Trends Open

CellPress OPEN ACCESS

Review

DNA methylation in cell state establishment and stability

Areej Khatib[1], Ofra Sabag[1], Yosef Buganim[1,*], and Howard Cedar[1,*]

Our genome includes the genetic text used to generate the protein building blocks and a second layer which interacts with DNA providing information for where and when the text should be read. This second type is mediated by transcription factors, histone modifications, and DNA methylation. While all annotations can influence transcription, the only one that provides cell-type stability is DNA methylation, notable for two characteristics. Unlike other modifications involving labile electrostatic interactions with DNA, methylation entails the covalent addition of a methyl group to cytosine, creating a highly stable epigenetic mark. Second, DNA methylation is unique in its faithful maintenance through cell division. In this review, we highlight DNA methylation as a key regulator of cell fate and identity.

Significance

DNA methylation functions as a molecular memory that enables cells to preserve their identity across divisions. Failures in this system can lead to major developmental abnormalities. Understanding how it works not only reveals a fundamental principle of biology but also opens the door to more precise and stable cell reprogramming, an essential step toward effective regenerative medicine and future disease treatments.

DNA methylation provides long-term cell-type memory

Early experiments in tissue culture have demonstrated that the mere presence of methyl groups on DNA has a strong effect on chromatin structure, making it resistant to DNase I digestion in a manner that is completely independent of sequence content [1]. Similarly, it was shown *in vivo* that DNA methylation operates by limiting local accessibility to factors in the surrounding nucleoplasm by means of different molecular mechanisms, including both histone modification and higher-order structural changes [2–4].

This important concept was confirmed during normal development and tissue adaptation, in which it was shown by Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAC-seq) that genomic **regulatory elements (REs;** see Glossary) are in an inaccessible conformation in most cell types but become accessible following **DNA demethylation**. When demethylation is prevented by **Ten-eleven translocation (TET) enzyme** depletion, these sites remain in a closed conformation, and proteins are no longer able to bind to their specific motifs within these elements [5–7]. In line with this, deletion of the DNA methylation machinery hyperactivates repetitive retrotransposons, initiates transcription from cryptic **promoters** within transposons and gene bodies, and creates ectopic protein-binding sites [8–11].

Notably, accumulating evidence suggests that viewing DNA methylation as a simple repressive mark is an oversimplification. For example, although DNA methylation restricts the binding of many protein factors, specific groups of **transcription factors (TFs)** are either insensitive to DNA methylation or even preferentially bind methylated CpG sites [12,13]. These include the well-studied methyl-CpG-binding domain proteins as well as several other factors. Furthermore, studies have suggested that methylated REs can gain accessibility through *in vitro* cell fate transitions even before the erasure of the methyl group [14]. Taken together, these results suggest that while DNA methylation lies at the heart of gene regulatory networks, its biological function remains controversial.

Here, we review accumulating evidence supporting a major role for DNA methylation in stabilizing transcriptional programs across development, adaptation, and cell fate transitions. This entire

[1]Department of Developmental Biology and Cancer Research, The Institute for Medical Research Israel-Canada, the Hebrew University-Hadassah Medical School, Jerusalem 91120, Israel

*Correspondence:
yossib@ekmd.huji.ac.il (Y. Buganim) and cedar@mail.huji.ac.il (H. Cedar).

system is based on the use of gene REs composed of **enhancers** [15,16] and **silencers** [17–19]. The methylation state of these sequences is programmed by information in the genetic text [20] that provides motifs capable of recruiting the *de novo*- or demethylation machinery [21], as well as a system of cell- and differentiation-specific trans-acting TFs, as we will further describe in this work. Knowledge of this system will greatly enhance our ability to design and produce cells for cell-replacement therapy.

General principles of gene regulation

Simple organisms control gene expression through a system of repressors and activators that can interact with gene promoters and affect the degree of expression, mainly by influencing the initiation of transcription. In higher organisms, regulation is much more complex, with every tissue having many genes that are inactive and some that are active, with a different set in each tissue. This requires a new mode of control, since the earlier mode would require too many TFs. This is accomplished through a large network of REs [22].

In this sophisticated system, the promoter is activated by basic TFs but can then be modulated through an array of distal elements that are able to interact in 3D space to affect transcription [23–25], either through activation or repression. Each of these elements is essentially nondescript, having multiple motifs for binding factors, most of which are not necessarily cell-type specific, but rather are expressed ubiquitously and, as a result, can potentially affect gene expression in every cell of the body. Furthermore, the fact that each element binds multiple TFs increases the local TF concentration over the promoter, thus chemically amplifying its effect on transcription [26,27].

DNA methylation as a marker of cell identity

Although these REs are part of the DNA text itself and therefore present in all cells of the body, it appears that they actually bind the trans-acting TFs to their motifs in a tissue- and developmental-specific manner. Thus, for example, the REs surrounding the albumin gene can only be engaged in the liver and not in any other tissue. In contrast, the elements surrounding the globin genes are only open for binding in blood cells but not in the liver or any other cell type. Network-level analysis strongly suggests that this cell-type specificity is controlled mainly by DNA methylation programming [28–33]. This has been determined by bioinformatically comparing tissue-specific demethylated tiles and TF binding data (Encyclopedia of DNA Elements (ENCODE)) in many different cell types in mice. For any given TF, binding is observed only within tissue-specific undermethylated sites but not at these exact loci in other tissues, where the DNA is methylated [7]. This picture is also consistent with data suggesting that most known TFs are nonspecific and these play the dominant role in tissue-specific gene expression [32,34]. While some TFs can access and bind methylated DNA, this epigenetic hallmark generally creates a restrictive environment by raising the threshold required for RE activation. Under physiological TF levels, this threshold is rarely reached, limiting inappropriate activation. When stable reactivation is required, cells can increase **pioneer transcription factors** levels to engage methylated elements and recruit DNA demethylation machinery.

