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

Nelly Olova[1†], Daniel J Simpson[1†], Riccardo Marioni[2], Tamir Chandra[1]

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

Abstract

Induced pluripotent stem cells (IPSCs), with their unlimited regenerative capacity, carry the promise for tissue replacement to counter age-related decline. However, attempts to realise *in vivo* iPSC have invariably resulted in the formation of teratomas. Partial reprogramming in prematurely aged mice has shown promising results in alleviating age-related symptoms without teratoma formation. Does partial reprogramming lead to rejuvenation (i.e. “younger” cells), rather than dedifferentiation, which bears the risk of cancer? Here we analyse cellular age during iPSC reprogramming and find that partial reprogramming leads to a reduction in the biological age of cells. We also find that the loss of somatic gene expression and epigenetic age follow different kinetics, suggesting that rejuvenation can be achieved with a minimised risk of cancer.

Introduction

The human ageing process is accompanied by multiple degenerative diseases. Our understanding of such ageing related disorders is, nevertheless, fragmented, and the existence and nature of a general underlying cause are still much debated (Faragher 2015; Gladyshev & Gladyshev 2016). A breakthrough technique, the generation of

induced pluripotent stem cells (iPSCs), allows the reprogramming of somatic cells back to an embryonic stem cell (ESC) like state with an unlimited regenerative capacity. This has led to multiple strategies for tissue replacement in degenerative diseases (Takahashi et al. 2007). Clinical application of iPSCs however, is at its infancy (Takahashi & Yamanaka 2016; Singh et al. 2015; Soria-Valles et al. 2015), and the potency of iPSCs bears risks, not least cancer induction. For example, *in vivo* experiments with iPSCs have shown that continuous expression of Yamanaka factors (Oct4, Sox2, Klf4 and c-Myc, thus OSKM) in adult mice invariably leads to cancer (Abad et al. 2013; Ohnishi et al. 2014).

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

Ocampo et al. demonstrated that partial reprogramming by transient cyclic induction of OSKM ameliorates signs of ageing and extends lifespan in progeroid mice, with no resulting teratoma formation (Ocampo et al. 2016). Ocampo et al. thus established partial reprogramming as a promising candidate intervention for age-related disease. However, accurately determining biological age in mice was not

possible at the time of the Ocampo study. Therefore, the nature (i.e. dedifferentiation/rejuvenation) of the cellular changes remain unexplored:

1) Does the epigenetic remodelling seen truly reflect rejuvenation (i.e. a reduction in cellular/tissue age)? If so, can we observe rejuvenation in human cells?
2) What is the extent of the rejuvenation after a cycle of partial reprogramming (e.g. years/cyclic induction)?
3) Are there signs of dedifferentiation in early reprogramming?

A major obstacle in understanding the relation between differentiation and ageing has been our inability to accurately measure cellular age with a high correlation to the chronological age of the organism. However, over the last five years a number of age predictors have been developed, the most accurate of which utilise DNA methylation (known as epigenetic clocks) (Horvath 2013; Hannum et al. 2013), with the “Horvath clock” being the most widely used (r=0.96). The Horvath clock predicts the age (or epigenetic age, eAge) of multiple tissues with a median error of 3.6 years (Horvath 2013). A predicted eAge older than the chronological age (“epigenetic age acceleration”) is associated with a higher risk of all-cause mortality (Marioni et al. 2015; Christiansen et al. 2016; Perna et al. 2016), premature ageing syndromes (Down and Werner) (Maierhofer et al. 2017; Horvath et al. 2015), frailty and menopause (Breitling et al. 2016; Levine et al. 2016). Epigenetic age is distinct from other biomarkers, such as senescence and telomere length (Lowe et al. 2016). All of these studies suggest that eAge truly measures biological age.

Does partial reprogramming lead to a reduction in cellular age? Calculating eAge dynamics over the course of iPSC reprogramming of human dermal fibroblasts

(HDFs) allows us to address this question. We observe onset of a continuous decline of eAge after day 3 from induction with OSKM. Our results suggest that partial reprogramming leads to a reduction in the eAge of cells and is therefore a rejuvenation mechanism. Comparing eAge decline to loss of somatic gene expression indicates a window within which rejuvenation might occur uncoupled from dedifferentiation.

