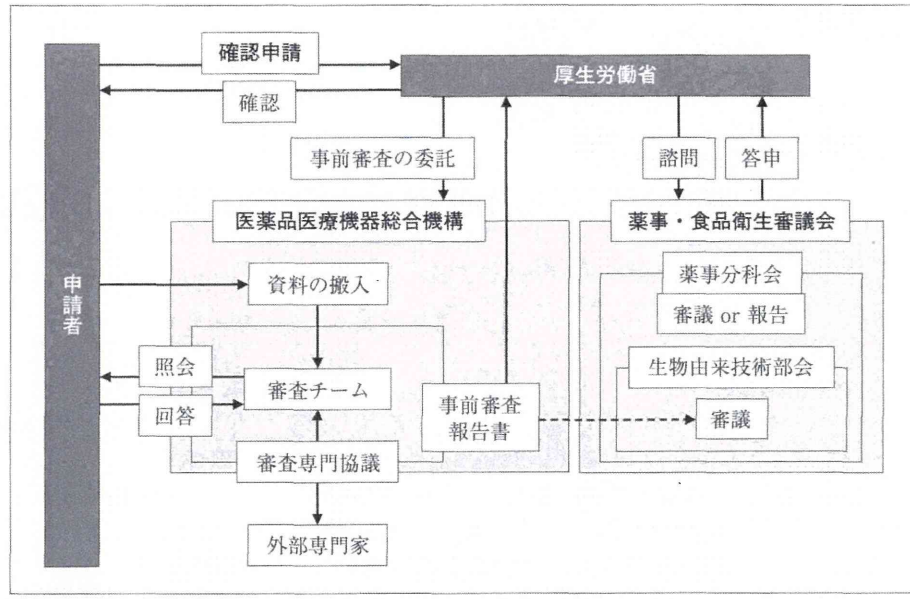
図1　確認申請の流れ



れる．

製造販売承認申請の審査および提出資料に関する調査は，申請者からの提出資料および申請者への照会・回答に基づき，PMDA関係各部（生物系審査第二部，信頼性保証部，品質管理部など）により実施されている．提出資料には，信頼性基準のほか，治験（臨床試験）についてはGCP（Good Clinical Practice），非臨床安全性試験についてはGLP（Good Laboratory Practice）に従って実施されたデータや情報が収集されている必要があり，承認申請後に審査と並行してPMDAにおいてGCP，GLP等の信頼性基準適合性調査（医療機関に対する実地調査も含む）が実施され，QMS（Quality Management System）（医療機器）/GMP（Good Manufacturing Practice）（医薬品）適合性調査が行われる．PMDAでの審査終了後，さらに薬事・食品衛生審議会において承認の可否が審議される．

新しい技術を応用した細胞・組織利用製品は，新規性が高く，過去の使用経験・情報の蓄積が少なく，技術的に発展途上である．また，生物由来原材料を使用するため，感染性物質混入のリスクに対して慎重な検討が必要である．したがって，細胞・組織利用製品の治験開始前（ヒトに初めて使用する前）に，製品の最低限の品質・安全性をチェックするため，治験依頼者は，厚生労働大臣に対して，品質・安全性が適当であるか「確認申請」（通知による行政指導）を行うよう指導されている（**図1**）．なお，開発のタイミングに応じて，PMDAでは各種対面助言が設定されている．対面助言には，手続，資料整備，品質，安全性，申請前など臨床試験計画以外に対する相談区分もある．

また，法第2条の9，10に基づき，保健衛生上特別の注意を要する医薬品，すなわち感染リスクの高い医薬品は生物由来製品あるいは特定生物由来製品に指定され，法第68条の7に基づき，特定生物由来製品を取り扱う医師らは，その適正な使用のために必要な事項について，使用の対象者に適切な説明を行い，同意を得るよう努めなければならない．このほか感染症定期報告や生物由来製品に関する製造などや特定生物由来製品の使用者の対象者らの記録を作成し，保存するなどの義務がある．

法第42条に基づいて生物由来原料基準が定められており，医薬品等の製造に用いる生物由来原材料に関して，感染性因子混入のリスクの視点から，ドナー，検査，記録等の必要な管理について示されている．

Ⅱ．細胞・組織利用製品の審査プロセスと留意点

1．基礎研究～非臨床試験

前述のとおり，承認を取得するための提出資料は各種基準に従って収集・作成されている必要があり，申請内容の根拠となるデータについては，これら基準への適合性が事後的に確認できるよう，適切に原資料を保管する必要がある．

他社からの導入の場合に，技術だけ導入して自社で提出資料を作成するのか，導入元の作成した申請資料

の利用契約を結び，これを国内申請の提出資料として利用するのかで扱いが異なる．後者の場合，導入元などがPMDAによる調査を受ける可能性がある．

技術導入とはいえ，例えば生細胞率などについて，導入元と国内での製造を比較したとき国内での生細胞率が大幅に低い場合などは，完全に技術が導入されていない等の懸念をもたれることもありうる．

動物を用いた試験についても，実施の意義など考え方について一律には適用できないので指針やQ＆Aを参照いただきたい．

2．確認申請

細胞・組織利用製品については平成11年7月30日医薬発第906号医薬安全局長通知「細胞・組織を利用した医療用具又は医薬品の品質及び安全性の確保について」により，確認申請の書式（様式1）および資料提出項目が示されると共に各種報告事項などが求められている．すなわち，関連情報の収集を求めると共に，厚生労働大臣に対し，イ．評価に影響を及ぼす知見についての報告，ロ．製造数量・出荷先・出荷数量の年度末報告，ハ．業者名称・住所などの変更の場合の報告である．

平成12年12月26日医薬発第1314号医薬安全局長通知「ヒト又は動物由来成分を原料として製造される医薬品等の品質及び安全性確保について」の別添1では「細胞・組織利用医薬品等の取扱い及び使用に関する基本的考え方（以下；基本的考え方）」が示されている．この「基本的考え方」は，“ヒト由来”か“動物由来”かにかかわらず，また“加工する”，“加工しない”にかかわらず，細胞・組織を取り扱う際の基本的要件を示すと共に，細胞・組織を“利用”した製品の品質および安全性，ならびに細胞・組織の取り扱いに関する科学的および倫理的妥当性を確保することを目的として作成されたものである．その背景には，①細胞・組織利用製品については，細胞・組織に由来する感染症の伝播の危険性が懸念されるため，細菌，真菌，ウイルス等に汚染されていない原料の使用，製造工程中における汚染の防止等を図ることが不可欠である，②また，不適切な製造等による不良製品の製造，不適切な製品の取り扱いや使用による問題の発生を防止する必要がある，③このような観点に立ち，細胞・組織の採取から，製造，使用まで一貫した方策が必要である，との考え方がある．内容は，大項目として，1）目的，基本原則および定義，2）細胞・組織採取について，3）製造段階における安全性確保対策，4）職員および組織ならびに管理体制等，5）使用段階における安全性確保対策，6）個人情報の保護，7）見直し，から構成されている．科学的な関心事と共に，倫理面で細胞・組織採取段階や使用段階で配慮すべきことを網羅しているのも，この「基本的考え方」の特徴である．さらに，必要な記録・資料や細胞試料等に関する保管，職員および組織ならびに管理体制や個人情報の保護等にも言及している．結局，「基本的考え方」はいわゆるGTP（Good Tissue Practice）といえる．「基本的考え方」に示された方法以外の方法を採る場合には，品質および安全性確保の観点からその必要性および妥当性を説明し，その根拠を示すことが必要であり，「基本的考え方」に示された事項は，細胞・組織利用製品の承認後のみならず，治験時においても適用されるとされている．

医薬発第1314号通知の別添2には，当面，製品が登場することが予測された，“ヒト由来”細胞・組織を“加工して”医薬品等を製造するケースに特化した指針として「ヒト由来細胞・組織加工医薬品等の品質及び安全性の確保に関する指針」が示されていた．すなわち，本指針は，ヒト由来の細胞・組織を加工した医薬品又は医療用具（「細胞・組織加工医薬品等」）の品質および安全性の確保のために必要な基本的要件を定めると共に，確認申請にあたって添付するべき資料の内容を示すことを目的とするものであった．

ヒト由来細胞・組織利用製品では，前述の906号通知および1314号通知に基づき確認申請が求められることとされていた．

1314号通知後6年を経て，細胞・組織加工医薬品等による再生医療について，研究の進歩に伴う技術的な実現可能性の高まりと共に，医療としての実用化推進を求める社会的要請が急激に強くなった．また，平成18年7月の総合科学技術会議の「科学技術の振興及び成果の社会への還元に向けた制度改革について（中間報告）」においても，細胞・組織利用製品についての安全評価基準の明確化や確認申請・治験計画届に係る調査における重複部分の簡素化について，早急に検討するよう要望が出された．これを受けて，我が国の再生医療を適正な規制のもとに推進していくために，急速に発展する学問・技術，倫理上の観点，国際的動向等を反映した安全性評価基準の作成など規制のあり方について至急検討することを目的とし，平成18年11月に筆者の一人（早川堯夫）を班長とする厚生労働科学研究班が設置された．

その結果を受けて，まず平成19年3月30日薬食発0330030号医薬食品局長通知「細胞・組織を利用した

医薬品等の品質及び安全性の確保に係る手続きの変更について」により，906号通知様式1(別添注2)で求める治験計画書届出書案の添付を取りやめると共に，治験計画の概要の項目を簡素化した．また研究班では，患者自身の細胞を用いる細胞・組織加工医薬品等は，同種(他人)や異種(動物)由来の細胞と感染症などの面でリスクが異なる点もあり，新たな基準を作成すべきとの総合科学技術会議の提言(平成18年12月25日「技術の振興及び成果の社会への還元に向けた制度改革について」)を受け，1314号通知別添2の「ヒト由来細胞・組織加工医薬品等の品質及び安全性の確保に関する指針」を見直すための検討を行った．その際の基本方針は，以下のとおりとした．すなわち，①自己由来のものと同種由来のものに分け，それぞれの製品における品質および安全性確保のために必要な基本的な技術要件を明確にすること，②基本的要件は承認申請をも念頭に置いたものであるのに対して，確認申請とは治験を開始するにあたって支障となる品質，安全性上の問題があるか否かを確認するためという趣旨を踏まえて，基本的な技術要件のうち確認申請までにどの程度の試験や評価をするべきかを明確にすること，③従来は必要な試験や評価に関する科学的考え方および申請に際して必要な情報や記載すべき事項が一つの指針に盛り込まれていたが，確認申請の記載要領に関することは別記事項として明確にすること，④指針の記述は理解しやすいものとすると共に，Q＆Aにより，必要な背景説明を行うこと．

この研究結果をベースにして，行政当局は，平成20年2月8日に薬食発第0208003号医薬食品局長通知(自己由来細胞指針)を，平成20年9月12日に薬食発第0912006号医薬食品局長通知(同種由来細胞指針)をそれぞれ公表した．両指針とも，1)総則(目的，定義)，2)製造方法，3)製品の安定性，4)非臨床安全性試験，5)効力または性能を裏づける試験，6)体内動態，7)臨床試験を柱として構成されていて，基本的留意事項や要件に関する内容の大半も当然のことながら共通している．一方で，自己由来細胞と同種由来細胞とでは異なる観点から対処すべき基本的要件については両指針で明確に区別して記載されている．例えば，同種由来細胞を用いる場合では，①目的とする細胞・組織の原材料となる細胞のドナーの選択基準や適格性に関しては，各種ウイルス感染に関する検査と否定，免疫適合性等の考慮，ならびに各種細菌感染症や重篤疾病への既往歴，輸血や移植医療を受けた経験などよりその適格性を判断すべきこと，②株化細胞の使用に関する記述，③ドナーに関する記録の整備と保存，などが自己由来細胞の場合と異なる対応が必要とされるところである．これらに加えて自己由来細胞Q＆Aに関して平成20年3月12日に事務連絡，および，同種由来細胞Q＆Aに関して平成20年10月3日に事務連絡がそれぞれ発出された．

これら両指針とも，確認申請を含む承認申請に必要な留意事項や基本的技術要件を示しているが，前文において「その時点の学問の進歩を反映した合理的根拠に基づき，ケースバイケースの原則で柔軟に対応することが必要」であり，確認申請時点における本指針への適合性の確認の趣旨は，「当該細胞・組織加工医薬品等の治験を開始するにあたって支障となる品質および安全性上の問題が存在するか否かの確認にある」「確認に必要とされる資料の範囲，程度については，当該製品の由来，対象疾患，適用部位，適用方法，加工方法等により異なり，本指針では具体的に明らかでないことも少なくないので，個別に独立行政法人医薬品医療機器総合機構の相談を受けることが望ましい」などと，その解釈，運用について考慮すべき点も強調している．なお，資料作成を円滑化するために治験相談の区分に「細胞・組織利用製品資料整備相談」が新設されたので利用することを推奨する．これまで，確認申請から確認終了まで，平均2年程度要しており，治験相談を利用することで結果として確認までの時間を短縮できることを期待している．

3. 治験届出調査

確認申請制度のない通常の新薬と異なり，治験開始のための支障となる品質，安全性の問題は確認申請で確認されているため，治験届出調査は，円滑に進むはずである．しかし，確認申請段階の専門協議や生物由来技術部会での議論を参考にしつつ，適宜，次のステップに備えておくことが必要な場合もありうる．必要に応じて対面助言の利用を勧める．また，対面助言を有益なものとするための事前面談などもある．詳細はPMDAホームページ(http://www.pmda.go.jp/)を参照いただきたい．

4. 治験実施～申請

医師主導治験の実施や，通常の企業主導の治験として実施するには，GCPの遵守が必要である．細胞・組織利用製品の場合は，感染への配慮だけでなく，細胞・組織の採取など通常の治験薬交付だけの治験と異なる手続きがあることに留意していただきたい．また，自己由来の細胞であっても，培養担当者や医療従事者には感染のリスクがあり，配慮する必要がある．

