# [Chemical reprogramming of human somatic cells to pluripotent stem cells](https://www.nature.com/articles/s41586-022-04593-5), 2022

***Abstract***

Es un desafío inducir células madre pluripotentes humanas mediante reprogramación química debido a que las células somáticas humanas son refractarias a la estimulación química por su epigenoma estable y su plasticidad reducida.

Se identifica la vía JNK como una barrera importante para la reprogramación química, cuya inhibición es indispensable para inducir la plasticidad celular y un programa similar a la regeneración mediante la supresión de las vías proinflamatorias.

Intentan sentar las bases para desarrollar estrategias terapéuticas regenerativas que utilicen sustancias químicas bien definidas para cambiar el destino de las células en humanos.

Es importante destacar que dicho estado plástico a menudo tiene una arquitectura de cromatina relativamente abierta con mayor accesibilidad, lo cual es esencial para inducir un nuevo destino celular

***Pluripotencia***

OCT4 fue activado por una combinación de moléculas pequeñas, incluidos reguladores epigenéticos (tranilcipromina, ácido valproico, DZNep, EPZ004777 y UNC0379) e inhibidores de la señalización celular (CHIR99021, 616452, Y27632 y PD0325901) (condición de etapa III). Para activar la red de genes de pluripotencia, seleccionamos pequeñas moléculas que facilitan el mantenimiento de las células hPS. El tratamiento con CHIR99021, PD0325901, SB590885, IWP2 y Y27632 (condición de etapa IV) indujo colonias compactas que coexpresan OCT4, SOX2 y NANOG (en adelante, OSN).

***Caracterización hCiPS (se hace con células embrionarias y adultas)***

*Resultados que sugieren que las líneas celulares OSN + se parecen a las células hES en los perfiles transcriptómicos y epigenéticos: -> son células PS humanas inducidas químicamente (hCiPS).*

* Derivación de líneas celulares estables de estas colonias OSN + de diferentes donantes. Estas líneas celulares pudieron expandirse durante más de 20 pases y el tiempo de duplicación celular fue similar al de las células hES. Expresaron los marcadores de superficie celular pluripotentes TRA-1-60, TRA-1-81 y SSEA-4, además de OCT4, SOX2 y NANOG.
* El análisis RT-qPCR mostró la expresión de un panel de genes de pluripotencia en estas células, cuyos niveles eran comparables con los de las células hES.
* La secuenciación de ARN (RNA-seq) y el análisis de metilación del ADN mostraron que los transcriptomas y el epigenoma eran similares entre estas células y las células hES
* Los loci de los promotores OCT4 y NANOG se desmetilaron con accesibilidad abierta a la cromatina.
* Las células hCiPS pudieron formar células de las tres capas germinales in vivo formando teratomas en ratones inmunodeficientes e in vitro formando cuerpos embrioides.
* La diferenciación dirigida de células hCiPS pudo generar células progenitoras hematopoyéticas y avanzar hacia el linaje de células T
* También generamos hepatocitos y células madre neurales
* Para la integridad genómica, el análisis de bandas G de alta resolución mostró que las células hCiPS tenían cariotipos diploides normales
* Análisis de repetición corta en tándem confirmó que las células hCiPS derivaban de sus fibroblastos parentales y eran distintas de otras líneas celulares hES establecidas

***Identificación de un estado intermedio plástico***

RNA-seq (scRNA-seq) unicelular de todo el proceso de reprogramación para comprender mejor la reprogramación química de las células somáticas humanas.

La trayectoria de reprogramación se reconstruyó utilizando el método Waddington-OT 28 , que identificó las células de las cuales era más probable que los descendientes fueran células hCiPS y su probabilidad de descendencia se determinó como puntuación de células hCiPS. Luego nos centramos en las células que siguieron la exitosa trayectoria de reprogramación, que se clasificaron además en tres estados celulares secuenciales distintos. En conjunto, identificamos tres fases clave secuenciales durante el proceso de reprogramación química humana: la adquisición de un programa genético similar a la regeneración en las primeras etapas (etapa I y etapa II), la activación del programa XEN (etapa III) y, finalmente, el establecimiento de una red de pluripotencia ingenua (etapa IV)

