

1 Partial reprogramming induces a steady decline in epigenetic 2 age before loss of somatic identity

3 Nelly Olova^{1†}, Daniel J Simpson^{1†}, Riccardo Marioni², Tamir Chandra^{1*}

4 ¹MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine,
5 University of Edinburgh, Edinburgh, United Kingdom; ²Centre for Cognitive Ageing and Cognitive
6 Epidemiology, and Centre for Genomic and Experimental Medicine, University of Edinburgh,
7 Edinburgh, United Kingdom; [†]These authors contributed equally to this work

8

9 * to whom correspondence should be addressed:

10 Dr Tamir Chandra

11 MRC Human Genetics Unit

12 Institute of Genetics & Molecular Medicine

13 University of Edinburgh EH4 2XU

14 E-mail: tamir.chandra@igmm.ed.ac.uk

15 Tel.: +44 (0) 131 651 8619

16

17 Non-corresponding e-mails: nelly.olova@igmm.ed.ac.uk; s1684303@sms.ed.ac.uk;

18 riccardo.marioni@ed.ac.uk

19

20 Keywords

21 Rejuvenation, epigenetic age, ageing clock, partial reprogramming, iPSC, ageing

22

23 Summary

24 Induced pluripotent stem cells (iPSCs), with their unlimited regenerative capacity,
25 carry the promise for tissue replacement to counter age-related decline. However,
26 attempts to realise *in vivo* iPSC have invariably resulted in the formation of
27 teratomas. Partial reprogramming in prematurely aged mice has shown promising
28 results in alleviating age-related symptoms without teratoma formation. Does partial
29 reprogramming lead to rejuvenation (i.e. “younger” cells), rather than
30 dedifferentiation, which bears the risk of cancer? Here we analyse the dynamics of
31 cellular age during human iPSC reprogramming and find that partial reprogramming
32 leads to a reduction in the epigenetic age of cells. We also find that the loss of
33 somatic gene expression and epigenetic age follow different kinetics, suggesting that
34 they can be uncoupled and there could be a safe window where rejuvenation can be
35 achieved with a minimised risk of cancer.

36

37 Introduction, Results, Discussion

38 The human ageing process is accompanied by multiple degenerative diseases. Our
39 understanding of such ageing related disorders is, nevertheless, fragmented, and the
40 existence and nature of a general underlying cause are still much debated (Faragher
41 2015; Gladyshev & Gladyshev 2016). The generation of induced pluripotent stem
42 cells (iPSCs) allows the reprogramming of somatic cells back to an embryonic stem
43 cell (ESC) like state with an unlimited regenerative capacity. This has led to multiple
44 strategies for tissue replacement in degenerative diseases (Takahashi et al. 2007).
45 Clinical application of iPSCs however, is at its infancy (V. K. Singh et al., 2015;
46 Soria-Valles et al., 2015; Takahashi & Yamanaka, 2016), and the potency of iPSCs
47 bears risks, not least cancer induction. For example, *in vivo* experiments with iPSCs
48 have shown that continuous expression of Yamanaka factors (Oct4, Sox2, Klf4 and
49 c-Myc, thus OSKM) in adult mice invariably leads to cancer (Abad et al. 2013;
50 Ohnishi et al. 2014).

51 To avoid this risk, a parallel concept of epigenetic rejuvenation has been proposed:
52 the ageing process in cells can be reversed whilst avoiding dedifferentiation (Singh &
53 Zacouto 2010; Manukyan & Singh 2012). In other words, an old dysfunctional heart
54 cell could be rejuvenated without the need for it to be passed through an
55 embryonic/iPSC state. The concept of epigenetic rejuvenation requires that
56 rejuvenation and dedifferentiation each follow a distinct pathway. Nevertheless, it is
57 not well understood whether rejuvenation and dedifferentiation are invariably
58 intertwined, or instead whether it is possible to manipulate age without risking
59 dedifferentiation.

60 The epigenetic rejuvenation potential of partial reprogramming with OSKM factors
61 was previously shown by the forced expression of OSKM+*LIN28* in senescent
62 human fibroblasts, which led to recovering the high mobility of histone protein 1 β by
63 day 9, a feature characteristic for young fibroblasts (Manukyan & Singh 2014).
64 Ocampo et al. further demonstrated that partial reprogramming by transient cyclic
65 induction of OSKM ameliorates signs of ageing and extends lifespan in progeroid
66 mice, with no resulting teratoma formation (Ocampo et al. 2016). This established
67 partial reprogramming as a promising candidate intervention for age-related disease.
68 Estimating epigenetic age, which is currently the most promising proxy for biological

69 age (Jylhävä et al. 2017; Wagner 2017), was, however, not possible to measure in
70 mice at the time of the Ocampo study. This has left the nature (i.e.

71 dedifferentiation/rejuvenation) of the described cellular changes unexplored:

72 1) Does the epigenetic remodelling seen truly reflect rejuvenation (i.e. a reduction
73 in cellular/tissue age)? If so, can we observe a decrease in epigenetic age in
74 partially reprogrammed human cells?

75 2) What is the extent of rejuvenation upon reaching a partially reprogrammed
76 state (e.g. years of epigenetic age decrease)?

77 3) What are the dynamics of dedifferentiation in early reprogramming?
78

79 A major obstacle in understanding the relation between differentiation and ageing
80 has been our inability to accurately measure cellular age with a high correlation to
81 the chronological age of the organism. However, over the last five years a number of
82 age predictors have been developed, the most accurate of which utilise DNA
83 methylation (known as epigenetic clocks) (Horvath 2013; Hannum et al. 2013;
84 Weidner et al. 2014; Levine et al. 2018; Horvath et al. 2018), with the first Horvath
85 multi-tissue age-predictor being the most widely applicable and used ($r=0.96$). This
86 “Horvath clock” shows the highest correlation to chronological age, predicting the
87 age (or epigenetic age, eAge) of multiple tissues with a median error of 3.6 years
88 (Horvath 2013). eAge is distinct from and poorly correlated with other age-related
89 biomarkers, such as senescence and telomere length, which have been shown to
90 correlate independently with the process of ageing (Lowe et al. 2016; Marioni et al.
91 2016). Moreover, an acceleration of epigenetic age as measured by the “Horvath
92 clock” is associated with a higher risk of all-cause mortality (Marioni et al. 2015;
93 Christiansen et al. 2016; Perna et al. 2016), premature ageing syndromes (Down
94 and Werner) (Maierhofer et al. 2017; Horvath et al. 2015), frailty and menopause
95 (Breitling et al. 2016; Levine et al. 2016). All of these studies suggest that eAge may
96 capture a degree of biological ageing.