申請資料の準備を円滑に進めるためにも，ぜひ治験相談を利用していただきたい．なお，次項に述べるが，審査報告書で治験実施や資料作成の不備について指摘されているので，参照いただきたい．

5．承認審査

承認審査は，承認された方法を遵守して製造していれば，品質・有効性・安全性が確保され，当該効能で医療上有用かどうかの観点で実施される．承認後，厚生労働大臣が出荷ごとに治験の実施を求め，有効性・安全性をその都度確認することはないので，製法・規格遵守＋市販後調査により，これを担保することになる．

本邦初の細胞・組織利用製品の承認品目であるジャパン・ティッシュ・エンジニアリング(J-TEC)社の自家培養表皮(ジェイス)の審査報告書が，医薬品医療機器情報提供ホームページ(http://www.info.pmda.go.jp/)で公表されており，品目の概要，審査における問題点など承認審査がどのような視点で行われるか理解するために参考にされることをお勧めする．さらに，厚生労働省ホームページ(http://www.mhlw.go.jp/)では本品目の部会審議の議事録も公表されているのでそちらも参照されたい．また，これら承認申請を踏まえ，承認書の製造方法，規格および試験方法などの整備がなされた．添付文書の整備についても並行して行われている．

6．承認条件

法第79条第1項では，「この法律に規定する許可，認定又は承認には，条件又は期限を付し，及びこれを変更することができる」，同条第2項では，「前項の条件又は期限は，保健衛生上の危害の発生を防止するため必要な最小限度のものに限り，かつ，許可，認定又は承認を受ける者に対し不当な義務を課することとなるものであってはならない」とされている．

前述の議論などを受け，ジェイスの審査報告書には，承認条件として以下の項目について記載されている．

- 使用する医師，医療機関
- 製造販売後臨床試験の実施
- 全症例調査
- 製造販売後調査結果の情報提供

このほか，フィーダー細胞に関して条件が付されている．

7．製造販売後調査

副作用/感染症報告や感染症定期報告など各種法令の遵守が必要である．ジェイスの再審査期間は7年とされたが，製造販売後調査を実施すると共に再審査期間終了後，再審査申請を行うことが求められている．

Ⅲ．実用化・産業化の促進に向けて

より有効でより安全な医薬品・医療機器をより迅速に患者のもとへ届け，より適正に，より広く使用されるのは全ての関係者の願いである．その目的を果たす一翼としてPMDAとしても，①量・質の両面からの審査・安全対策の体制強化と透明性，②対面助言の強化，③通知，ガイドラインおよびQ＆Aの整備での協力や記載事例の作成，④学会・業界・各国規制当局などとの連携を推進している．また，長期的視野でのトランスレーショナルリサーチの促進が重要であるとの観点より，今後さらなる大学・研究機関等との人事交流が必要であると考えている．

細胞・組織利用製品については，特にベンチャー企業からの申請が多い．それゆえ，申請品目についての専門性は高いものの，申請資料の作成に時間がかかり，申請までに時間を要する場合がある．また，照会事項への対応に手間どっている面も多々見受けられる．前述のとおり，対面助言を有効に活用いただくことで，審査時間の短縮の一助になればと考えている．また，ベンチャー企業と大手製薬企業との連携を一層円滑にするような政策も今後必要と考える．

優良な細胞・組織利用製品については円滑な確認申請，治験届出，承認申請を経て，薬事法に基づき，国内さらには世界の患者が使えるよう企業により実用化(薬事法承認による国内医療機関への供給および保険適用)されるべきである．細胞・組織利用製品の実用化を目指す企業にとって重要なことは，製品のリスク・ベネフィットを勘案し，期待される臨床的意義を明確にすることである．また，治験を実施するにあたってはその臨床的意義を確認できるよう適切な投与対象を設定し，有効性および安全性を適切に評価できるプロトコールを作成する等，開発戦略を明確にすることに加え，承認申請審査の際に，データに基づいて製品の臨床的意義を明確に説明する必要があると考える．基礎研究を支える大学・研究機関等には，さらに良いものあるいは治療法のない領域の研究を推進すべく尽力いただくことを期待する．

2008年11月，先端技術の早期実用化に向けた政策の第一弾として，先端医療開発特区(スーパー特区)が開始された．再生医療の分野で5課題，iPS細胞の分野で2課題が採択された．スーパー特区は革新的技

術の早期実用化のため，研究資金の特例や規制を担当する部局との並行協議などを試行的に行うことを目的としたものである．これまでは，文科省，経産省，厚労省独自に再生医療実用化促進プロジェクトが推進されてきたが，規制緩和により，予算の弾力使用が可能となった．PMDAも相談対応において協力する予定であるため，ぜひ課題採択された企業，大学・研究機関等は細胞・組織利用製品の実用化・産業化に向けて利用していただきたい．

おわりに

病に苦しむ患者にとってのなによりの希望，期待は，疾病治療のための有用な医薬品・医療機器や医療技術の1日も早い開発であり，実用化である．全ての医療関係者の役割は，それぞれの立場でこの期待に応えることである．

細胞・組織利用製品には，他の治療法では得られない効果が期待できる場合も多々あると考えられるので，その適正な実用化の推進を図る必要があることは論を待たない．良いシーズをいかに見出すかは，基本的には，医学・薬学等の研究者の役割であり，それをいかに実用化に向けて開発するかは企業の役割である．一方で，細胞・組織利用製品の品質，安全性，有効性のセーフガードを維持しながら，その実用化をいかに合理的，効率的，効果的に進めるかの道筋を示し，必要に応じてサポートして推進を図ることは行政の重要な役割の一つである．規制環境を整備し，ガイドラインの作成や各種対面助言，合理的な審査などにより，製品がより早く，患者のもとに届くよう方策を講ずるのも極めて重要な行政的役割である．行政がこのような役割を全うするには，当該分野の科学や技術，関連情報に対する認識や理解度を常に高めていくことが必須の要件である．同様に，産・学が行政の仕組みや考え方をよく理解して活用する，あるいは必要な有機的連携をとおして，それぞれの活動の一層の進展を図ることも極めて重要である．

一般に医薬品等にあっては，品質，安全性，有効性の確保が求められる．一方，細胞・組織利用製品の場合，もとより細胞・組織の全貌が明らかにされているわけではないので，現在の科学的知見や技術で，全ての安全性上の関心事に応えることは困難である．しかし，既存のウイルス感染に関わる安全性，造腫瘍性，同種由来細胞における免疫原性など，想定される安全性上の主な関心事に応えることは可能であり，不可欠なことである．問題は，このような安全性上の主要問題がクリアされた細胞・組織利用製品にあって，なお，リスクゼロとはいえないが，臨床上のベネフィットが示唆される場合に，どう考え，取り扱うかということである．他の治療法がもはや望めない患者の立場に思いを馳せると，その状況を何らかの形でブレークスルーできる先端医療としての再生医療の可能性が考えられる場合は，治療法に関する情報を早期に提供し，治療の機会を逸することのないようにすることが重要であると考える．その時点でのリスク・ベネフィットについて，全てを明らかにする十分な説明により患者に選択肢を与え，最終的に同意が得られれば，適正に使用していくという方策が大切であると考える．

参考資料

(細胞・組織利用製品に関連する通知・ガイドラインなど)

1) 細胞・組織を利用した医療用具又は医薬品の品質及び安全性の確保について．平成11年7月30日，医薬発第906号
2) 細胞・組織を利用した医薬品等の品質及び安全性の確保に係る手続きの変更について．平成19年3月30日，薬食発0330030号
3) ヒト又は動物由来成分を原料として製造される医薬品等の品質及び安全性確保について．平成12年12月26日，医薬発第1314号
4) ヒト(自己)由来細胞や組織を加工した医薬品又は医療機器の品質及び安全性の確保について．平成20年2月8日，薬食発第0208003号(平成20年9月12日事務連絡にて一部訂正)
5) ヒト(自己)由来細胞・組織加工医薬品等の品質及び安全性の確保に関する指針に係るQ&Aについて．平成20年3月12日，事務連絡
6) ヒト(自己)由来細胞・組織加工医薬品等の製造管理・品質管理の考え方について．平成20年3月27日，薬食監麻発第0327025号
7) ヒト(同種)由来細胞や組織を加工した医薬品又は医療機器の品質及び安全性の確保について．平成20年9月12日，薬食発第0912006号
8) ヒト(同種)由来細胞・組織加工医薬品等の品質及び安全性の確保に関する指針に係るQ&Aについて．平成20年10月3日，事務連絡
9)「異種移植の実施に伴う公衆衛生上の感染症問題に関する指針」に基づく3T3J2株及び3T3NIH株をフィーダー細胞として利用する上皮系の再生医療への指針について，平成16年7月2日，医政研発第0702001号

Cell Stem Cell
Article

A Small-Molecule Inhibitor of Tgf-β Signaling Replaces *Sox2* in Reprogramming by Inducing *Nanog*

Justin K. Ichida,[1,3,6] Joel Blanchard,[1,6] Kelvin Lam,[1,6] Esther Y. Son,[1,2,3,5,6] Julia E. Chung,[1,2,3] Dieter Egli,[1,3] Kyle M. Loh,[1] Ava C. Carter,[1,3] Francesco P. Di Giorgio,[1,3] Kathryn Koszka,[1,3] Danwei Huangfu,[1] Hidenori Akutsu,[4] David R. Liu,[5] Lee L. Rubin,[1,*] and Kevin Eggan[1,2,3,*]

[1]Harvard Stem Cell Institute, Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA
[2]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA
[3]Stowers Medical Institute, 1000 East 50th Street. Kansas City, MO 64110, USA
[4]Department of Reproductive Biology, National Research Institute for Child Health and Development, 2-10-1 Okura, Setagaya, Tokyo 157-8535
[5]Howard Hughes Medical Institute and the Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA
[6]These authors contributed equally to this work
*Correspondence: lee_rubin@harvard.edu (L.L.R.), eggan@mcb.harvard.edu (K.E.)
DOI 10.1016/j.stem.2009.09.012

SUMMARY

The combined activity of three transcription factors can reprogram adult cells into induced pluripotent stem cells (iPSCs). However, the transgenic methods used for delivering reprogramming factors have raised concerns regarding the future utility of the resulting stem cells. These uncertainties could be overcome if each transgenic factor were replaced with a small molecule that either directly activated its expression from the somatic genome or in some way compensated for its activity. To this end, we have used high-content chemical screening to identify small molecules that can replace *Sox2* in reprogramming. We show that one of these molecules functions in reprogramming by inhibiting Tgf-β signaling in a stable and trapped intermediate cell type that forms during the process. We find that this inhibition promotes the completion of reprogramming through induction of the transcription factor *Nanog*.

INTRODUCTION

Retroviral transduction with three genes, *Sox2*, *Oct4*, and *Klf4*, can directly reprogram somatic cells to a pluripotent stem cell state (Okita et al., 2007; Takahashi et al., 2007b). Unfortunately, the resulting induced pluripotent stem cells (iPSCs) are suboptimal for applications in transplantation medicine and disease modeling because both the viral vectors used for gene transfer and the reprogramming factors they encode are oncogenic (Hacein-Bey-Abina et al., 2003; Nakagawa et al., 2008; Thrasher and Gaspar, 2007).

One potential solution is to identify small molecules that can efficiently reprogram cells and produce unmodified iPSC lines better suited for downstream applications as a result. Identification of such compounds would allow reprogramming that would not be impeded by the laborious nature of protein transduction or the safety concerns surrounding transgenic approaches (Kaji et al., 2009; Kim et al., 2009; Okita et al., 2008).

Several small molecules that catalyze reprogramming have already been described. Compounds that alter chromatin structure, including the DNA methyltransferase inhibitor 5-aza-cytidine (AZA) and the histone deacetylase (HDAC) inhibitor valproic acid (VPA), can increase reprogramming efficiency and even reduce the number of factors required for reprogramming (Huangfu et al., 2008a; Huangfu et al., 2008b; Mikkelsen et al., 2008; Shi et al., 2008b) . Treatment with these inhibitors presumably lowers the barrier to activation of endogenous pluripotency-associated genes. However, *Oct4* and *Sox2* not only activate genes required for pluripotency, they also function to repress genes promoting differentiation. It is therefore unlikely that this class of small molecules would be sufficient to completely replace the transgenic factors. As a result, there remains a need to identify novel small molecules that can function in reprogramming.

Here, we report the discovery of compounds that can replace the central reprogramming factor *Sox2*. We demonstrate that one of these chemicals specifically acts by inhibiting Tgf-β signaling. Interestingly, this compound does not act by inducing *Sox2* expression in the target fibroblasts. Instead, we show that it enables reprogramming through the induction of *Nanog* transcription in a stable, partially reprogrammed cell type that accumulates in the absence of *Sox2*.

RESULTS

A Screen for Chemical Mediators of Reprogramming

To identify small molecules that function in reprogramming, we transduced fibroblasts with viral vectors encoding *Oct4*, *Klf4*, and *cMyc* and then screened for compounds that allowed for reprogramming in the absence of *Sox2*. We favored this approach because it was unbiased with respect to the mechanism by which a given chemical could function and would not only deliver chemical compounds with translational utility but also provide novel insights into the mechanisms controlling reprogramming.