* un ***estado plástico intermedio (grupo R9): final estado II***
  + regulación negativa de los programas somáticos, seguida de la regulación positiva de un panel de genes que están involucrados en el desarrollo embrionario
  + aumento de la proliferación celular
  + células también adquirieron un estado epigenético hipometilado en la etapa II y las regiones promotoras de genes relacionados con el desarrollo embrionario, el ciclo celular y la proliferación de células madre se desmetilaron, lo que sugiere características de desdiferenciación celular y un estado celular plástico
  + análisis de ontología genética (GO) mostró que los genes que se activaron en este estado plástico intermedio se enriquecieron en términos de desarrollo de extremidades y apéndices
  + El ensayo unicelular para la cromatina accesible a la transposasa mediante análisis de secuenciación (scATAC-seq) reveló un estado de cromatina abierta de estas células intermedias con mayor accesibilidad en ciertos loci genéticos relacionados con el desarrollo de las extremidades
  + CONCLUSIÓN: programa regulador de genes similar a la regeneración que subyace al estado plástico intermedio, que es indispensable para adquirir la pluripotencia celular.
* un ***estado extraembrionario similar al endodermo (similar a XEN)*** (grupo R11): estadio III
* ***estado de pluripotencia ingenua*** (grupo R13): etapa IV
  + células hCiPS transitó un estado de pluripotencia ingenuo antes de mantenerse como líneas celulares hCiPS estables

Como vemos encontramos conjuntos de datos asociados a la reprogramación química en diferentes estados:

* Estado I: generan células epiteliales. Tiene una duración aproximada de entre 8 y 10 días.
* Estado II: inducen hipometilación y un estado de proliferación celular (reactivación de genes relacionados con el desarrollo, desdiferenciación, estado de cromatina abierta) Se considera que se encuentran en un estado de plasticidad intermedio. Tiene una duración aproximada de 16-20 días. **ESTADO PARCIAL, 10-15 DÍAS DESDE EL INICIO DE LA REPROGRAMACIÓN**
* Estado III: OCT4 (gen de pluripotencia) es activado por las moléculas, reguladores epigenéticos e inhibidores celulares de señal. Tiene una duración aproximada de 8-12 días.
* Estado IV: OCT4, SOX2 y NANOG, se consigue la pluripotencia de las células llevándolas a una morfología de células madre embrionarias. Tiene una duración aproximada de 7-10 días. Al final de la etapa IV, se recomienda la tinción inmunofluorescente de coexpresión de OCT4 y NANOG para confirmar la generación de colonias de células hCiPS primarias

**Conclusiones**

* Reprogramación química en 4 fases con moléculas no integrables en el genoma y controlables para generar células IPS equivalentes a hES.
* Durante la reprogramación se llega a un estado plástico (desdiferenciación y capacidad proliferación). El programa genético de dicho estado expresa SALL4 , LIN28A , MSX1 , MSX2 y HOXB9 + patrón de expresión génica desdiferenciado (regenerativo)
* ATAC sugirió que este estado plástico poseía un estado de cromatina abierta y una metilación del ADN reducida a nivel global.
* La inhibición de la vía JNK es indispensable para la inducción de la regeneración mediante la supresión de genes proinflamatorios. respuesta proinflamatoria reducida puede restaurar la capacidad regenerativa

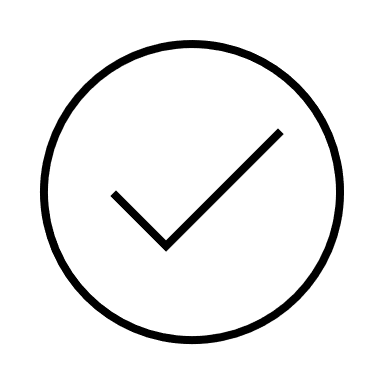
En este experimento se **reprograman químicamente** diferentes tipos de células somáticas para llevarlas a un estado de pluripotencia. Concretamente, en scRNA las muestras son de 32.738 genes en cada uno de los siguientes tipos celulares:

* **0330**: Línea celular de fibroblastos humanos embrionarios.
  + [GSM5387808](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387808), ***scRNA-Seq HEFs-0330***: fibroblastos embrionarios humanos. Contiene datos de 4.335 células.
  + [GSM5387813](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387813), ***scRNA-Seq stageII-0330***: células somáticas humanas durante la fase II de reprogramación química.
  + [GSM5387814](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387814), **scRNA-Seq StageII-5azac-0330**: células somáticas humanas durante la fase II de reprogramación química sin 5azac (regula la metilación global del ADN que salvaguarda la identidad de las células somáticas , lo que induce un estado hipometilado, aunque vieron que afecta a la proliferación celular)
  + [GSM5387815](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387815), **scRNA-Seq StageII-JNKIN8-0330**: células somáticas humanas durante la fase II de reprogramación química sin JNKIN8 (vieron que junto con 5azac y tranyl induce hipometilación y estado de proliferación celular pero no activa el gen de la pluripotencia OCT4).
  + [GSM5387816](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387816), **scRNA-Seq StageII-Tranyl-0330**: células somáticas humanas durante la fase II de reprogramación química sin Tranyl. (regula la metilación global del ADN que salvaguarda la identidad de las células somáticas , lo que induce un estado hipometilado, aunque vieron que afecta a la proliferación celular)
  + [GSM5534156](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534156), ***hCiPSCs-1117***: células madre pluripotentes humanas inducidas químicamente a partir de fibroblastos embrionarios humanos. Contiene datos de 18.823 células.
* **0605**: Línea celular de fibroblastos dérmicos de un humano adulto.
  + [GSM5534140](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534140), **hASFs-0605**: fibroblastos dérmicos de un humano adulto. Contiene datos de 5.475 células.
  + [GSM5534142](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534142)**, StageI-0605**: células somáticas humanas durante la fase I de reprogramación química
  + [GSM5534143](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534143)**, StageII-0605**: células somáticas humanas durante la fase II de reprogramación química
  + [GSM5534157](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534157), **hCiPSCs-1230**: células madre pluripotentes humanas inducidas químicamente a partir de fibroblastos de piel humana adulta. Contiene datos de 20.414 células.
* **0618**: Línea celular de adipocitos.
  + [GSM5387809](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387809), **scRNA-Seq hADSCs-0618**: células madre mesenquimales derivadas de tejido adiposo humano adulto. Contiene datos de 4.507 células.
  + [GSM5683317](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5683317)**, StageI\_Day0-0618**: células somáticas humanas durante la fase I día 0 de reprogramación química.
  + [GSM5683318](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5683318)***,* StageI\_Day0.5-0618**: células somáticas humanas durante la fase I día 0.5 de reprogramación química.
  + [GSM5683319](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5683319)**, StageI\_Day1-0618**: células somáticas humanas durante la fase I día 1 de reprogramación química.
  + [GSM5683320](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5683320), **StageI\_Day2-0618**: células somáticas humanas durante la fase I día 2 de reprogramación química.
  + [GSM5534144](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534144), **StageI\_Day4-0618**: células somáticas humanas durante la fase I día 4 de reprogramación química.
  + [GSM5387810](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387810)**, StageI\_Day8-0618**: células somáticas humanas durante la fase I día 8 de reprogramación química.
  + [GSM5534145](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534145)**, StageII\_Day4-0618**: células somáticas humanas durante la fase II día 4 de reprogramación química.
  + [GSM5534146](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534146)**, StageII\_Day8-0618**: células somáticas humanas durante la fase II día 8 de reprogramación química.
  + [GSM5534147](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534147), **StageII\_Day12-0618**: células somáticas humanas durante la fase II día 12 de reprogramación química.
  + [GSM5387811](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387811)**, StageII\_Day20-0618**: células somáticas humanas durante la fase II día 20 de reprogramación química.
  + [GSM5534148](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534148), **StageIII\_Day4-0618**: células somáticas humanas durante la fase III día 4 de reprogramación química.
  + [GSM5534149](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534149), **StageIII\_Day8-0618**: células somáticas humanas durante la fase III día 8 de reprogramación química.
  + [GSM5387812](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387812), **StageIII\_Day12-0618**: células somáticas humanas durante la fase III día 12 de reprogramación química.
  + [GSM5534150](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534150), **StageIV\_Day1-0618**: células somáticas humanas durante la fase IV día 1 de reprogramación química.
  + [GSM5534151](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534151)**, StageIV\_Day2-0618**: células somáticas humanas durante la fase IV día 2 de reprogramación química.
  + [GSM5534152](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534152), **StageIV\_Day4-0618**: células somáticas humanas durante la fase IV día 4 de reprogramación química.
  + [GSM5534153](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534153), **StageIV\_Day10-0618**: células somáticas humanas durante la fase IV día 10 de reprogramación química.
  + [GSM5534154](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534154)**, hCiPSCs-0618**: células madre pluripotentes humanas inducidas químicamente a partir de células madre mesenquimales derivadas de tejido adiposo humano adulto. Contiene datos de 18.792 células.
* **0809**: Línea celular -> relacionado con 0618?
  + [GSM5534141](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534141)**, hADSCs-0809**: células madre mesenquimales derivadas de tejido adiposo humano adulto. Contiene datos de 6.129 células.
  + [GSM5534155](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534155), **hCiPSCs-0809**: células madre pluripotentes humanas inducidas químicamente a partir de células estromales mesenquimales derivadas de tejido adiposo humano adulto. Contiene datos de 5.524 células.
* [GSM5534158](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5534158), **H1**: células madre embrionarias (control de cómo debería ser la expresión, es decir, las células madre pluripotentes deberían tener los mismos genes que estas). Contiene datos de 16.530 células.