97

98 To understand the dynamics of eAge during reprogramming, we applied Horvath’s
99 multi-tissue age predictor over a previously published reprogramming time-course on
100 human dermal fibroblasts (HDFs) (Ohnuki et al. 2014; Horvath 2013). After OSKM
101 transfection, successfully transformed subpopulations were isolated and analysed at
102 regular time points during 49-days for gene expression and DNA methylation

103 (detailed schematic shown in Supplementary Figure 1). Epigenetic rejuvenation, i.e.
104 decrease of eAge, commenced between days 3 and 7 after OSKM transduction in
105 the partially reprogrammed TRA-1-60 (+) cells (characterised in Tanabe et al. 2013)
106 and continued steadily until day 20, when eAge was stably reset to zero (Fig. 1a). A
107 broken stick model (comprising two linear regressions joined at a break-point),
108 showed a good fit to the observed data starting from day 3, and measured a steady
109 decrease with 3.8 years per day until day 20 (SE 0.27, $P = 3.8 \times 10^{-7}$) (Fig. 1a). The
110 TRA-1-60 (+) cell populations at days 7 and 11 have been previously characterised
111 as 'partially reprogrammed' for their high expression of pluripotency markers but also
112 high reversion rates towards somatic state (Tanabe et al. 2013). Therefore, the
113 observed eAge decline at days 7 and 11 suggests that partial reprogramming can
114 indeed be considered a rejuvenation mechanism in human cells.

115 Horvath's multi-tissue age predictor is the most accurate and widely used for various
116 cell types and tissues (Wagner 2017). Nevertheless, we calculated eAge from
117 alternative DNA methylation-based age predictors: four tissue-specific clocks
118 (Hannum et al. 2013; Weidner et al. 2014; Horvath et al. 2018), one that incorporates
119 clinical measures, called PhenoAge (Levine et al. 2018), and individual CpGs
120 previously correlated with age (Garagnani et al. 2012). All clocks consistently
121 reached the point of reset to their iPSC eAge at day 20, despite the cells not being
122 fully reprogrammed before day 28 (Ohnuki et al. 2014) (Supplementary Figure 2).
123 Again, eAge showed a steady decline from day 3 to day 20 in the skin & blood and
124 Weidner 99 CpG clocks, PhenoAge declined from day 7 to day 20, while the
125 Hannum and Weidner 3 CpG clocks did not produce informative trajectories. Overall,
126 eAge values and 'years' of decrease varied between the clocks (actual chronological
127 age of HDF donors is not available for reference) (Supplementary Figure 2). The
128 highest age associated individual CpG (*ELOVL2*'s cg16867657) showed a similar
129 trajectory to the Horvath eAge decline, however, the remaining CpGs produced
130 inconsistent trajectories (Supplementary Figure 2). The observed differences are not
131 surprising, given the alternative clocks were validated for blood (Hannum et al. 2013;
132 Weidner et al. 2014), forensic applications (Horvath et al. 2018), whole organisms
133 (Levine et al. 2018) or various tissues as for the individual CpGs (Garagnani et al.
134 2012).

135 In Ocampo et al. partial reprogramming was achieved after just two days of OKSM
136 induction in mice carrying an inducible OSKM transgene (Ocampo et al. 2016).

137 However, such 'secondary' systems for direct reprogramming are known to have up
138 to 50-fold higher efficiency and accelerated kinetics in comparison to virally
139 transduced *in vitro* systems (Wernig et al. 2008). To facilitate comparison to other
140 systems and associate eAge with intermediate states in the reprogramming
141 trajectory we compared it to gene expression measured in the same samples. We
142 analysed corresponding microarray expression data for 19 well-established
143 pluripotency marker genes (Table 1 and Supplementary fig.3) as a proxy for
144 reaching a mature pluripotent state (Ginis et al. 2004; Cai et al. 2006; Mallon et al.
145 2013; Galan et al. 2013; Boyer et al. 2005). We statistically clustered the expression
146 patterns of those genes (Genolini et al. 2015), which resulted in two composite
147 trajectories. These followed previously described expression dynamics of early
148 (cluster 1) and late (cluster 2) activated pluripotency genes (Fig. 1a) (Tanabe et al.
149 2013; Chung et al. 2014; Buganim et al. 2012; Takahashi & Yamanaka 2016).
150 Pluripotency gene cluster 1 included *NANOG*, *SALL4*, *ZFP42*, *TRA-1-60*, *UTF1*,
151 *DPPA4* and *LEFTY2*, and their expression increased dramatically within the first 10
152 days and then established stable pluripotency expression levels by day 20. In
153 contrast, pluripotency gene cluster 2 (containing late expressing genes such as
154 *LIN28*, *ZIC3* and *DNMT3B*) elevated expression more slowly and reached stable
155 pluripotency levels by day 28 (Tanabe et al. 2013; Chung et al. 2014). Interestingly,
156 eAge reset to zero at the same time that the genes in cluster 1 reached their
157 pluripotent state levels, which temporally precedes full pluripotency. This also
158 coincided with a peak in expression of a number of embryonic developmental genes
159 between days 15 and 20, and might suggest that the reset marks a point where the
160 cells reach an embryonic-like state but are not yet fully pluripotent (Table 1 and
161 Supplementary Figure 4). In summary, eAge decline is observed well within the first
162 wave of pluripotency gene expression.

