Rackham, O. J. L., Firas, J., Fang, H., Oates, M. E., Holmes, M. L., Knaupp, A. S., the FANTOM Consortium, Suzuki, H., Nefzger, C. M., Daub, C. O., Shin, J. W., Petretto, E., Forrest, A. R. R., Hayashizaki, Y., Polo, J. M., & Gough, J. (2016). A predictive computational framework for direct reprogramming between human cell types. *Nature Genetics*, *48*(3), 331-335. https://doi.org/10.1038/ng.3487

A predictive computational framework for direct reprogramming between human cell types

Authors: Owen J.L. Rackham[1,2,†,*], Jaber Firas[3,4,9,†], Hai Fang[1], Matt E. Oates[1], Melissa Holmes[3,4,9], Anja Knaupp[2,3,9], the FANTOM consortium, Harukazu Suzuki[5,6], Christian Nefzger[3,4,9], Carsten O. Daub[5,6,7], Jay W. Shin[5,6], Enrico Petretto[2], Alistair R.R. Forrest[5,6,8], Yoshihide Hayashizaki[6], Jose M. Polo[3,4,9,*] and Julian Gough[1,*]

Affiliations:
[1]Department of Computer Science, University of Bristol, Bristol, UK
[2]Duke-NUS Graduate Medical School, Centre for Computational Biology, 8 College Road, Singapore
[3]Department of Anatomy and Developmental Biology, Monash University, Clayton, Victoria 3800, Australia
[4]Australian Regenerative Medicine Institute, Monash University, Clayton, Victoria 3800, Australia
[5]RIKEN Omics Science Center, Yokohama, Kanagawa, 230-0045 Japan
[6]RIKEN Center for Life Science Technologies, Division of Genomic Technologies, Yokohama, Kanagawa, 230-0045 Japan
[7]Department of Biosciences and Nutrition, Karolinska Institutet, NOVUM S-14183 Huddinge, Stockholm, Sweden
[8]Harry Perkins Institute of Medical Research, QEII Medical Centre and Centre for Medical Research, the University of Western Australia, Nedlands, Western Australia, Australia.
[9]Development and Stem Cells Program, Monash Biomedicine Discovery Institute

§RIKEN Omics Science Center ceased to exist as of April 1st, 2013, due to RIKEN reorganization

† These authors contributed equally
* Correspondence to: owen.rackham@duke-nus.edu.sg; jose.polo@monash.edu and julian.gough@bristol.ac.uk

Introduction

Trans-differentiation, the process of converting from one cell type to another without going through a pluripotent state, has great promise for regenerative medicine. However, the identification of key transcription factors to directly reprogram the identity of cell types is currently limited by the cost of exhaustive experimental testing of plausible sets of factors, an approach that is inefficient and unscalable. Here we present a predictive computational framework (Mogrify) that combines gene expression data with regulatory network information to predict the reprogramming factors necessary to induce cell conversion. We have applied Mogrify to 173 human cell types and 134 tissues, defining an atlas of cellular reprogramming. Mogrify correctly predicts transcription factors used in known trans-differentiations. Furthermore we validated two novel trans-differentiations predicted by Mogrify. We provide a practical and efficient mechanism for systematically implementing novel cell conversions, facilitating the generalization of the reprogramming of human cells. All predictions are made openly accessible to the community to help rapidly further the field of cell conversion (www.mogrify.net).

We now know that it is possible to switch the phenotype of one somatic cell type to another. This epigenetic re-wiring process can be artificially managed and even reversed with the use of transcription factors (TFs)[1]. The best-known example is the reprogramming of somatic cells into induced pluripotent stem (iPS) cells by the introduction of four exogenous factors (Oct3/4, Sox2, c-Myc and Klf4)[2,3]. Previous and subsequent reports have demonstrated that other cell types can also be obtained by direct trans-differentiation using the same strategy[4–9]. These discoveries came about through a process of exhaustive testing of large sets of TFs selected with expert knowledge. With roughly 2000 different TFs[10–12] and approximately 400 unique cell types in humans[13], the space of possible sets is very large ($> 10^{11}$ combinations of 3 factors across 400 cell types) and discovery will advance slowly using an educated trial and error approach. There are a number of existing algorithms that identify TFs that might assist in cell-to-cell conversions considering both epigenetic factors and transcription factor activation[14–17]. More recently approaches such as CellNet[18,19] and a new entropy based method[20] have provided larger scale predictions for transcription factors to convert into many cell types, as well as showing experimentally that these predictions are able to control cell identity. Herein, we present a comprehensive atlas of predictions for a large number of cell conversions, which we have implemented using a network-based computational framework (Mogrify) applied to the FANTOM5 datasets[21] which includes ~300 different cell and tissue types. We have shown that we are able to independently recover (via prediction), the human conversion factors that were previously discovered experimentally and more importantly we have predicted and validated two new conversions.