Generation of tissue-specific methylation patterns during development

How are tissue-specific methylation patterns formed during development? Almost all REs become methylated from the time of **implantation** and remain at this baseline level through germ-layer formation. Then, in the mid-embryo stage, when lineage differentiation reaches its peak (~E15–16), region-specific morphogens interact with local cells and induce specific TFs that can interact with motifs within the REs [35]. These bound complexes then serve as an anchor to recruit the demethylation machinery, which can actively remove nearby methyl groups from this element. Once demethylated, these regions remain accessible because local concentrations of any *de novo* methylases present in the nucleus are too low to remethylate these sites and because

Glossary

BAM factors: BRN2, ASCL1, and MYT1L are a set of TFs capable of directly converting fibroblasts into induced neuronlike cells.
Blastocyst: an early-stage embryo characterized by the first lineage segregation, resulting in the formation of the outer trophectoderm (TE) and the inner cell mass (ICM).
Chimeric contribution: the ability of transplanted pluripotent stem cells to integrate into a developing embryo and contribute to the formation of its tissues and organs.
4N competence: the ability of pluripotent stem cells to generate an entire embryo following tetraploid (4N) complementation, a stringent in vivo test of pluripotency.
Dedifferentiation: the process by which a cell transitions from a terminally differentiated state to a more progenitorlike or less specialized state.
Differentially methylated regions: genomic regions that exhibit differences in DNA methylation levels between two or more biological states or conditions.
DNA demethylation: the removal of a methyl group from a methylated CpG site, occurring either passively or actively.
DNA methyltransferases: enzymes that catalyze the transfer of a methyl group to cytosine residues, primarily at CpG dinucleotides, including both *de novo* methylation (e.g., DNMT3A and DNMT3B) and maintenance methylation (e.g., DNMT1).
Enhancer: a distal DNA RE that positively regulates transcription of target promoters, often through three-dimensional chromatin interactions.
Epigenetic memory: the heritable transmission of gene expression states without changes in the underlying DNA sequence, enabling cells to retain information about past signals or environmental exposures. In the context of reprogramming and transdifferentiation, it refers to the persistence of epigenetic signatures that remain incompletely reset in reprogrammed cells.
Implantation: the process by which the early embryo attaches to and invades the uterine lining (endometrium).
Induced pluripotent stem cells (iPSCs): embryonic stem cell-like cells that represent the *in vitro* counterpart of the epiblast of the ICM and are generated from somatic cells

there are no preprogrammed sequence instructions for recruiting the *de novo* methylases. It should be noted that while the demethylation itself is highly specific (because of local morphogens), most of the trans-acting factors that bind to the REs are ubiquitously expressed in many tissues. This arrangement facilitates the accumulation of multiple TFs at high concentrations, thus increasing the probability of transcriptional activation when they chemically interact with the promoter.

RE organization

REs provide a great deal of flexibility regarding the control of gene expression. As noted above, these elements can be located either 5′ or 3′ of the gene promoter or even within introns and may be found at large distances from the gene, while still being within the same **topologically associating domain** [25]. It is not uncommon for genes to be associated with multiple different REs, with some being opened during tissue development and others programmed to become demethylated as an adaptive response to circulating factors. Other REs are designed to become inaccessible through the specific recruitment of **DNA methyltransferases**, thus preventing the nearby gene from being reactivated. Finally, it should be noted that this system is completely modular: some specific elements may be located in multiple gene domains while some gene regions may be associated with regulatory sequences covering a wide range of different specificities.

Role of DNA methylation in the dynamics of development

Many studies have demonstrated that the presence of methyl groups on promoters and other REs decreases chromatin accessibility and thus inhibits interactions with TFs [36]. In contrast, there are only a few studies that address the role of DNA methylation in development. One good example is the inactivation of pluripotency genes such as *Oct4* and *Nanog* at the time of implantation (Figure 1). Careful studies both *in vitro* using ESCs and *in vivo* clearly show that this is a multistep process [37,38]. First, a repressor protein binds to the gene's REs, thus halting transcription. In the second step, a histone methylase (G9a) binds and causes heterochromatinization of the gene. Only after this has happened a specific domain on the G9a protein serves to recruit DNA methylases, which then bring about *de novo* methylation of this region. Thus, it is not the addition of DNA methylation that inactivates these genes, but rather specific developmentally regulated factors that recognize the genes and prevent transcription. Nonetheless, it is the continued presence of DNA methylation that prevents these genes from becoming activated again [37,38]. This is in contrast to regulation by protein–DNA interactions (e.g., repressors or heterochromatin), which are only transiently associated with the genes and are usually no longer present at later stages of development.

This same principle also defines the role of DNA demethylation during development. In this case, genes designated for activation in a particular tissue are first recognized and bound by high concentrations of cell-type-specific pioneer factors that cause transcriptional activation and induce chromatin opening (Figure 1 [39,40]), thus making these genes accessible to the DNA demethylation machinery [41,42]. Once these regions have undergone active demethylation, they will be maintained in this state throughout life even though the initial pioneer factors have been diluted out (Figure 1 [43]). These mechanistic dynamics have been confirmed genetically using conditional knockouts of the TET enzymes required for demethylation, thereby demonstrating that, without the removal of methyl groups, these tissue-specific genes quickly revert to a closed configuration [7]. It thus appears that while DNA methylation changes do not actually drive development, it is the final methylation pattern that defines stable long-term cell identity.

Postnatal adaptation

The same principles that govern DNA methylation programming during development also enable a variety of adaptive changes that occur after birth. However, whereas prenatal cell fates are

(e.g., fibroblasts) through the ectopic expression of key pluripotency-associated TFs (e.g., OSKM).

Induced trophoblast stem cells: trophoblast stemlike cells that represent the *in vitro* counterpart of the TE of the blastocyst and are generated from somatic cells (e.g., fibroblasts) through ectopic expression of key trophoblast-associated TFs (e.g., GATA-binding protein 3 (GATA3), Eomesodermin (EOMES), transcription factor AP-2 gamma (TFAP2C), and MYC proto-oncogene, bHLH transcription factor (MYC) [GETM] in the mouse).

Mesenchymal-to-epithelial transition: an early step in cellular reprogramming that involves the loss of mesenchymal characteristics (e.g., those of fibroblasts) and the acquisition of epithelial features.

MyoD: a myogenic master TF that can directly convert fibroblasts into myoblastlike cells.

Pioneer transcription factor: a TF that can bind to and engage target DNA sequences within closed (condensed) chromatin.

Promoter: a cis-regulatory DNA element, typically located proximal to the transcription start site, that recruits and coordinates the transcriptional machinery to initiate gene transcription.

Regulatory element (RE): a cis-regulatory DNA sequence that influences transcription of a target gene, frequently acting distal to the promoter through mechanisms such as chromatin looping.

Silencer: a cis-regulatory DNA element that reduces or inhibits transcription of a target gene, often by recruiting repressive factors or chromatin-modifying complexes.