Results

Epigenetic age shows a steady decline in early reprogramming

To understand the dynamics of eAge within a reprogramming time-course, we calculated eAge using Horvath's multi-tissue age predictor over a previously published 49-day reprogramming time-course on HDFs (Ohnuki et al. 2014; Horvath 2013). Epigenetic rejuvenation, i.e. decrease of eAge, commenced between days 3 and 7 after OSKM transduction and continued until day 20, when it was stably reset to zero (Fig. 1a). A broken stick model with two linear sections starting from day 3 showed a good fit to the observed data and measured a steady decrease with 3.8 years per day until day 20 (SE 0.27, $P = 3.8 \times 10^{-7}$) (Fig. 1a). Our data suggest that partial reprogramming does indeed result in a reduction of eAge in human cells and can be considered a rejuvenation mechanism.

Partial reprogramming in Ocampo et al. was achieved after just two days of OKSM induction in mice carrying an inducible OSKM transgene. However, kinetics are likely to be different with *in vitro* fibroblasts where OSKM is induced by viral transduction. To associate the eAge with intermediate states in the reprogramming trajectory we compared it to gene expression measured in the same samples. We analysed

corresponding microarray gene expression data for 19 well-established pluripotency genes (Table 1 and Supplementary fig.1) as a proxy for reaching a mature pluripotent state (Ginis et al. 2004; Cai et al. 2006; Mallon et al. 2013; Galan et al. 2013; Boyer et al. 2005). We clustered expression patterns of those genes (Genolini et al. 2015), which resulted in two composite trajectories. These followed previously described expression dynamics of early (cluster 1) and late (cluster 2) activated pluripotency genes (Fig. 1a) (Tanabe et al. 2013; Chung et al. 2014; Buganim et al. 2012; Takahashi & Yamanaka 2016). Pluripotency gene cluster 1 included *NANOG*, *SALL4*, *ZFP42*, *TRA-1-60, UTF1, DPPA4* and *LEFTY2*, and their expression increased dramatically within the first 10 days and then established stable pluripotency expression levels by day 20. In contrast, pluripotency gene cluster 2 (containing late expressing genes such as *LIN28*, *ZIC3* and *DNMT3B*) elevated expression more slowly and reached stable pluripotency levels by day 30 (Tanabe et al. 2013; Chung et al. 2014). eAge is reset to zero at the same time that the genes in cluster 1 reach their pluripotent state levels, which temporally precedes completion of the full iPSC time course. In summary, eAge decline is observed well within the first wave of pluripotency gene expression.

Loss of somatic gene expression is uncoupled from rejuvenation dynamics and occurs step-wise

Therapeutic partial reprogramming will depend on rejuvenation with minimal dedifferentiation, which carries the risk of malignancies. Here we studied the dynamics of fibroblast gene down-regulation as a proxy for the loss of somatic cell identity. The individual trajectories of commonly used fibroblast marker genes (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) (Table 1 and Supplementary Fig. 2) clustered into three composite expression patterns, two of which (clusters 2 and 3) went into an immediate decline after OSKM induction (Fig. 1b). However, one fibroblast-specific cluster (cluster 1) remained stable in its expression for the first 15 days. Interestingly, after day 7, fibroblast-specific gene expression in clusters 2 and 3 stopped declining and plateaued until day 15. Vimentin (*VIM*), for example, remained at 60% of maximal expression until day 15 of reprogramming, similarly to *FAP*, *CD248* and *COL1A2* in cluster 2. After day 15, fibroblast gene expression declined rapidly in all three clusters, and only by day 35 had all reached ESC expression levels, marking a complete loss of somatic identity. Of the three fibroblast gene clusters, cluster 1 showed the slowest decline, and was also the last to reach ESC expression levels. This cluster contains *FSP1*, *COL3A1* and *TGFB2/3* (Supplementary fig. 2), which are well described indicators of fibroblast identity (Kalluri & Zeisberg 2006). In summary, we found several fibroblast specific genes (cluster 1) that maintained fibroblast expression levels until day 15 (by which time a significant drop in eAge has been observed), and the pattern of the decline observed in all three fibroblast clusters indicates that loss of tissue specific expression occurs in a step-wise manner.