In order to predict the sets of TFs required for each cell conversion we identify those TFs that are not only differentially expressed between cell types, but also exert regulatory influence on other differentially expressed genes in the local network (see Figure 1a). A single score that captures the differential expression for every gene in every cell type is defined by combining the log-fold change and adjusted *p*-value. The regulatory influence of each TF in each cell type is calculated by performing a weighted sum of the differential expression scores over the known interactome (as defined by STRING and MARA, see Figure 1c). This sum is weighted by two factors: (1) by the directness of the regulation, i.e. how many intermediates between the TF and a downstream

gene, and (2) the specificity, i.e. the number of other genes the upstream TF also regulates. This weighted sum allows TFs to be ranked in each cell type according to their influence. The final step is to select the optimal set of TFs with the greatest combined influence over genes differentially expressed in the target cell type compared to the donor. This is done by adding TFs to the set in order of rank by differential influence, omitting those which don't increase the influence of the set, until the combined influence reaches 98% of expressed target cell genes (see Figure 1d and online materials and methods). Biologically speaking, Mogrify identifies TFs which control the parts of the regulatory network most responsible for the identity of the target cell type.

In order to assess the predictive power of Mogrify we first determined how Mogrify performs against well-known, previously published direct cell conversions, focusing on those involving human cells. These should not be considered as absolute perfect combinations, but as positive example reference points useful for comparison. As shown in Figure 2, in almost every case Mogrify predicts the complete set of TFs previously demonstrated to work, but sometimes includes an upstream TF in lieu of the published factor. For example, it is known that human fibroblasts can be converted to iPS cells by introducing *OCT4 (*also known as *POU5F1), SOX2, KLF4* and *MYC*[3] or *OCT4, SOX2, NANOG* and *LIN28*[22]. Mogrify predicts *NANOG, OCT4* and *SOX2* as the top 3 TFs for this conversion, a combination that has also been experimentally validated[23]. Seminal work by the Graf lab demonstrated that the conversion of B-cells and fibroblasts into macrophage-like cells was possible by the expression of CEBPa and PU.1 (also known as SPI1)[24,25] which Mogrify perfectly predicts. For the conversion of human dermal fibroblasts into cardiomyocytes, we chose to not use the data in the FANTOM5 set since it lacks many key cardiomyocyte genes (indicating a deficiency in the origin of the sample). Nevertheless using the heart sample, which is a cellularly heterogeneous tissue and not ideal, Mogrify's predicted list includes four out of the five TFs (or a closely related factor) used in the human conversion[26]. There are a number of reports in the literature of transdifferentiations from various cell types to neurons in both mouse and human (Table S1). The sets of TFs used vary, probably due to the heterogeneity and complexity of neurons, however factors common to all experiments[27] are predicted by Mogrify (Table S2). Finally between human fibroblasts and hepatocytes, Mogrify predicts a combination of TFs highly similar to that required for conversion and maturation (Figure 2)[6,28,29]. Using the conversions shown in Figure 2 we assessed the ability of Mogrify, CellNet and the entropy-based approach from D'Alessio *et al*[20] to recover these known factors. The average recovery rate of the published transcription factors for Mogrify was 84%, for CellNet 31% and D'Alessio *et al* 51% (see supplementary Figure S2 for details). In six out of the ten conversions in Figure 2 Mogrify recovered 100% of the required TFs, meaning that if Mogrify had been used to provide the TF set for these conversions, the experiment could have been a success first time. On the other hand CellNet and D'Allesio *et al* only recovered all factors for one of the ten conversions. It is important to note that the conversions proposed by CellNet and D'Allesio *et al* may also work, as the published conversions represent only one positive example of success.

In order to empirically demonstrate the predictive capabilities of Mogrify we conducted two novel cell conversions using human cells. **1) Human Fibroblast to Keratinocyte (iKer) conversion:** For this conversion, cells were transduced with FOXQ1*, SOX9, MAFB, CDH1, FOS* and *REL*, predicted by Mogrify (Figure 3A and supplementary Table S3). By day 16 post-