163

164 Therapeutic partial reprogramming will depend on rejuvenation with minimal
165 dedifferentiation, which carries the risk of malignancies. We studied the dynamics of
166 fibroblast gene down-regulation as a proxy for the loss of somatic cell identity. The
167 individual trajectories of 19 commonly used fibroblast marker genes (Kalluri &
168 Zeisberg 2006; Zhou et al. 2016; Janmaat et al. 2015; Pilling et al. 2009; Chang et
169 al. 2014; Goodpaster et al. 2008; MacFadyen et al. 2005) (Table 1 and
170 Supplementary Fig. 5) clustered into three composite expression patterns, two of

171 which (clusters 2 and 3) went into an immediate decline after OSKM induction (Fig.
172 1b). However, one fibroblast-specific cluster (cluster 1) remained stable in its
173 expression for the first 15 days. Interestingly, after day 7, fibroblast-specific gene
174 expression in clusters 2 and 3 stopped declining and plateaued until day 15,
175 coinciding with a peak in expression of senescence markers between days 11 and
176 15 (Supplementary Figure 6). Vimentin (*VIM*), for example, remained at 60% of
177 maximal expression until day 15 of reprogramming, similarly to *FAP*, *CD248* and
178 *COL1A2* in cluster 2 (Supplementary fig. 5). After day 15, fibroblast gene expression
179 declined rapidly in all three clusters, and only by day 35 had all reached ESC
180 expression levels, marking a complete loss of somatic identity (Fig. 1b). Cluster 1,
181 which contains the well described indicators of fibroblast identity *FSP1*, *COL3A1* and
182 *TGFB2/3* (Kalluri & Zeisberg 2006), showed the slowest decline, and was also the
183 last to reach ESC expression levels. In summary, we found that a number of
184 fibroblast specific genes maintained high expression levels until day 15, by which
185 time a substantial drop in eAge has been observed.

186

187 Epigenetic rejuvenation or the reversal of cellular age, is a promising concept as it
188 could avoid the oncogenic risks associated with dedifferentiation. Here, we analysed
189 a reprogramming time-course on HDFs and show that eAge declines in partially
190 reprogrammed cells before their somatic identity is entirely lost.

191 It is well established that partial reprogramming happens within an early, reversible
192 phase during the iPSC reprogramming time-course, which involves the stochastic
193 activation of pluripotency genes. It is followed by a more deterministic maturation
194 phase with predictable order of gene expression changes, where cell fate is firmly
195 bound towards pluripotency (Takahashi & Yamanaka 2016; Smith et al. 2016).
196 Indeed, it has been shown that mouse fibroblasts fail to become iPSC and revert to
197 their original somatic state if OSKM expression is discontinued during the initial
198 stochastic phase (Brambrink et al. 2008; Stadtfeld et al. 2008). Previously, Tanabe et
199 al. showed that TRA-1-60 (+) cells at reprogramming days 7 and 11 have not yet
200 reached maturation and are partially reprogrammed (Tanabe et al. 2013) but our
201 analysis already shows a decrease in their eAge according to multiple age predictors
202 (Fig. 1a and Supplementary Figure 2). We have also shown that a large proportion of
203 fibroblast marker genes maintain relatively high levels of expression until day 15
204 (Fig. 1b and Supplementary Figure 5). Nearly unchanged levels of expression on

205 day 15 were previously also shown for a large proportion of somatic genes (Tanabe
206 et al. 2013). Together with increased senescence gene expression between days 11
207 and 15 (Supplementary Figure 6), this likely contributes to the high propensity of
208 partially reprogrammed TRA-1-60 (+) cells to revert back to somatic phenotype
209 before day 15 in the time-course (Tanabe et al. 2013). Interestingly, the step-wise
210 decline of fibroblast gene expression coinciding with a peak in expression of
211 senescence genes seems to delay the loss of somatic identity but not the expression
212 of pluripotency genes. Taken together, the different dynamics between the step-wise
213 fibroblast expression and the linear decline in eAge further indicate that
214 dedifferentiation and epigenetic rejuvenation can be uncoupled.

215 Our data suggest a window of opportunity within the uncommitted reprogramming
216 phase, where a decline of eAge happens alongside partial maintenance of fibroblast
217 gene expression. A deeper understanding of the kinetics of rejuvenation will be
218 required to master therapeutic partial reprogramming, since any progress of
219 dedifferentiation, even in a small subpopulation, carries the risk of malignancies. Our
220 bulk expression analysis does not allow for a precise definition of the safe
221 rejuvenation boundaries, and further experiments on a single cell level and in *in vivo*
222 conditions are needed to determine a safe epigenetic rejuvenation window in
223 different reprogramming systems. Upon defining safe boundaries, consideration
224 should also be given to the steep decline of eAge, which resets to zero well ahead of
225 the establishment of a pluripotent state, according to a number of age predictors
226 (Supplementary Figure 2). Most likely this marks the point of reaching prenatal or
227 embryonic stage, as suggested by the peak in expression of key developmental
228 genes (Supplementary Figure 4).

229 The extent of epigenetic rejuvenation in years (human) or months (mouse), which
230 can be achieved through partial reprogramming, also needs further attention and will
231 most likely differ with the different reprogramming systems. The ‘Horvath clock’
232 shows up to 10 years of rejuvenation in Ohnuki et al.’s system by day 7 and another
233 10+ years by day 11. However, the intrinsic median estimation error of 3.6 years in
234 this age predictor, the varying eAge rejuvenation values between the different age
235 predictors, and the intra-replicate biological variation seen from the large error bars,
236 highlight the need for more experiments and repetitions before this is established
237 with a higher certainty.

