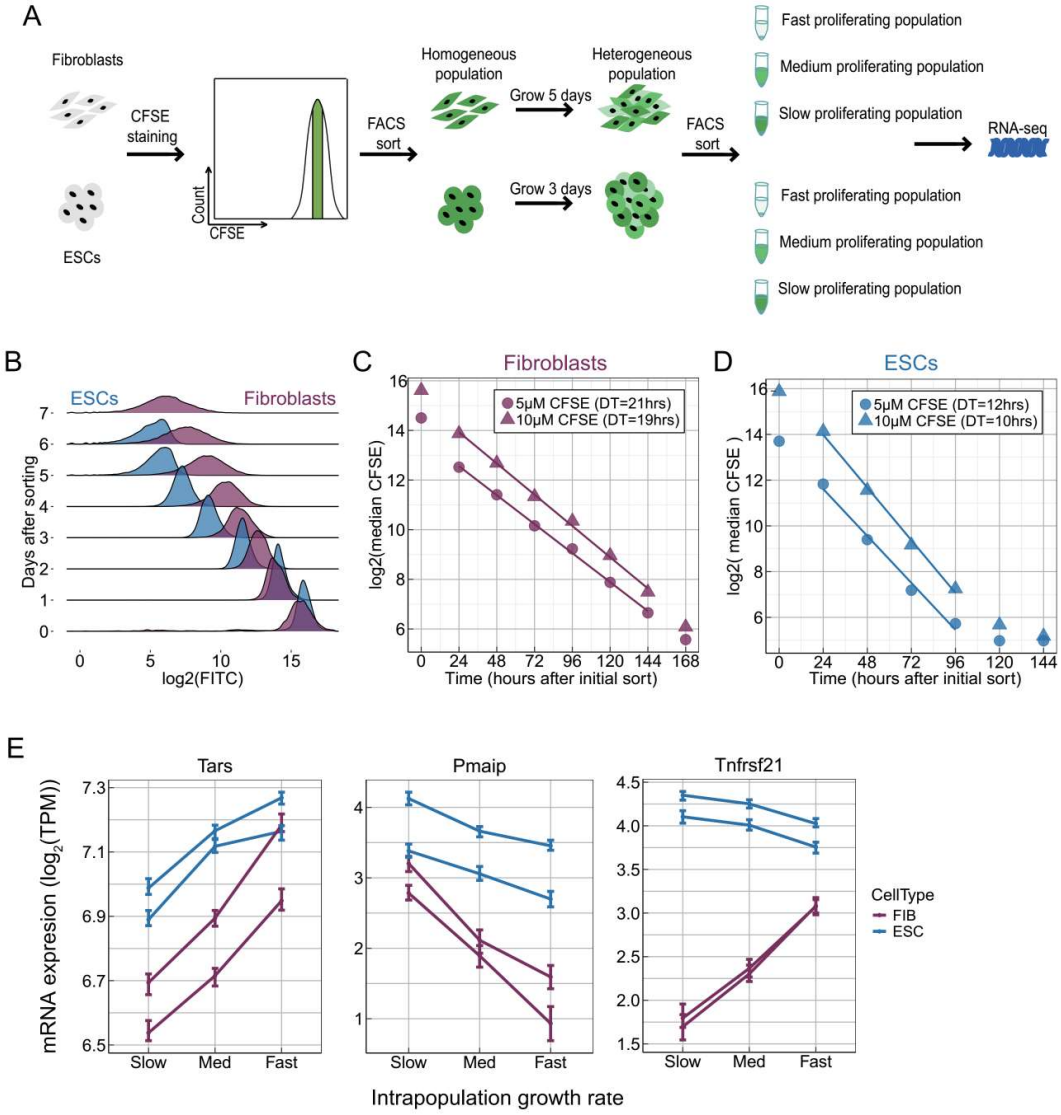
Z.J. and S.F.G. made the figures. Z.J. and M.B. analyzed the data. S.F.G. and M.B. did the experiments. L.B.C. and B.P. supervised the project. L.B.C. with help from Z.J. and B.P. wrote the manuscript. All authors read and approve of the final manuscript.

ACKNOWLEDGMENTS

We thank the CRG/UPF flow-cytometry core and CRG Genomics facilities for help with experiments, and Yang Zhao for comments on the manuscript.

This work has been funded by the Spanish Ministry of Science, Innovation and Universities (BFU2014-55275-P and BFU2017-88407-P to B.P. and BFU2015-68351-P to L.B.C.), the AXA Research Fund and the Agencia de Gestio d'Ajuts Universitaris i de Recerca (AGAUR, 2017 SGR 346 to B.P. and 2014 SGR 0974 & 2017 SGR 1054 to L.B.C.). We would like to thank the Spanish Ministry of Economy, Industry and Competitiveness (MEIC) to the EMBL partnership, to the 'Centro de Excelencia Severo Ochoa', and and the Unidad de Excelencia María de Maeztu, funded by the MINECO (MDM-2014-0370). We also acknowledge support of the CERCA Programme of the Generalitat de Catalunya.

561 FIGURE LEGENDS AND TABLE



562

563 **Figure 1. A CFSE-based method to sort mammalian cells by proliferation rate. (A)** Cells were
564 stained with CFSE and a subpopulation of cells with identical CFSE levels was collected by
565 FACS. Growth for several generations resulted in a heterogeneous cell population with a
566 broad CFSE distribution, and cells with high, medium, and low CFSE signal (slow, medium and
567 fast proliferation, respectively) were sorted by FACS for RNA-sequencing. **(B)** The change in
568 the CFSE distribution over time, for fibroblasts and ESCs. **(C, D)** The population-level doubling
569 time can be calculated by fitting a line to the median of the log2(CFSE) signal. We discard
570 data from time 0, cells immediately after the sort, because the CFSE signal decreases in the
571 initial hours, even in the absence of cell division, likely due to efflux pumps. **(E)** Examples of
572 genes whose expression positively or negatively correlated with proliferation rate. Each line
573 is one biological replicate, and the error bars are 95% confidence intervals for each
574 expression value.