Discussion

Ground-breaking work by Ocampo et al. showed that partial reprogramming can alleviate age-related pathologies in prematurely ageing mice, highlighting it as a rejuvenation strategy to counteract age-related disease (Ocampo et al. 2016). The authors suspect that epigenetic rejuvenation is the driver behind the improved age-associated phenotype both in their *in vivo* and *in vitro* experiments. Epigenetic

rejuvenation, reversal of cellular age, is a promising concept as it could avoid the oncogenic risks associated with dedifferentiation. However, to determine whether partial reprogramming indeed leads to epigenetic rejuvenation (or other cellular changes that might improve age-related pathologies) requires an accurate measure of biological age. This is currently only feasible through profiling DNA methylation, which remained unexplored in Ocampo et al. Here, analysing a reprogramming time-course on HDFs, we show that eAge indeed declines early in reprogramming suggesting that the improvements Ocampo et al. observed might be due to epigenetic rejuvenation.

A deep understanding of the kinetics of rejuvenation will be required to master therapeutic partial reprogramming, since dedifferentiation carries the risk of malignancies. We analysed the dynamics of rejuvenation along the iPSC time-course and compared it to fibroblast specific gene expression as a proxy for dedifferentiation. Within the iPSC reprogramming time-course, partial reprogramming happens within an early, reversible phase involving stochastic activation of pluripotency genes. It is followed by a more deterministic maturation phase with predictable order of gene expression changes, where cell fate is firmly bound towards pluripotency (Takahashi & Yamanaka 2016; Smith et al. 2016). For example, Tanabe et al. showed that ~50% of cells with TRA-1-60 expression, a marker of human ESC/iPSC cells, spontaneously revert to a TRA-1-60 negative state at reprogramming days 7 and 11 (Tanabe et al. 2013). Transcriptome clustering in Tanabe et al. also suggests that a reversion for 9 days within the uncommitted phase places the cells closer to the original fibroblast than day 4 OSKM (Tanabe et al. 2013). Indeed, it has been shown that mouse fibroblasts fail to

become iPSC and revert to their original state if OSKM expression is discontinued during the initial stochastic phase (Brambrink et al. 2008; Stadtfeld et al. 2008). Our data suggest a window of opportunity (3-7 days in the iPSC experiment) within the uncommitted reprogramming phase where a pronounced decline of eAge happens alongside a partial maintenance of fibroblast gene expression (Fig. 1b) (Tanabe et al. 2013). Further experiments are needed to determine the safe rejuvenation boundary in different reprogramming systems.

Different dynamics between the step-wise decline in fibroblast expression and the linear decline in eAge further indicate that dedifferentiation and epigenetic rejuvenation can be uncoupled. It remains to be shown how stable the rejuvenated phenotype is. How long or how often would cyclic induction of OSKM need to be conducted to stabilise the young phenotype? Further analysis is also needed regarding the effect of partial reprogramming on adult stem cells or premalignant cells, which have already shown a higher propensity of transforming to malignancy (Abad et al. 2013; Ohnishi et al. 2014). It is possible that a premalignant phenotype could be attenuated or amplified by partial reprogramming.

Methods

Overview of the Ohnuki et al dataset

450K DNA methylation array and gene expression microarray data of full HDF reprogramming time-course was obtained from GSE54848. Microarray data (LOG2 transformed) was available for three to four replicates per data point, whilst DNA methylation data was available for three replicates.

Predicting eAge

For each time point, methylation data for 26987 CpG sites was uploaded to the online DNA methylation age calculator to assess eAge: https://labs.genetics.ucla.edu/horvath/dnamage/ (Horvath 2013).