Activation of an *Oct4*::GFP reporter gene in colonies with an embryonic stem cell (ESC) morphology has been shown to be a stringent assay for reprogramming (Meissner et al., 2007). In

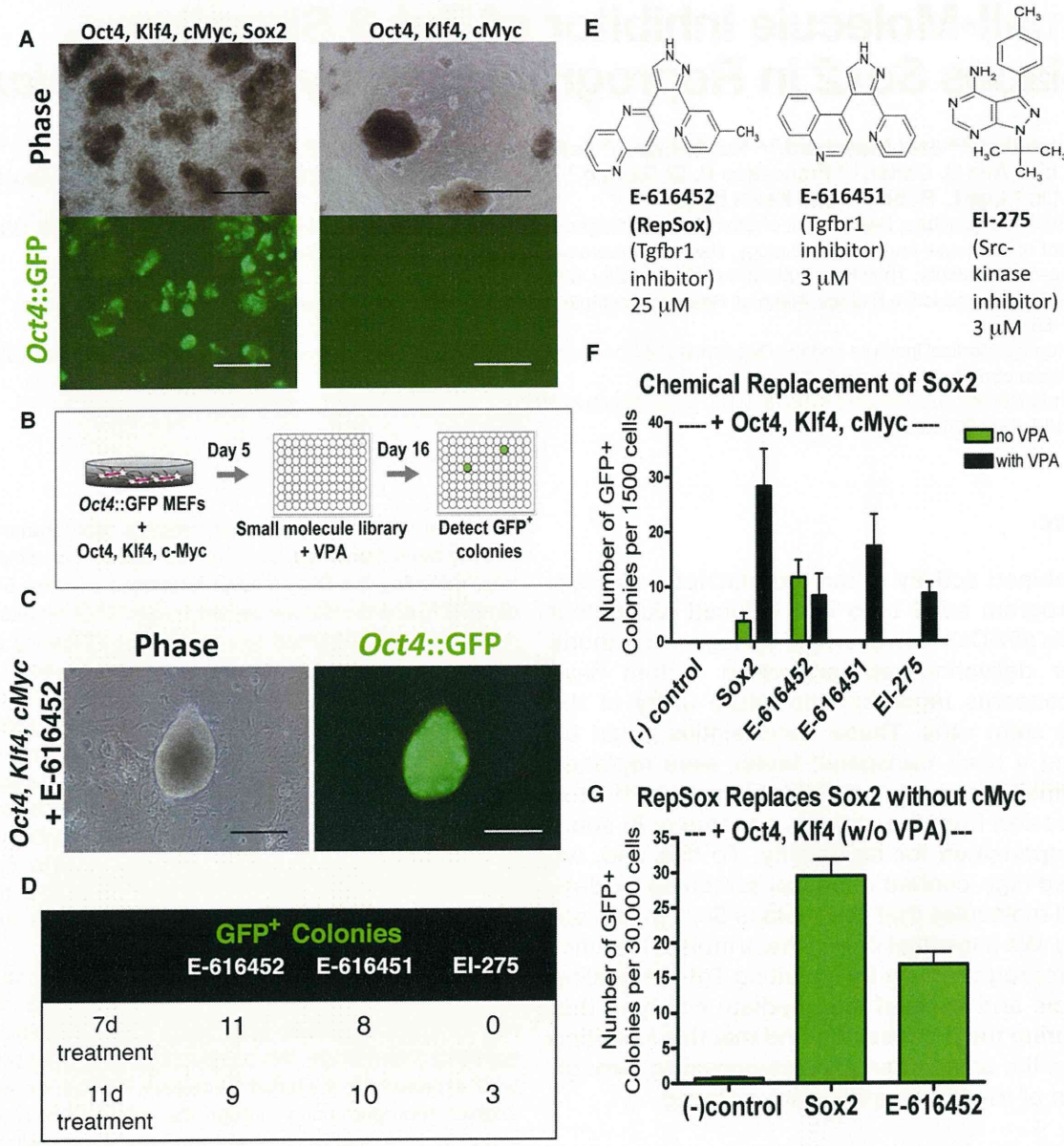


Figure 1. Identification of Small Molecules That Replace of *Sox2*

(A) *Oct4*::GFP$^+$ colonies form readily in *Oct4*, *Klf4*, *cMyc*, and *Sox2*-infected MEF cultures and do not form in *Oct4*-, *Klf4*-, and *cMyc*-infected MEF cultures. Scale bars represent 500 μm.

(B) Overview of chemical screen for replacement of *Sox2*.

(C) A P0 colony from *Oct4*-, *Klf4*-, and *cMyc*-infected MEFs + RepSox that displays a mESC-like morphology and is *Oct4*::GFP$^+$. Scale bars represent 200 μm.

(D) Number of *Oct4*::GFP$^+$ colonies detected for each hit in the primary screen after transduction of *Oct4*, *Klf4*, and *cMyc* and VPA treatment.

(E) Chemical structures of E-616452, E-616451, and EI-275, with the optimal concentrations for reprogramming listed.

(F) Quantification of small-molecule replacement of *Sox2* in *Oct4*-, *Klf4*-, and *cMyc*-infected MEFs with and without VPA treatment. The error bars denote the standard error derived from quantification of three separate wells of cells.

(G) *Sox2* replacement by RepSox is not dependent on *cMyc* (no VPA treatment). The error bars denote the standard error derived from quantification of three separate wells of cells.

mouse embryonic stem cell (mESC) culture medium supplemented with VPA, retroviral transduction of 7500 *Oct4*::GFP transgenic mouse embryonic fibroblasts (MEFs) with *Oct4*, *Klf4*, *cMyc*, and *Sox2* (Boiani et al., 2004) routinely generated 100–200 GFP$^+$ colonies (Figure 1A). In contrast, we observed no GFP$^+$ colonies when *Sox2* was omitted (Figure 1A). We used this robust difference to identify small molecules that can replace *Sox2*.

To facilitate the identification of cellular targets and signaling pathways affected by any compounds we discovered, we utilized a library of molecules with known pharmacological targets. We transduced *Oct4*::GFP MEFs with *Oct4*, *Klf4*, and *cMyc* and

then plated 2000 cells per well in a 96-well format. To each well, we added one of 800 distinct compounds for 7–11 days and also treated each well with 2 mM VPA for the first 7 days (Figure 1B). It was our hope that this approach would allow us to identify both compounds that required chromatin remodeling to induce reprogramming (Huangfu et al., 2008a) and compounds that did not. After 16 days, we scored each well for the presence of GFP$^+$ colonies with a mESC-like morphology (Figure 1C) and identified three independent hit compounds (Figure 1D). Two of these compounds were distinct transforming growth factor-β receptor 1 (Tgfbr1) kinase inhibitors (E-616452 and E-616451 [Figure 1E] [Gellibert et al., 2004]), whereas the third was a Src-family kinase inhibitor (EI-275 [Figure 1E] [Hanke et al., 1996]).

Efficient Small-Molecule Replacement of *Sox2*

Next, we optimized the effective concentration for each hit molecule (Figure S1 available online) and quantified the efficiency at which it synergized with VPA to replace *Sox2*. When 1500 MEFs were transduced with only *Oct4*, *Klf4*, and *cMyc* and then treated with VPA, we did not observe GFP$^+$ colonies (Figure 1F). However, the addition of E-616452 (25 μM), E-616451 (3 μM), or EI-275 (3 μM) led to the formation of GFP$^+$ colonies with an ESC morphology at a rate that was comparable to transduction with *Sox2* (Figure 1F).

Given that the three compounds were identified in the presence of VPA, we next determined whether these molecules were dependent on this HDAC inhibitor for their reprogramming activities. We found that E-616451 and EI-275 could not induce the appearance of GFP$^+$ colonies in the absence of VPA (Figure 1F), whereas E-616452 could do so and at a rate that was similar to a positive control transduced with the *Sox2* retrovirus (Figure 1F).

Although *cMyc* does increase the efficiency of reprogramming, it is not required for the generation of iPSCs (Nakagawa et al., 2008). Because the elimination of *cMyc* is an important step toward reducing the risk of tumor formation, we tested whether E-616452 could function in the absence of this oncogene. When added to MEFs transduced with only *Oct4* and *Klf4*, E-616452 induced the formation of GFP$^+$ colonies with an efficiency similar to viral *Sox2* (Figure 1G).

Previous reports on small molecules that affect reprogramming have focused on MEFs or neural stem cells (NSCs). These cells may be reprogrammed more easily because of either their proliferative capacity or their expression of iPSC factors (Huangfu et al., 2008a; Shi et al., 2008a; Shi et al., 2008b). However, it may be that chemical modulation of gene expression is cell-type specific, and we therefore determined whether the reprogramming compound we identified functioned in a more patient-relevant cell type. When we infected adult tail tip fibroblasts with *Oct4*, *Klf4*, and *cMyc* alone, we did not observe *Oct4*::GFP$^+$ colonies. However, when we added E-616452, we readily observed reprogramming (Figure S2A). The resulting *Oct4*::GFP$^+$ colonies could be expanded into cell lines that maintained homogeneous *Oct4*::GFP expression and self-renewed similarly to mESC and 4-factor control iPSC lines (Figure S2B). Because it could efficiently replace transgenic *Sox2* in the absence of VPA and *cMyc*, as well as in both embryonic and adult fibroblasts, we chose to further characterize E-616452 and named it RepSox, for Replacement of *Sox2*.

RepSox-Reprogrammed Cells Are iPSCs

Investigation of self-renewal capacity (Figure 2A), gene expression program, and pluripotency demonstrated that *Oct4::*GFP$^+$ cells induced by the RepSox replacement of *Sox2* were bona fide iPSCs. PCR with primers specific to the *Oct4*, *Klf4*, *cMyc*, and *Sox2* transgenes confirmed that this cell line did not harbor transgenic *Sox2* (Figure S3A). Chromosomal analysis indicated it was karyotypically normal (Figure S3B).

The *Oct4*::GFP$^+$ cells coexpressed alkaline phosphatase (Figure S3C) and the endogenous alleles of the *Nanog* and *Sox2* genes, suggesting pluripotency had been established (Figure 2B). The global transcriptional profile of cells reprogrammed with RepSox was similar to that of an iPSC line produced with all four transgenes and as similar to those of mESCs (Pearson correlation coefficient = 0.95–0.97) as two distinct mESC lines profiles were to each other (Pearson correlation coefficient = 0.96) (Figure 2C, Figure S3D, and Table S1). The profile differed significantly from that of the somatic MEFs (Figure 2C).

Cells produced with RepSox could readily form both embryoid bodies and teratomas that contained differentiated cell types of the three distinct embryonic germ layers (Figure 2E and Figure S4A). In addition, we observed that these cells could respond to directed differentiation signals in vitro and robustly differentiate into Hb9+/Tuj1+ motor neurons (Figure 2D and Figure S5).

In order to more definitively confirm the pluripotency of cells reprogrammed with RepSox, we tested their ability to contribute to chimeric embryos in vivo. We labeled cells with a lentiviral transgene encoding the dTomato red fluorescent protein and injected them into blastocysts. Both embryos and adult mice with significant contribution from the iPSCs were obtained (Figures 2F and 2G). Although adult mice with high contribution from the iPSCs were observed, we found it difficult to assess the contribution of these cells to the germline because the majority of animals developed tumors at or before the time of sexual maturity. However, we did observe that the reprogrammed cells could contribute *Oct4::*GFP$^+$ cells to the genital ridges of embryonic chimeras, demonstrating contribution of these pluripotent cells to the germline (Figure 2H). Together, these results demonstrate that the RepSox-reprogrammed cells are indeed iPSCs.

RepSox Can Replace *Sox2* and *c-Myc* by Inhibiting Tgf-β Signaling

Previous studies with RepSox suggest that it can act as an inhibitor of the Tgfbr1 kinase (Gellibert et al., 2004). Therefore, we investigated whether the mechanism by which RepSox functions to replace *Sox2* is through the inhibition of Tgf-β signaling. If Tgfbr1 is the functional target of RepSox, then a structurally unrelated inhibitor of Tgf-β signaling or depletion of Tgf-β ligands from the culture medium might also replace *Sox2*. The small molecule SB431542 (Figure 3A) is known to inhibit Tgfbr1 kinase and is structurally distinct from RepSox (Inman et al., 2002). When we treated fibroblasts transduced with *Oct4*, *Klf4*, and *cMyc* with 25 μM SB431542, we observed ~10 GFP$^+$ colonies per 7500 cells plated (Figure 3B). Likewise, when we transduced fibroblasts in the presence of either an antibody that neutralized a variety of Tgf-β ligands (R&D Systems, AB-100-NA) or an antibody specific to Tgf-β II (R&D Systems, AB-12-NA), *Oct4*::GFP$^+$ colonies were generated (Figure 3B). In contrast, we observed no GFP$^+$ colonies in transductions without these Tgf-β inhibitors.