Somatic cell reprogramming: the process of inducing a differentiated somatic cell to revert to a pluripotent state, typically through the ectopic expression of defined TFs.

Ten-eleven translocation (TET) enzymes: dioxygenases that oxidize 5-methylcytosine to 5-hydroxymethylcytosine and further-oxidized derivatives, thereby facilitating DNA demethylation.

Topologically associating domain: a self-interacting genomic region in which DNA sequences preferentially interact with one another, contributing to the organization of chromatin architecture and the regulation of gene expression.

Transcription factor (TF): a DNA-binding protein that recognizes specific

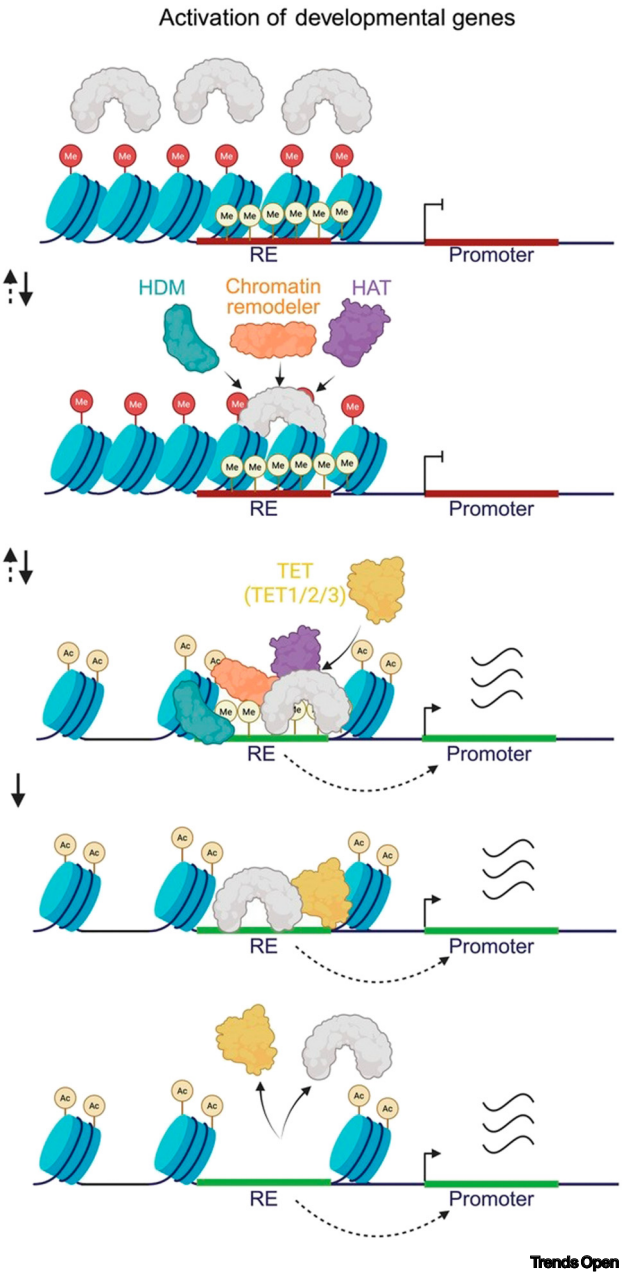

Figure 1. Demethylation and the long-term activation of developmental genes. During development, cells obtain tissue-specific patterns of gene expression. In the case of gene activation, a local morphogen/hormone induces the expression of a specific pioneer factor (brown) that binds inaccessible chromatin and reactivates the RE through the recruitment of chromatin remodelers or modifiers. An active RE then interacts with a promoter sequence and induces transcription. In the last step, demethylation machinery is recruited to maintain the long-term opening of the chromatin regardless of the expression of the pioneer factor. Note that while RE activity is regulated by chromatin structure and DNA methylation, promoters lie within CpG islands and are usually nucleosome-free. Ac: Acetylation; HAT: histone acetyl transferase; HDM: histone demethylase; Me: DNA methylation; RE: regulatory element; TET: ten-eleven translocation enzymes.

sequence motifs and regulates gene transcription, typically by activating or repressing the expression of target genes.

Transdifferentiation: the process by which a differentiated somatic cell is directly reprogrammed into another differentiated cell type without transitioning through an intermediate pluripotent or progenitor state.

Waddington's proposal: a conceptual model in which cell differentiation is represented as a ball rolling down a branching landscape of valleys, symbolizing the progressive restriction of developmental potential and commitment to distinct cell fates.

determined by local morphogens, postnatal decisions are determined by hormones. One example is the selection of sex-specific expression patterns. In this case, both males and females are born with the same expression pattern in a variety of different tissues, but as sexual maturity is attained, males undergo additional changes in DNA methylation [44,45]. These changes are apparently mediated by the secretion of testosterone, which then serves as a blood-borne inducer of specific factors within the target cells that bind to designated motifs and recruit the DNA methylation machinery, resulting in the activation of new REs [44]. These new patterns are then maintained for the life of the organism. Another example of sex-related adaptation has been observed in the mammary gland as a result of pregnancy, and in dendritic cells as a response to infections [46,47].

Other examples of postnatal adaptation include the inherent response of the liver to dietary changes, as occurs with the transition to lipid-rich milk following the onset of nursing immediately after birth. Additional alterations in expression occur 3 weeks later, following weaning. In this later case, the extracellular signal that induces demethylation appears to be dependent on the blood insulin concentration [7]. Furthermore, specific templates are designed to be differentially methylated in response to physiological and environmental changes, covering a variety of events such as chronic inflammation [48], tissue damage [49], or as part of normal homeostasis and aging [50,51].

Evidence suggests that at least part of the environmentally adaptive changes in DNA methylation are made to provide an **epigenetic memory** that profoundly strengthens tissue responses to similar stimuli and enhances tissue fitness. For example, acute inflammation triggers demethylation and the opening of specific DNA regions in epidermal stem cells, changes that are preserved for at least a year postinflammation [52]. The same is true for muscle injury [53].

Somatic cell reprogramming and transdifferentiation

The discovery by Yamanaka and Takahashi that forced expression of four genes, *Oct4*, *Sox2*, *Klf4*, and *Myc* (OSKM), is sufficient to convert somatic cells into pluripotent stem cells [**induced pluripotent stem cells (iPSCs)**; Figure 1] fundamentally transformed our understanding of how new cell states can be achieved and how TFs regulate cell identity [54,55]. Building on this breakthrough, numerous studies have since identified alternative factor combinations capable of generating iPSCs with varying efficiencies and qualities [56–62].