238 Despite the obvious differences in reprogramming kinetics, our results also suggest
239 that the improvements observed by Ocampo et al. in their OSKM-inducible
240 secondary reprogramming system, might be due to epigenetic rejuvenation. It
241 remains to be shown how stable in time the rejuvenated phenotype is in either of the
242 systems. Further analysis is also needed regarding the effect of partial
243 reprogramming on adult stem cells or premalignant cells, which have already shown
244 a higher propensity of transforming to malignancy (Abad et al. 2013; Ohnishi et al.
245 2014). It is possible that a premalignant phenotype could be attenuated or amplified
246 by partial reprogramming. In summary, our findings reveal exciting possibilities but
247 also open a number of questions and highlight areas that need further attention.

248

249 Acknowledgements

250 We thank Chris Ponting, Steve Horvath and Keisuke Kaji for their helpful advice and
251 comments on the manuscript.

252

253 Conflict of interest

254 The authors of this paper have no conflicts of interest to declare.

255

256 Supplemental Experimental procedures

257 *Overview of the Ohnuki et al. experimental setup and datasets*

258 450K DNA methylation array and gene expression microarray data of full HDF
259 reprogramming time-course was obtained from GSE54848. A schematic of
260 experimental setup and time points is provided in Supplementary Figure 1. Briefly,
261 HDF cells were transfected with EGFP-labelled OSKM on day 0 and cultured in
262 virus-containing medium for 24 hours, then replaced by 10% FBS-containing
263 medium for 8 days before replacing with human ESC medium. EGFP (+) cells,
264 representing the population of successfully transfected cells, which permanently
265 express the OSKM factors, were sorted by flow cytometry on day 3. Intermediate
266 reprogrammed cells positive for the pluripotency marker TRA-1-60 were sorted by
267 magnetic activated cell sorting on days 7, 11, 15, 20 and 28 post-transfection. Day
268 28-sorted TRA-1-60 (+) cells were further expanded and samples collected three
269 more times on each seventh day, i.e. on days 35, 42 and 49. Thus, sorted and
270 collected cells at each time point were subjected to both gene expression and DNA

271 methylation array analysis. Microarray gene expression (data available as LOG2
272 transformed) was performed for three to four replicates per data point, whilst DNA
273 methylation data was performed for two to three replicates per time point.

274

275 *Predicting eAge*

276 The pre-processed 450K DNA methylation array matrix of average methylation per
277 CpG site of the full HDF reprogramming time-course was obtained from GSE54848
278 (downloaded using getGEO function from GEOquery package) and uploaded to the
279 online DNA methylation age calculator to assess eAge:
280 <https://labs.genetics.ucla.edu/horvath/dnamage/> (Horvath 2013). Data processing
281 including Horvath's normalisation was performed according to tutorial guidelines.
282 Missing CpG values were imputed by Horvath's online DNAm age calculator. During
283 QC, around 1600 CpGs were lost, therefore methylation data for each time point
284 contained 26,987 CpG sites out of the suggested 28,587 CpGs, a fact unlikely to
285 have any significant impact on the normalisation or age prediction. PhenoAge, skin &
286 blood, Hannum, Weidner 99 and 3 CpG age predictors were applied to average
287 methylation values. Missing CpG values were imputed as zero before applying these
288 age predictors.
289 All ages presented in the manuscript are calculated eAges, no actual ages of HDF
290 donors were available.

291

292 *Methylation Age Trajectories*

293 For the Horvath multi-tissue age predictor, a 'broken stick' model with two linear
294 sections was constructed to chart overall change in DNA methylation age over time
295 between the three HDF cell lines. A linear mixed model was then specified with a
296 random intercept term for each replicate. A variable break point was set between the
297 minimum and maximum day, plus and minus a small constant (3 days), respectively.
298 The predicted values from the regression models were plotted against the
299 measurement day. For the all other age predictor plots (Supplementary Figure 2),
300 mean eAge was calculated for all samples at each time point (2-3 samples
301 depending on the time point) and plotted against time during the time-course.
302 Standard deviation for eAge was also calculated and plotted as error bars at each
303 time point.

304

305 *Gene clusters and trajectories*

306 For each gene in a category (e.g. pluripotent gene list), a loess curve with a span of
307 0.5 was fitted with the predicted values extracted at each time point. The predicted
308 values were then normalised within each gene to a value of 1 at the first time point
309 and a value of 0 and the last time point (and vice versa for the pluripotent genes). K-
310 means clustering for longitudinal data was applied to determine the optimal number
311 of trajectories within each gene category.

312 All analyses were performed in R, using the kml (Genolini et al. 2015), lme4 (Bates
313 et al. 2014), and lmerTest (Kuznetsova et al. 2016) packages.

314 References

315

- 316 Abad M, Mosteiro L, Pantoja C, Cañamero M, Rayon T, Ors I, Graña O, Megías D,
317 Domínguez O, Martínez D, Manzanares M, Ortega S & Serrano M (2013)
318 Reprogramming in vivo produces teratomas and iPS cells with totipotency
319 features. *Nature* 502, 340–345.
- 320 Bates D, Mächler M, Bolker B & Walker S (2014) Fitting Linear Mixed-Effects Models
321 using lme4. 67.
- 322 Boyer L a L a., Lee TITI, Cole MFMF, Johnstone SESE, Stuart S, Zucker JPJP,
323 Guenther MGMG, Kumar RMRM, Murray HLHL, Jenner RGRG, Gifford DK,
324 Melton D a, Jaenisch R, Young R a, Levine SS & Others (2005) Core
325 Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. *Young*
326 122, 947–956.
- 327 Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, Suh H & Jaenisch
328 R (2008) Sequential Expression of Pluripotency Markers during Direct
329 Reprogramming of Mouse Somatic Cells. *Cell Stem Cell* 2, 151–159.
- 330 Breitling LP, Saum K-U, Perna L, Schöttker B, Holleczeck B & Brenner H (2016)
331 Frailty is associated with the epigenetic clock but not with telomere length in a
332 German cohort. *Clin. Epigenetics* 8, 21.
- 333 Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, Klemm SL,
334 Van Oudenaarden A & Jaenisch R (2012) Single-cell expression analyses
335 during cellular reprogramming reveal an early stochastic and a late hierarchic
336 phase. *Cell* 150, 1209–1222.
- 337 Cai J, Chen J, Liu Y, Miura T, Luo Y, Loring JF, Freed WJ, Rao MS & Zeng X (2006)
338 Assessing self-renewal and differentiation in human embryonic stem cell lines.
339 *Stem Cells* 24, 516–30.
- 340 Chang Y, Li H & Guo Z (2014) Mesenchymal stem cell-like properties in fibroblasts.
341 *Cell. Physiol. Biochem.* 34, 703–714.
- 342 Christiansen L, Lenart A, Tan Q, Vaupel JW, Aviv A, McGue M & Christensen K