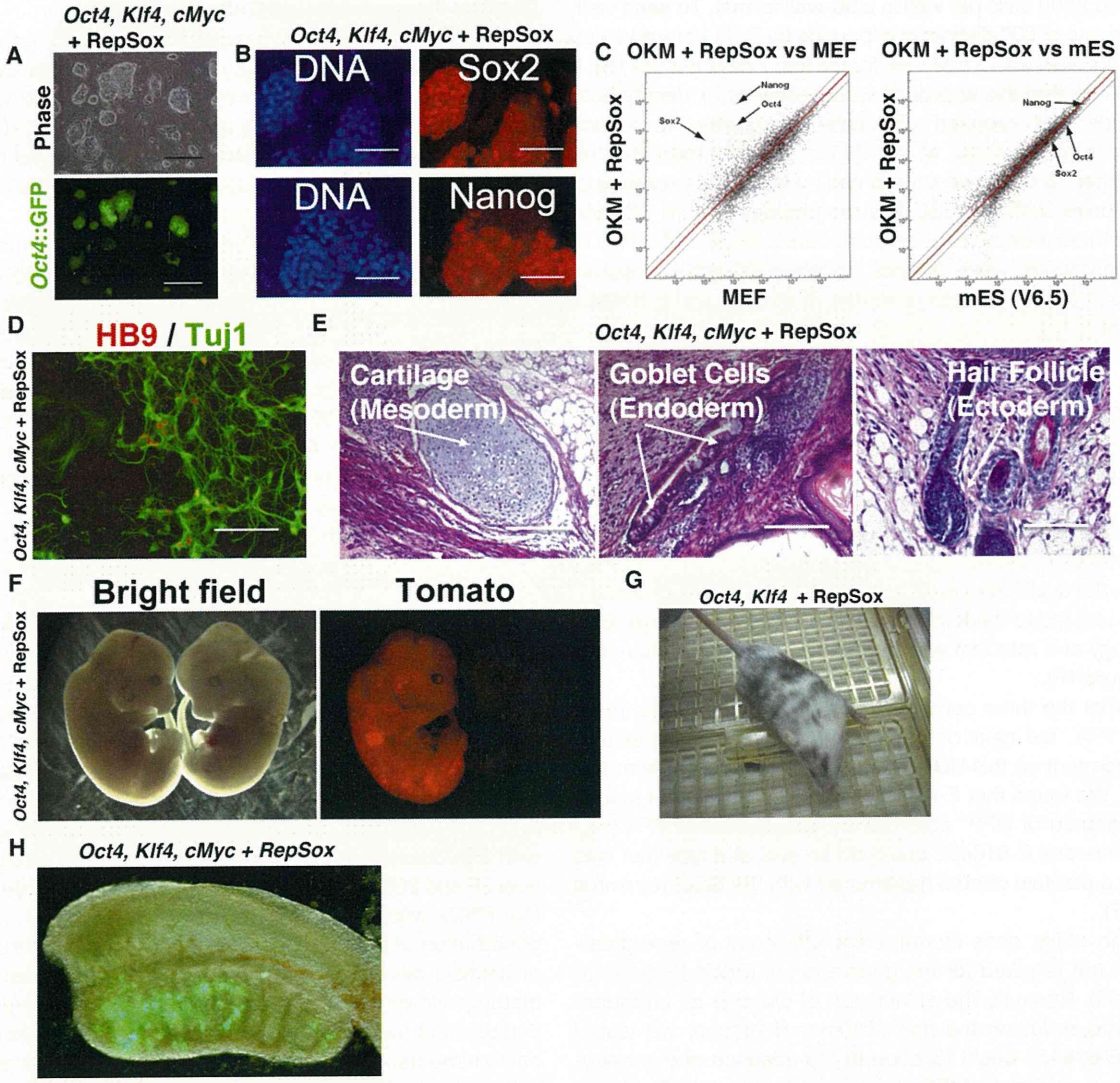


Figure 2. RepSox-Reprogrammed Cells Are Pluripotent

(A) An *Oct4*::GFP$^+$ iPSC line that was derived from a culture of RepSox-treated *Oct4*, *Klf4*, and *cMyc*-infected MEFs (OKM + RepSox line 1) displays the characteristic mESC-like morphology and self-renewal properties. Passage 11 is shown. Scale bars represent 500 μm.

(B) Antibody staining of OKM + RepSox line 1 cells shows that they express markers of pluripotent stem cells Sox2 and Nanog. Scale bars represent 100 μm.

(C) Microarray scatter plots showing that the global gene expression profile of OKM + RepSox line 1 is highly similar to that of mESC line V6.5 and very different from that of somatic MEFs.

(D) Motor neurons differentiated in vitro from OKM + RepSox line 1. The scale bar represents 200 μm.

(E) Teratomas containing cells of all three germ layers formed by injection of OKM + RepSox line 1 cells into nude mice.

(F) E12.5 chimeric mouse embryo (left, versus nonchimeric littermate on the right) showing a high amount of contribution from OKM + RepSox line 1 cells constitutively expressing the dTomato red fluorescent protein.

(G) Eight-week-old chimeric mouse formed by injection of OK + RepSox line 1 cells (C57BL6 genetic background) into an ICR blastocyst.

(H) *Oct4*::GFP$^+$ cells derived from an OKM + RepSox cell line are present in the genital ridge of a male embryo at 13.5 days postcoitum (dpc).

These results are consistent with the notion that at least part of the mechanism by which RepSox replaces *Sox2* in reprogramming is through the inhibition of Tgf-β signaling.

Our goal was to identify molecules that specifically replace *Sox2* instead of generally increasing reprogramming efficiency. If RepSox acts specifically to replace *Sox2*, then we would not expect it to stimulate reprogramming in the presence of transgenic *Sox2*. When RepSox- or Tgf-β antibody-treated MEFs were transduced with *Oct4*, *Klf4*, *cMyc*, and *Sox2*, we observed less than a 2-fold increase in the number of GFP$^+$ colonies over the untreated controls (Figures 3C and 3D). The magnitude by which RepSox stimulated reprogramming in this context was significantly less than the 10-fold increase that we observed after treatment with VPA, a compound thought to increase reprogramming efficiency (Figure 1F).

In order to further investigate the specificity of *Sox2* replacement by RepSox, we tested the ability of this molecule to individually replace *Oct4*, *Klf4*, and *cMyc* in reprogramming. We found

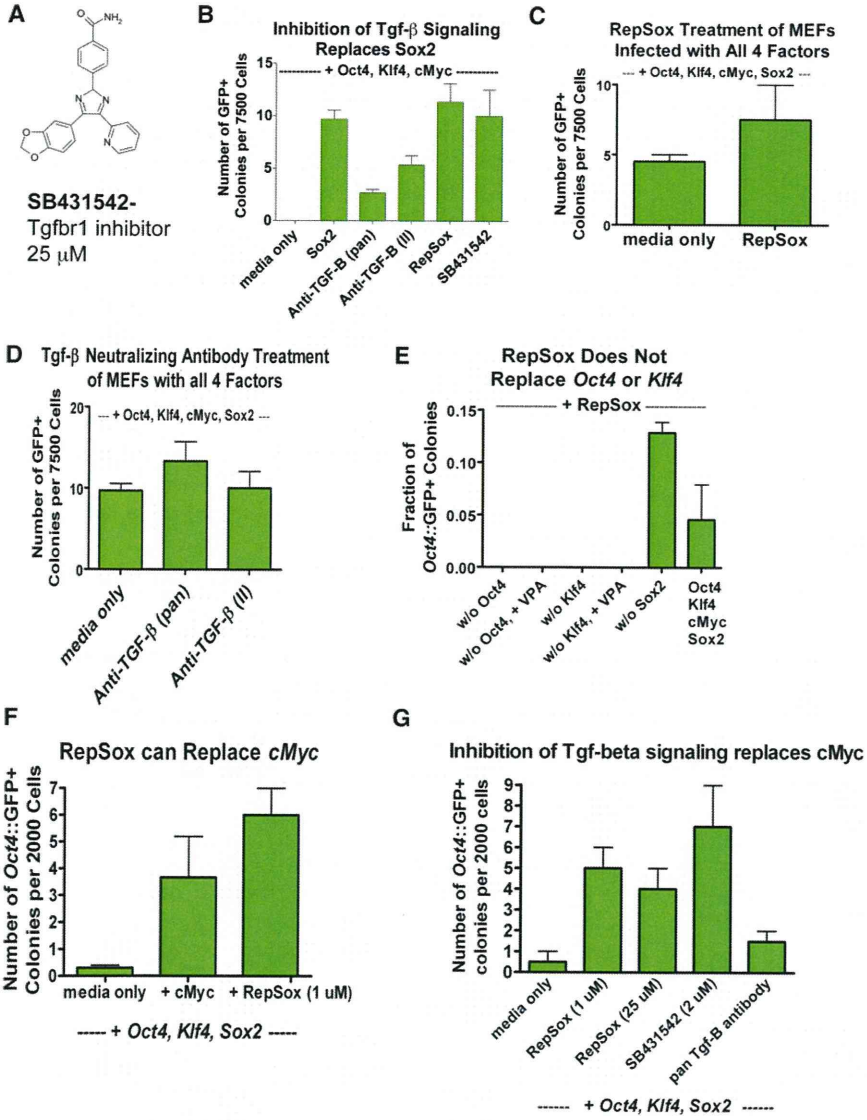


Figure 3. RepSox Specifically Replaces *Sox2* by Inhibiting Tgf-β Signaling

(A) Chemical structure of SB431542, an inhibitor of Tgfbr1 activity.

(B) Inhibition of Tgf-β signaling by treatment of *Oct4*-, *cMyc*-, and *Sox2*-infected MEFs with SB431542 or TGF-β neutralizing antibodies replaces *Sox2*.

(C) RepSox does not increase the efficiency of *Oct4*::GFP$^+$ colony induction in *Oct4*-, *Klf4*-, *cMyc*-, and *Sox2*-infected MEFs. The error bars denote the standard error derived from quantification of three separate wells of cells.

(D) Inhibition of Tgf-β signaling by TGF-β neutralizing antibodies does not increase the efficiency of *Oct4*::GFP$^+$ colony induction in *Oct4*-, *Klf4*-, *cMyc*-, and *Sox2*-infected MEFs. The error bars denote the standard error derived from quantification of three separate wells of cells.

(E) RepSox does not replace transgenic *Oct4* or transgenic *Klf4* in reprogramming. We observed no *Oct4*::GFP$^+$ colonies in RepSox-treated *Klf4*-, *cMyc*-, and *Sox2*-infected MEFs or *Oct4*-, *cMyc*-, *Sox2*-infected MEFs out of 30,000 cells plated both with and without VPA treatment. We routinely observe 30–40 *Oct4*::GFP$^+$ colonies when we plate the same number of *Oct4*-, *Klf4*-, and *cMyc*-infected MEFs and treat them with RepSox. The error bars denote the standard error derived from quantification of three separate wells of cells.

(F) RepSox can replace *cMyc* in reprogramming. Cells were transduced with *Oct4*, *Klf4*, and *cMyc* and treated with RepSox continuously starting at day 5 postinfection. The error bars denote the standard error derived from quantification of two separate wells of cells.

(G) Inhibition of Tgf-β signaling can replace *cMyc* in reprogramming. Cells were transduced with *Oct4*, *Klf4*, and *cMyc* and treated with inhibitors of Tgf-β signaling continuously starting at day 5 postinfection. The error bars denote the standard error derived from quantification of two separate wells of cells.

that RepSox could not induce GFP$^+$ colonies in the absence of either *Oct4* or *Klf4*, even in the presence of VPA (Figure 3E). In contrast, we found that RepSox did increase the number of *Oct4*::GFP$^+$ colonies by 20-fold in the absence of *cMyc*, thereby fully replacing it in reprogramming (Figure 3F). In addition, the structurally distinct Tgf-β inhibitor SB431542 and a Tgf-β-specific neutralizing antibody both increased reprogramming efficiency in the absence of *cMyc* (Figure 3G). From these experiments, we conclude that RepSox enables the replacement of the reprogramming activities provided by both transgenic *Sox2* and *cMyc*. In both cases, these complementing activities seem to be mediated through the inhibition of Tgf-β signaling.

RepSox Replaces *Sox2* by Acting on Intermediates Formed During the Reprogramming Process

The development of cocktails of small molecules that can effectively reprogram somatic cells may require a detailed knowledge of the mechanism and kinetics by which each compound acts. Therefore, we determined the optimal duration of time by which inhibition of Tgf-β signaling with RepSox can help induce reprogramming.

Initially, we pretreated MEFs with RepSox, applying the chemical for 3 days, and then removed it at the time of transduction with *Oct4*, *Klf4*, and *cMyc*. In these experiments, no *Oct4*::GFP$^+$ colonies were formed (Figure 4A), suggesting that RepSox does not act on the initial somatic cells to replace *Sox2*. Consistent with this result, we did not detect a significant increase in the expression of endogenous *Sox2* or closely related *Sox* family members upon RepSox treatment (Figure S6A). In addition, RepSox treatment did not decrease the expression of the mesenchymal gene *Snai1* (Figure S6B), which is downregulated 5- to 40-fold by transduction of the four reprogramming factors (Mikkelsen et al., 2008). Thus, RepSox does not destabilize the pre-existing MEF transcriptional program.

In contrast, we found that RepSox did increase by 5-fold the expression of *L-Myc*, a close homolog of *cMyc* that can functionally replace it in reprogramming (Nakagawa et al., 2008) (Figure S6C). Together, these data suggest that although RepSox

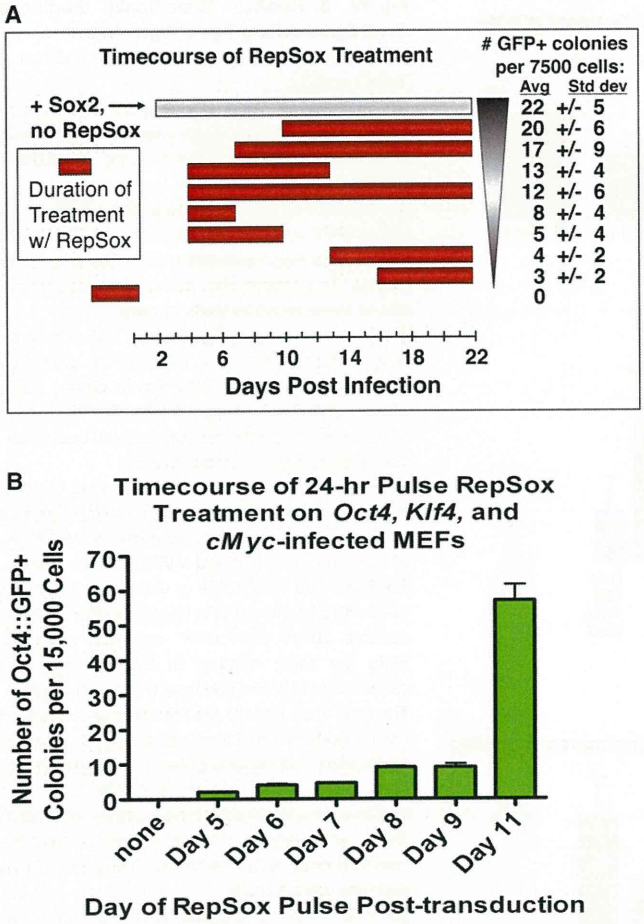


Figure 4. A Short Pulse of RepSox Is Sufficient for *Sox2* Replacement and Most Effective at Later Time Points after Infection

(A) Graph showing the number of *Oct4*::GFP$^+$ colonies induced by various timings of RepSox treatment of *Oct4*-, *Klf4*-, and *Sox2*-infected MEFs in mESC medium. Colonies were counted at 24 days postinfection. Shown are average colony numbers ± the standard deviation. The standard deviation was derived from quantification of two separate wells.