Changes in DNA methylation are indispensable for reprogramming

Interestingly, a central insight emerging from these and related studies [63–66] is that DNA demethylation is indispensable for **somatic cell reprogramming** (Figure 3A). For example, Hu *et al*. demonstrated that TET triple-knockout mouse embryonic fibroblasts fail to reprogram due to an inability to undergo the early **mesenchymal-to-epithelial transition** [64]. Consistent with this, TET1 was shown to functionally substitute for OCT4 [66] or to replace SOX2 and KLF4 when paired with OCT4 [67], emphasizing DNA methylation remodeling as a critical determinant of successful reprogramming. In line with these findings, single-cell tracking studies revealed that the extent of DNA demethylation at key REs accurately predicts which cells will successfully complete reprogramming [68].

However, given that the promoter and enhancer elements of the master pluripotency regulator *Oct4* are normally unmethylated in Embryonic stem cells (ESCs) and throughout development, these observations raise an important question: does the role of TET proteins in reprogramming primarily reflect their DNA demethylation activity, or do TETs also exert methylation-independent functions that contribute to cell-fate conversion?

The ability of the reprogramming factors to reset the methylome is essential for full developmental potential. Proper methylome re-establishment enables reprogrammed cells to follow normal developmental trajectories and differentiate into any lineage, defining their quality and capacity to replace or model endogenous cells. Supporting this view, aberrant imprinting or silencing of the *Dlk1–Dio3* imprinted locus correlates with poor **chimeric contribution** and failure to generate 'all-iPSC' mice [69,70]. Consistently, iPSCs generated using only OCT4 and TET1 exhibit markedly improved **4N competence** compared with conventional OSKM-derived iPSCs [67].

Trophoblast stem cell reprogramming

Interestingly, the same capacity to obtain near-complete epigenetic reprogramming of the DNA methylome was observed during the conversion of human and mouse fibroblasts into

trophoblast stem cells (TSCs) [71–74]. In a manner similar to iPSCs, **induced trophoblast stem cells** represent stable, highly functional, and faithful counterparts to **blastocyst**-derived TSCs. Central to this remarkable fidelity is the capacity of early embryonic pioneer factors to accurately regenerate 95–99% of the endogenous DNA methylation landscape [60,71–75].

Direct reprogramming (transdifferentiation)

Another strategy for generating specific cell identities is **transdifferentiation**, in which one lineage is directly converted into another without passing through a less-differentiated or pluripotent state (Figure 2). Classic examples include the conversion of fibroblasts into myoblastlike cells by enforced expression of ***MyoD*** [76], fibroblasts into neuronlike cells using *Brn2, Ascl1*, and *Myt1l* (the **BAM factors**) [77], fibroblasts into hepatocytelike cells [78], and fibroblasts into cardiomyocytelike cells [79]. Although transdifferentiated cells often resemble their endogenous counterparts in morphology and share significant transcriptional similarities, they generally remain unstable and dependent on continued transgene expression, highlighting that complete epigenetic reprogramming is not achieved during transdifferentiation. A recent study from the Stanger, Cedar, and Buganim laboratories examined DNA methylation across multiple transdifferentiation systems and demonstrated that at least part of this incomplete cell-fate transition is due to defects in recapitulating the original DNA methylation landscape.

Unlike reprogrammed iPSCs, transdifferentiated cells retain substantial epigenetic memory and fail to acquire an appropriate DNA methylation landscape at key tissue-specific regulatory regions (Figure 3B [80]). This finding reinforces earlier work showing that directly converted cells inadequately silence the lineage programs of the starting population [81]. Similar to transdifferentiated somatic cells *in vitro*, attempts to directly convert ESCs into TSC-like cells through *Oct4* downregulation or *Cdx2* overexpression fail to fully recapitulate the methylome of bona fide TSCs, thereby limiting stable expression of key trophoblast-specific genes such as *Elf5* and *Hand1* (Figure 3B [82]). Together, these observations highlight that although morphological and transcriptional changes can be induced to a significant extent without complete methylome

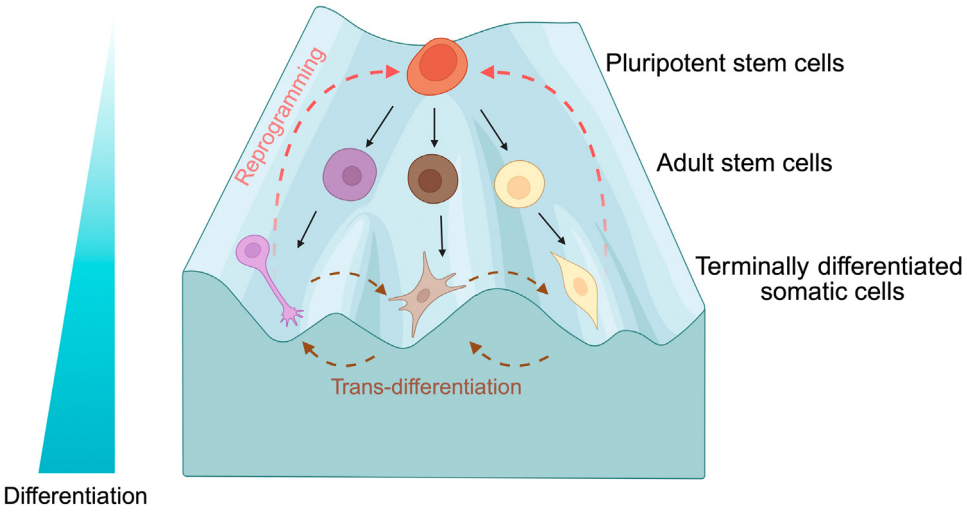

Figure 2. Differentiation, transdifferentiation, and reprogramming on Waddington's proposal. Cells are organized along a differentiation continuum where, in typical developmental trajectories (indicated by black arrows), they transition from a pluripotent stem cell state at the apex to terminally differentiated somatic cells. Alternative cell fate transitions are depicted, including reprogramming to an embryonic pluripotent state (red dashed arrows) or transdifferentiation between different somatic cell types (brown dashed arrows). These atypical, plastic cell-type transitions (dashed arrows) are usually prompted by external factors, such as injury or ectopic expression of transcription factors. Figure created using BioRender (http://biorender.com/).