- 343 (2016) DNA methylation age is associated with mortality in a longitudinal Danish
344 twin study. *Aging Cell* 15, 149–154.
- 345 Chung KM, Kolling FW, Gajdosik MD, Burger S, Russell AC & Nelson CE (2014)
346 Single cell analysis reveals the stochastic phase of reprogramming to
347 pluripotency is an ordered probabilistic process. *PLoS One* 9.
- 348 Faragher RGA (2015) Should we treat aging as a disease? The consequences and
349 dangers of miscategorisation. *Front. Genet.* 6, 1–7.
- 350 Galan A, Diaz-Gimeno P, Poo ME, Valbuena D, Sanchez E, Ruiz V, Dopazo J,
351 Montaner D, Conesa A & Simon C (2013) Defining the Genomic Signature of
352 Totipotency and Pluripotency during Early Human Development. *PLoS One* 8,
353 20–23.
- 354 Garagnani P, Bacalini MG, Pirazzini C, Gori D, Giuliani C, Mari D, Di AM, Gentilini D,
355 Vitale G, Rezzi S, Castellani G, Capri M, Salvioli S & Franceschi C (2012)
356 Methylation of ELOVL2 gene as a new epigenetic marker of age *Aging Cell*.
357 *Aging Cell* 11, 1132–1134.
- 358 Genolini C, Alacoque X & Marianne Sentenac CA (2015) kml and kml3d: R
359 Packages to Cluster Longitudinal Data. *J. Stat. Softw.* 65, 1–34. Available at:
360 <http://www.jstatsoft.org/v65/i04/>.
- 361 Ginis I, Luo Y, Miura T, Thies S, Brandenberger R, Gerecht-Nir S, Amit M, Hoke A,
362 Carpenter MK, Itskovitz-Eldor J & Rao MS (2004) Differences between human
363 and mouse embryonic stem cells. *Dev. Biol.* 269, 360–380.
- 364 Gladyshev T V. & Gladyshev VN (2016) A Disease or Not a Disease? Aging As a
365 Pathology. *Trends Mol. Med.* 22, 995–996.
- 366 Goodpaster T, Legesse-Miller A, Hameed MR, Aisner SC, Randolph-Habecker J &
367 Coller HA (2008) An Immunohistochemical Method for Identifying Fibroblasts in
368 Formalin-fixed, Paraffin-embedded Tissue. *J. Histochem. Cytochem.* 56, 347–
369 358.
- 370 Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sada S, Klotzle B, Bibikova M,
371 Fan J, Gao Y, Deconde R, Chen M, Rajapakse I & Friend S (2013) Genome-
372 wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates.
373 *Mol. Cell* 49, 359–367.
- 374 Horvath S (2013) DNA methylation age of human tissues and cell types. *Genome*
375 *Biol.* 14, R115.
- 376 Horvath S, Garagnani P, Bacalini MG, Pirazzini C, Salvioli S, Gentilini D, Di Blasio
377 AM, Giuliani C, Tung S, Vinters H V. & Franceschi C (2015) Accelerated
378 epigenetic aging in Down syndrome. *Aging Cell* 14, 491–495.
- 379 Horvath S, Oshima J, Martin GM, Lu AT, Quach A, Cohen H, Felton S, Matsuyama
380 M, Lowe D, Kabacik S, Wilson JG, Reiner AP, Maierhofer A, Flunkert J, Aviv A,
381 Hou L, Baccarelli AA, Li Y, Stewart JD, Whitsel EA, Ferrucci L, Matsuyama S &
382 Raj K (2018) Epigenetic clock for skin and blood cells applied to Hutchinson
383 Gilford Progeria Syndrome and ex vivo studies. *Aging (Albany, NY)*. 10, 1758–
384 1775.
- 385 Janmaat CJ, De Rooij KE, Locher H, De Groot SC, De Groot JCMJ, Frijns JHM &