575

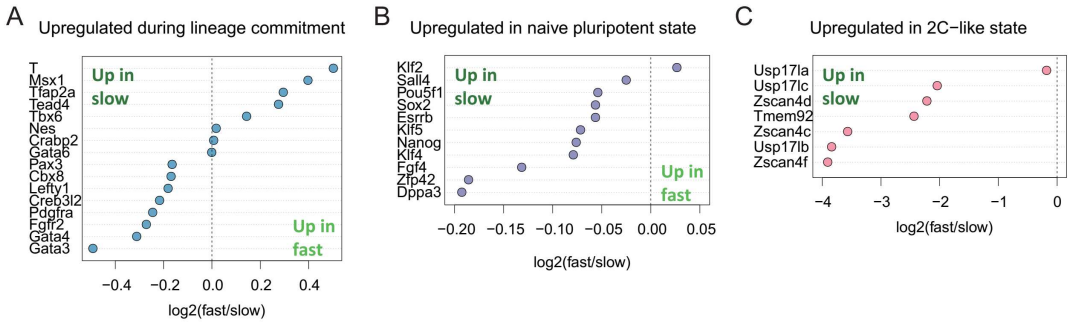


Figure 2. Slow-proliferating ESCs display a more naïve pluripotent stemness character than fast-proliferating ESCs. **(A)** Comparison of lineage commitment-related gene expression between fast and slow proliferating sub-populations. **(B)** Comparison of pluripotency-associated gene expression between fast and slow proliferating sub-populations. **(C)** Comparison of 2C-like state markers expression between fast proliferating subpopulation and slow proliferating sub-population. The dashed line in panels **(A-C)** separates genes expressed preferentially in slow- (left of dashed line) or in fast-proliferating (right of dashed line) ESCs.

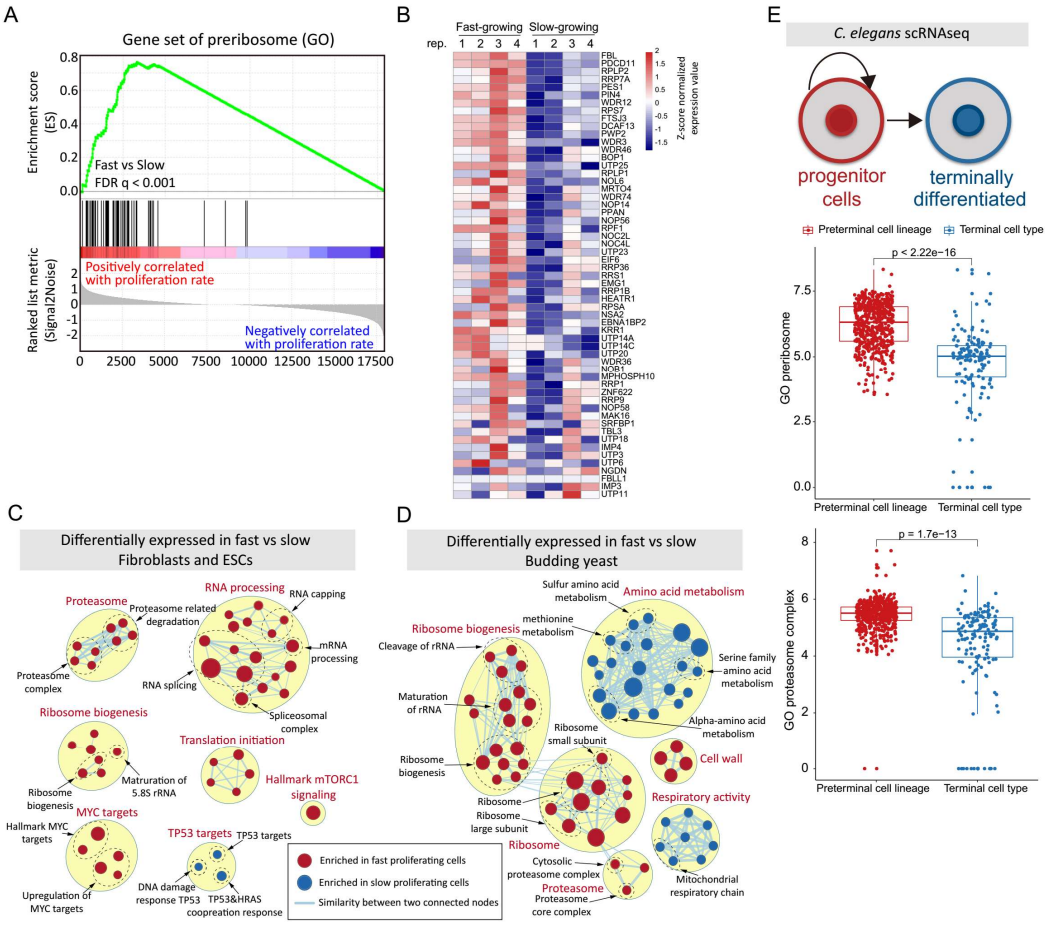


Figure 3. Functional pathways for which cell-to-cell heterogeneity in proliferation correlates with expression rate across cell types and species. (A) In Gene Set Enrichment Analysis, genes are sorted by their fast/slow expression value (left panel, bottom), and each gene is represented by a single black line (left panel middle). The enrichment score is calculated as follows: for each gene not in the gene set, the value of the green line decreases, and for each gene in the gene set, the value of the green line increases. The ES score will be near zero if the genes in a gene set are randomly distributed across the sorted list of genes, positive if most genes are to the left, and negative if most genes are to the right. **(B)** The heatmap (right panel) shows the expression (z-scored read counts) of preribosome genes in fibroblasts across four biological replicates of the CFSE sorting experiment. **(C)** Gene sets enriched (FDR<0.1) in both fibroblasts and ESCs were mapped as a network of gene sets (nodes) related by mutual overlap (edges), where the color (red or blue) indicates if the gene set is more highly expressed in fast (red) or slow (blue) proliferating cells. Node size is proportional to the total number of genes in each set and edge thickness represents the number of overlapping genes between sets. **(D)** GSEA results (FDR<0.1) of *S. cerevisiae* (van Dijk et al., 2015) sorted by cell-to-cell heterogeneity in proliferation rate. **(E)** Comparison of ribosome biogenesis and proteasome genes expression in preterminal cell lineage and terminal cell type.