transduction, keratinocyte-associated markers keratin1, keratin14 and involucrin, were markedly up-regulated in the transdifferentiated cells (Figure 3C). Moreover, within three weeks, the majority of transduced cells exhibited cobblestone morphology, a classic characteristic displayed by keratinocytes. Adjacent un-transduced GFP negative cells or control cells transduced with GFP-only viruses maintained their fibroblastic morphology (arrow in Figure 3D). This morphological and molecular characterization of the reprogrammed cells indicates that Mogrify successfully predicts the TFs necessary to induce the conversion from human fibroblasts to keratinocyte-like cells. **2) Adult Human Keratinocyte (HEKa) to Microvascular Endothelial cells (iECs)**: For this conversion we selected *SOX17, TAL1, SMAD1, IRF1* and *TCF7L1* to be used from the seven TFs suggested by Mogrify (Figure 3E and supplementary Table S4). These five TFs are predicted to regulate ~92% of the required genes for iECs. Once these TFs were over-expressed in the HEKa cells we determined that the cells needed to be kept in their media until day four (Figure 3F). We used FACS to follow the kinetics of the cell reprogramming, using the well-established endothelial marker CD31 (Figure 3G), and by day 14 after transduction we detected that more than 2% of the infected cells had up-regulated CD31 and by day 18 almost 10% had up-regulated CD31. At that point we isolated those CD31 cells and evaluated the expression of the endothelial-associated genes (*CD31*, VE-Cadherin, and *VEGFR2*) by qPCR which resulted in a clear reactivation of all the assessed genes (Figure 3H). Finally, we performed immunofluorescence (IF) to verify the morphology and expression of the trans-differentiated cells. As shown in Figure 3I, only the cells transduced with the predicted TFs -and not the control cells- presented the right morphology and expressed CD31 and VE-Cadherin on the surface. This morphology and molecular characterization of the reprogrammed cells indicates the successful transition of human keratinocytes into human endothelial-like cells.

There have been several reports supporting the idea that the Yamanaka factors can be used to initiate transdifferentiation without traversing the pluripotent state (reviewed in Firas Et al 2015[1]). This has been recently challenged by the groups of Hochedlinger and Hanna[30,31]. We observe that Mogrify did not predict the use of Yamanaka factors for the transdifferentiations mentioned in this paper (except to iPS cells). Mogrify prediction however is based on the source and target regulatory networks and does not have the capacity to detect factors only transiently expressed during the reprogramming process.

Since Conrad Waddington depicted the first epigenetic landscape, several attempts have been made to produce a more representative cellular landscape[32–34] but have focused on one or two cell types and are based on path-integral quasi-potentials, mechanistic modeling or probability landscapes. We hypothesised that comparing all-against-all TF network differences as determined by Mogrify in combination with the transcriptional profiles would allow the creation of a 3D landscape representing human cell type (Figures S3 and S4). The landscape places those cell types that are molecularly similar close together in the *x*-*y* plane, and adjusts the height (*z* direction) according to how likely a cell type is to be a good starting cell source (see online materials and methods for details). Interestingly, we observe that different stem cells are placed in the highest locations. This may suggest that the transcriptional networks of those cells at the highest points in the landscape are controlled by fewer TFs, and that the more differentiated the cell becomes (in the valleys) the more TFs are needed to fine tune the transcriptional network.

Having mapped the landscape of human cell type in terms of naturally-occurring states and the transitions between them, we note that a core control set of TFs that describe the individual cell

types is captured, even though the primary aim of Mogrify is to predict TFs for cellular conversions. We believe that this *per se* could aid researchers to unveil the role of different TFs in their favourite cell type. In practice Mogrify provides a significant advance over the strategies currently being applied in laboratories for cell reprogramming, helping in the prediction of TFs whose over-expression will induce directed cell conversion. Mogrify has been pre-calculated on conversions between all possible combinations of the 307 FANTOM5 tissue/cell types resulting in 93,942 directed conversions and provided online (www.mogrify.net) via an interface for guiding experimentation and exploring the cellular landscape. Whilst it is likely that some trial and error will still be involved for some conversions, Mogrify provides a starting point and systematic means to explore new conversions in human. Because Mogrify incorporates a TF redundancy step, it is able to give a finite set of TFs as a prediction for the cell conversion, which is of more utility than just the ranking of all TFs. Although Mogrify has taken advantage of the rich FANTOM5 data, MARA and STRING, these databases have their own limitations which will impart some restrictions on Mogrify's predictions. For example, the FANTOM5 is in some instances limited to few replicates and there is a possible heterogeneity of some samples. MARA relies on known DNA binding motifs to estimate binding to target genes, knowledge that is incomplete. STRING is incomplete in other ways but future rises in the abundance of empirical data on TF interactions and binding in diverse cell types will help to improve Mogrify. It should be noted that Mogrify as well as other methods finds positive regulators of the target cell and does not interrogate the extinction of the source cell signature. This may result in less faithful conversions, mitigated by the downregulation of source genes that has been observed after introduction of core target TFs (e.g. in Polo el al, 2012[35]). Mogrify predictions will not always guarantee conversion, but will certainly aid transdifferentiation protocols. Other players like non-coding RNAs, small molecules, epigenetic factors and signaling pathways provide a rich source of improvements for the future. At present the major challenge to progress in the field of reprogramming is in increasing the number of successful cell conversions. That is what this resource makes possible, paving the way for the routine manipulation of cells, an understanding of the processes involved, and the immediate translation of any breakthroughs in the clinical delivery techniques under heavy development in academia and in industry.

