Figure S1

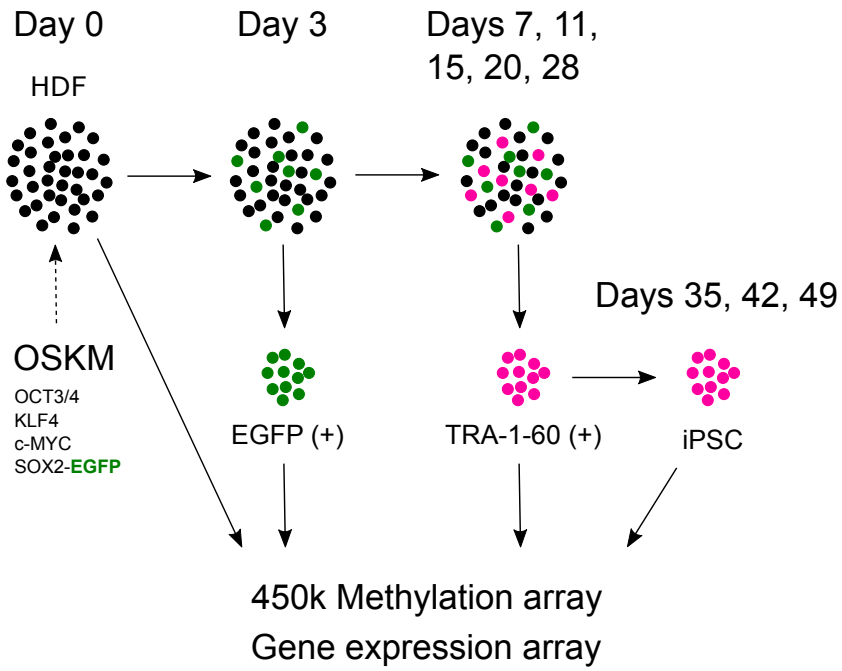

Figure S1. Schematic overview of the experimental setup of Ohnuki et al iPSC time-course and dataset time points. HDF cells were transfected with EGFP-labelled OSKM on day 0 and cultured in virus-containing medium for 24 hours, then replaced by 10%-FBS containing medium for 8 days before replacing with human ESC medium. EGFP (+) cells, representing the population of successfully transfected cells, were sorted by flow cytometry on day 3. Intermediate reprogrammed cells positive for the human pluripotency marker TRA-1-60 were sorted by magnetic activated cell sorting on days 7, 11, 15, 20 and 28 post-transfection. Day 28-sorted TRA-1-60 (+) cells were further expanded and samples collected three more times on each seventh day, i.e. on days 35, 42 and 49. The sorted and collected cells at each time point were subjected to both gene expression and CpG methylation array sequencing. Microarray gene expression was performed for three to four replicates per data point, whilst DNA methylation was performed for two to three replicates per time point. Both 450K DNA methylation array and gene expression microarray datasets were obtained from GSE54848 (Ohnuki et al, 2014).