Methylation Age Trajectories

A 'broken stick' model with two linear sections was constructed to chart overall change in DNA methylation age over time between the three HDF cell lines. A linear mixed model was specified with a random intercept term for each replicate. A variable break point was set between the minimum and maximum day, plus and minus a small constant (3 days), respectively. The predicted values from the regression models were plotted against the measurement day.

Gene clusters and trajectories

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

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

References

Abad M, Mosteiro L, Pantoja C, Cañamero M, Rayon T, Ors I, Graña O, Megías D,

Domínguez O, Martínez D, Manzanares M, Ortega S & Serrano M (2013) Reprogramming in vivo produces teratomas and iPS cells with totipotency features. *Nature* 502, 340–345.

Bates D, Mächler M, Bolker B & Walker S (2014) Fitting Linear Mixed-Effects Models using lme4. 67.

Boyer L a L a., Lee TITI, Cole MFMF, Johnstone SESE, Stuart S, Zucker JPJP, Guenther MGMG, Kumar RMRM, Murray HLHL, Jenner RGRG, Gifford DK, Melton D a, Jaenisch R, Young R a, Levine SS & Others (2005) Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. *Young* 122, 947–956.

Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, Suh H & Jaenisch R (2008) Sequential Expression of Pluripotency Markers during Direct Reprogramming of Mouse Somatic Cells. *Cell Stem Cell* 2, 151–159.

Breitling LP, Saum K-U, Perna L, Schöttker B, Holleczek B & Brenner H (2016) Frailty is associated with the epigenetic clock but not with telomere length in a German cohort. *Clin. Epigenetics* 8, 21.

Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, Klemm SL, Van Oudenaarden A & Jaenisch R (2012) Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 150, 1209–1222.

Cai J, Chen J, Liu Y, Miura T, Luo Y, Loring JF, Freed WJ, Rao MS & Zeng X (2006) Assessing self-renewal and differentiation in human embryonic stem cell lines. *Stem Cells* 24, 516–30.

Chang Y, Li H & Guo Z (2014) Mesenchymal stem cell-like properties in fibroblasts. *Cell. Physiol. Biochem.* 34, 703–714.

Christiansen L, Lenart A, Tan Q, Vaupel JW, Aviv A, McGue M & Christensen K (2016) DNA methylation age is associated with mortality in a longitudinal Danish twin study. *Aging Cell* 15, 149–154.

Chung KM, Kolling FW, Gajdosik MD, Burger S, Russell AC & Nelson CE (2014) Single cell analysis reveals the stochastic phase of reprogramming to pluripotency is an ordered probabilistic process. *PLoS One* 9.

Faragher RGA (2015) Should we treat aging as a disease? The consequences and dangers of miscategorisation. *Front. Genet.* 6, 1–7.

Galan A, Diaz-Gimeno P, Poo ME, Valbuena D, Sanchez E, Ruiz V, Dopazo J, Montaner D, Conesa A & Simon C (2013) Defining the Genomic Signature of Totipotency and Pluripotency during Early Human Development. *PLoS One* 8, 20–23.

Genolini C, Alacoque X & Marianne Sentenac CA (2015) kml and kml3d: R Packages to Cluster Longitudinal Data. *J. Stat. Softw.* 65, 1–34. Available at: http://www.jstatsoft.org/v65/i04/.

Ginis I, Luo Y, Miura T, Thies S, Brandenberger R, Gerecht-Nir S, Amit M, Hoke A, Carpenter MK, Itskovitz-Eldor J & Rao MS (2004) Differences between human and mouse embryonic stem cells. *Dev. Biol.* 269, 360–380.

269 Gladyshev T V. & Gladyshev VN (2016) A Disease or Not a Disease? Aging As a
270 Pathology. *Trends Mol. Med.* 22, 995–996.

271 Goodpaster T, Legesse-Miller A, Hameed MR, Aisner SC, Randolph-Habecker J &
272 Coller HA (2008) An Immunohistochemical Method for Identifying Fibroblasts in
273 Formalin-fixed, Paraffin-embedded Tissue. *J. Histochem. Cytochem.* 56, 347–
274 358.