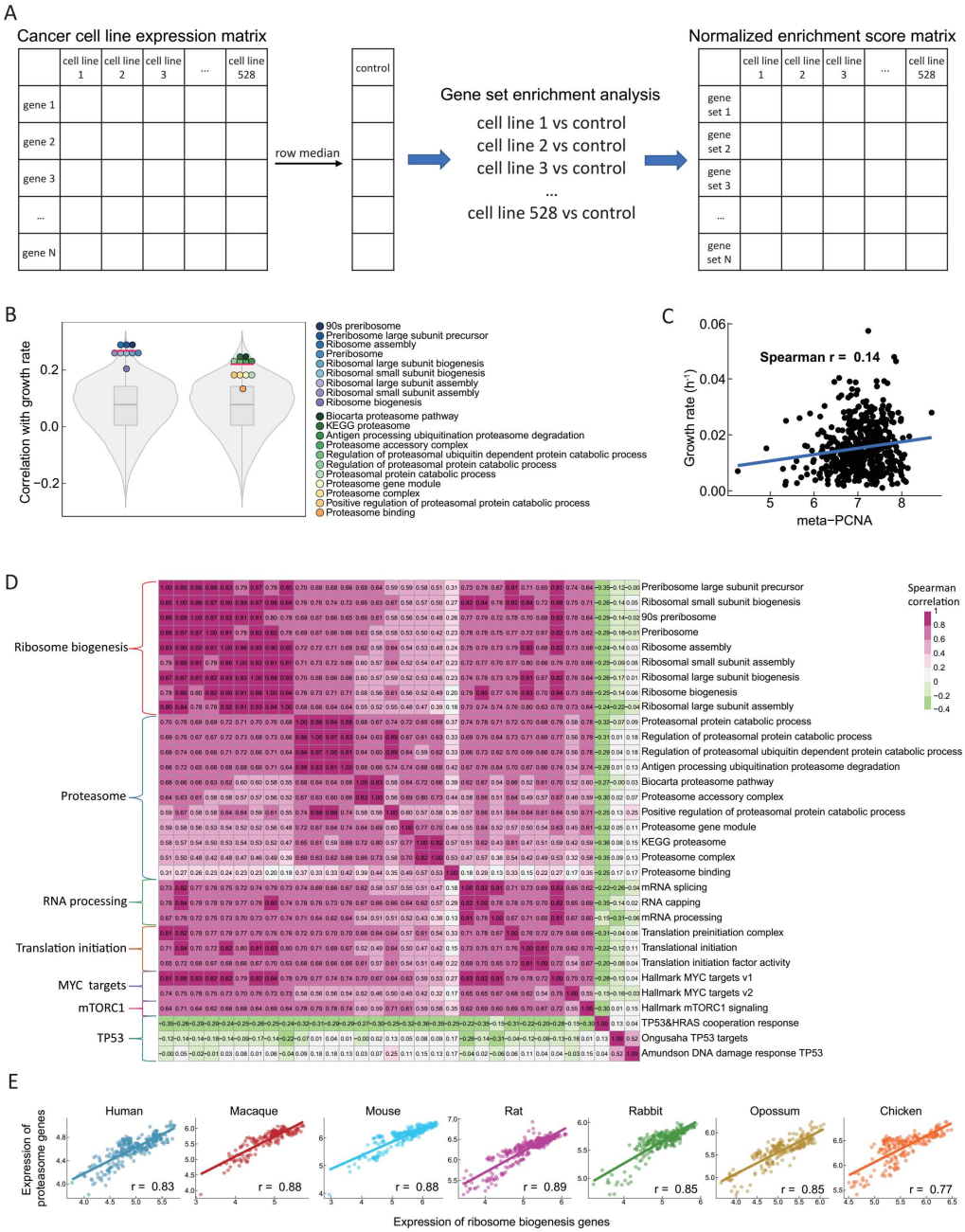
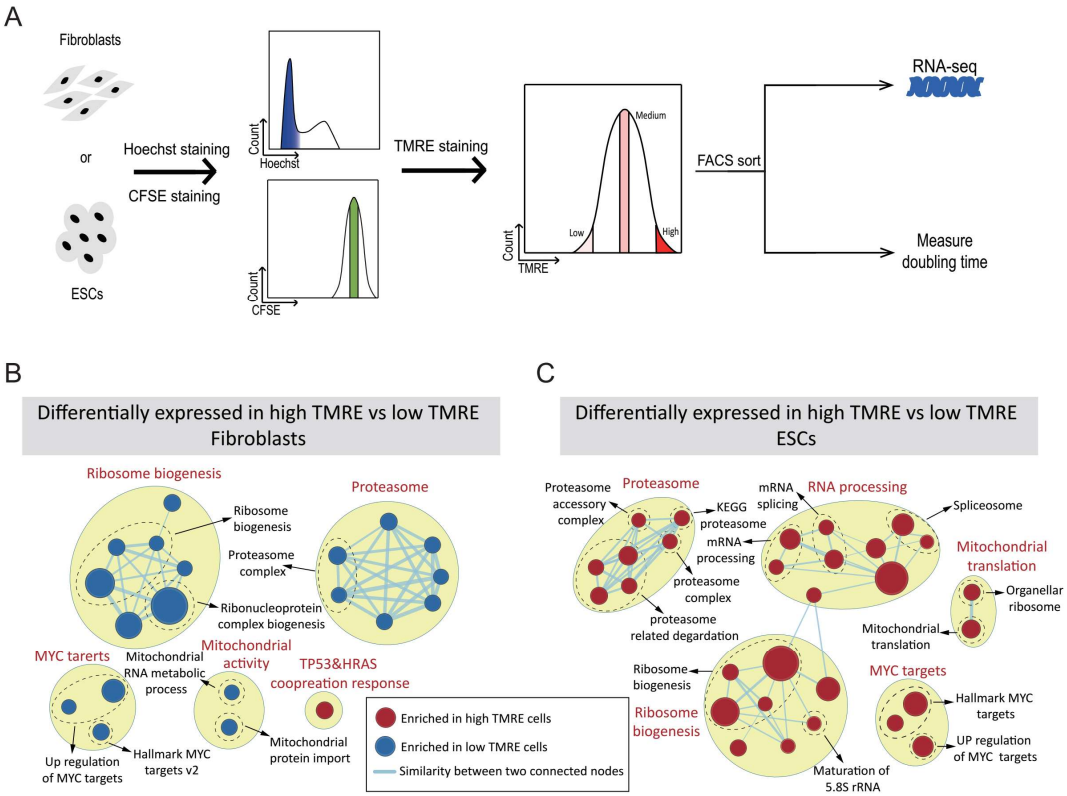