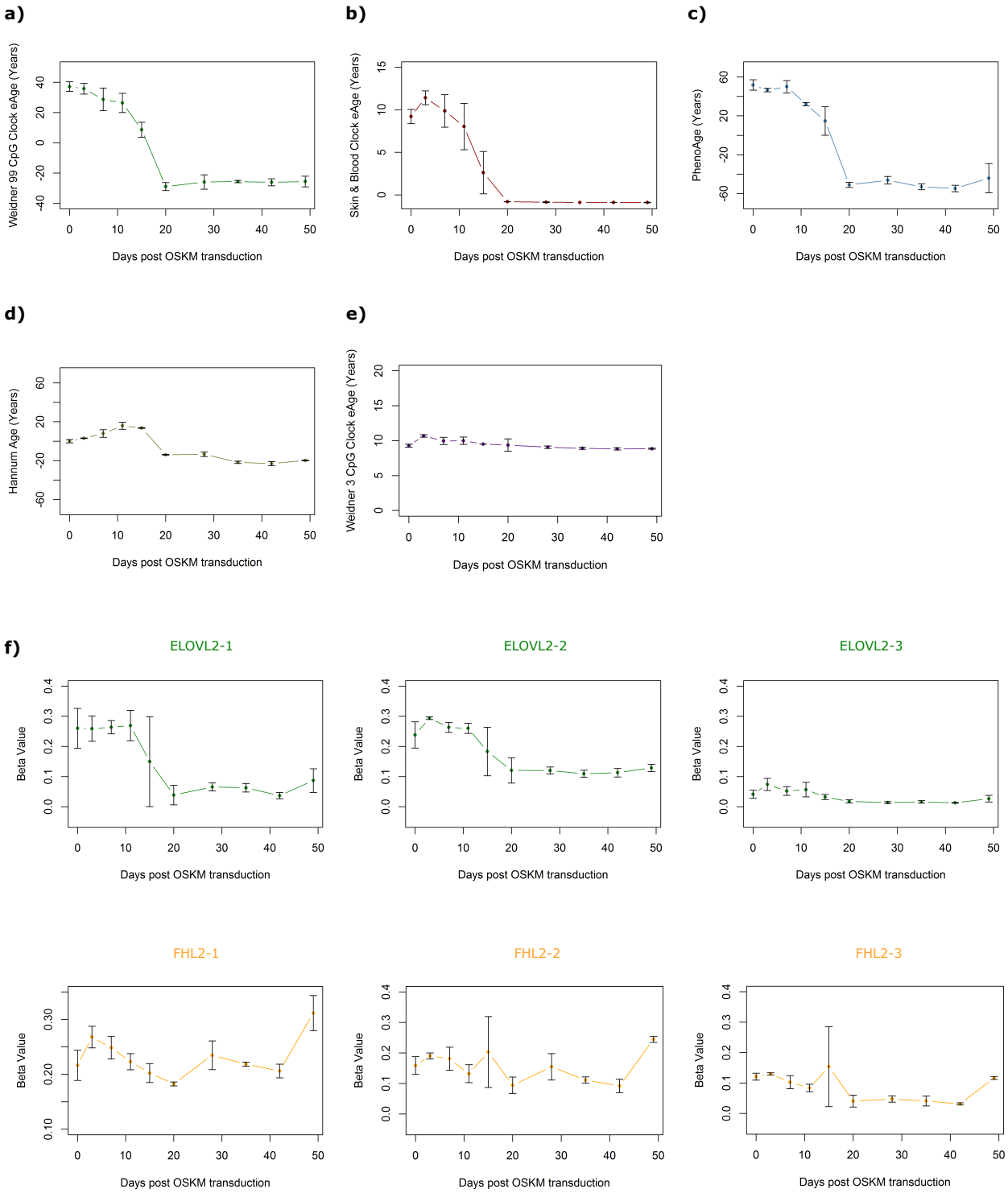

Figure S2. eAge trajectories of different DNA methylation-based epigenetic clocks. (a) Weidner 99 CpG blood-based epigenetic clock (Weidner et al. 2014); **(b)** Skin & blood clock (Horvath et al. 2018); **(c)** PhenoAge (Levine et al. 2018); **(d)** Hannum blood-based epigenetic clock (Hannum et al. 2013); **(e)** Weidner 3 CpG blood-based epigenetic clock (Weidner et al. 2014); **(f)** Individual CpG age predictors found in CpG islands within the *ELOVL2* and *FHL2* genes (Garagnani et al. 2012).