(B) Time course of RepSox treatment showing the number of *Oct4*::GFP$^+$ colonies induced by a 24 hr pulse of RepSox on *Oct4*-, *Klf4*-, and *Sox2*-infected MEFs in serum-free mESC medium with knockout serum replacement (KSR mESC). Colonies were counted at 24 days postinfection. The error bars denote the standard error derived from quantification of three separate wells of cells.

probably functions at the level of the initial somatic cell population to replace *cMyc*, it does not act on the starting MEF population to replace *Sox2*.

Because RepSox did not seem to act directly on the fibroblasts to replace *Sox2*, we investigated whether it functioned on intermediates that arose during reprogramming. To address this question, we varied both the duration and timing of RepSox treatment in order to determine when it was most effective. First, we transduced 7500 MEFs with *Oct4*, *Klf4*, and *cMyc*, waited for 4 days, and subsequently treated cultures with RepSox for 3, 6, 9, or 18 additional days. Although a short 3 day treatment from days 4–7 induced a small number of *Oct4*::GFP$^+$ colonies, the 9-day treatment from days 4–13 yielded the most *Oct4*::GFP$^+$ colonies (Figure 4A).

Next, we varied the timing at which we initiated RepSox treatment, by administering the compound beginning at day 4, 7, 10, 13, or 16 after transduction. We found that delaying the start of RepSox treatment increased its reprogramming potency, with optimal treatment beginning at 10 days posttransduction (Figure 4A). Together, these results suggest that RepSox treatment is most effective between days 7–12 posttransduction.

To more precisely define the optimal treatment window, we determined the minimal duration of treatment required to induce reprogramming. We found that a treatment as short as only 1 day was sufficient for inducing detectable reprogramming (Figure 4B). Delaying this short treatment yielded more reprogrammed colonies, with a sharp increase at day 11 (Figure 4B). These results indicate that RepSox is most effective at replacing *Sox2* during days 10–11 after transduction and that therefore cultures of *Oct4*, *Klf4*, and *cMyc*-transduced MEFs give rise to intermediates capable of responding to RepSox treatment. These intermediates appear at day 4 posttransduction and peak at days 10–11.

Interestingly, when we tracked the timing of the initial appearance of reprogrammed colonies as a function of the timing of RepSox administration, we found that regardless of whether we began treatment at day 7 or day 10 posttransduction, *Oct4*::GFP$^+$ colonies first appeared at day 14 (Figure S7). This suggests that RepSox may not always be the rate-limiting step in this reprogramming process and that other, RepSox-independent events take place during the formation of the RepSox-responsive intermediates.

RepSox-Responsive Cell Lines

Our finding that a 24 hr pulse of RepSox can replace *Sox2* (Figure 4B) differs strikingly from the 5–10 day period of transgene expression normally required (Sridharan et al., 2009; Wernig et al., 2007) and suggests that RepSox could trigger a switch activating reprogramming. If RepSox acts to flip a switch in semi-stable intermediate cell types that accumulate in the absence of retroviral *Sox2* expression, we reasoned that it might also be possible to culture these responsive intermediates for prolonged periods of time. In contrast, if RepSox acts during a critical window on very transient intermediates, this might not be possible. To distinguish between these models, we transduced *Oct4*::GFP MEFs with *Oct4*, *Klf4*, and *cMyc*, waited 10-14 days, and subsequently clonally expanded ten iPSC-like, GFP-negative colonies (Figure 5A). These cell lines continued to proliferate for at least four passages and often maintained an iPSC-like morphology (Figure 5A), but never further activated expression of *Oct4*::GFP. However, when we treated these cell lines with a 48 hr pulse of RepSox, 5%–10% of the colonies in two of the ten lines became *Oct4*::GFP$^+$ (Figures 5A and 5B). These results demonstrate that partially reprogrammed cells can accumulate in the absence of *Sox2* and that some, but not all, of these cells can be clonally expanded and cultured for prolonged periods while maintaining responsiveness to RepSox.

As we had shown that this particular reprogramming molecule seems to replace *Sox2* through the inhibition of Tgf-β signaling, we sought to determine whether RepSox treatment affected Tgf-β signal transduction pathways in these responsive cell lines. To this end, we determined the levels of phosphorylated Smad3 by western blotting in cell line OKM 10 both with and without

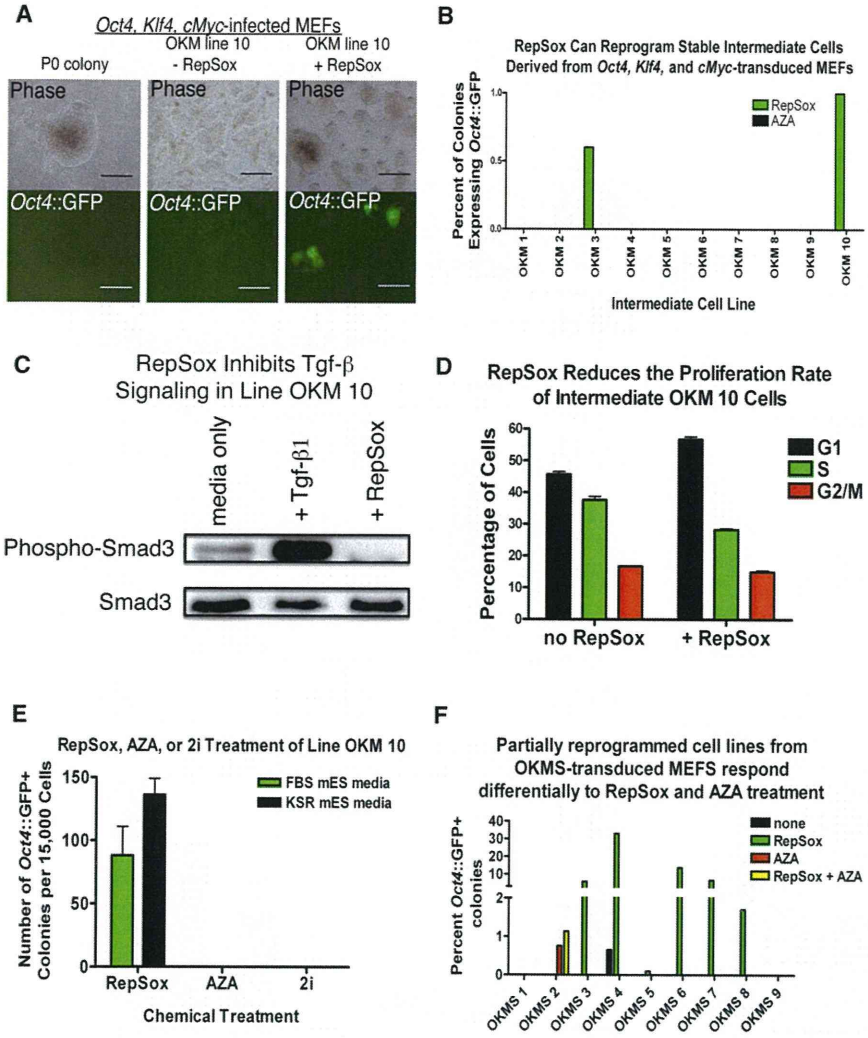


Figure 5. Stable Intermediates Can Be Reprogrammed by RepSox

(A) Stable *Oct4*::GFP-negative cell lines derived from *Oct4*::GFP-negative colonies in *Oct4*-, *Klf4*-, and *cMyc*-infected MEF cultures can be reprogrammed by RepSox. Scale bars in "OKM line 10 + RepSox" panels represent 500 μm; all other scale bars represent 200 μm.

(B) Two of ten stable, nonpluripotent intermediate cell lines derived from MEFs transduced with *Oct4*, *Klf4*, and *cMyc* can be reprogrammed with RepSox treatment, but none can be reprogrammed with AZA treatment.

(C) Western blot for phospho-Smad3 showing that RepSox inhibits Tgf-β signaling in line OKM 10 (OKM 10) cells.

(D) RepSox does not increase the proliferation of OKM 10 cells. The error bars denote the standard error derived from quantification of two separate wells of cells.

(E) Line OKM 10 can be reprogrammed with RepSox treatment but not with AZA or 2i, indicating it is distinct from cell lines that can be reprogrammed by AZA or 2i. The error bars denote the standard error derived from quantification of three separate wells of cells.

(F) Stable *Oct4*::GFP-negative cell lines derived from *Oct4*::GFP-negative colonies in *Oct4*-, *Klf4*-, *cMyc*-, and *Sox2*-infected MEF cultures can be reprogrammed by RepSox or by AZA, but lines responsive to RepSox are not responsive to AZA alone and lines responsive to AZA are not responsive to RepSox alone, indicating the presence of two different types of stable intermediates in the reprogramming cultures.

RepSox treatment. Without RepSox treatment, we detected relatively high levels of phosphorylated Smad3, suggesting that Tgf-β signaling was active (Figure 5C). In contrast, treatment with 25 μM RepSox almost completely eliminated Smad3 phosphorylation (Figure 5C), indicating that RepSox strongly inhibited Tgf-β signaling in these cells.

Because an increase in cell proliferation can also increase reprogramming efficiency (Hong et al., 2009) and possibly contribute to the replacement of transgenic *Sox2*, we measured the proliferation rate of partially reprogrammed OKM 10 cells both with and without RepSox. Treatment with RepSox decreased the proportion of cells in G2/M phase of the cell cycle (Figure 5D), indicating it does not increase the proliferation rate of these partially reprogrammed cells.

Cells that Respond to RepSox Treatment Are Distinct from Previously Described Intermediates

It has been shown that certain nonpluripotent, partially reprogrammed cell lines derived from MEFs transduced with *Oct4*, *Klf4*, *cMyc*, and *Sox2* can be fully reprogrammed with AZA or a combination of chemical inhibitors of glycogen synthase kinase 3β (GSK-3β) and the Mek signaling pathway (2i conditions) (Mikkelsen et al., 2008; Silva et al., 2008). If the RepSox-responsive cell lines generated by overexpression of *Oct4*, *Klf4*, and *cMyc* are similar to these four-factor cell lines, then they should also be reprogrammed by AZA or 2i. However, when we treated the ten stable intermediate lines with either AZA or 2i for 48 hr, we found that none became reprogrammed (Figure 5B), indicating that the RepSox-responsive stable intermediates are distinct from partially reprogrammed cell lines described previously (Mikkelsen et al., 2008; Silva et al., 2009). Consistent with these results, in vitro assays of kinase activity revealed that RepSox does not inhibit the targets of the 2i cocktail (Table S2).

It occurred to us that some nonpluripotent cells derived from MEFs transduced with *Oct4*, *Klf4*, *cMyc*, and *Sox2* could potentially be held in a nonpluripotent state because of inappropriate levels of transgene expression and therefore might also be responsive to RepSox treatment. To test this hypothesis, we transduced *Oct4*::GFP MEFs with *Oct4*, *Klf4*, *cMyc*, and *Sox2*, then picked and clonally expanded nine GFP-negative colonies at day 14 after transduction (Figure S8). After treatment with RepSox, five of the nine cell lines yielded reprogrammed colonies, with 2%–33% of the colonies in each line becoming *Oct4*::GFP$^+$ (Figure 5F and Figure S8). These results indicate

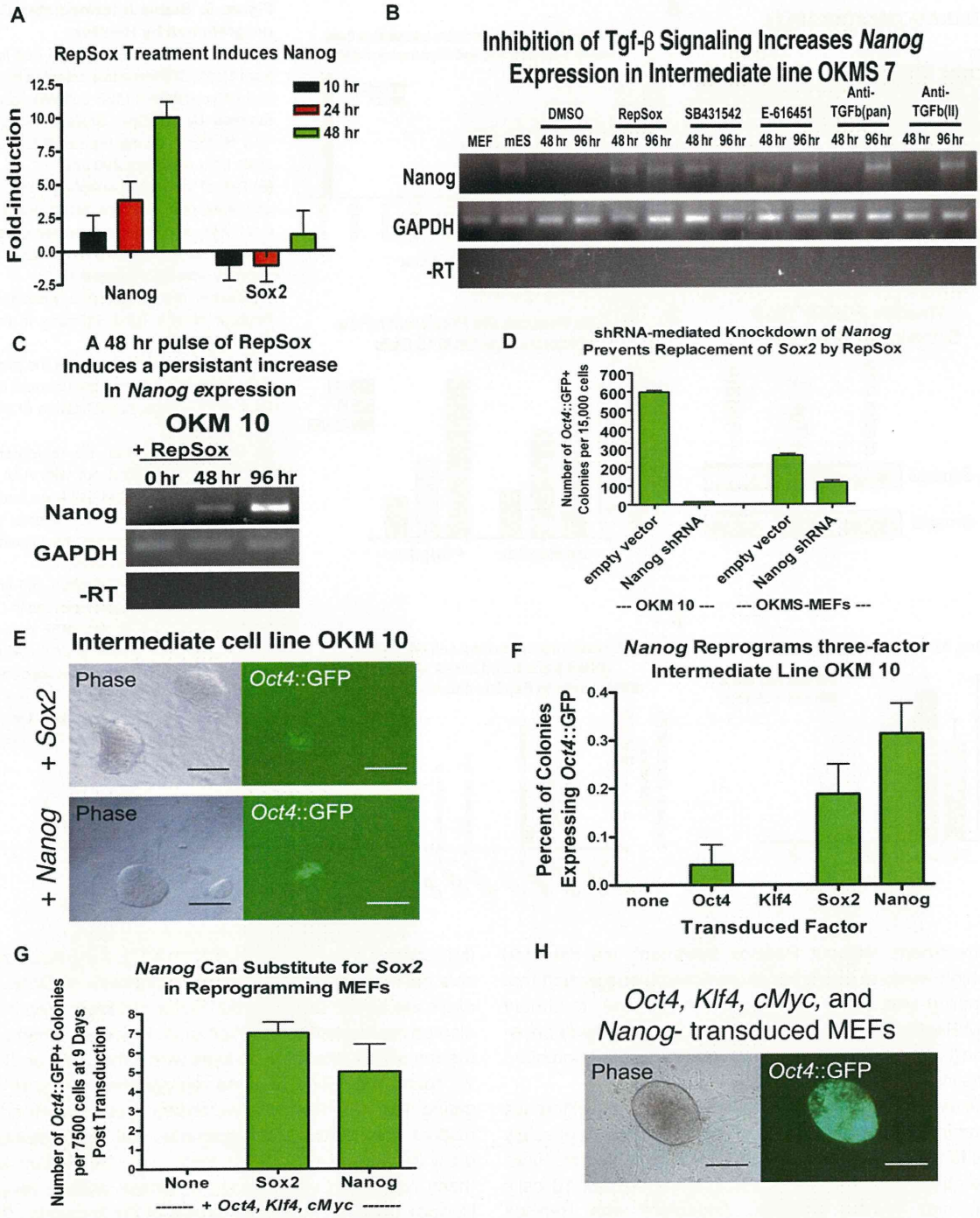


Figure 6. RepSox Replaces *Sox2* by Inducing *Nanog* Expression

(A) RepSox treatment of RepSox-responsive line OKMS 6 strongly increases *Nanog* mRNA levels. Data were generated by microarray analysis and are relative to untreated controls. *Nanog* is induced more quickly and more significantly than *Sox2*, indicating it is upregulated before fully reprogrammed cells form. The error bars denote the standard error derived from quantification of three separate wells of cells.