275 Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, Klotzle B, Bibikova M,
276 Fan J-B, Gao Y, Deconde R, Chen M, Rajapakse I, Friend S, Ideker T & Zhang
277 K (2013) Genome-wide Methylation Profiles Reveal Quantitative Views of
278 Human Aging Rates. *Mol. Cell* 49, 359–367.

279 Horvath S (2013) DNA methylation age of human tissues and cell types. *Genome*
280 *Biol.* 14, R115.

281 Horvath S, Garagnani P, Bacalini MG, Pirazzini C, Salvioli S, Gentilini D, Di Blasio
282 AM, Giuliani C, Tung S, Vinters H V. & Franceschi C (2015) Accelerated
283 epigenetic aging in Down syndrome. *Aging Cell* 14, 491–495.

284 Janmaat CJ, De Rooij KE, Locher H, De Groot SC, De Groot JCMJ, Frijns JHM &
285 Huisman MA (2015) Human dermal fibroblasts demonstrate positive
286 immunostaining for neuron- and glia-specific proteins. *PLoS One* 10, 1–14.

287 Kalluri R & Zeisberg M (2006) Fibroblasts in cancer. *Nat. Rev. Cancer* 6, 392–401.

288 Kuznetsova A, Brockhoff PB & Bojesen Christensen RH (2016) lmerTest: Tests in
289 Linear Mixed Effects Models. R package version 2.0-33. Available at:
290 https://cran.r-project.org/web/packages/lmerTest/index.html.

291 Levine ME, Lu AT, Chen BH, Hernandez DG, Singleton AB, Ferrucci L, Bandinelli S,
292 Salfati E, Manson JE, Quach A, Kusters CDJ, Kuh D, Wong A, Teschendorff
293 AE, Widschwendter M, Ritz BR, Absher D, Assimes TL & Horvath S (2016)
294 Menopause accelerates biological aging. *Proc. Natl. Acad. Sci.* 113, 9327–9332.

295 Lowe D, Horvath S & Raj K (2016) Epigenetic clock analyses of cellular senescence
296 and ageing. *Oncotarget* 7, 8524–31.

297 MacFadyen JR, Haworth O, Roberston D, Hardie D, Webster MT, Morris HR, Panico
298 M, Sutton-Smith M, Dell A, Van Der Geer P, Wienke D, Buckley CD & Isacke
299 CM (2005) Endosialin (TEM1, CD248) is a marker of stromal fibroblasts and is
300 not selectively expressed on tumour endothelium. *FEBS Lett.* 579, 2569–2575.

301 Maierhofer A, Flunkert J, Oshima J, Martin GM, Haaf T & Horvath S (2017)
302 Accelerated epigenetic aging in Werner syndrome. *Aging (Albany. NY).* 9,
303 1143–1152.

304 Mallon BS, Chenoweth JG, Johnson KR, Hamilton RS, Tesar PJ, Yavatkar AS,
305 Tyson LJ, Park K, Chen KG, Fann YC & McKay RDG (2013) StemCellDB: The
306 Human Pluripotent Stem Cell Database at the National Institutes of Health.
307 *Stem Cell Res.* 10, 57–66.

308 Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, Gibson J, Henders
309 AK, Redmond P, Cox SR, Pattie A, Corley J, Murphy L, Martin NG, Montgomery
310 GW, Feinberg AP, Fallin M, Multhaup ML, Jaffe AE, Joehanes R, Schwartz J,
311 Just AC, Lunetta KL, Murabito JM, Starr JM, Horvath S, Baccarelli AA, Levy D,

312 Visscher PM, Wray NR & Deary IJ (2015) DNA methylation age of blood
313 predicts all-cause mortality in later life. *Genome Biol.* 16, 25.

314 Ocampo A, Reddy P, Martinez-Redondo P, Platero-Luengo A, Hatanaka F, Hishida
315 T, Li M, Lam D, Kurita M, Beyret E, Araoka T, Vazquez-Ferrer E, Donoso D,
316 Roman JL, Xu J, Rodriguez Esteban C, Nuñez G, Nuñez Delicado E, Campistol
317 JM, Guillen I, Guillen P & Izpisua Belmonte JC (2016) In Vivo Amelioration of
318 Age-Associated Hallmarks by Partial Reprogramming. *Cell* 167, 1719–
319 1733.e12.