Acknowledgements

JG and OR were supported by grants from the Biotechnology and Biological Sciences research council and the Japan Society for the Promotion of Science. JMP was supported by a Silvia and Charles Senior Medical Viertel Fellowship, the Metcalf award from the National Stem Cell Foundation of Australia, an NHMRC project grant APP1085302 and the Australia Research Council's special initiative Stem Cells Australia. FANTOM5 was made possible by a Research Grant for RIKEN Omics Science Center from MEXT to Yoshihide Hayashizaki and a grant of the Innovative Cell Biology by Innovative Technology (Cell Innovation Program) from the MEXT, Japan to Y.H.. We would like to thank all members of the FANTOM5 consortium for contributing to generation of samples and analysis of the dataset and thank GeNAS for data production.

Author Contributions:

O.J.L.R. and J.G. initiated the project on the basis of discussions with Y.H. about FANTOM5. J.M.P. led the experimental contribution and helped further develop the Mogrify algorithm. J.F. performed all the experimental validations with the contribution of M.H., A.K and C.N. O.J.L.R. performed the data

analysis and interpretation, with significant input by J.G. in the early stages of the work. O.L.J.R., J.M.P. and J.G. prepared the manuscript with input from all named authors at various stages. M.E.O and H.F. provided help and advice with technical aspects of the implementation. H.S. and J.W.S. were involved in early discussion of cell conversion concepts. A.R.R.F. and C.O.D were involved in the FANTOM5 management.

Online Methods

References:

1. Firas, J., Liu, X., Lim, S. M. & Polo, J. M. Transcription factor-mediated reprogramming: epigenetics and therapeutic potential. *Immunol. Cell Biol.* **93,** 284–289 (2015).

2. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126,** 663–76 (2006).

3. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131,** 861–72 (2007).

4. Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463,** 1035–41 (2010).

5. Ieda, M. *et al.* Direct Reprogramming of Fibroblasts into Functional Cardiomyocytes by Defined Factors. *Cell* **142,** 375–386 (2010).

6. Du, Y. *et al.* Human hepatocytes with drug metabolic function induced from fibroblasts by lineage reprogramming. *Cell Stem Cell* **14,** 394–403 (2014).

7. Sekiya, S. & Suzuki, A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature* **475,** 390–3 (2011).

8. Pfisterer, U. *et al.* Direct conversion of human fibroblasts to dopaminergic neurons. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 10343–8 (2011).

9. Addis, R. C. *et al.* Optimization of direct fibroblast reprogramming to cardiomyocytes using calcium activity as a functional measure of success. *J. Mol. Cell. Cardiol.* **60,** 97–106 (2013).

10. Wilson, D., Charoensawan, V., Kummerfeld, S. K. & Teichmann, S. A. DBD--taxonomically broad transcription factor predictions: new content and functionality. *Nucleic Acids Res.* **36,** D88–92 (2008).

11. Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. a & Luscombe, N. M. A census of human transcription factors: function, expression and evolution. *Nat. Rev. Genet.* **10,** 252–63 (2009).

12. Fulton, D. L. *et al.* TFCat: the curated catalog of mouse and human transcription factors. *Genome Biol.* **10,** R29 (2009).

13. Vickaryous, M. K. & Hall, B. K. Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. *Biol. Rev. Camb. Philos. Soc.* **81,** 425–55 (2006).

14. Heinäniemi, M. *et al.* Gene-pair expression signatures reveal lineage control. *Nat. Methods* **10,** 577–83 (2013).

15. Lang, A. H., Li, H., Collins, J. J. & Mehta, P. Epigenetic landscapes explain partially reprogrammed cells and identify key reprogramming genes. *PLoS Comput. Biol.* **10,** e1003734 (2014).

16. Del Sol, I. C. A. A general strategy for cellular reprogramming: The importance of transcription factor cross-repression. *Stem Cells* **31,** 2127–2135 (2013).

17. Davis, F. P. & Eddy, S. R. Transcription Factors That Convert Adult Cell Identity Are Differentially Polycomb Repressed. *PLoS One* **8,** 1–8 (2013).