(B) RT-PCR analysis showing that inhibition of Tgf-β signaling increases *Nanog* expression in the RepSox-responsive intermediate line OKMS 7.

(C) A pulse of RepSox induces a persistent increase in *Nanog* expression in the RepSox-responsive intermediate line OKM 10. OKM 10 cells were treated with 25 μM RepSox for 48 hr and RNA samples were taken at 0, 48, and 96 hr (48 hr after removal of RepSox) and analyzed by RT-PCR.

(D) shRNA-mediated knockdown of *Nanog* in OKM 10 cells inhibits replacement of *Sox2* by RepSox. The error bars denote the standard error derived from quantification of two separate wells of cells.

(E) Pictures of reprogrammed *Oct4*::GFP$^+$ colonies induced by *Sox2* (A) or *Nanog* (B) transduction of line OKM 10. Scale bars represent 200 μm.

(F) *Nanog* transduction can reprogram line OKM 10 at an efficiency similar to that of *Sox2* transduction. The error bars denote the standard error derived from quantification of three separate wells of cells.

that like the stable intermediate cells generated with only *Oct4*, *Klf4*, and *cMyc*, certain incompletely reprogrammed cells generated by *Oct4*, *Klf4*, *cMyc*, and *Sox2* transduction can also be reprogrammed by RepSox.

Next, in order to determine whether these RepSox-responsive intermediate cell lines derived after *Oct4*, *Klf4*, *cMyc*, and *Sox2* transduction were similar to or distinct from previously described partially reprogrammed cell lines (Mikkelsen et al., 2008), we applied AZA to all nine lines. After 48 hr of AZA treatment and 12 subsequent days in culture, none of the RepSox-responsive cell lines expressed *Oct4*::GFP (Figure 5F). However, one of the lines that had been refractory to RepSox treatment did express *Oct4*::GFP after AZA treatment, indicating that it had undergone complete reprogramming (Figure 5F). Together, these results show that there are a variety of intermediates that can form after retroviral transduction and that they vary in their responsiveness to reprogramming molecules.

RepSox Replaces *Sox2* by Inducing *Nanog* Expression

The causal molecular events that drive reprogramming are difficult to detect because of the low efficiency at which somatic cells are successfully reprogrammed (Amabile and Meissner, 2009). However, when we administered RepSox to cell lines that had been partially reprogrammed by retroviral transduction, *Oct4*::GFP expression was induced in up to 33% of the resulting colonies (Figure 5F). We used this more efficient reprogramming system to identify the changes in gene expression induced by RepSox that enable it to bypass the requirement for transgenic *Sox2* expression.

We treated an *Oct4*::GFP-negative, partially reprogrammed cell line (OKMS 6) with RepSox and performed global gene expression analysis at 10, 24, and 48 hr after the initiation of treatment. To confirm that RepSox was inhibiting Tgf-β signaling in this intermediate cell line, we investigated expression changes in known Tgf-β-responsive genes after RepSox treatment. The *Inhibition of Differentiation* genes *Id1*, *Id2*, and *Id3* are repressed by Tgf-β signaling in mESCs (Ying et al., 2003). After treating the RepSox-responsive intermediate line OKM 10 with RepSox for 24 hr, we observed increased expression of *Id1*, *Id2*, and *Id3* (Figure S9A).

One way that RepSox could function to replace transgenic *Sox2* would be to induce the expression of endogenous *Sox2* or a *Sox* family member, such as *Sox1* or *Sox3*, that can substitute for it in reprogramming (Nakagawa et al., 2008). However, we again did not observe a significant increase in the expression of *Sox1*, *Sox2*, *Sox3*, or any of the remaining *Sox* family transcription factors within the first 48 hr of RepSox treatment (Figure S9B). Additionally, shRNA-mediated depletion of *Sox1*, the most potent *Sox* family member other than *Sox2* itself (Nakagawa et al., 2008), did not affect the rate of reprogramming in the presence of RepSox (Figure S9C). These results show that RepSox does not replace *Sox2* by directly activating endogenous *Sox2* or other closely related genes.

Next, we more broadly investigated changes in transcription factor expression after chemical treatment. We did not observe an increase in endogenous *Oct4* or *Klf4* expression at early time points after RepSox treatment. However, we found that the expression of the homeodomain factor *Nanog* was among the most increased after RepSox treatment. Relative to untreated controls, *Nanog* transcription increased 4-fold within 24 hr and 10-fold after 48 hr of RepSox treatment (Figure 6A). In contrast, we did not observe a rapid increase in *Nanog* expression in two *Oct4*::GFP-negative intermediate cell lines that could not be fully reprogrammed with RepSox (Figure S10). Therefore, we hypothesized that RepSox might replace *Sox2* by inducing *Nanog* expression.

Because we had determined that inhibition of Tgf-β signaling by several different small molecules and antibodies can replace *Sox2*, we reasoned that if the increase in *Nanog* expression was critical for *Sox2* replacement, the alternative inhibitors of Tgf-β signaling should also upregulate *Nanog*. To test this hypothesis, we treated the RepSox-responsive cell lines with RepSox, SB431542, or neutralizing antibodies and analyzed *Nanog* expression after 48 hr. In all cases, *Nanog* expression was strongly induced within 48–96 hr (Figure 6B).

If RepSox functions by increasing *Nanog* expression, then a short pulse of RepSox should induce a persistent increase in *Nanog* expression. To test this, we treated the RepSox-responsive intermediate cell line OKM 10 with RepSox for 48 hr, withdrew RepSox, and analyzed *Nanog* expression 48 hr later. A control time point taken just before RepSox withdrawal showed a significant increase in *Nanog* transcription (Figure 6C). Forty-eight hours after RepSox removal (96 hr after the initiation of treatment), *Nanog* expression continued to increase (Figure 6C).

If RepSox replaces *Sox2* by increasing *Nanog* expression, then a forced reduction of *Nanog* expression should inhibit or even prevent reprogramming by RepSox. To test this hypothesis, we transduced the RepSox-responsive cell line with a lentivirus encoding a short-hairpin RNA specific for *Nanog*. The *Nanog*-knockdown cells reprogrammed at a frequency that was 50-fold lower than cells transduced with an empty control vector (Figure 6D). This effect was not caused by a general decrease in reprogramming efficiency or differentiation of reprogrammed cells due to *Nanog* depletion because MEFs transduced with *Oct4*, *Klf4*, *cMyc*, *Sox2*, and the *Nanog* shRNA construct only suffered a 50% loss in reprogramming efficiency (Figure 6D). These results demonstrate that increased *Nanog* expression in this context was only necessary for the replacement of *Sox2* by RepSox.

Previous reports have shown that chemical inhibition of Tgf-β signaling by SB431542 increases bone morphogenetic protein (Bmp) signaling in embryonic stem cells (Xu et al., 2008). It has separately been shown that Bmp signaling in the presence of Stat3 induces *Nanog* expression in mESCs (Suzuki et al., 2006). The crosstalk between the Tgf-β and Bmp signaling pathways may be the result of a common requirement for Smad 4, which mediates transcriptional events in the nucleus (Attisano and Wrana, 2002). Similarly, we observed an increase in the levels of phosphorylated Smad1 protein and *Bmp-3* mRNA in incompletely reprogrammed intermediates after RepSox

(G) *Nanog* can substitute for *Sox2* in defined-factor reprogramming of somatic fibroblasts. The error bars denote the standard error derived from quantification of three separate wells of cells.

(H) Picture of a reprogrammed *Oct4*::GFP$^+$ colony induced by *Oct4*, *Klf4*, *cMyc*, and *Nanog* transduction of MEFs. Scale bars represent 100 μm.

treatment (Figure S11). Furthermore, the stable, partially reprogrammed cells that responded to RepSox expressed the LIF receptor at levels equivalent to those found in mESCs (Figure S12A). Expression of this receptor suggests that its downstream signal transduction pathway could be active in these cells, thereby resulting in the presence of activated Stat3, which is known to induce *Nanog* expression in conjunction with Bmp signaling.

Because RepSox does not act on the initial population of fibroblasts to replace *Sox2*, we would not expect *Nanog* to be upregulated in RepSox-treated MEFs. Indeed, within 7 days of transduction of MEFs with *Oct4*, *Klf4*, and *cMyc*, we did not observe an increase in *Nanog* expression upon RepSox treatment (Figure S12B). This may be explained in part by the observation that the LIF receptor, and thus activated Stat3, was not highly expressed in these cells (Figure S12A). Because *Nanog* plays a key role in maintaining ESCs in an undifferentiated state (Chambers et al., 2003) and has been shown to enhance the efficiency of reprogramming (Silva et al., 2006; Silva et al., 2009; Yu et al., 2007), we decided to test whether *Nanog* could directly replace *Sox2* in reprogramming.

If RepSox replaces *Sox2* by inducing *Nanog* expression, then retroviral transduction of RepSox-responsive intermediate cells (line OKM 10, Figures 5A and 5B) with *Nanog* should reprogram them. When we transduced line OKM 10 with *Sox2* as a control, 0.2% of the colonies expressed *Oct4*::GFP after 10 days, indicating that reprogramming could be induced in this cell line by *Sox2* (Figures 6E and 6F). When we transduced the same stable intermediate cell line with *Nanog*, it could also be reprogrammed, with 0.3% of the colonies expressing *Oct4*::GFP$^+$ after 10 days (Figures 6E and 6F). In contrast, transductions with *Oct4* or *Klf4* resulted in only 0.04% and 0% reprogramming efficiencies (Figure 6F). These results suggest that *Nanog* can indeed functionally replace *Sox2* and induce reprogramming in these stable intermediates formed from *Oct4*-, *Klf4*-, and *cMyc*-transduced MEFs.

If *Nanog* can compensate for the omission of *Sox2* in defined-factor reprogramming, then MEFs transduced with *Oct4*, *Klf4*, *cMyc*, and *Nanog* might be as efficiently reprogrammed as MEFs transduced with *Oct4*, *Klf4*, *cMyc*, and *Sox2*. When we transduced MEFs with *Oct4*, *Klf4*, *cMyc,* and *Sox2* and scored cultures 9 days later, an average of seven *Oct4*::GFP$^+$ colonies appeared for every 7500 cells plated (Figure 6G). A control transduction with only *Oct4*, *Klf4*, and *cMyc* yielded no *Oct4*::GFP$^+$ colonies (Figure 6G). Similar to the positive control transduction, MEFs transduced with *Oct4*, *Klf4*, *cMyc*, and *Nanog* gave rise to an average of five *Oct4*::GFP$^+$ colonies for every 7500 cells plated (Figures 6G, H). These colonies could be picked and expanded and remained *Oct4*::GFP$^+$ for at least five passages (Figure S13A). Immunocytochemistry indicated that these cells strongly activated *Sox2* expression from the endogenous allele (Figure S13B). Importantly, QPCR analysis demonstrated that they also transcribed endogenous *Oct4*, *Klf4*, *Nanog*, and *Rex1* (Figure S13C), indicating that a pluripotent gene expression program had been established. Furthermore, transgene-specific QPCR analysis showed that these cells had silenced the retroviral *Oct4*, *Klf4*, and *cMyc* transgenes (Figure S13D). Additionally, *Oct4*-, *Klf4*-, *cMyc*-, and *Nanog*-reprogrammed cells could readily form embryoid bodies in vitro (Figure S13E). However, we found that leaky expression of transgenic *Nanog*, which is a potent inhibitor of ESC differentiation (Chambers et al., 2003; Chambers et al., 2007), reduced the amount of differentiation in vitro (Figure S13D). We anticipate that efficient differentiation of cells created with *Oct4*, *Klf4*, *cMyc*, and *Nanog* will eventually require the use of an excisable transgenic *Nanog* cassette to completely remove ectopic *Nanog* expression. Although definitive proof of the pluripotency of these cells will be required to conclude that *Nanog* expression is sufficient for replacing *Sox2* in defined factor reprogramming, our results suggest that this may be the case. Taken together, our results demonstrate that RepSox inhibition of Tgf-β signaling bypasses the need for *Sox2* in defined-factor reprogramming through the induction of *Nanog*.