320 Ohnishi K, Semi K, Yamamoto T, Shimizu M, Tanaka A, Mitsunaga K, Okita K,
321 Osafune K, Arioka Y, Maeda T, Soejima H, Moriwaki H, Yamanaka S, Woltjen K
322 & Yamada Y (2014) Premature termination of reprogramming in vivo leads to
323 cancer development through altered epigenetic regulation. *Cell* 156, 663–677.

324 Ohnuki M, Tanabe K, Sutou K, Teramoto I, Sawamura Y, Narita M, Nakamura M,
325 Tokunaga Y, Nakamura M, Watanabe A, Yamanaka S & Takahashi K (2014)
326 Dynamic regulation of human endogenous retroviruses mediates factor-induced
327 reprogramming and differentiation potential. *Proc. Natl. Acad. Sci.* 111, 12426–
328 12431.

329 Perna L, Zhang Y, Mons U, Holleczek B, Saum K-U & Brenner H (2016) Epigenetic
330 age acceleration predicts cancer, cardiovascular, and all-cause mortality in a
331 German case cohort. *Clin. Epigenetics* 8, 64.

332 Pilling D, Fan T, Huang D, Kaul B & Gomer RH (2009) Identification of markers that
333 distinguish monocyte-derived fibrocytes from monocytes, macrophages, and
334 fibroblasts. *PLoS One* 4, 31–33.

335 Singh PB & Zacouto F (2010) Nuclear reprogramming and epigenetic rejuvenation.
336 *J. Biosci.* 35, 315–319.

337 Singh VK, Kalsan M, Kumar N, Saini A & Chandra R (2015) Induced pluripotent stem
338 cells: applications in regenerative medicine, disease modeling, and drug
339 discovery. *Front. Cell Dev. Biol.* 3, 1–18.

340 Smith ZD, Sindhu C & Meissner A (2016) Molecular features of cellular
341 reprogramming and development. *Nat. Rev. Mol. Cell Biol.* 17, 139–154.

342 Soria-Valles C, Osorio FG, Gutiérrez-Fernández A, De Los Angeles A, Bueno C,
343 Menéndez P, Martín-Subero JI, Daley GQ, Freije JMP & López-Otín C (2015)
344 NF-κB activation impairs somatic cell reprogramming in ageing. *Nat. Cell Biol.*
345 17, 1004–13.

346 Stadtfeld M, Maherali N, Breault DT & Hochedlinger K (2008) Defining Molecular
347 Cornerstones during Fibroblast to iPS Cell Reprogramming in Mouse. *Cell Stem*
348 *Cell* 2, 230–240.

349 Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K & Yamanaka S
350 (2007) Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by
351 Defined Factors. *Cell* 131, 861–872.

352 Takahashi K & Yamanaka S (2016) A decade of transcription factor-mediated
353 reprogramming to pluripotency. *Nat. Rev. Mol. Cell Biol.* 17, 183–193.

354 Tanabe K, Nakamura M, Narita M, Takahashi K & Yamanaka S (2013) Maturation,

355 not initiation, is the major roadblock during reprogramming toward pluripotency
356 from human fibroblasts. *Proc. Natl. Acad. Sci.* 110, 12172–12179.

357 Zhou L, Yang K, Randall Wickett R & Zhang Y (2016) Dermal fibroblasts induce cell
358 cycle arrest and block epithelial–mesenchymal transition to inhibit the early
359 stage melanoma development. *Cancer Med.* 5, 1566–1579.