Figure 4. Expression of genes involved in ribosome-biogenesis and the proteasome are correlated with proliferation rate in cancer cell lines. (A) A cartoon of GSEA on 528 cancer cell lines. **(B)** Grey violins show the distribution of Spearman correlation coefficients of NES and growth rate for all genes sets across all 528 cancer cell lines. Points show the correlation of growth rate and the NES of gene sets involved in protein synthesis (left), or protein degradation (right). **(C)** Correlation of measured growth rate and predicted growth rate using meta-PCNA. **(D)** Spearman correlations of NES values among representative functional groups of gene sets. **(E)** Pearson correlations of mean expression (average of log2(TPM+1)) of ribosome biogenesis genes vs proteasome genes across organ developmental time course (see also Fig S1).



617

618 **Figure 5. Expression of proliferation-related gene sets in cells sorted by intra-population**
619 **heterogeneity in mitochondria membrane potential. (A)** Cells were stained with Hoechst
620 and CFSE and a homogenous population of equally sized cells cells in G1 with equal CFSE was
621 obtained by FACS. These cells were stained with TMRE sorted by TMRE, and then used for
622 RNA-seq, or allowed to proliferate to measure the doubling time of each TMRE
623 sub-population. **(B,C)** Enrichment maps of fibroblasts and ESCs sorted by TMRE.
624

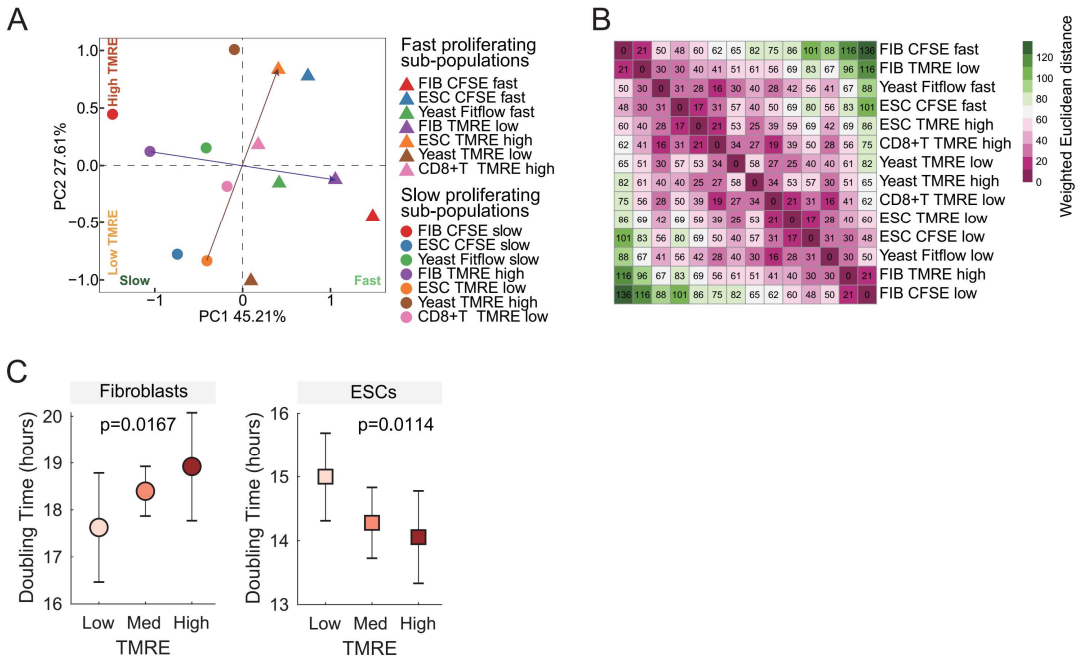


Figure 6. The relation between mitochondria and proliferation is highly cell-type specific

(A) A biplot of principal component analysis (PCA) of the Normalized Enrichment Scores (NES) values from the Gene Set Enrichment Analysis (GSEA) for all cell types sorted by intra-population heterogeneity in either proliferation or mitochondria state. Fast and slow were determined experimentally for all samples except CD8+ T cells, for which Ki67 staining was used as a proxy for proliferation. **(B)** The weighted Euclidean distance in PC space between all sorted populations. **(C)** The doubling time for fibroblasts and ESCs sorted by TMRE. Points (circle or square) show the mean doubling time for high, medium, and low TMRE cells. Error bars are the standard deviation across the eight samples for each TMRE level.