DISCUSSION

We have used a phenotypic chemical screen to identify compounds that can replace the reprogramming transcription factor *Sox2* and have confirmed the mechanism by which the most potent compound acts: RepSox replaces *Sox2* by inhibiting the broadly expressed Tgf-β signaling pathway (Attisano and Wrana, 2002) in cultures containing stable intermediate cells that are trapped in a partially reprogrammed state. This inhibition in turn leads to sustained transcription of *Nanog*, through which reprogramming is achieved in the absence of *Sox2*. These results demonstrate the feasibility of replacing the central reprogramming transgenes with small molecules that modulate discrete cellular pathways or processes rather than by globally altering chromatin structure. Furthermore, they show that the mechanisms by which these molecules act in reprogramming can be distinct from those of the factor(s) that they replace.

Importantly, and unlike many other studies (Mikkelsen et al., 2008; Shi et al., 2008a; Shi et al., 2008b; Utikal et al., 2009), the approach that we report here for replacing *Sox2* did not rely on procurement of a highly specialized or rare cell type that already expresses *Sox2*. Furthermore, treatment with RepSox allowed the generation of iPSCs from both adult and embryonic fibroblasts with a frequency comparable to that of transduction with *Sox2*. Thus, reprogramming efficiency does not need to be compromised by small-molecule replacement of transgenic factors.

We observed that instead of working on the initial fibroblast population to replace *Sox2*, RepSox acts on cellular intermediates formed by overexpression of *Oct4*, *Klf4,* and *cMyc*. Without RepSox treatment, these intermediates are trapped in an unproductive state. Unlike previously described partially reprogrammed cells (Mikkelsen et al., 2008; Silva et al., 2009), the RepSox-responsive intermediates could not be reprogrammed with AZA or 2i treatment, suggesting that they are distinct. In addition, we found that RepSox does not target any of the kinases inhibited by the 2i cocktail, indicating that it works through a different mechanism. Furthermore, four-factor intermediates that reprogram with RepSox treatment are not responsive to AZA, indicating that they also are distinct.

These findings demonstrate that reprogramming can proceed in a stepwise fashion through different intermediates. Just as in a geographical setting in where there are multiple routes to travel from point A to point B, there exist different intermediate states

or "way stations" that somatic cells can transit through on the way to complete reprogramming. Interestingly, although our results indicate that defined-factor reprogramming with *Oct4*, *Klf4*, *cMyc*, and *Sox2* can occur in the absence of *Nanog*, its induction is required for chemical reprogramming of both our RepSox-responsive intermediates and the recently described 2i-responsive intermediates made from *Oct4*, *Klf4*, and *cMyc* transduction of cells that express *Sox2* endogenously (Silva et al., 2009). This indicates that commonalities can exist in the reprogramming routes used by some sets of distinct intermediates.

Originally, we found it surprising that *Nanog* was not included in the initial set of defined reprogramming factors (Takahashi and Yamanaka, 2006) given its critical role in maintaining pluripotency in ESCs (Boyer et al., 2005; Chambers et al., 2003) and its ability to stimulate reprogramming by cell fusion (Silva et al., 2006). However, Takahashi and Yamanaka reported that a combination of nine factors that included *Oct4*, *Klf4*, *cMyc*, and *Nanog*, but not *Sox2*, generated iPSC colonies at a detectable rate (Takahashi and Yamanaka, 2006). This combination of factors included other genes that may have inadvertently lowered the rate of reprogramming, thereby causing the combination of *Oct4*, *Klf4*, *cMyc*, and *Nanog* to be overlooked. Consistent with these data, work by Niwa and coworkers with inducible *Sox2* null mESCs demonstrated that *Sox2* is dispensable for modulation of the Oct-Sox enhancers that regulate pluripotent-specific gene expression and instead mainly governs pluripotency in ESCs by regulating the expression of *Oct4* through other factors (Masui et al., 2007). Therefore, it is possible that *Nanog* may alleviate the requirement for *Sox2* in reprogramming by stimulating or maintaining *Oct4* expression. Indeed, *Nanog* is capable of maintaining *Oct4* expression in mESCs (Chambers et al., 2003). Thompson and coworkers also reported that *NANOG* expression enhanced the reprogramming of human fibroblasts but that it was not able to replace *SOX2* in the presence of only *OCT4* and *LIN-28* (Yu et al., 2007). This may indicate that *Klf4* is required for *Nanog* to function optimally in reprogramming and suggests that either they or the genes they modulate interact during the reprogramming process.

It is well known that ~90% of genes with promoters bound by OCT4 and SOX2 in human ESCs are also bound by NANOG (Boyer et al., 2005). Our result suggests that either *Nanog* or *Sox2* may be sufficient for collaborating with *Oct4* to modulate these genes and drive reprogramming. Although *Nanog* is not required for pluripotency, it safeguards ESCs against neuroectodermal and, to a more limited extent, mesodermal differentiation (Chambers et al., 2007; Vallier et al., 2009). Therefore, it is possible that *Nanog* functions in reprogramming by repressing differentiation signals, thereby assisting in the transition to an undifferentiated state.

Interestingly, we found that RepSox is also able to functionally replace *cMyc* in reprogramming. Together, these observations highlight the fact that small molecules may functionally replace reprogramming transcription factors at either early or late stages of the process and that they can act by different mechanisms—by inducing the expression of the gene itself, a closely related family member, or an unrelated gene that can functionally rescue the omission of the reprogramming transcription factor.

Our observation that a 1 day treatment with RepSox can relieve the requirement for transgenic *Sox2* indicates that unlike reprogramming with transgenic *Oct4*, *Klf4*, and *Sox2*, in which each transgene must be expressed for several days (Sridharan et al., 2009; Stadtfeld et al., 2008), small molecules can act as switches to induce stable gene expression changes that promote the completion of reprogramming. This could be an important concept for achieving purely chemical reprogramming given that our data show that chemicals such as RepSox can affect cellular processes differently depending on the timing of administration.

As we have shown here, there need not always be a discrete, one-to-one mapping between the functions of the reprogramming factors and their chemical replacements. Thus, it may be that reiterative screening in the presence of *Sox2* replacement molecules will be required for identifying compounds that can act in concert to replace *Oct4* and *Klf4*. However, it will be of significant interest to determine whether the novel reprogramming compounds we have identified can collaborate with those previously described (Marson et al., 2008; Shi et al., 2008a; Silva et al., 2008) to replace the remaining reprogramming genes, opening a route to purely chemical reprogramming.

EXPERIMENTAL PROCEDURES

Retroviral Infection

Retroviral infections were performed as previously described with the pMXs vector (Takahashi et al., 2007a). MEFs were infected with two to three pools of viral supernatant during a 72 hr period. The first day that viral supernatant was added was termed "day 1 post-infection." For quantification, *Oct4*::GFP$^+$ colonies were counted at day 30 postinfection unless otherwise stated. All animal research was performed under the oversight of the Office of Animal Resources at Harvard University.

Small-Molecule Screens

On day 4 postinfection, infected MEFs were trypsinized and reseeded on irradiated feeders in 96-well plates at 2000 cells/well and cultured in mouse ESC media (Knockout DMEM,15% Hyclone FBS, L-glutamine, penicillin/streptomycin, nonessential amino acids, β-mercaptoethanol, and 1000 U/ml LIF). The next day, compound stock solutions diluted in DMSO and VPA (Sigma) were added at a final concentration of 1 μM and 2 mM, respectively. VPA was removed after 1 week, and the compound was reapplied every other day with each media change. Plates were scored for GFP$^+$ colonies after 11 days of compound treatment.

Quantification of *Oct4*::GFP$^+$ iPSCs Generated with Small-Molecule Hit Compounds, SB431542, and Tgf-β Antibodies

Retroviral infection and compound or antibody treatment was performed as in the original chemical screen. For quantification of the numbers of GFP$^+$ colonies produced in different conditions, the number of colonies in each well was counted and at least two different wells were counted and averaged. Concentrations of compounds and antibodies were as follows: VPA (Sigma), 2 mM; RepSox (Calbiochem), 25 μM or 1 μM as noted; E-616451 (Calbiochem), 3 μM; EI-275 (Biomol), 3 μM; SB431542 (Sigma), 25 μM or 2 μM as noted; TgfβII-specific antibody (R&D Systems, AB-12-NA), 10 μg/ml; and pan-Tgfβ antibody (R&D Systems, AB-100-NA), 10 μg/ml. Unless otherwise noted, all chemical treatments were continuous from initial administration at day 4–5 postinfection until GFP$^+$ colonies were scored at day 30 posttransduction. Fresh chemical was added at each media change.

Chemical Reprogramming of Stable Intermediate Cell Lines

Oct4::GFP-negative colonies in *Oct4*, *Klf4*, and *cMyc* or *Oct4*-, *Klf4*-, *cMyc*-, and *Sox2*-infected MEF cultures were picked and plated on irradiated feeders, and single colonies were picked after 1 week. The resulting cell lines were

passaged with trypsin and grown in mESC media on feeders until passage 4, at which time they were treated with RepSox (25 μM), AZA (500 μM), or both for 48 hr. For 2i treatment, CHIR99021 (Stemgent) was used at 3 μM and PD0325901 (Stemgent) was used at 1 μM. *Oct4*::GFP$^+$ colonies were scored 12 days after the beginning of chemical treatment. Treatments were performed in mESC media containing FBS unless otherwise noted.

SUPPLEMENTAL DATA

Supplemental Data include Supplemental Experimental Procedures, 13 figures, and 2 tables and can be found with this article online at http://www.cell.com/cell-stem-cell/supplemental/S1934-5909(09)00508-6.

ACKNOWLEDGMENTS

We thank S. Noggle, E. Kiskinis, A. Arvanites, S. Bobrowicz, L. Weisenthal, R. Maehr, A. Kweudjeu, R. Gali, M. Yamaki, E. Massassa, R. Martinez, and K. Rosowski for technical assistance and S. Sullivan, K. Niakan, K. Rodolfa, S. Mekhoubad, I. Tabansky, C. Sasaki, D. Melton, and S. Chen for helpful discussions. This work was made possible by support provided by the Harvard Stem Cell Institute to L.R. and K.E., by support from the NIH grant R01 HD046732-01A1 to K.E., and by an innovation grant from the NYSCF to L.R. E.S. and D.L. acknowledge support from the NIH (GM065400) and from the Howard Hughes Medical Institute. J.K.I. and F.P.D. are New York Stem Cell Foundation postdoctoral fellows. D.H. is a Helen Hay Whitney postdoctoral fellow. K.E. is a fellow of the John D. and Catherine T. MacArthur Foundation. H.A. was funded by the Sankyo Foundation of Life Science. The authors are filing a patent based on the results reported in this paper. L.L.R. is a founder of iPierian and a member of its scientific advisory board. K.E. is a member of the iPierian scientific advisory board.

Received: June 10, 2009
Revised: September 2, 2009
Accepted: September 23, 2009
Published online: October 8, 2009

REFERENCES

Amabile, G., and Meissner, A. (2009). Induced pluripotent stem cells: Current progress and potential for regenerative medicine. Trends Mol. Med. *15*, 59–68.

Attisano, L., and Wrana, J.L. (2002). Signal transduction by the TGF-beta superfamily. Science *296*, 1646–1647.

Boiani, M., Kehler, J., and Scholer, H.R. (2004). Activity of the germline-specific Oct4-GFP transgene in normal and clone mouse embryos. Methods Mol. Biol. *254*, 1–34.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643–655.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230–1234.

Gellibert, F., Woolven, J., Fouchet, M.H., Mathews, N., Goodland, H., Lovegrove, V., Laroze, A., Nguyen, V.L., Sautet, S., Wang, R., et al. (2004). Identification of 1,5-naphthyridine derivatives as a novel series of potent and selective TGF-beta type I receptor inhibitors. J. Med. Chem. *47*, 4494–4506.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E., et al. (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science *302*, 415–419.

Hanke, J.H., Gardner, J.P., Dow, R.L., Changelian, P.S., Brissette, W.H., Weringer, E.J., Pollok, B.A., and Connelly, P.A. (1996). Discovery of a novel, potent, and Src family-selective tyrosine kinase inhibitor. Study of Lck- and FynT-dependent T cell activation. J. Biol. Chem. *271*, 695–701.

Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka, S. (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature *460*, 1132–1135.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008a). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat. Biotechnol. *26*, 795–797.

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D.A. (2008b). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat. Biotechnol. *26*, 1269–1275.

Inman, G.J., Nicolas, F.J., Callahan, J.F., Harling, J.D., Gaster, L.M., Reith, A.D., Laping, N.J., and Hill, C.S. (2002). SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. Mol. Pharmacol. *62*, 65–74.

Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P., and Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature *458*, 771–775.

Kim, D., Kim, C., Moon, J., Chung, Y., Chang, M., Han, B., Ko, S., Yang, E., Cha, K.Y., Lanza, R., and Kim, K.S. (2009). Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. Cell Stem Cell *4*, 472–476.

Marson, A., Foreman, R., Chevalier, B., Bilodeau, S., Kahn, M., Young, R.A., and Jaenisch, R. (2008). Wnt signaling promotes reprogramming of somatic cells to pluripotency. Cell Stem Cell *3*, 132–135.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., et al. (2007). Pluripotency governed by *Sox2* via regulation of *Oct3/4* expression in mouse embryonic stem cells. Nat. Cell Biol. *9*, 11.

Meissner, A., Wernig, M., and Jaenisch, R. (2007). Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. Nat. Biotechnol. *25*, 1177–1181.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49–55.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat. Biotechnol. *26*, 101–106.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature *448*, 313–317.

Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., and Yamanaka, S. (2008). Generation of mouse induced pluripotent stem cells without viral vectors. Science *322*, 949–953.

Shi, Y., Desponts, C., Do, J.T., Hahm, H.S., Scholer, H.R., and Ding, S. (2008a). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell Stem Cell *3*, 568–574.

Shi, Y., Do, J.T., Desponts, C., Hahm, H.S., Scholer, H.R., and Ding, S. (2008b). A combined chemical and genetic approach for the generation of induced pluripotent stem cells. Cell Stem Cell *2*, 525–528.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol. *6*, e253.