- 386 Huisman MA (2015) Human dermal fibroblasts demonstrate positive
387 immunostaining for neuron- and glia-specific proteins. *PLoS One* 10, 1–14.
- 388 Jylhävä J, Pedersen NL & Hägg S (2017) Biological Age Predictors. *EBioMedicine*
389 21, 29–36.
- 390 Kalluri R & Zeisberg M (2006) Fibroblasts in cancer. *Nat. Rev. Cancer* 6, 392–401.
- 391 Kuznetsova A, Brockhoff PB & Bojesen Christensen RH (2016) lmerTest: Tests in
392 Linear Mixed Effects Models. R package version 2.0-33. Available at:
393 <https://cran.r-project.org/web/packages/lmerTest/index.html>.
- 394 Levine ME, Lu AT, Chen BH, Hernandez DG, Singleton AB, Ferrucci L, Bandinelli S,
395 Salfati E, Manson JE, Quach A, Kusters CDJ, Kuh D, Wong A, Teschendorff
396 AE, Widschwendter M, Ritz BR, Absher D, Assimes TL & Horvath S (2016)
397 Menopause accelerates biological aging. *Proc. Natl. Acad. Sci.* 113, 9327–9332.
- 398 Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Hou L, Baccarelli AA, Stewart
399 JD, Li Y, Whitsel EA, Wilson G, Reiner AP, Aviv A, Lohman K, Liu Y & Ferrucci
400 L (2018) An epigenetic biomarker of aging for lifespan and healthspan. 10, 573–
401 591.
- 402 Lowe D, Horvath S & Raj K (2016) Epigenetic clock analyses of cellular senescence
403 and ageing. *Oncotarget* 7, 8524–31.
- 404 MacFadyen JR, Haworth O, Roberston D, Hardie D, Webster MT, Morris HR, Panico
405 M, Sutton-Smith M, Dell A, Van Der Geer P, Wienke D, Buckley CD & Isacke
406 CM (2005) Endosialin (TEM1, CD248) is a marker of stromal fibroblasts and is
407 not selectively expressed on tumour endothelium. *FEBS Lett.* 579, 2569–2575.
- 408 Maierhofer A, Flunkert J, Oshima J, Martin GM, Haaf T & Horvath S (2017)
409 Accelerated epigenetic aging in Werner syndrome. *Aging (Albany. NY)*. 9,
410 1143–1152.
- 411 Mallon BS, Chenoweth JG, Johnson KR, Hamilton RS, Tesar PJ, Yavatkar AS,
412 Tyson LJ, Park K, Chen KG, Fann YC & McKay RDG (2013) StemCellDB: The
413 Human Pluripotent Stem Cell Database at the National Institutes of Health.
414 *Stem Cell Res.* 10, 57–66.
- 415 Manukyan M & Singh PB (2012) Epigenetic rejuvenation. *Genes to Cells* 17, 337–
416 343.
- 417 Manukyan M & Singh PB (2014) Epigenome rejuvenation: HP1 β mobility as a
418 measure of pluripotent and senescent chromatin ground states. *Sci. Rep.* 4, 1–
419 8.
- 420 Marioni RE, Harris SE, Shah S, McRae AF, von Zglinicki T, Martin-Ruiz C, Wray NR,
421 Visscher PM & Deary IJ (2016) The epigenetic clock and telomere length are
422 independently associated with chronological age and mortality. *Int. J. Epidemiol.*
423 45, 424–432.
- 424 Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, Gibson J, Henders
425 AK, Redmond P, Cox SR, Pattie A, Corley J, Murphy L, Martin NG, Montgomery
426 GW, Feinberg AP, Fallin M, Multhaup ML, Jaffe AE, Joehanes R, Schwartz J,
427 Just AC, Lunetta KL, Murabito JM, Starr JM, Horvath S, Baccarelli AA, Levy D,
428 Visscher PM, Wray NR & Deary IJ (2015) DNA methylation age of blood

- 429 predicts all-cause mortality in later life. *Genome Biol.* 16, 25.
- 430 Ocampo A, Reddy P, Martinez-Redondo P, Platero-Luengo A, Hatanaka F, Hishida
431 T, Li M, Lam D, Kurita M, Beyret E, Araoka T, Vazquez-Ferrer E, Donoso D,
432 Roman JL, Xu J, Rodriguez Esteban C, Nuñez G, Nuñez Delicado E, Campistol
433 JM, Guillen I, Guillen P & Izpisua Belmonte JC (2016) In Vivo Amelioration of
434 Age-Associated Hallmarks by Partial Reprogramming. *Cell* 167, 1719–
435 1733.e12.
- 436 Ohnishi K, Semi K, Yamamoto T, Shimizu M, Tanaka A, Mitsunaga K, Okita K,
437 Osafune K, Arioka Y, Maeda T, Soejima H, Moriwaki H, Yamanaka S, Woltjen K
438 & Yamada Y (2014) Premature termination of reprogramming in vivo leads to
439 cancer development through altered epigenetic regulation. *Cell* 156, 663–677.
- 440 Ohnuki M, Tanabe K, Sutou K, Teramoto I, Sawamura Y, Narita M, Nakamura M,
441 Tokunaga Y, Nakamura M, Watanabe A, Yamanaka S & Takahashi K (2014)
442 Dynamic regulation of human endogenous retroviruses mediates factor-induced
443 reprogramming and differentiation potential. *Proc. Natl. Acad. Sci.* 111, 12426–
444 12431.
- 445 Perna L, Zhang Y, Mons U, Holleczek B, Saum K-U & Brenner H (2016) Epigenetic
446 age acceleration predicts cancer, cardiovascular, and all-cause mortality in a
447 German case cohort. *Clin. Epigenetics* 8, 64.
- 448 Pilling D, Fan T, Huang D, Kaul B & Gomer RH (2009) Identification of markers that
449 distinguish monocyte-derived fibrocytes from monocytes, macrophages, and
450 fibroblasts. *PLoS One* 4, 31–33.
- 451 Singh PB & Zacouto F (2010) Nuclear reprogramming and epigenetic rejuvenation.
452 *J. Biosci.* 35, 315–319.
- 453 Singh VK, Kalsan M, Kumar N, Saini A & Chandra R (2015) Induced pluripotent stem
454 cells: applications in regenerative medicine, disease modeling, and drug
455 discovery. *Front. Cell Dev. Biol.* 3, 1–18.
- 456 Smith ZD, Sindhu C & Meissner A (2016) Molecular features of cellular
457 reprogramming and development. *Nat. Rev. Mol. Cell Biol.* 17, 139–154.
- 458 Soria-Valles C, Osorio FG, Gutiérrez-Fernández A, De Los Angeles A, Bueno C,
459 Menéndez P, Martín-Subero JI, Daley GQ, Freije JMP & López-Otín C (2015)
460 NF-κB activation impairs somatic cell reprogramming in ageing. *Nat. Cell Biol.*
461 17, 1004–13.
- 462 Stadtfeld M, Maherali N, Breault DT & Hochedlinger K (2008) Defining Molecular
463 Cornerstones during Fibroblast to iPS Cell Reprogramming in Mouse. *Cell Stem*
464 *Cell* 2, 230–240.
- 465 Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K & Yamanaka S
466 (2007) Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by
467 Defined Factors. *Cell* 131, 861–872.
- 468 Takahashi K & Yamanaka S (2016) A decade of transcription factor-mediated
469 reprogramming to pluripotency. *Nat. Rev. Mol. Cell Biol.* 17, 183–193.
- 470 Tanabe K, Nakamura M, Narita M, Takahashi K & Yamanaka S (2013) Maturation,
471 not initiation, is the major roadblock during reprogramming toward pluripotency