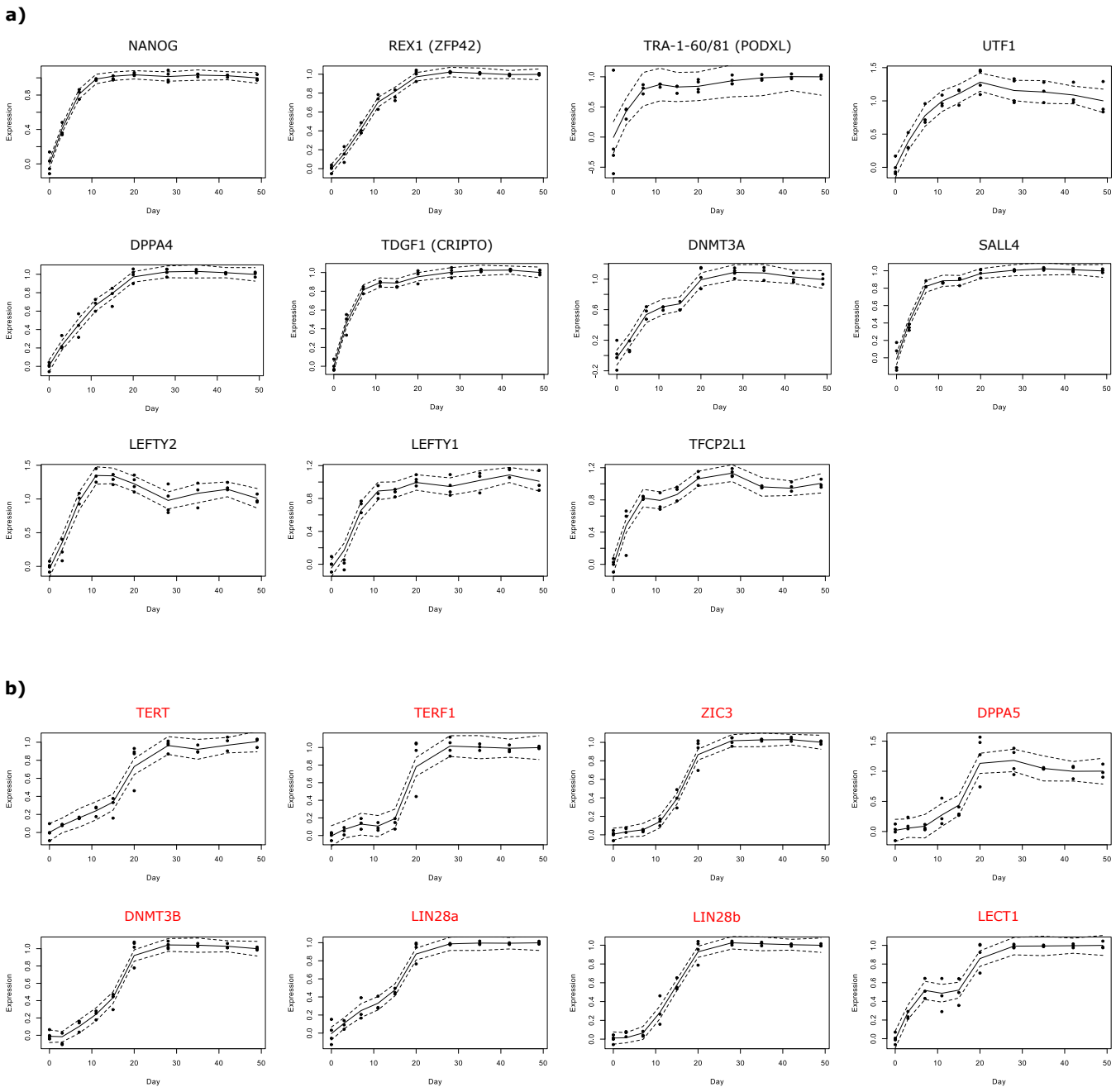

Figure S3. Expression of key pluripotency markers in a 49-day HDF reprogramming time-course. Individual expression dynamics of Cluster 1 genes (early pluripotency markers) in **(a)** and Cluster 2 genes (late expressing pluripotency markers) in **(b)**. Values are LOG2 transformed and normalised between 0 and 1 for 'day 0' and 'day 49', respectively, based on the average values between biological replicates for each time point. Dotted line marks CI. Gene label colours correspond to cluster colours in Fig. 1A.