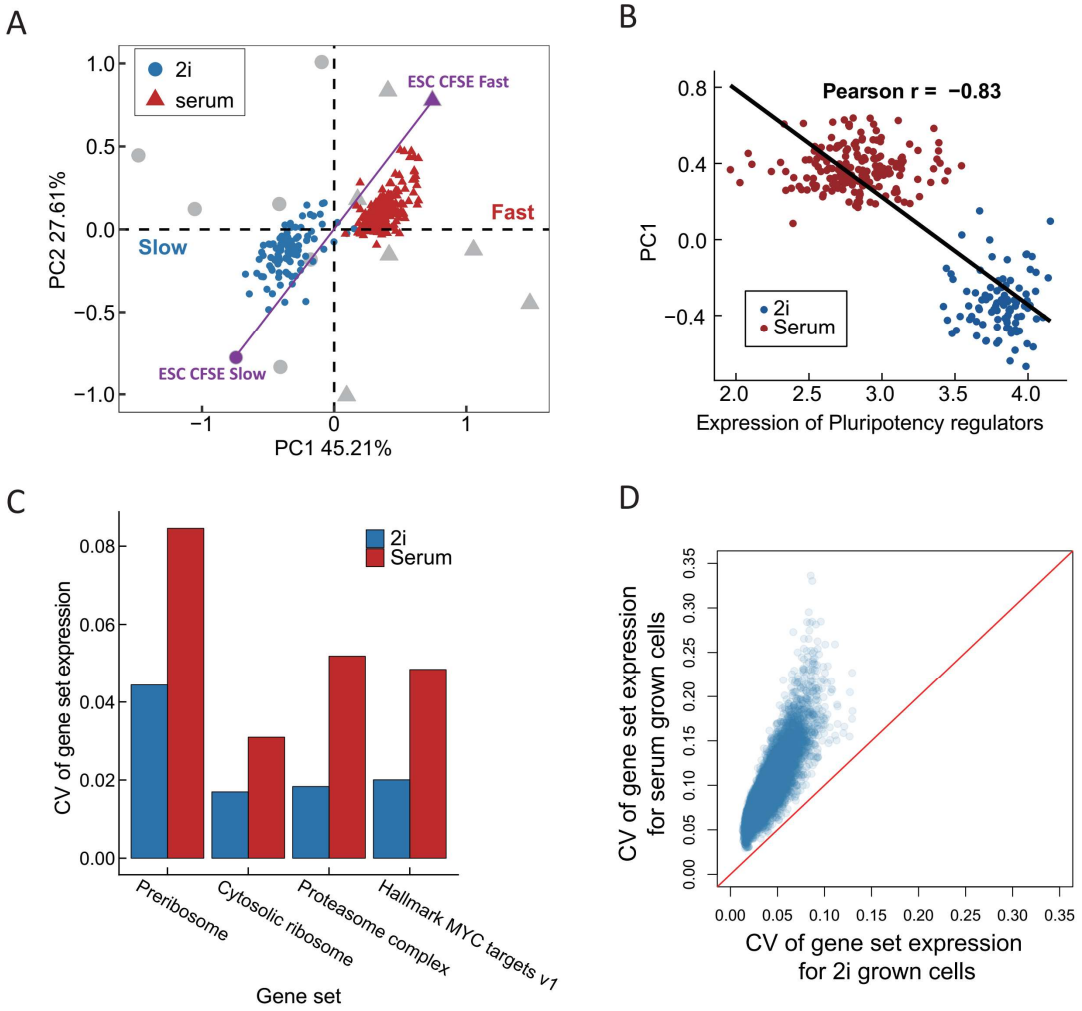


Figure 7. Prediction of proliferation in single cells using data from sorted bulk populations. **(A)** Single cell RNA-seq data of mESCs grown in either serum+LIF (fast) or 2i+LIF (slow) were projected into the PCA space from Figure 6. **(B)** Scatter plot showing the mean expression (average of log2(TPM+1)) of pluripotency markers (genes in Figure 2B) vs predicted proliferation rate (the PC1 described in Figure 6A) for each single cell. **(C)** Barplot show higher coefficient of variation (CV) in serum+LIF grown cells compare with 2i+LIF grown cells in four example gene sets. **(D)** CV across of all single cells for the mean expression (average of log2(TPM+1)) of genes in each gene set for cells grown in serum+LIF (y-axis) vs 2i+LIF (x-axis).

	Gene set name	Gene set size	NES of Fibroblasts	NES of ESCs
Mitochondria	Inner mitochondrial membrane protein complex	101	2.52	-0.39
	Mitochondrial membrane part	164	2.26	-0.44
	Mitochondrial respiratory chain complex assembly	74	2.19	-0.49
	Mitochondrial respiratory chain complex I biogenesis	54	2.12	-0.49
	Mitochondrial matrix	404	1.97	-0.49
Metabolism	Metabolism of proteins	377	2.47	-0.54
	Glycolysis gluconeogenesis	60	2.03	-1.35
	Monosaccharide biosynthetic process	52	1.94	-0.76
	Monosaccharide catabolic process	56	1.67	-0.93
	Hallmark fatty acid metabolism	157	1.52	-0.71
Differentiation	Dopaminergic neuron differentiation	28	-1.66	1.27
	Hematopoietic progenitor cell differentiation	97	-1.59	1.09
	Regulation of cardiac muscle cell differentiation	19	-1.57	0.92
	Regulation of smooth muscle cell differentiation	20	-1.45	1.66
	Glial cell differentiation	136	-1.00	1.63
Cell cycle	Cell cycle G1 S phase transition	104	-1.95	2.03
	Hallmark E2F targets	195	-2.09	2.43
	Fischer G1 S cell cycle	177	-2.03	1.90
	Cell cycle checkpoints	110	-0.84	1.82
	Cell cycle phase transition	247	-1.99	1.31

NES>0 (higher expression in fast) NES<0 (higher expression in slow) FDR q ≤ 0.001 FDR q ≤ 0.05 FDR q ≤ 0.1 FDR q > 0.1

Table 1. Gene sets whose expression exhibits opposite correlations with growth between fibroblasts and ESCs. Shown are representative gene sets whose expression is significantly correlated with proliferation in either fibroblasts or ESCs, but whose expression changes with proliferation in opposing directions.