- 472 from human fibroblasts. *Proc. Natl. Acad. Sci.* 110, 12172–12179.
- 473 Wagner W (2017) Epigenetic aging clocks in mice and men. *Genome Biol.* 18, 107.
- 474 Weidner C, Lin Q, Koch C, Eisele L, Beier F, Ziegler P, Bauerschlag D, Jöckel K-H,
475 Erbel R, Mühleisen T, Zenke M, Brümmendorf T & Wagner W (2014) Aging of
476 blood can be tracked by DNA methylation changes at just three CpG sites.
477 *Genome Biol.* 15, R24.
- 478 Wernig M, Lengner CJ, Hanna J, Lodato MA, Steine E, Foreman R, Staerk J,
479 Markoulaki S & Jaenisch R (2008) A drug-inducible transgenic system for direct
480 reprogramming of multiple somatic cell types. *Nat. Biotechnol.* 26, 916–924.
- 481 Zhou L, Yang K, Randall Wickett R & Zhang Y (2016) Dermal fibroblasts induce cell
482 cycle arrest and block epithelial–mesenchymal transition to inhibit the early
483 stage melanoma development. *Cancer Med.* 5, 1566–1579.

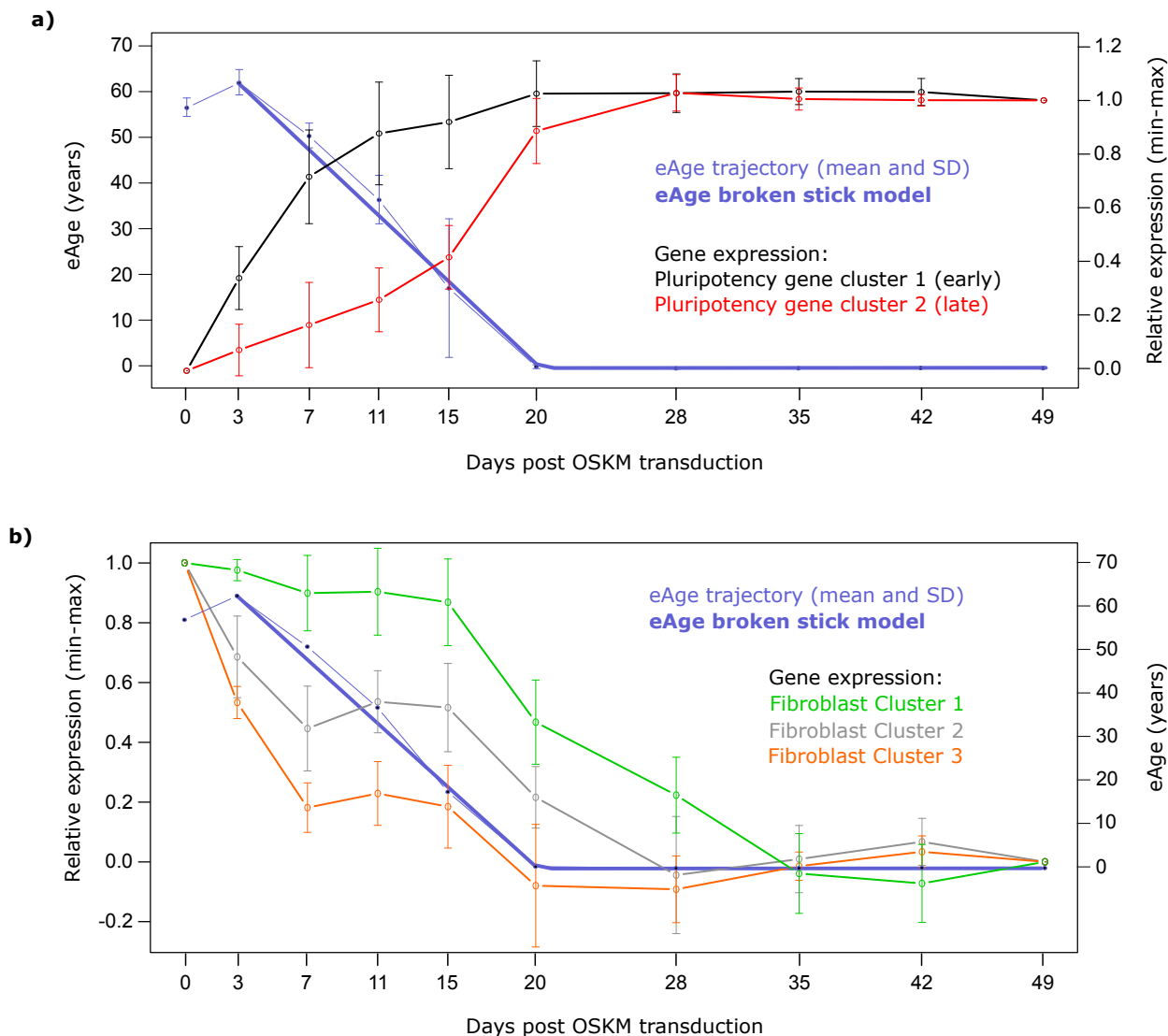


Figure 1. Dynamics of eAge and gene expression in a 49-day HDF reprogramming time-course. (a) Left Y axis: eAge trajectory of Horvath's multi-tissue age predictor calculated from DNA methylation arrays from the following cell populations: day 0 (HDFs), day 3 (OSKM-expressing EGFP (+) HDFs), day 7, 11, 15, 20 and 28 (human pluripotency marker TRA-1-60 (+) cells at intermediate stages of reprogramming), and fully reprogrammed iPSCs from days 35, 42 and 49. Data was fit with a broken stick model composed of two linear sections. Error bars represent SD. Measured rate (years per day) of eAge decrease [day 3 - day 20] = -3.8, SE 0.27, $P = 3.8 \times 10^{-7}$. **Right Y axis:** Composite gene expression trajectories of key pluripotency markers, statistically clustered as per Genolini et al. 2016. Microarray expression data was obtained for the same time points and cell populations as for eAge. Relative expression values were LOG2 transformed and presented as arbitrary units starting from '0' for 'day 0' to '1' for 'day 49'. Error bars represent SD. **(b) Left Y axis:** Composite gene expression trajectories of key fibroblast markers statistically clustered as described for the pluripotency markers in (a). Relative expression values are presented as arbitrary units starting from '1' for 'day 0' to '0' for 'day 49'. **Right Y axis:** eAge as in (a) left Y axis, without SD.

Table 1. List of pluripotency and fibroblast marker genes used in gene expression clusters. Key pluripotent marker genes were selected from Ginis et al. 2004; Cai et al. 2006; Mallon et al. 2013; Galan et al. 2013; Boyer et al. 2005. Fibroblast marker genes were selected from Kalluri & Zeisberg 2006; Zhou et al. 2016; Janmaat et al. 2015; Pilling et al. 2009; Chang et al. 2014; Goodpaster et al. 2008; MacFadyen et al. 2005.