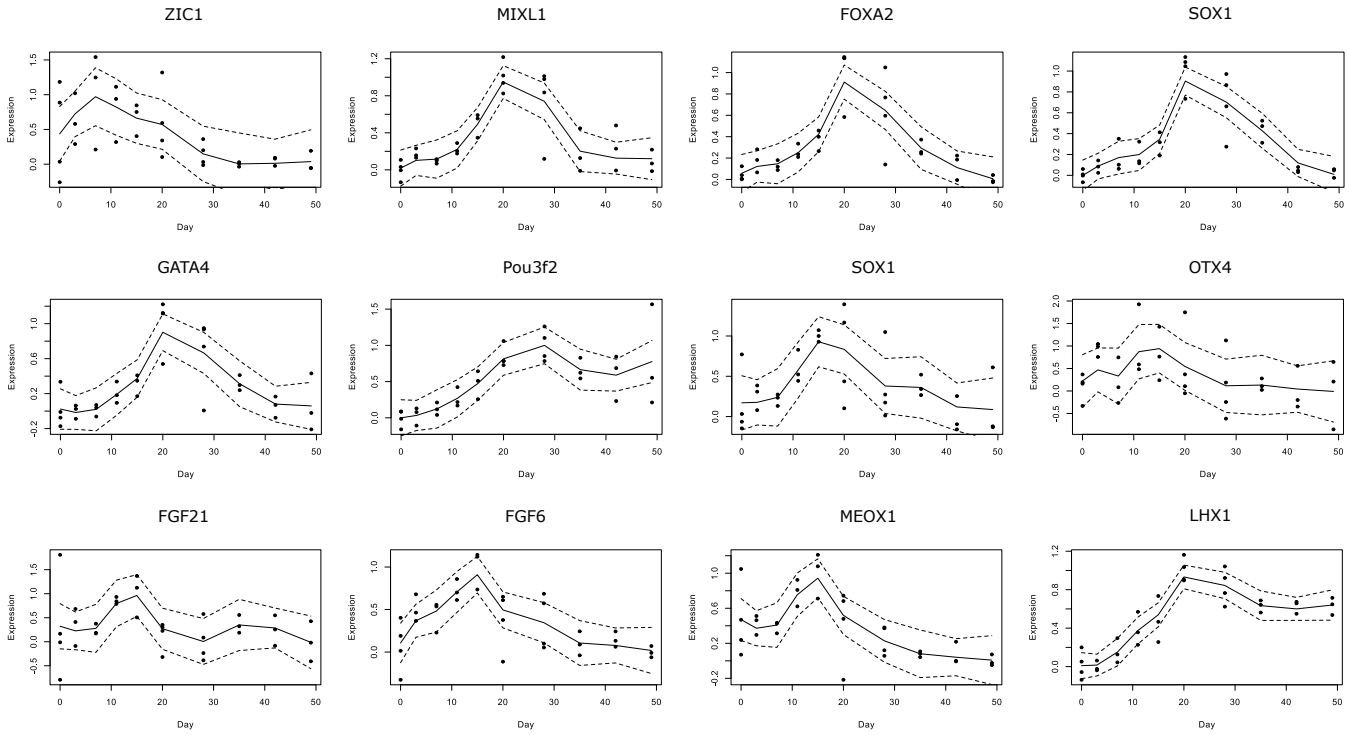

Figure S4. Expression trajectories of key developmental genes in a 49-day HDF reprogramming time-course. Values are LOG2 transformed and normalised between 0 and 1 for the ‘minimum’ and ‘maximum’ value, respectively, based on the average values between biological replicates for each time point. Dotted line marks CI.