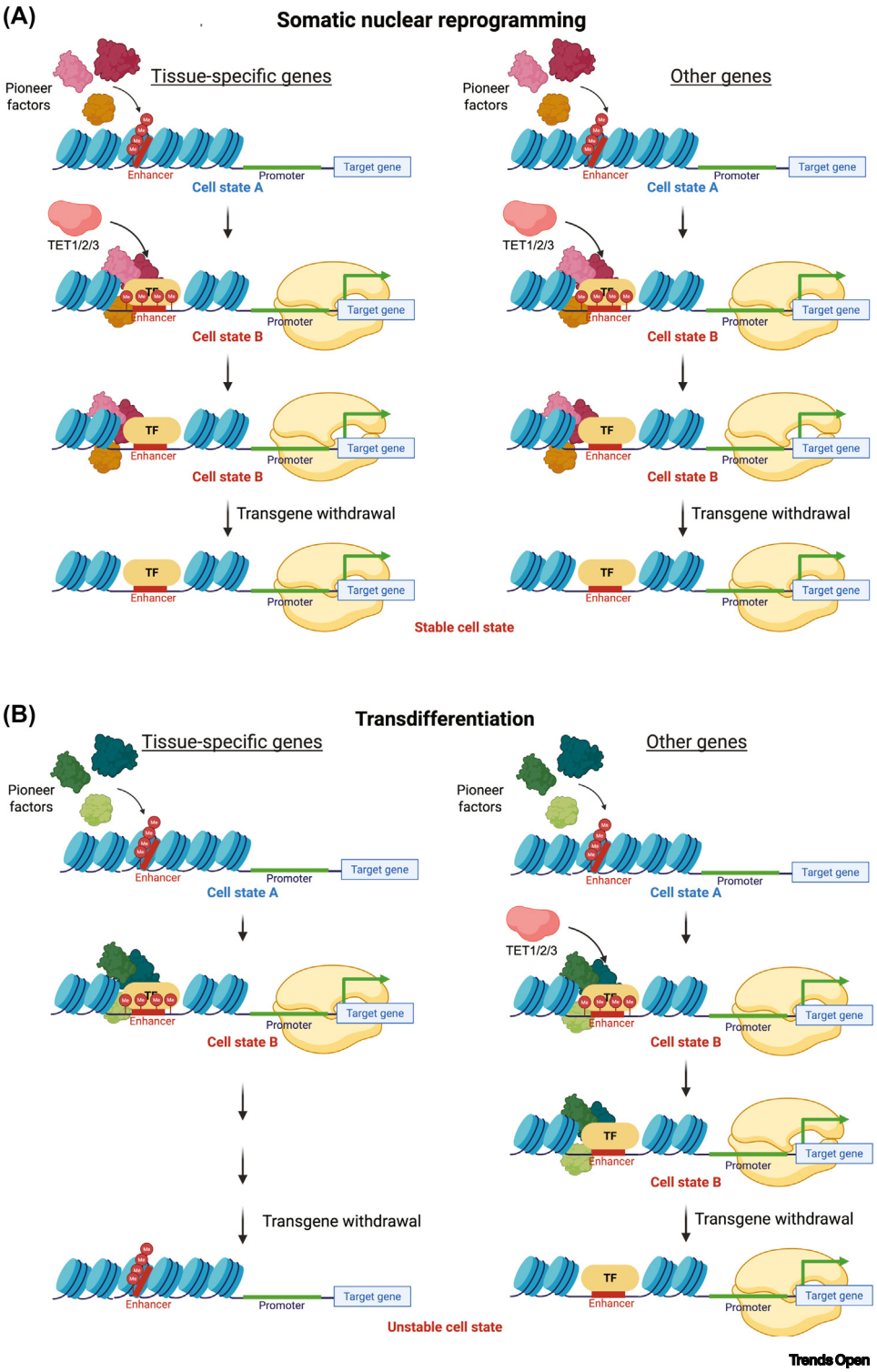

Figure 3. Epigenetic differences between transdifferentiation and somatic nuclear reprogramming. (A,B) Somatic nuclear reprogramming (A) involves TET1/2/3-mediated DNA demethylation on both tissue-specific and general enhancers,

(Figure legend continued at the bottom of the next page.)

remodeling, proper DNA methylation reprogramming, mainly at tissue-specific **differentially methylated regions**, is essential for establishing and maintaining long-term cell identity.

A similar pattern is observed in some examples of *in vivo* transdifferentiation. For example, during hepatocyte-to-cholangiocyte conversion following injury, cells acquire a new identity without extensively demethylating cholangiocyte-specific unmethylated REs [80]. However, in another reported case of transdifferentiation, the transition of astrocytes into neural stem cells (NSCs) after ischemic injury did involve active DNA methylation remodeling at NSC-specific REs [83]. These contrasting examples raise the possibility that certain adult TFs may function as pioneerlike regulators capable of mediating DNA methylation reprogramming. Supporting this notion, a study examining DNA methylation during fibroblast-to-neuron conversion by the BAM factors reported that *Ascl1* alone induces robust demethylation at its binding sites, which are enriched for REs, but the addition of *Brn2* markedly diminishes this demethylation activity [84]. These findings suggest that replacing *Brn2* with an alternative factor may enable a more complete remodeling of the methylome and promote the generation of more bona fide neurons.

Whether an embryonic or a stem-cell state is required for the demethylation of tissue-specific REs remains unclear. A study from the Hochedlinger group investigating two MyoD-driven transdifferentiation pathways—direct conversion into myocytes compared with conversion into myogenic progenitor cells (with the addition of forskolin, RepSox, and CHIR99021; FRC)—indicates that extensive and necessary DNA demethylation occurs only in the latter process [85]. However, it could be that the addition of the FRC cocktail causes the expression of pioneer factors required for this DNA demethylation, rather than this demethylation requiring that specific cell state.

To conclude, based on current models of reprogramming and transdifferentiation, it seems that early embryonic and some stem cell TF networks carry powerful (and possibly unique) epigenetic remodeling capabilities that are largely absent, or at least undiscovered, in adult cell TF networks.