Silva, J., Chambers, I., Pollard, S., and Smith, A. (2006). Nanog promotes transfer of pluripotency after cell fusion. Nature *441*, 997–1001.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. Cell *138*, 722–737.

Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. Cell *136*, 364–377.

Stadtfeld, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell *2*, 230–240.

Suzuki, A., Raya, A., Kawakami, Y., Morita, M., Matsui, T., Nakashima, K., Gage, F.H., Rodriguez-Esteban, C., and Izpisua Belmonte, J.C. (2006). Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. Proc. Natl. Acad. Sci. USA *103*, 10294–10299.

Takahashi, K., Okita, K., Nakagawa, M., and Yamanaka, S. (2007a). Induction of pluripotent stem cells from fibroblast cultures. Nat. Protoc. *2*, 3081–3089.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007b). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872.

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.

Thrasher, A.J. and Gaspar, B. (2007). Severe adverse event in clinical trial of gene therapy for X-SCID. http://wwwasgtorg/UserFiles/XSCIDstatementpdf.

Utikal, J., Maherali, N., Kulalert, W., and Hochedlinger, K. (2009). Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells. J. Cell Sci. *122*, 3502–3510.

Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L.E., Trotter, M.W., Cho, C.H., Martinez, A., Rugg-Gunn, P., et al. (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. Development *136*, 1339–1349.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318–324.

Xu, R.H., Sampsell-Barron, T.L., Gu, F., Root, S., Peck, R.M., Pan, G., Yu, J., Antosiewicz-Bourget, J., Tian, S., Stewart, R., et al. (2008). NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. Cell Stem Cell *3*, 196–206.

Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281–292.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917–1920.

ARTICLES

Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells

Matthias Stadtfeld[1,2,3]*, Effie Apostolou[1,2,3]*, Hidenori Akutsu[4], Atsushi Fukuda[5], Patricia Follett[1], Sridaran Natesan[6], Tomohiro Kono[5], Toshi Shioda[2] & Konrad Hochedlinger[1,2,3]

Induced pluripotent stem cells (iPSCs) have been generated by enforced expression of defined sets of transcription factors in somatic cells. It remains controversial whether iPSCs are molecularly and functionally equivalent to blastocyst-derived embryonic stem (ES) cells. By comparing genetically identical mouse ES cells and iPSCs, we show here that their overall messenger RNA and microRNA expression patterns are indistinguishable with the exception of a few transcripts encoded within the imprinted *Dlk1–Dio3* gene cluster on chromosome 12qF1, which were aberrantly silenced in most of the iPSC clones. Consistent with a developmental role of the *Dlk1–Dio3* gene cluster, these iPSC clones contributed poorly to chimaeras and failed to support the development of entirely iPSC-derived animals ('all-iPSC mice'). In contrast, iPSC clones with normal expression of the *Dlk1–Dio3* cluster contributed to high-grade chimaeras and generated viable all-iPSC mice. Notably, treatment of an iPSC clone that had silenced *Dlk1–Dio3* with a histone deacetylase inhibitor reactivated the locus and rescued its ability to support full-term development of all-iPSC mice. Thus, the expression state of a single imprinted gene cluster seems to distinguish most murine iPSCs from ES cells and allows for the prospective identification of iPSC clones that have the full development potential of ES cells.

Induced pluripotent stem cells (iPSCs), generated by the overexpression of transcription factors such as Oct4 (also called Pou5f1), Sox2, Klf4 and c-Myc in somatic cells[1,2], have enormous therapeutic potential as they enable the derivation of patient-specific pluripotent cell lines to study and possibly treat degenerative diseases. Although the generation of iPSCs is technically simple and ethically uncontroversial, it remains unclear whether iPSCs are molecularly and functionally different from ES cells derived from blastocysts, which are considered the gold standard for pluripotent cells. Previously published reports indicate high similarities between ES cells and iPSCs, including indistinguishable global histone modification and DNA methylation patterns[3,4]. In addition, iPSCs, like ES cells, give rise to numerous differentiated cell types, including the germ line, in the context of chimaeric animals[5,6]. More recently, iPSCs have been shown to support the development of all-iPSC mice using tetraploid (4n) embryo complementation[7–9], the most stringent assay for developmental potential[10,11].

Despite these similarities, there is emerging evidence for substantial differences between ES cells and iPSCs. For example, most iPSC clones give rise to low-grade chimaeras after injection into diploid blastocysts and fail to support the development of postnatal all-iPSC mice upon 4n embryo complementation[12–14]. At the molecular level, major differences in mRNA and microRNA (miRNA) expression[15–17], as well as in DNA methylation[18–20], have been reported between ES cells and iPSCs. These observations indicate that factor-mediated reprogramming results in abnormalities in resultant iPSCs, which could impede their therapeutic utility. In contrast, nuclear-transfer-mediated reprogramming gives rise to nuclear transfer ES cells that are molecularly and functionality indistinguishable from ES cells derived from fertilized embryos[21,22], raising the possibility that nuclear transfer generates cells that are more completely reprogrammed than iPSCs.

A potential limitation of the aforementioned studies is that ES cells were compared with iPSCs of different genetic backgrounds and harbouring viral transgenes, which are known to affect gene expression patterns[21,23] and the functionality[1,6] of cells. We therefore revisited the question of whether ES cells and iPSCs are equivalent by comparing genetically matched cell lines.

Transcriptional comparison of ES cells and iPSCs

To circumvent the potentially confounding effects of genetic background and viral integrations on gene expression patterns and developmental potential, we used a novel transgenic reprogramming system to generate genetically matched mouse ES cells and iPSCs[24]. Briefly, a polycistronic cassette expressing Oct4, Klf4, Sox2 and c-Myc[25] (OKSM) under the control of a doxycycline-inducible promoter was inserted into the collagen type I α1 (*Col1a1*) locus of ES cells expressing the reverse tetracycline-dependent transactivator (rtTA) from the *ROSA26* promoter[26]. These collagen-OKSM ES cells were then used to generate mice from which different somatic cell types were isolated and induced with doxycycline to derive genetically matched iPSCs for molecular and functional comparisons (Fig. 1a, b).

[1]Howard Hughes Medical Institute at Massachusetts General Hospital, Center for Regenerative Medicine; Harvard Stem Cell Institute, 185 Cambridge Street, Boston, Massachusetts 02114, USA. [2]Massachusetts General Hospital Cancer Center and Harvard Medical School, 149 13th Street, Charlestown, Massachusetts 02129, USA. [3]Department of Stem Cell and Regenerative Biology, Harvard University and Harvard Medical School, 42 Church Street, Cambridge, Massachusetts 02138, USA. [4]Department of Reproductive Biology, National Institute for Child Health and Development, Tokyo 157-8535, Japan. [5]Department of BioScience, Tokyo University of Agriculture, Tokyo 156-8502, Japan. [6]Sanofi-Aventis, 270 Albany Street, Cambridge, Massachusetts 02139, USA.
*These authors contributed equally to this work.

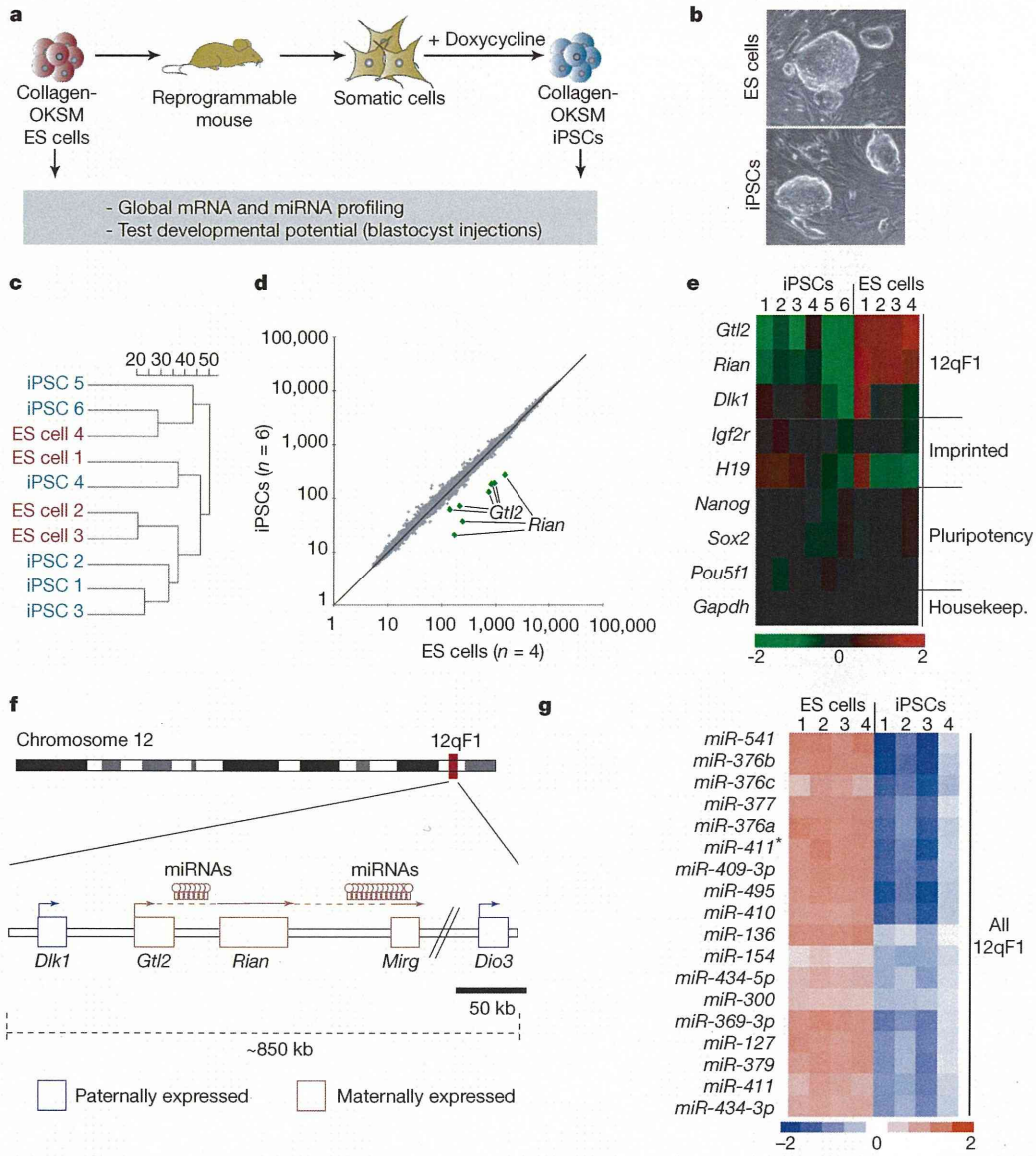


Figure 1 | Aberrant silencing of the *Dlk1–Dio3* gene cluster in mouse iPSCs. **a**, Strategy for comparing genetically matched ES cells and iPSCs generated with the doxycycline-controllable collagen-OKSM system. **b**, Morphology of collagen-OKSM ES cells and iPSCs. **c**, Unsupervised clustering of four ES cell and six derivative iPSC lines based on microarray expression data. **d**, Scatter plot of microarray data comparing iPSCs and ES cells with differentially expressed genes highlighted in green (twofold, $P < 0.05$, *t*-test with Benjamini–Hochberg correction). **e**, Heat map showing relative expression levels of selected mRNAs in ES cells and iPSCs. **f**, Schematic representation of the *Dlk1–Dio3* gene cluster with maternally and paternally expressed transcripts shown in red and blue, respectively. **g**, Heat map showing miRNAs that are differentially expressed between ES cells and iPSCs (twofold, $P < 0.01$, *t*-test).

We first compared the abilities of parental collagen-OKSM ES cells and iPSCs derived from mouse embryonic fibroblasts (MEFs) to support the development of all-iPSC mice using 4n embryo complementation. The two tested ES cell lines gave rise to viable mice at expected frequencies (13–20%)[11], demonstrating that the OKSM transgene per se does not adversely affect development (Supplementary Table 1). In contrast, all four tested iPSC lines repeatedly failed to support the development of all-iPSC mice, indicating qualitative differences between these iPSCs and ES cells (Supplementary Table 1).

We reasoned that a transcriptional comparison of the iPSC lines that failed 4n complementation with matched 4n-complementation-competent ES cell lines might reveal molecular changes that explain the developmental deficits of iPSCs. Global mRNA profiling showed marked similarities in the overall transcriptional patterns of four collagen-OKSM ES cells and six derivative iPSCs and did not separate these cell lines using unsupervised clustering or principal component analysis (Fig. 1c and data not shown). In fact, only two transcripts were identified as differentially expressed (>2-fold difference, *t*-test, $P < 0.05$) between ES cells and iPSCs. These were the non-coding RNA *Gtl2* (also know as *Meg3*) and the small nucleolar RNA *Rian* (Fig. 1d, e).

Repression of *Dlk1–Dio3* transcripts in iPSCs

Gtl2 and *Rian* localize to the imprinted *Dlk1–Dio3* gene cluster on mouse chromosome 12qF1 and are maternally expressed in mammals (Fig. 1f)[27]. Both genes were strongly repressed in iPSC clones compared to ES cell clones, whereas expression of pluripotency and housekeeping genes remained unaffected (Fig. 1e). Quantitative PCR (qPCR) analysis of *Gtl2*, *Rian* and *Mirg*, another maternally expressed imprinted gene in the *Dlk1–Dio3* cluster, confirmed transcriptional silencing in iPSCs (Supplementary Fig. 1a). Expression of other imprinted genes showed clone-to-clone variations, as was previously seen for ES cells[28], but no consistent differences between ES cells and iPSCs (Fig. 1e and Supplementary Table 2). This shows that imprinted gene silencing is not a genome-wide phenomenon. Of note, we failed to detect differences with the collagen-OKSM system