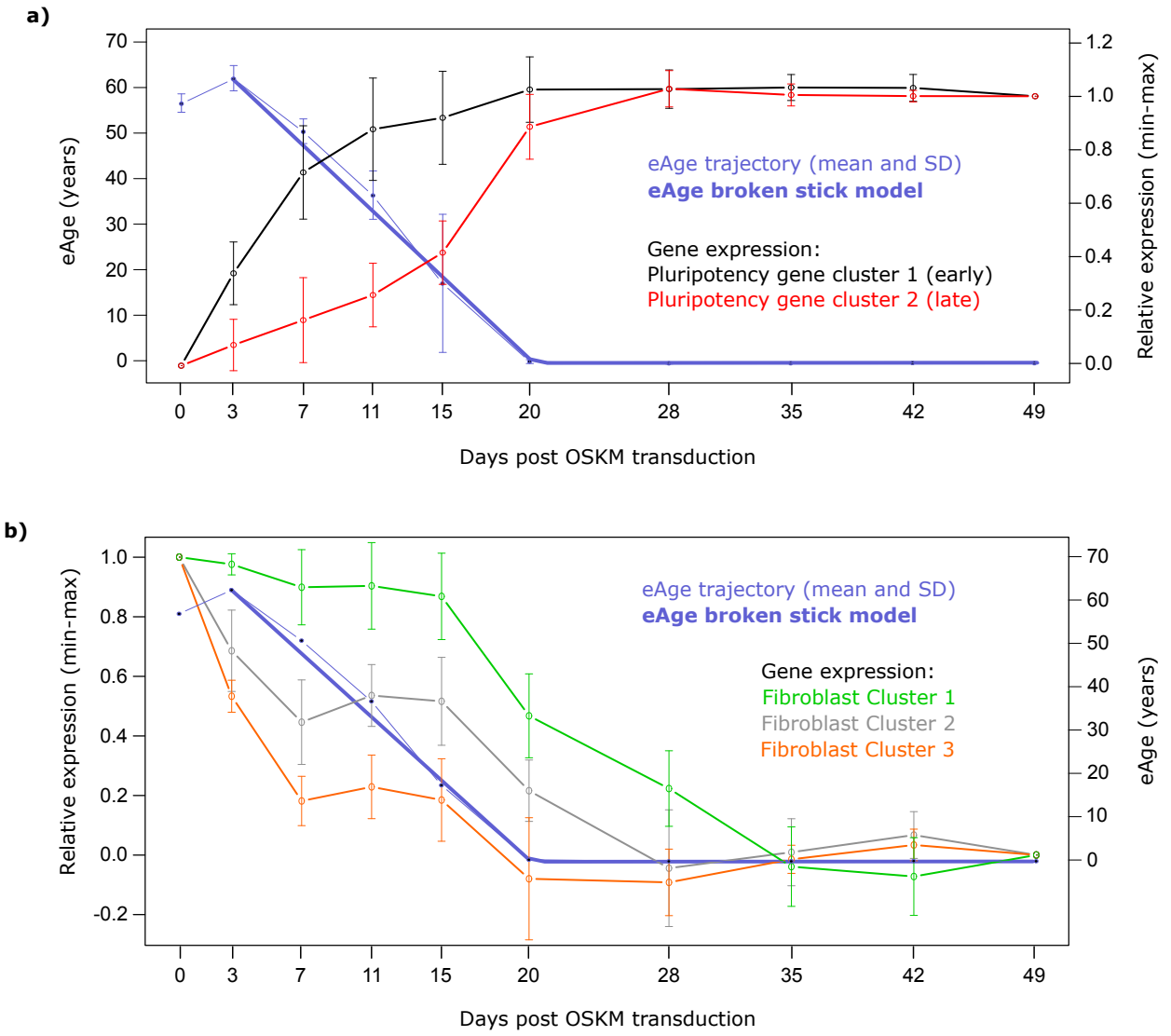

Figure 1. Dynamics of eAge and gene expression in a 49-day HDF reprogramming time-course. **(a) Left Y axis**: eAge trajectory fitted with a broken stick model with two linear sections, error bars represent SD. Measured rate (years per day) of eAge decrease: [day 3 -day 20] = -3.8, SE 0.27, $P = 3.8 \times 10^{-7}$. **Right Y axis**: Composite gene expression trajectories of of key pluripotency markers, clustered as per Genolini et al. 2016. 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 generated as described in (a). **Right Y axis**: same as left Y axis in (a), 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	*NANOG*	Nanog homeobox	A_23_P204640	1 (early)
Pluripotency	*REX1 (ZFP42)*	Zinc Finger Protein 42	A_23_P395582	1 (early)
Pluripotency	*TRA-1-60/81 (PODXL)*	Podocalyxin	A_23_P215060	1 (early)
Pluripotency	*UTF1*	Undifferentiated embryonic cell transcription factor 1	A_33_P3294217	1 (early)
Pluripotency	*DPPA4*	Developmental pluripotency associated 4	A_23_P380526	1 (early)
Pluripotency	*TDGF1 (CRIPTO)*	Teratocarcinoma-derived growth factor 1	A_23_P366376	1 (early)
Pluripotency	*SALL4*	Spalt like transcription factor 4	A_23_P109072	1 (early)
Pluripotency	*LEFTY1*	Left-right determination factor 1	A_23_P160336	1 (early)
Pluripotency	*LEFTY2*	Left-right determination factor 2	A_23_P137573	1 (early)
Pluripotency	*DNMT3A*	DNA methyl-transferase 3A	A_23_P154500	1 (early)
Pluripotency	*TFCP2L1*	Transcription factor CP2 like 1	A_23_P5301	1 (early)
Pluripotency	*TERF1*	Telomeric repeat binding factor (NIMA-interacting) 1	A_23_P216149	2 (late)