SUPPLEMENTAL FIGURE LEGENDS

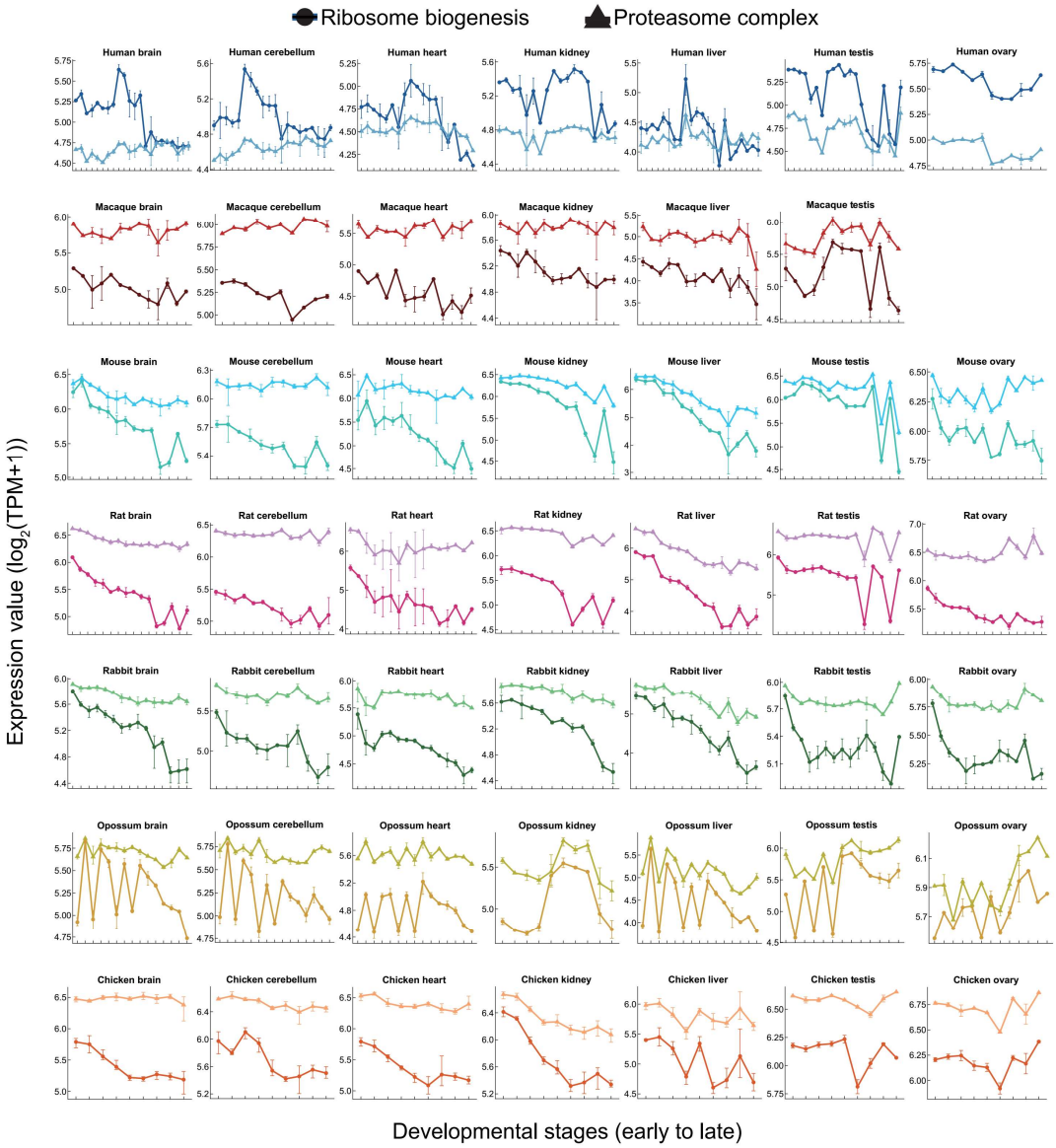


Figure S1. Correlated changes in the expression of ribosome biogenesis and proteasome related genes during organ development.

Change of average expression of log2(TPM+1) of genes in ribosome biogenesis (Go preribosome) gene set and proteasome complex (Go proteasome complex) gene set with developmental stages across different organs in seven species [16]. Points (circle and triangle) are the mean expression of replicates, error bars represent the maximum and minimum value in the replicates.

669 References

670 1. Min, M. and S.L. Spencer, *Spontaneously slow-cycling subpopulations of human cells*
671 *originate from activation of stress-response pathways.* PLOS Biology, 2019. **17**(3): p.
672 e3000178.
673 2. Wakamoto, Y., et al., *Dynamic persistence of antibiotic-stressed mycobacteria.* Science,
674 2013. **339**(6115): p. 91-5.
675 3. Fridman, O., et al., *Optimization of lag time underlies antibiotic tolerance in evolved*
676 *bacterial populations.* Nature, 2014. **513**(7518): p. 418-21.
677 4. Balaban, N.Q., et al., *A problem of persistence: still more questions than answers?* Nature
678 Reviews Microbiology, 2013. **11**(8): p. 587-591.
679 5. Gupta, P.B., et al., *Stochastic state transitions give rise to phenotypic equilibrium in*
680 *populations of cancer cells.* Cell, 2011. **146**(4): p. 633-44.
681 6. Brown, R., et al., *Poised epigenetic states and acquired drug resistance in cancer.* Nat Rev
682 Cancer, 2014. **14**(11): p. 747-53.
683 7. Marusyk, A., V. Almendro, and K. Polyak, *Intra-tumour heterogeneity: a looking glass for*
684 *cancer?* Nat Rev Cancer, 2012. **12**(5): p. 323-34.
685 8. van Dijk, D., et al., *Slow-growing cells within isogenic populations have increased RNA*
686 *polymerase error rates and DNA damage.* Nat Commun, 2015. **6**: p. 7972.
687 9. Dhar, R., et al., *Single cell functional genomics reveals the importance of mitochondria in*
688 *cell-to-cell phenotypic variation.* eLife, 2019. **8**: p. e38904.
689 10. Yaakov, G., et al., *Coupling phenotypic persistence to DNA damage increases genetic*
690 *diversity in severe stress.* Nat Ecol Evol, 2017. **1**(1): p. 16.
691 11. Paek, A.L., et al., *Cell-to-Cell Variation in p53 Dynamics Leads to Fractional Killing.* Cell,
692 2016. **165**(3): p. 631-42.
693 12. Brauer, M.J., et al., *Coordination of Growth Rate, Cell Cycle, Stress Response, and*
694 *Metabolic Activity in Yeast.* Molecular Biology of the Cell, 2007. **19**(1): p. 352-367.
695 13. Regenberg, B., et al., *Growth-rate regulated genes have profound impact on*
696 *interpretation of transcriptome profiling in Saccharomyces cerevisiae.* Genome biology,
697 2006. **7**(11): p. R107-R107.
698 14. Im, H.K., et al., *Mixed effects modeling of proliferation rates in cell-based models:*
699 *consequence for pharmacogenomics and cancer.* PLoS genetics, 2012. **8**(2): p.
700 e1002525-e1002525.
701 15. Choy, E., et al., *Genetic Analysis of Human Traits In Vitro: Drug Response and Gene*
702 *Expression in Lymphoblastoid Cell Lines.* PLOS Genetics, 2008. **4**(11): p. e1000287.
703 16. Cardoso-Moreira, M., et al., *Gene expression across mammalian organ development.*
704 Nature, 2019.
705 17. Venet, D., J.E. Dumont, and V. Detours, *Most random gene expression signatures are*
706 *significantly associated with breast cancer outcome.* PLoS computational biology, 2011.
707 **7**(10): p. e1002240-e1002240.
708 18. Levy, S.F., N. Ziv, and M.L. Siegal, *Bet hedging in yeast by heterogeneous, age-correlated*
709 *expression of a stress protectant.* PLoS Biol, 2012. **10**(5): p. e1001325.
710 19. Johnston, I.G., et al., *Mitochondrial Variability as a Source of Extrinsic Cellular Noise.* PLOS
711 Computational Biology, 2012. **8**(3): p. e1002416.
712 20. das Neves, R.P., et al., *Connecting variability in global transcription rate to mitochondrial*