The Waddington model

Our overall concept of the developmental pathway from pluripotency to tissue specification is largely shaped by **Waddington's proposal,** which visualized this as a one-directional process [86]. According to this idea, development proceeds 'passively' along prescribed valleys over a sloped terrain, with built-in barriers making it difficult to reverse direction or jump from one lineage stream to another (Figure 2 [86]). It has been suggested that these limitations may be mediated at the molecular level through epigenetic programming. In a sense, DNA methylation dynamics, as we understand them, can indeed explain the basic features of this model, since this system makes it possible to define, in a stable, fixed manner, the accessibility of regulatory regions to the many interacting protein factors that drive development—thereby setting up the rules of permissibility. Thus, for example, cells at one particular stage in development cannot undergo **dedifferentiation** to an earlier stage simply because there is no built-in program to reverse the methylation changes. In apparent contradiction to this principle, it was recently found that regeneration of the liver following partial hepatectomy is accomplished by first reverting the remaining cells to an earlier stage of development [49]. However, careful analysis of this system showed that these cells are indeed programmed with the ability to reverse their methylation state following liver injury [49]. This same idea may also explain what happens in muscle repair [53].

enabling stable gene activation and maintenance of cell identity upon transgene withdrawal. However, pioneer TFs driving transdifferentiation (B) fail to remodel the DNA methylation state of tissue-specific REs, resulting in an unstable cell state that collapses shortly after transgene withdrawal. Enhancers (red), promoters (green), DNA methylation (Me), and transcription factors (TF) are indicated. Note: The activity of TFs on promoters is not discussed in this illustration for the sake of simplicity. RE: regulatory element; TET: ten-eleven translocation enzymes.

Concluding remarks

As outlined in this review, the special chemical characteristics of DNA methylation, together with the existence of a modular RE network, have enabled higher organisms to develop and differentiate in a reliable and long-term stable manner. By understanding the genetic programming rules used to operate this system and providing answers to key questions in the field (such as those highlighted in the Outstanding questions), it should then be possible to artificially control this process and, in this way, enable both stable cell-type reprogramming and tissue regeneration.

Author contributions

H.C. conceived the idea for the review. H.C., A.K., O.S., and Y.B. wrote the manuscript. H.C., O.S., and A.K. focused on the general role of DNA methylation during development and under physiological conditions, whereas A.K. and Y.B. focused on the role of DNA methylation during reprogramming and transdifferentiation.

Acknowledgments

H.C. was supported by research grants from the Howard Jonas Foundation and The Kibbeh Foundation. Y.B. was supported by research grants from the EMBO Young Investigator Program, the Israel Science Foundation (161/23), the Leo Foundation (LF-OC-24-001624), the Israel Cancer Research Fund (1276940), and by a generous gift from Ms Nadia Guth Biasini.

Declaration of interests

The authors declare no competing interests.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors declare that generative AI was used solely for language editing and grammar improvement. No AI tools were used for scientific content generation, data analysis, or interpretation. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Author biographies

Howard (Chaim) Cedar is a professor in the Faculty of Medicine at the Hebrew University of Jerusalem. He received his undergraduate degree from MIT and an MD/PhD from New York University. He carried out postdoctoral studies at the NIH in the laboratory of Dr Gary Felsenfeld. Cedar has been studying DNA methylation since 1975. He is a member of both the US and Israel National Academies of Sciences. His work on DNA methylation has been recognized by a number of awards in biology and medicine.

Yosef (Yossi) Buganim is a full professor at the Faculty of Medicine, the Hebrew University of Jerusalem. He received his undergraduate degrees from Bar-Ilan University and his PhD from the Weizmann Institute of Science. He conducted his postdoctoral training at the Whitehead Institute for Biomedical Research at MIT. Buganim is renowned for being the first to generate iPSCs without using any of Yamanaka's factors. His research group pioneered the conversion of skin cells into placental stem cells and was the first to induce all three early embryonic cell types with a single combination of factors. Buganim is an Howard Hughes Medical Institute (HHMI) International Scholar and has received numerous prestigious international grants and awards, including the European Research Council (ERC) Starting Grant, the European Molecular Biology Organization (EMBO) Young Investigator Program, the 2016 'Stem Cell and Regenerative Medicine' Award by Science and Science Translational Medicine, and the 2016 'KY Cha Award' by the American Society for Reproductive Medicine, among others.

Outstanding questions

Can tissue-specific demethylation or methylation occur without transitioning through a stem cell or an embryonic state?

Can we identify a transdifferentiation combination that do manage to induce DNA demethylation on tissue-specific REs?

Do tissue-specific pioneer factors inherently distinguish lineage-specific regulatory elements to enable their closure?

What downstream players mediate the actual recruitment of TETs to tissue-specific REs during development and cell maturation? Can we stabilize transdifferentiation models using such factors? Can they be used to enhance tissue regeneration?

Can we modulate cellular memory by artificially changing the methylation status of specific regulatory elements?

Does the diminished ability of aged tissues to regenerate related to changes in the activity of the DNA demethylation machinery?