18. Morris, S. A. *et al.* Dissecting Engineered Cell Types and Enhancing Cell Fate Conversion via CellNet. *Cell* **158,** 889–902 (2014).

19. Cahan, P. *et al.* CellNet: Network Biology Applied to Stem Cell Engineering. *Cell* **158,** 903–915 (2014).

20. D'Alessio, A. C. *et al.* A Systematic Approach to Identify Candidate Transcription Factors that Control Cell Identity. *Stem Cell Reports* (2015). doi:10.1016/j.stemcr.2015.09.016

21. Forrest, A. R. R. *et al.* A promoter-level mammalian expression atlas. *Nature* **507,** 462–470 (2014).

22. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318,** 1917–20 (2007).

23. Huangfu, D. *et al.* Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat. Biotechnol.* **26,** 1269–1275 (2008).

24. Xie, H., Ye, M., Feng, R. & Graf, T. Stepwise Reprogramming of B Cells into Macrophages. *Cell* **117,** 663–676 (2004).

25. Rapino, F. *et al.* C/EBPα induces highly efficient macrophage transdifferentiation of B lymphoma and leukemia cell lines and impairs their tumorigenicity. *Cell Rep.* **3,** 1153–63 (2013).

26. Fu, J.-D. *et al.* Direct reprogramming of human fibroblasts toward a cardiomyocyte-like state. *Stem cell reports* **1,** 235–47 (2013).

27. Zou, Q. *et al.* Direct conversion of human fibroblasts into neuronal restricted progenitors. *J. Biol. Chem.* **289,** 5250–60 (2014).

28. Huang, P. *et al.* Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature* **advance on,** (2011).

29. Kogiso, T., Nagahara, H., Otsuka, M., Shiratori, K. & Dowdy, S. F. Transdifferentiation of human fibroblasts into hepatocyte-like cells by defined transcriptional factors. *Hepatol. Int.* **7,** 937–944 (2013).

30. Bar-Nur, O. *et al.* Lineage conversion induced by pluripotency factors involves transient passage through an iPSC stage. *Nat. Biotechnol.* **33,** 1–11 (2015).

31. Maza, I. *et al.* Transient acquisition of pluripotency during somatic cell transdifferentiation with iPSC reprogramming factors. *Nat. Biotechnol.* **33,** 769–774 (2015).

32. Qiu, X., Ding, S. & Shi, T. From understanding the development landscape of the canonical fate-switch pair to constructing a dynamic landscape for two-step neural differentiation. *PLoS One* **7,** e49271 (2012).

33. Bhattacharya, S., Zhang, Q. & Andersen, M. E. A deterministic map of Waddington's epigenetic landscape for cell fate specification. *BMC Syst. Biol.* **5,** 85 (2011).

34. Flöttmann, M., Scharp, T. & Klipp, E. A stochastic model of epigenetic dynamics in somatic cell reprogramming. *Front. Physiol.* **3,** 216 (2012).

35. Polo, J. M. *et al.* A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* **151,** 1617–1632 (2012).

36. Anders, S. & Huber, W. Differential expression analysis for sequence count data. (2010). doi:10.1038/npre.2010.4282.2

37. Lattanzi, L. *et al.* High efficiency myogenic conversion of human fibroblasts by adenoviral vector-mediated MyoD gene transfer. An alternative strategy for ex vivo gene therapy of primary myopathies. *J. Clin. Invest.* **101,** 2119–28 (1998).