Marker	Gene	Protein name	Accession	Cluster
Pluripotency	<i>NANOG</i>	Nanog homeobox	A_23_P204640	1 (early)
Pluripotency	<i>REX1 (ZFP42)</i>	Zinc Finger Protein 42	A_23_P395582	1 (early)
Pluripotency	<i>TRA-1-60/81 (PODXL)</i>	Podocalyxin	A_23_P215060	1 (early)
Pluripotency	<i>UTF1</i>	Undifferentiated embryonic cell transcription factor 1	A_33_P3294217	1 (early)
Pluripotency	<i>DPPA4</i>	Developmental pluripotency associated 4	A_23_P380526	1 (early)
Pluripotency	<i>TDGF1 (CRIPTO)</i>	Teratocarcinoma-derived growth factor 1	A_23_P366376	1 (early)
Pluripotency	<i>SALL4</i>	Spalt like transcription factor 4	A_23_P109072	1 (early)
Pluripotency	<i>LEFTY1</i>	Left-right determination factor 1	A_23_P160336	1 (early)
Pluripotency	<i>LEFTY2</i>	Left-right determination factor 2	A_23_P137573	1 (early)
Pluripotency	<i>DNMT3A</i>	DNA methyl-transferase 3A	A_23_P154500	1 (early)
Pluripotency	<i>TFCP2L1</i>	Transcription factor CP2 like 1	A_23_P5301	1 (early)
Pluripotency	<i>TERF1</i>	Telomeric repeat binding factor (NIMA-interacting) 1	A_23_P216149	2 (late)
Pluripotency	<i>DPPA5</i>	Developmental pluripotency associated 5	A_32_P233950	2 (late)
Pluripotency	<i>TERT</i>	Telomerase reverse transcriptase	A_23_P110851	2 (late)
Pluripotency	<i>ZIC3</i>	Zic family member 3	A_23_P327910	2 (late)
Pluripotency	<i>LIN28a</i>	LIN28 homolog A	A_23_P74895	2 (late)
Pluripotency	<i>LIN28b</i>	LIN28 homolog B	A_33_P3220615	2 (late)
Pluripotency	<i>LECT1</i>	Leukocyte cell derived chemotaxin 1	A_23_P25587	2 (late)
Pluripotency	<i>DNMT3B</i>	DNA methyl-transferase 3B	A_23_P28953	2 (late)
Fibroblast	<i>COL3A1</i>	Pro-collagen α 2(III)	A_24_P935491	1
Fibroblast	<i>FSP-1</i>	Fibroblast surface protein	A_23_P94800	1
Fibroblast	<i>TGFB3</i>	Transforming growth factor beta 3	A_23_P88404	1
Fibroblast	<i>TGFB2</i>	Transforming growth factor beta 2	A_24_P402438	1
Fibroblast	<i>COL1A2</i>	Pro-collagen α 2(I)	A_24_P277934	2
Fibroblast	<i>ITGA1</i>	Integrin α 1b1 (VLA-1)	A_33_P3353791	2
Fibroblast	<i>DDR2</i>	Discoidin-domain-receptor-2	A_23_P452	2
Fibroblast	<i>P4HA3</i>	Prolyl 4-hydroxylase	A_24_P290286	2
Fibroblast	<i>THY1</i>	Thy-1 cell surface antigen; CD90	A_33_P3280845	2
Fibroblast	<i>FAP</i>	Fibroblast activation protein	A_23_P56746	2
Fibroblast	<i>CD248</i>	Endosialin, TEM1	A_33_P3337485	2
Fibroblast	<i>VIM</i>	Vimentin	A_23_P161190	2
Fibroblast	<i>COL1A1</i>	Pro-collagen α 1(I)	A_33_P3304668	3
Fibroblast	<i>ITGA5</i>	Integrin α 5b1	A_23_P36562	3
Fibroblast	<i>P4HA1</i>	Prolyl 4-hydroxylase	A_33_P3214481	3
Fibroblast	<i>P4HA2</i>	Prolyl 4-hydroxylase	A_33_P3394933	3
Fibroblast	<i>TGFB1</i>	Transforming growth factor beta 1	A_24_P79054	3
Fibroblast	<i>HSP47</i>	Serpin family H member 1, SERPINH1	A_33_P3269203	-
Fibroblast	<i>CD34</i>	Hematopoietic progenitor cell antigen	A_23_P23829	-