References

1. Keshet, I. *et al.* (1986) DNA methylation affects the formation of active chromatin. *Cell* 44, 535–543
2. Hashimshony, T. *et al.* (2003) The role of DNA methylation in setting up chromatin structure during development. *Nat. Genet.* 34, 187–192
3. Jones, P.L. and Wolffe, A.P. (1999) Relationships between chromatin organization and DNA methylation in determining gene expression. *Semin. Cancer Biol.* 9, 339–347
4. Siegfried, Z. *et al.* (1999) DNA methylation represses transcription *in vivo*. *Nat. Genet.* 22, 203–206
5. Fang, S. *et al.* (2019) Tet inactivation disrupts YY1 binding and long-range chromatin interactions during embryonic heart development. *Nat. Commun.* 10, 4297
6. Lio, C.-W. *et al.* (2016) Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *eLife* 5, e18290
7. Reizel, Y. *et al.* (2018) Postnatal DNA demethylation and its role in tissue maturation. *Nat. Commun.* 9, 2040
8. Dahlet, T. *et al.* (2020) Genome-wide analysis in the mouse embryo reveals the importance of DNA methylation for transcription integrity. *Nat. Commun.* 11, 3153
9. Kaluscha, S. *et al.* (2022) Evidence that direct inhibition of transcription factor binding is the prevailing mode of gene and repeat repression by DNA methylation. *Nat. Genet.* 54, 1895–1906
10. D'anna, F. *et al.* (2020) DNA methylation repels binding of hypoxia-inducible transcription factors to maintain tumor immunotolerance. *Genome Biol.* 21, 182
11. Domcke, S. *et al.* (2015) Competition between DNA methylation and transcription factors determines binding of NRF1. *Nature* 528, 575–579
12. Kribelbauer, J.F. *et al.* (2017) Quantitative analysis of the DNA methylation sensitivity of transcription factor complexes. *Cell Rep.* 19, 2383–2395
13. Yin, Y. *et al.* (2017) Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* 356, eaaj2239
14. Barnett, K.R. *et al.* (2020) ATAC-Me captures prolonged DNA methylation of dynamic chromatin accessibility loci during cell fate transitions. *Mol. Cell* 77, 1350–1364.e6
15. Visel, A. *et al.* (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457, 854–858
16. Hallikas, O. *et al.* (2006) Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* 124, 47–59
17. Doni Jayavelu, N. *et al.* (2020) Candidate silencer elements for the human and mouse genomes. *Nat. Commun.* 11, 1061
18. Pang, B. and Snyder, M.P. (2020) Systematic identification of silencers in human cells. *Nat. Genet.* 52, 254–263
19. Segert, J.A. *et al.* (2021) Transcriptional silencers: driving gene expression with the brakes on. *Trends Genet.* 37, 514–527
20. Rosenski, J. *et al.* (2026) The genetic basis for DNA methylation variation across tissues and development. *Nat. Commun.* 17, 5194
21. Lienert, F. *et al.* (2011) Identification of genetic elements that autonomously determine DNA methylation states. *Nat. Genet.* 43, 1091–1097
22. Lee, T.I. and Young, R.A. (2013) Transcriptional regulation and its misregulation in disease. *Cell* 152, 1237–1251
23. Yu, M. and Ren, B. (2017) The three-dimensional organization of mammalian genomes. *Annu. Rev. Cell Dev. Biol.* 33, 265–289
24. Rao, S.S.P. *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680
25. Oudelaar, A.M. and Higgs, D.R. (2021) The relationship between genome structure and function. *Nat. Rev. Genet.* 22, 154–168
26. Jindal, G.A. and Farley, E.K. (2021) Enhancer grammar in development, evolution, and disease: dependencies and interplay. *Dev. Cell* 56, 575–587
27. Long, H.K. *et al.* (2016) Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* 167, 1170–1187
28. Cedar, H. *et al.* (2022) The role of DNA methylation in genome-wide gene regulation during development. *Development* 149, dev200118
29. Hon, G.C. *et al.* (2013) Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nat. Genet.* 45, 1198–1206
30. Ziller, M.J. *et al.* (2013) Charting a dynamic DNA methylation landscape of the human genome. *Nature* 500, 477–481
31. Roadmap Epigenomics Consortium et al (2015) Integrative analysis of 111 reference human epigenomes. *Nature* 518, 317–330
32. Loyfer, N. *et al.* (2023) A DNA methylation atlas of normal human cell types. *Nature* 613, 355–364
33. Zhou, J. *et al.* (2025) Human body single-cell Atlas of 3D genome organization and DNA methylation. *bioRxiv* https://doi.org/10.1101/2025.03.23.644697
34. Sonawane, A.R. *et al.* (2017) Understanding tissue-specific gene regulation. *Cell Rep.* 21, 1077–1088
35. He, Y. *et al.* (2020) Spatiotemporal DNA methylome dynamics of the developing mouse fetus. *Nature* 583, 752–759
36. Razin, A. and Cedar, H. (1991) DNA methylation and gene expression. *Microbiol. Rev.* 55, 451–458
37. Feldman, N. *et al.* (2006) G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat. Cell Biol.* 8, 188–194
38. Bergman, Y. and Cedar, H. (2010) Epigenetic control of recombination in the immune system. *Semin. Immunol.* 22, 323–329
39. Cirillo, L.A. *et al.* (2002) Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* 9, 279–289
40. Zaret, K.S. (2020) Pioneer transcription factors initiating gene network changes. *Annu. Rev. Genet.* 54, 367–385
41. Donaghey, J. *et al.* (2018) Genetic determinants and epigenetic effects of pioneer-factor occupancy. *Nat. Genet.* 50, 250–258
42. Mayran, A. *et al.* (2018) Pioneer factor Pax7 deploys a stable enhancer repertoire for specification of cell fate. *Nat. Genet.* 50, 259–269
43. Reizel, Y. *et al.* (2021) FoxA-dependent demethylation of DNA initiates epigenetic memory of cellular identity. *Dev. Cell* 56, 602–612.e4
44. Reizel, Y. *et al.* (2015) Gender-specific postnatal demethylation and establishment of epigenetic memory. *Genes Dev.* 29, 923–933
45. Grimm, S.A. *et al.* (2019) DNA methylation in mice is influenced by genetics as well as sex and life experience. *Nat. Commun.* 10, 305
46. Dos Santos, C.O. *et al.* (2015) An epigenetic memory of pregnancy in the mouse mammary gland. *Cell Rep.* 11, 1102–1109
47. Pacis, A. *et al.* (2019) Gene activation precedes DNA demethylation in response to infection in human dendritic cells. *Proc. Natl. Acad. Sci. U. S. A.* 116, 6938–6943
48. Abu-Remaileh, M. *et al.* (2015) Chronic inflammation induces a novel epigenetic program that is conserved in intestinal adenomas and in colorectal cancer. *Cancer Res.* 75, 2120–2130
49. Michaeli, T.F. *et al.* (2024) Hepatocyte regeneration is driven by embryo-like DNA methylation reprogramming. *Proc. Natl. Acad. Sci. U. S. A.* 121, e2314885121
50. Seale, K. *et al.* (2022) Making sense of the ageing methylome. *Nat. Rev. Genet.* 23, 585–605
51. Masika, H. *et al.* (2026) Cell-to-cell variability and gain of methylation at polycomb CpG islands as a hallmark of aging. *bioRxiv* https://doi.org/10.64898/2026.03.09.710505
52. Cowley, C.J. *et al.* (2026) Distinctive DNA sequence features define epigenetic longevity of inflammatory memory. *Science* 391, eadz6830
53. Falick Michaeli, T. *et al.* (2022) Muscle injury causes long-term changes in stem-cell DNA methylation. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2212306119
54. Takahashi, K. and Yamanaka, S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676
55. Takahashi, K. *et al.* (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–872
56. Huang, P. *et al.* (2021) Identification of new transcription factors that can promote pluripotent reprogramming. *Stem Cell Rev. Rep.* 17, 2223–2234

Key references

1. Radwan, A. *et al.* (2024) Transdifferentiation occurs without resetting development-specific DNA methylation, a key determinant of full-function cell identity. Proc. Natl. Acad. Sci. U S A 121, e2411352121. 10.1073/pnas.2411352121

This study compares DNA methylation dynamics at REs during reprogramming and transdifferentiation. It demonstrates that reprogramming factors can induce demethylation at tissue-specific REs, whereas transdifferentiation factors largely fail to do so. These differences likely contribute to the reduced stability and incomplete identity acquisition observed in transdifferentiated cells.