Por otro lado, los datos de sc-ATAC:

* **0330**: Línea celular de fibroblastos humanos embrionarios.
  + [GSM5387804](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387804), **scATAC-Seq fibroblast**: fibroblastos embrionarios humanos.
  + [GSM5387805](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387805), **scATAC-Seq StageII**: células somáticas humanas durante la fase II de reprogramación química.
  + [GSM5387806](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387806), **scATAC-Seq StageII-5azac**: células somáticas humanas durante la fase II de reprogramación química sin 5azac.
  + [GSM5387807](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387807)**, scATAC-Seq StageII-JNKIN8**: células somáticas humanas durante la fase II de reprogramación química sin JNKIN8.
* **0618**: Línea celular de adipocitos.
  + [GSM5683314](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5683314) , **scATAC-Seq hADSCs-0618**: células madre mesenquimales derivadas de tejido adiposo humano adulto
  + [GSM5683315](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5683315), **scATAC-Seq StageI-0618**: células somáticas humanas durante la fase II de reprogramación química. LA DESCRIPCIÓN NO COINCIDE CON EL NOMBRE.
  + [GSM5683316](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5683316), **scATAC-Seq StageII-0618**: células somáticas humanas durante la fase II de reprogramación química sin 5azac. LA DESCRIPCIÓN NO COINCIDE CON EL NOMBRE.

También contamos con los datos de secuenciación de bisulfito del genoma completo.

Distribución de datos sc-RNA () y sc-ATAC () por fase de reprogramación química:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Tipo celular** | **Control** | **Stage I**  **(1-8 días)** | **Stage II**  **(8-28 días)** | **Stage III**  **(28-40 días)** | **Stage IV**  **(40-50 días)** | **Pluripotentes**  **(hCiPSC)** |
| Fibroblastos embrionarios humanos (0330) | Marca de insignia1 con relleno sólido  Marca de insignia1 contorno |  | Marca de insignia1 con relleno sólido  Marca de insignia1 contorno  (ambos sin 4/2 moléculas diferentes) |  |  | Marca de insignia1 con relleno sólido |
| Fibroblastos dérmicos humano adulto (0605) | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido |  |  | Marca de insignia1 con relleno sólido |
| Células madre mesenquimales derivadas de adipocitos humano adulto (0618) | Marca de insignia1 con relleno sólidoMarca de insignia1 contorno | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólidoMarca de insignia1 contorno | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido |
| Células madre embrionarias (H1) | Marca de insignia1 con relleno sólido |  |  |  |  |  |

Los datos se pueden descargar de [GEO](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE178325).

***Preprocesamiento***:

Extraemos las matrices de conteos, barcodes y genes\_id para cada tipo celular (fibroblastos embrionarios, células madre mesenquimales adultas derivadas de tejido adiposo y fibroblastos dérmicos adultos).

Para dar la etiqueta nos vamos a basar en si las células están reprogramadas o no. Aquellas células en el primer estado (sin estar en cultivo todavía o llevando 1 día) se consideran no reprogramadas, y le damos una etiqueta de valor 0. Por otro lado, las células en el último estado de cultivo se consideran reprogramadas, y le damos una etiqueta 1.

En el One Drive, dentro de cada línea celular encontramos dos matrices y dos genes\_id asociadas a las células reprogramadas y sin reprogramar. La etiqueta Y se encuentra como una columna más dentro de las matrices X.

*-> También podemos sacar datos de reprogramación parcial (estado II) -> Probar si la máquina puede con la concatenación de scRNAseq+scATACseq (por separado y luego, si se puede concatenar —es decir, que se hayan hecho las dos técnicas en las mismas células—, junto)*