713 *variability.* PLoS biology, 2010. **8**(12): p. e1000560-e1000560.
714 21. Sukumar, M., et al., *Mitochondrial Membrane Potential Identifies Cells with Enhanced*
715 *Stemness for Cellular Therapy.* Cell Metabolism, 2016. **23**(1): p. 63-76.
716 22. Mathieu, J. and H. Ruohola-Baker, *Metabolic remodeling during the loss and acquisition*
717 *of pluripotency.* Development, 2017. **144**(4): p. 541.
718 23. Nichols, J. and A. Smith, *Naive and Primed Pluripotent States.* Cell Stem Cell, 2009. **4**(6): p.
719 487-492.
720 24. Wray, J., T. Kalkan, and Austin G. Smith, *The ground state of pluripotency.* Biochemical
721 Society Transactions, 2010. **38**(4): p. 1027.
722 25. Kolodziejczyk, Aleksandra A., et al., *Single Cell RNA-Sequencing of Pluripotent States*
723 *Unlocks Modular Transcriptional Variation.* Cell Stem Cell, 2015. **17**(4): p. 471-485.
724 26. ter Huurne, M., et al., *Distinct Cell-Cycle Control in Two Different States of Mouse*
725 *Pluripotency.* Cell Stem Cell, 2017. **21**(4): p. 449-455.e4.
726 27. Nair, G., et al., *Heterogeneous lineage marker expression in naive embryonic stem cells is*
727 *mostly due to spontaneous differentiation.* Scientific Reports, 2015. **5**: p. 13339.
728 28. Abranches, E., et al., *Stochastic NANOG fluctuations allow mouse embryonic stem cells to*
729 *explore pluripotency.* Development (Cambridge, England), 2014. **141**(14): p. 2770-2779.
730 29. Smith, A., *10 Embryonic Stem Cells.* Cold Spring Harbor Monograph Archive; Volume 40
731 (2001): Stem Cell Biology, 2001: p. 205-230.
732 30. Parish, C.R., *Fluorescent dyes for lymphocyte migration and proliferation studies.*
733 Immunology & Cell Biology, 1999. **77**(6): p. 499-508.
734 31. Weston, S.A. and C.R. Parish, *New fluorescent dyes for lymphocyte migration studies:*
735 *Analysis by flow cytometry and fluorescence microscopy.* Journal of Immunological
736 Methods, 1990. **133**(1): p. 87-97.
737 32. Romano, P., et al., *Cell Line Data Base: structure and recent improvements towards*
738 *molecular authentication of human cell lines.* Nucleic acids research, 2009. **37**(Database
739 issue): p. D925-D932.
740 33. Tamm, C., S. Pijuan Galitó, and C. Annerén, *A Comparative Study of Protocols for Mouse*
741 *Embryonic Stem Cell Culturing.* PLOS ONE, 2013. **8**(12): p. e81156.
742 34. Ying, Q.-L., et al., *The ground state of embryonic stem cell self-renewal.* Nature, 2008.
743 **453**: p. 519.
744 35. Hayashi, K., et al., *Dynamic equilibrium and heterogeneity of mouse pluripotent stem*
745 *cells with distinct functional and epigenetic states.* Cell stem cell, 2008. **3**(4): p. 391-401.
746 36. Toyooka, Y., et al., *Identification and characterization of subpopulations in*
747 *undifferentiated ES cell culture.* Development, 2008. **135**(5): p. 909.
748 37. Chambers, I., et al., *Nanog safeguards pluripotency and mediates germline development.*
749 Nature, 2007. **450**: p. 1230.
750 38. Nair, G., et al., *Heterogeneous lineage marker expression in naive embryonic stem cells is*
751 *mostly due to spontaneous differentiation.* Scientific reports, 2015. **5**: p. 13339-13339.
752 39. Subramanian, A., et al., *Gene set enrichment analysis: A knowledge-based approach for*
753 *interpreting genome-wide expression profiles.* Proceedings of the National Academy of
754 Sciences, 2005. **102**(43): p. 15545.
755 40. Liberzon, A., et al., *Molecular signatures database (MSigDB) 3.0.* Bioinformatics (Oxford,
756 England), 2011. **27**(12): p. 1739-1740.