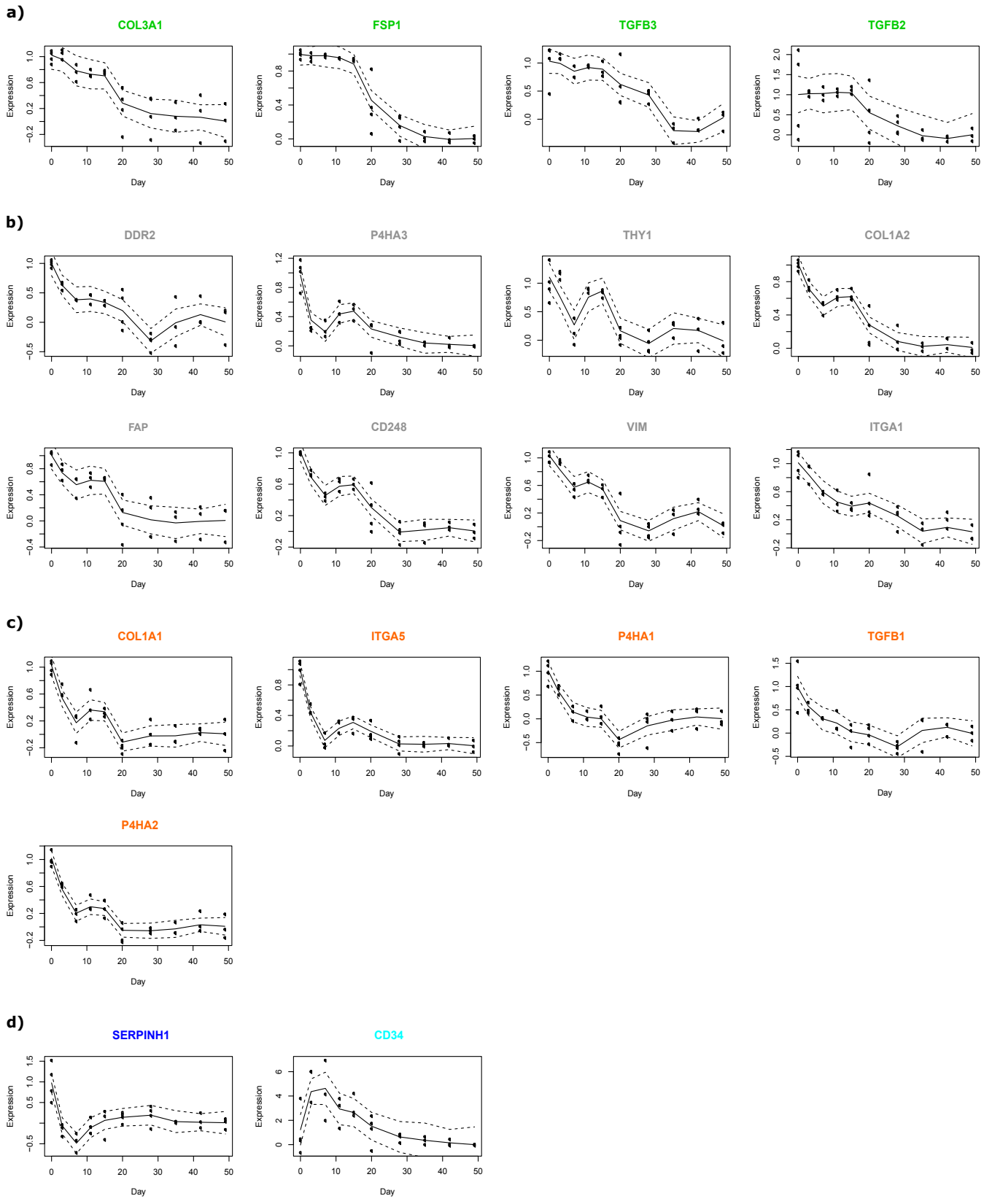

Figure S5. Expression of key fibroblast somatic markers in a 49-day HDF reprogramming time-course. Individual expression dynamics of Cluster 1 fibroblast genes **(a)**, Cluster 2 **(b)** and Cluster 3 genes **(c)**. Values are LOG2 transformed and normalised between 1 and 0 for 'day 0' and 'day 49', respectively, based on the average values between biological replicates for each time point. Dotted line marks CI. Gene label colours correspond to cluster colours in Fig. 1B. SERPINH1 and CD34 expression could not be fit in any of the above clusters and are presented separately **(d)**.