Figures

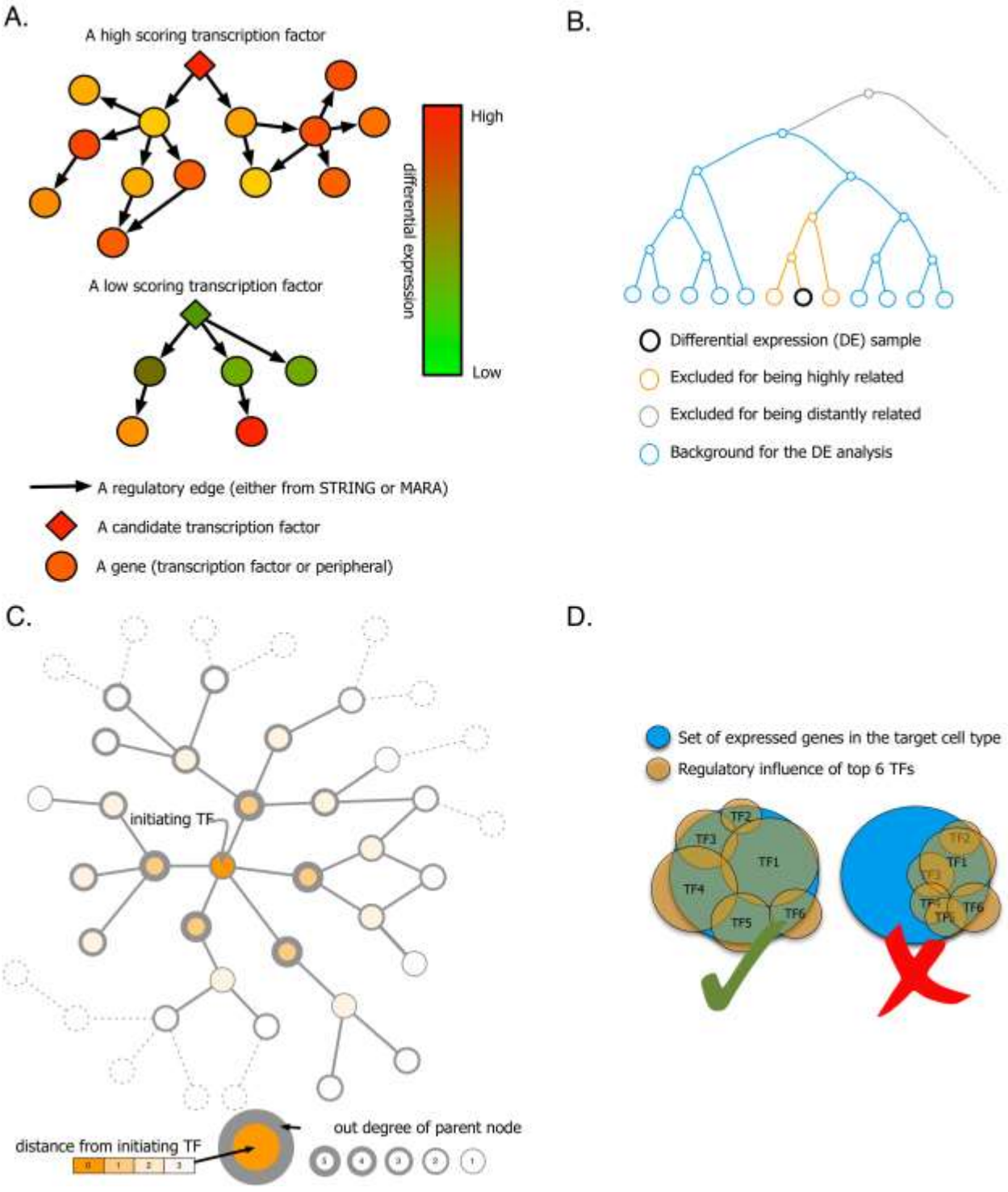

Figure 1. The Mogrify algorithm for predicting TFs for cell conversion. This is done as follows: (**A**) Mogrify aims to find those TFs that not only are differentially expressed but appear

to be responsible for the regulation of many differentially expressed genes in a given cell type. (**B**) We use the cell type ontology tree created as part of the FANTOM5 consortium [21] to select an appropriate background for DESeq [36] to calculate the adjusted *p*-value and log fold change for genes in the sample. (**C**) For each TF we construct a local network neighborhood of influence weighting the downstream effect on a gene by its connected distance and the out-degree of its parent. (**D**) We maximise regulatory coverage by removing TFs which are redundant in their influence over other factors.