2. Zaret, K.S. (2020) Pioneer transcription factors initiating gene network changes. Annu. Rev. Genet. 54, 367–385. 10.1146/annurev-genet-030220-015007

This review discusses how pioneer TFs access closed chromatin and initiate gene expression programs.

3. Reizel, Y. *et al.* (2021) FoxA-dependent demethylation of DNA initiates epigenetic memory of cellular identity. Dev. Cell 56, 602–612.e4. 10.1016/j.devcel.2021.02.005

Reizel and colleagues show that pioneer TFs are not required to maintain DNA methylation patterns once they are established. Instead, the intrinsic stability of DNA methylation preserves these states. For example, FoxA induces demethylation at liver-specific REs, and these changes persist even after conditional ablation of the FoxA genes, indicating a role in establishing epigenetic memory.

4. Cowley, C.J. *et al.* (2026) Distinctive DNA sequence features define epigenetic longevity of inflammatory memory. Science 391, eadz6830. 10.1126/science.adz6830

This study investigates the molecular mechanisms by which skin stem cells maintain inflammatory memory, highlighting DNA methylation as a central component of this process.

57. Wang, B. *et al.* (2023) The NuRD complex cooperates with SALL4 to orchestrate reprogramming. *Nat. Commun.* 14, 2846
58. Heng, J.-C.D. *et al.* (2010) The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell* 6, 167–174
59. Montserrat, N. *et al.* (2013) Reprogramming of human fibroblasts to pluripotency with lineage specifiers. *Cell Stem Cell* 13, 341–350
60. Buganim, Y. *et al.* (2014) The developmental potential of iPSCs is greatly influenced by reprogramming factor selection. *Cell Stem Cell* 15, 295–309
61. Benchetrit, H. *et al.* (2019) Direct induction of the three pre-implantation blastocyst cell types from fibroblasts. *Cell Stem Cell* 24, 983–994.e7
62. Jaber, M. *et al.* (2017) Acquisition of the pluripotent and trophectoderm states in the embryo and during somatic nuclear reprogramming. *Curr. Opin. Genet. Dev.* 46, 37–43
63. Costa, Y. *et al.* (2013) NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* 495, 370–374
64. Hu, X. *et al.* (2014) Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. *Cell Stem Cell* 14, 512–522
65. Doege, C.A. *et al.* (2012) Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* 488, 652–655
66. Gao, Y. *et al.* (2013) Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell* 12, 453–469
67. Chen, J. *et al.* (2015) The combination of Tet1 with Oct4 generates high-quality mouse-induced pluripotent stem cells. *Stem Cells* 33, 686–698
68. Zviran, A. *et al.* (2019) Deterministic somatic cell reprogramming involves continuous transcriptional changes governed by Myc and epigenetic-driven modules. *Cell Stem Cell* 24, 328–341.e9
69. Stadtfeld, M. *et al.* (2010) Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* 465, 175–181
70. Xu, X. *et al.* (2015) *Dppa3* expression is critical for generation of fully reprogrammed iPS cells and maintenance of *Dlk1-Dio3* imprinting. *Nat. Commun.* 6, 6008
71. Benchetrit, H. *et al.* (2015) Extensive nuclear reprogramming underlies lineage conversion into functional trophoblast stem-like cells. *Cell Stem Cell* 17, 543–556
72. Kubaczka, C. *et al.* (2015) Direct induction of trophoblast stem cells from murine fibroblasts. *Cell Stem Cell* 17, 557–568
73. Naama, M. *et al.* (2023) Pluripotency-independent induction of human trophoblast stem cells from fibroblasts. *Nat. Commun.* 14, 3359
74. Jaber, M. *et al.* (2022) Comparative parallel multi-omics analysis during the induction of pluripotent and trophectoderm states. *Nat. Commun.* 13, 3475
75. Polo, J.M. *et al.* (2012) A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* 151, 1617–1632
76. Davis, R.L. *et al.* (1987) Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51, 987–1000
77. Vierbuchen, T. *et al.* (2010) Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 463, 1035–1041
78. Huang, P. *et al.* (2011) Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature* 475, 386–389
79. Ieda, M. *et al.* (2010) Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* 142, 375–386
80. Radwan, A. *et al.* (2024) Transdifferentiation occurs without resetting development-specific DNA methylation, a key determinant of full-function cell identity. *Proc. Natl. Acad. Sci. U. S. A.* 121, e2411352121
81. Morris, S.A. *et al.* (2014) Dissecting engineered cell types and enhancing cell fate conversion via CellNet. *Cell* 158, 889–902
82. Cambuli, F. *et al.* (2014) Epigenetic memory of the first cell fate decision prevents complete ES cell reprogramming into trophoblast. *Nat. Commun.* 5, 5538
83. Kremer, L.P.M. *et al.* (2024) DNA methylation controls stemness of astrocytes in health and ischaemia. *Nature* 634, 415–423
84. Luo, C. *et al.* (2019) Global DNA methylation remodeling during direct reprogramming of fibroblasts to neurons. *eLife* 8, e40197
85. Yagi, M. *et al.* (2021) Dissecting dual roles of MyoD during lineage conversion to mature myocytes and myogenic stem cells. *Genes Dev.* 35, 1209–1228
86. Waddington, C.H. (1957) *The Strategy of Genes*, George Allen & Unwin Ltd

5. Kremer, L.P.M. *et al*. (2024) DNA methylation controls stemness of astrocytes in health and ischaemia. Nature 634, 415–423. 10.1038/s41586-024-07898-9

This study investigates the acquisition of stem cell-like properties by astrocytes following brain injury. The authors show that DNMT3A-mediated remodeling of DNA methylation at REs is essential for this transition, highlighting a key role for DNA methylation in maintaining and reshaping cell identity in the adult brain.