Pluripotency	*DPPA5*	Developmental pluripotency associated 5	A_32_P233950	2 (late)
Pluripotency	*TERT*	Telomerase reverse transcriptase	A_23_P110851	2 (late)
Pluripotency	*ZIC3*	Zic family member 3	A_23_P327910	2 (late)
Pluripotency	*LIN28a*	LIN28 homolog A	A_23_P74895	2 (late)
Pluripotency	*LIN28b*	LIN28 homolog B	A_33_P3220615	2 (late)
Pluripotency	*LECT1*	Leukocyte cell derived chemotaxin 1	A_23_P25587	2 (late)
Pluripotency	*DNMT3B*	DNA methyl-transferase 3B	A_23_P28953	2 (late)
Fibroblast	*COL3A1*	Pro-collagen α2(III)	A_24_P935491	1
Fibroblast	*FSP-1*	Fibroblast surface protein	A_23_P94800	1
Fibroblast	*TGFB3*	Transforming growth factor beta 3	A_23_P88404	1
Fibroblast	*TGFB2*	Transforming growth factor beta 2	A_24_P402438	1
Fibroblast	*COL1A2*	Pro-collagen α2(I)	A_24_P277934	2
Fibroblast	*ITGA1*	Integrin a1b1 (VLA-1)		2
Fibroblast	*DDR2*	Discoidin-domain-receptor-2	A_23_P452	2
Fibroblast	*P4HA3*	Prolyl 4-hydroxylase	A_24_P290286	2
Fibroblast	*THY1*	Thy-1 cell surface antigen; CD90	A_33_P3280845	2
Fibroblast	*FAP*	Fibroblast activation protein	A_23_P56746	2
Fibroblast	*CD248*	Endosialin, TEM1	A_33_P3337485	2
Fibroblast	*VIM*	Vimentin	A_23_P161190	2
Fibroblast	*COL1A1*	Pro-collagen α1(I)	A_33_P3304668	3
Fibroblast	*ITGA5*	Integrin a5b1	A_23_P36562	3
Fibroblast	*P4HA1*	Prolyl 4-hydroxylase	A_33_P3214481	3
Fibroblast	*P4HA2*	Prolyl 4-hydroxylase	A_33_P3394933	3
Fibroblast	*TGFB1*	Transforming growth factor beta 1	A_24_P79054	3
Fibroblast	*HSP47*	Serpin family H member 1, SERPINH1	A_33_P3269203	-
Fibroblast	*CD34*	Hematopoietic progenitor cell antigen	A_23_P23829	-