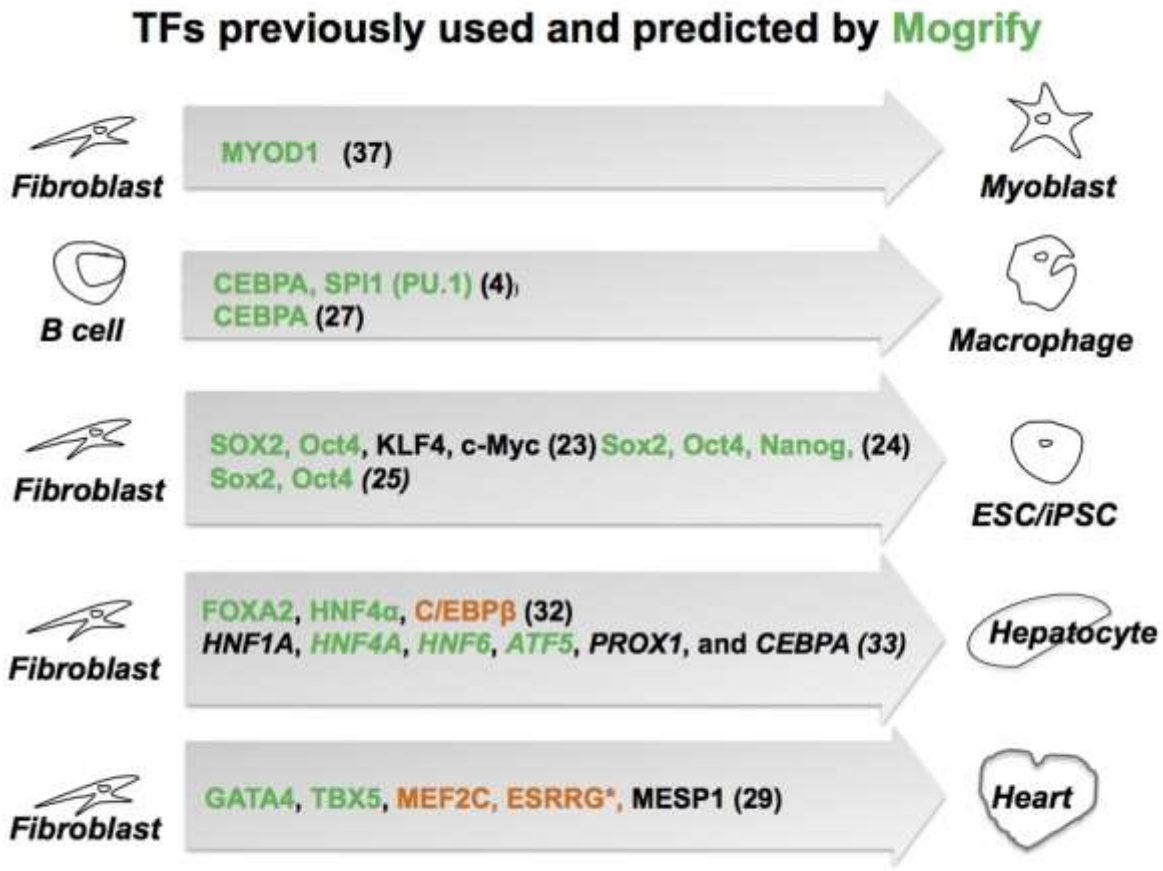

Figure 2. Mogrify predictions for some of the known trans-differentiations that are published in the literature. TFs that Mogrify correctly identifies from the published list are highlighted. Samples are grouped using the FANTOM cell ontology [21]. For each publication the transcription factors that are in the initial maximum coverage set are shown in green and in the overall predicted Mogrify set in orange. For instance the transdifferentiation between fibroblast and myoblast [37] required only MYOD and this was identified by Mogrify.