41. Athanasiadou, R., et al., *Growth Rate-Dependent Global Amplification of Gene Expression.* bioRxiv, 2016: p. 044735.
42. Kumatori, A., et al., *Abnormally high expression of proteasomes in human leukemic cells.* Proceedings of the National Academy of Sciences of the United States of America, 1990. **87**(18): p. 7071-7075.
43. Chen, L. and K. Madura, *Increased Proteasome Activity, Ubiquitin-Conjugating Enzymes, and eEF1A Translation Factor Detected in Breast Cancer Tissue.* Cancer Research, 2005. **65**(13): p. 5599.
44. Arlt, A., et al., *Increased proteasome subunit protein expression and proteasome activity in colon cancer relate to an enhanced activation of nuclear factor E2-related factor 2 (Nrf2).* Oncogene, 2009. **28**: p. 3983.
45. Cetin, B. and D.W. Cleveland, *How to survive aneuploidy.* Cell, 2010. **143**(1): p. 27-29.
46. Iadevaia, V., R. Liu, and C.G. Proud, *mTORC1 signaling controls multiple steps in ribosome biogenesis.* Semin Cell Dev Biol, 2014. **36**: p. 113-20.
47. Zhang, Y., et al., *Coordinated regulation of protein synthesis and degradation by mTORC1.* Nature, 2014. **513**(7518): p. 440-3.
48. Lempiainen, H. and D. Shore, *Growth control and ribosome biogenesis.* Curr Opin Cell Biol, 2009. **21**(6): p. 855-63.
49. Choi, J.-H., et al., *mTORC1 accelerates retinal development via the immunoproteasome.* Nature Communications, 2018. **9**(1): p. 2502.
50. Uprety, B., A. Kaja, and S.R. Bhaumik, *TOR Facilitates the Targeting of the 19S Proteasome Subcomplex To Enhance Transcription Complex Assembly at the Promoters of the Ribosomal Protein Genes.* Mol Cell Biol, 2018. **38**(14).
51. Yun, Y.S., et al., *mTORC1 Coordinates Protein Synthesis and Immunoproteasome Formation via PRAS40 to Prevent Accumulation of Protein Stress.* Mol Cell, 2016. **61**(4): p. 625-639.
52. Zhang, Y., et al., *Rapamycin extends life and health in C57BL/6 mice.* The journals of gerontology. Series A, Biological sciences and medical sciences, 2014. **69**(2): p. 119-130.
53. Zhao, J., G.A. Garcia, and A.L. Goldberg, *Control of proteasomal proteolysis by mTOR.* Nature, 2016. **529**: p. E1.
54. Gearhart, J., E.E. Pashos, and M.K. Prasad, *Pluripotency redux--advances in stem-cell research.* N Engl J Med, 2007. **357**(15): p. 1469-72.
55. Dang, C.V., *MYC, metabolism, cell growth, and tumorigenesis.* Cold Spring Harbor perspectives in medicine. **3**(8): p. a014217.
56. *Drosophila myc Regulates Cellular Growth during Development.* Cell, 1999. **98**(6): p. 779 - 790.
57. van Riggelen, J., A. Yetil, and D.W. Felsher, *MYC as a regulator of ribosome biogenesis and protein synthesis.* Nat Rev Cancer, 2010. **10**(4): p. 301-9.
58. Csibi, A., et al., *The mTORC1/S6K1 pathway regulates glutamine metabolism through the eIF4B-dependent control of c-Myc translation.* Curr Biol, 2014. **24**(19): p. 2274-80.
59. Liu, P., et al., *A functional mammalian target of rapamycin complex 1 signaling is indispensable for c-Myc-driven hepatocarcinogenesis.* Hepatology, 2017. **66**(1): p. 167-181.
60. Packer, J.S., et al., *A lineage-resolved molecular atlas of C. elegans*

embryogenesis at single-cell resolution. Science, 2019: p. eaax1971.

61. Barretina, J., et al., *The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity.* Nature, 2012. **483**(7391): p. 603-7.
62. Chen, Y., et al., *Selection for synchronized replication of genes encoding the same protein complex during tumorigenesis.* bioRxiv, 2018: p. 496059.
63. Kumar, R.M., et al., *Deconstructing transcriptional heterogeneity in pluripotent stem cells.* Nature, 2014. **516**: p. 56.
64. Conn, C.S. and S.B. Qian, *Nutrient signaling in protein homeostasis: an increase in quantity at the expense of quality.* Sci Signal, 2013. **6**(271): p. ra24.
65. Valvezan, A.J. and B.D. Manning, *Molecular logic of mTORC1 signalling as a metabolic rheostat.* Nature Metabolism, 2019. **1**(3): p. 321-333.
66. Teslaa, T. and M.A. Teitell, *Pluripotent stem cell energy metabolism: an update.* The EMBO journal, 2015. **34**(2): p. 138-153.
67. Lu, V. and M.A. Teitell, *Alpha-ketoglutarate: a "magic" metabolite in early germ cell development.* 2019. **38**(1): p. e100615.
68. Tischler, J., et al., *Metabolic regulation of pluripotency and germ cell fate through α-ketoglutarate.* 2019. **38**(1): p. e99518.
69. Brown, M., et al., *A recombinant murine retrovirus for simian virus 40 large T cDNA transforms mouse fibroblasts to anchorage-independent growth.* Journal of virology, 1986. **60**(1): p. 290-293.
70. Minajigi, A., et al., *A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation.* Science, 2015: p. aab2276.
71. Lee, J.T. and N. Lu, *Targeted Mutagenesis of Tsix Leads to Nonrandom X Inactivation.* Cell, 1999. **99**(1): p. 47-57.
72. Crowley, L.C., M.E. Christensen, and N.J. Waterhouse, *Measuring Mitochondrial Transmembrane Potential by TMRE Staining.* Cold Spring Harb Protoc, 2016. **2016**(12).
73. Lê, S., J. Josse, and F. Husson, *FactoMineR: An R Package for Multivariate Analysis.* 2008, 2008. **25**(1): p. 18 %J Journal of Statistical Software.