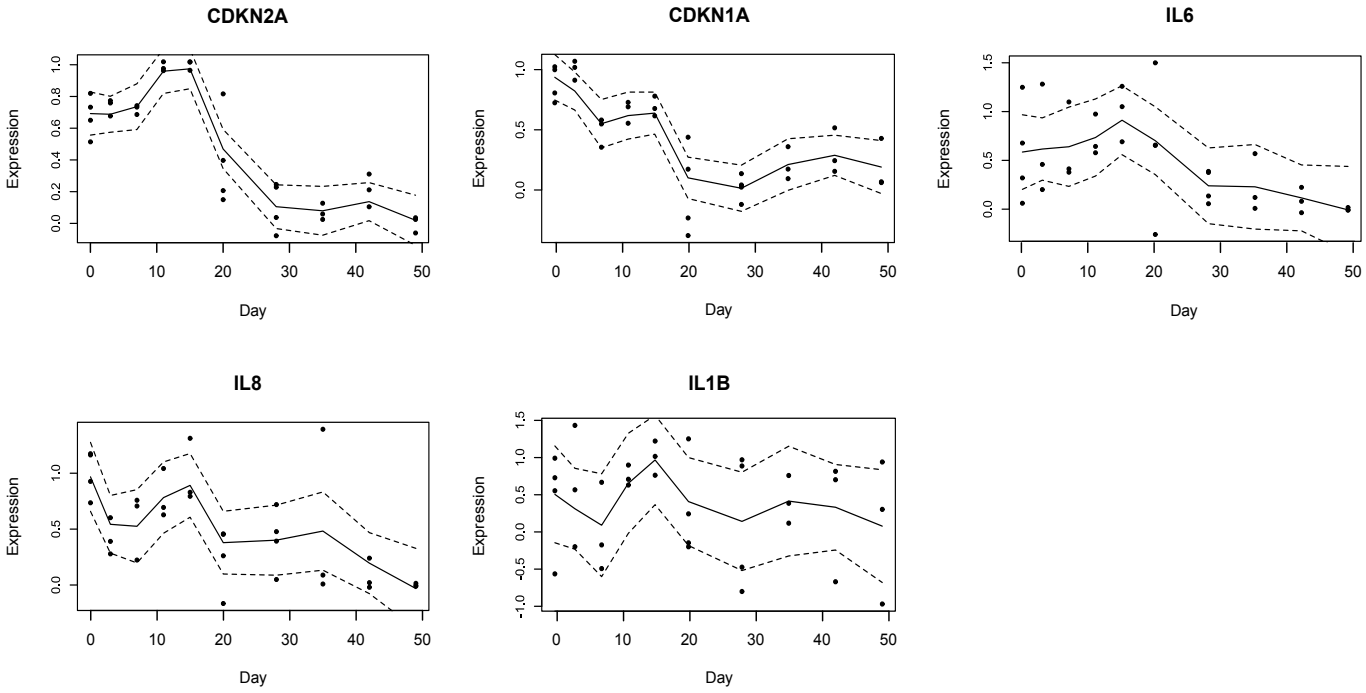

Figure S6. Expression trajectories of key senescence markers in a 49-day HDF reprogramming time-course. Individual expression trajectories of 5 senescence associated genes. Values are LOG2 transformed and normalised between 0 and 1 for the ‘minimum’ and ‘maximum’ value, respectively, based on the average values between biological replicates for each time point. Dotted line marks CI.