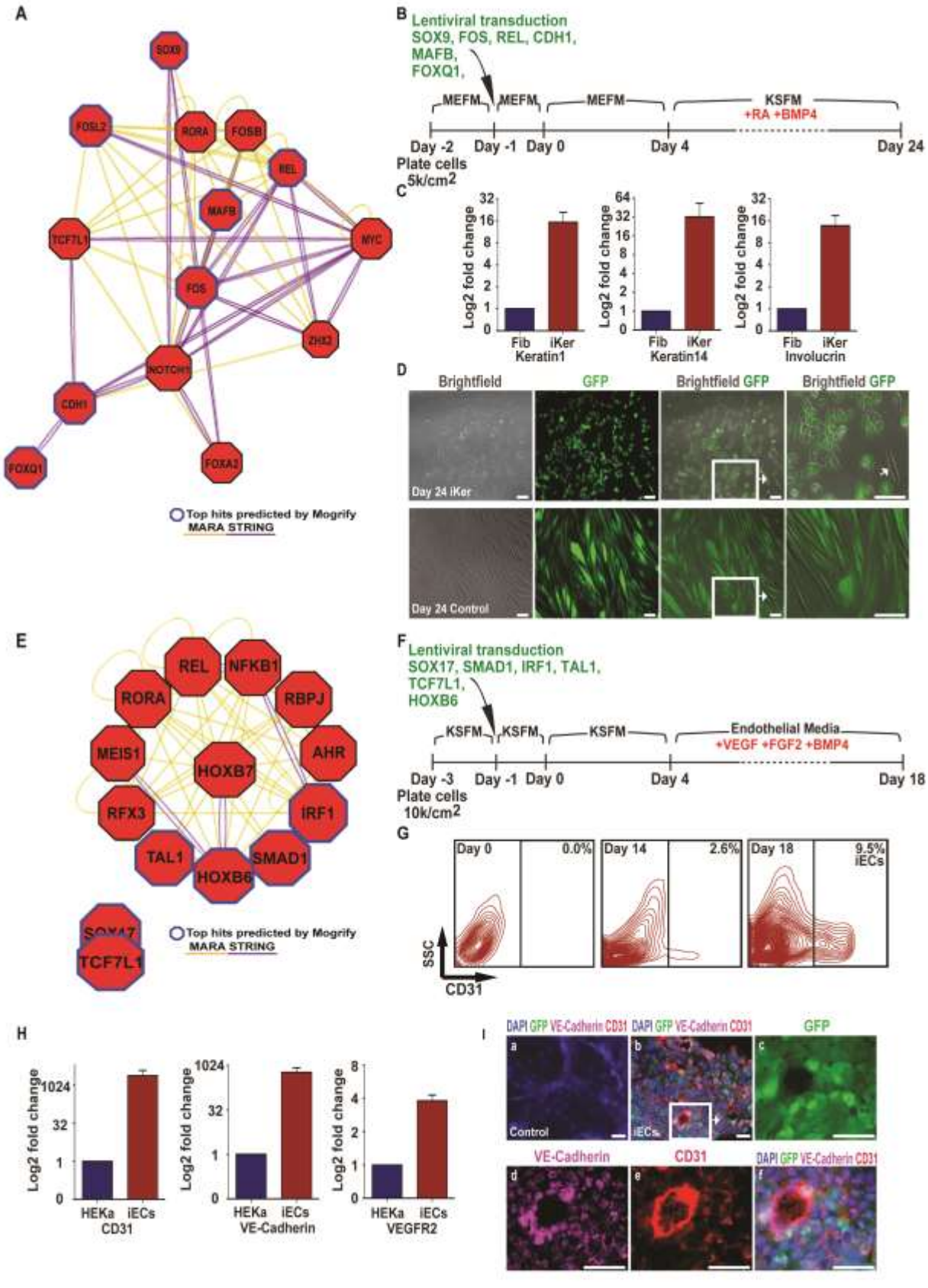
A
SOX9
FOSL2
RORA
FOSB
REL
MAFB
TCF7L1
MYC
FOS
ZHX3
NOTCH1
CDH1
FOXQ1
FOXA2
Top hits predicted by Mogrify
MARA STRING
B
Lentiviral transduction
SOX9, FOS, REL, CDH1,
MAFB,
FOXQ1,
MEFM
MEFM
MEFM
KSFM
+RA +BMP4
Day -2
Day -1
Day 0
Day 4
Day 24
Plate cells
5k/cm2
C
Log2 fold change
Fib
iKer
Keratin1
Keratin14
Involucrin
D
Brightfield
GFP
Brightfield GFP
Day 24 iKer
Day 24 Control
E
REL
NFKB1
RORA
RBPJ
MEIS1
AHR
HOXB7
RFX3
IRF1
TAL1
SMAD1
HOXB6
SOX17
TCF7L1
Top hits predicted by Mogrify
MARA STRING
F
Lentiviral transduction
SOX17, SMAD1, IRF1, TAL1,
TCF7L1,
HOXB6
KSFM
KSFM
KSFM
Endothelial Media
+VEGF +FGF2 +BMP4
Day -3
Day -1
Day 0
Day 4
Day 18
Plate cells
10k/cm2
G
Day 0
0.0%
Day 14
2.6%
Day 18
9.5%
iECs
SSC
CD31
H
Log2 fold change
HEKa
iECs
CD31
VE-Cadherin
VEGFR2
I
DAPI GFP VE-Cadherin CD31
GFP
Control
iECs
VE-Cadherin
CD31

Figure 3. Empirical validation of novel conversions predicted by Mogrify

(**A**) The transcription factor network predicted by Mogrify to be involved in the dermal fibroblast to keratinocyte transdifferentiation. (**B**) An outline of the method used for the transdifferentiation assay. (**C**) qPCR analysis of the indicated markers in cells harvested at days 12-16 during transdifferentiation. All values are experimental replicates and are relative to gene expression in dermal fibroblasts (*n*=3). (**D**) Brightfield and GFP images at day 24 showing the cobblestone morphology of transdifferentiated cells (upper panel) and GFP+ control cells (lower panel). (**E**) A schematic representation of the transcription factor network predicted by Mogrify to be involved in the keratinocyte to microvascular endothelial cell transdifferentiation. (**F**) An outline of the method used for the transdifferentiation assay. (**G**) Flow cytometric analysis of CD31 expression at day 0, 14 and 18 of transdifferentiation. (**H**) qPCR analysis of the indicated expression markers in CD31+ cells harvested at day 18 of transdifferentiation. All values are experimental replicates and are relative to gene expression in keratinocytes (*n*=3). (**I**) Immunofluorescence analysis of endothelial markers CD31 and VE-Cadherin at day 18 for vector free control cells (**a**) and transdifferentiating cells (**b-f**). Scale bar = 50μm.