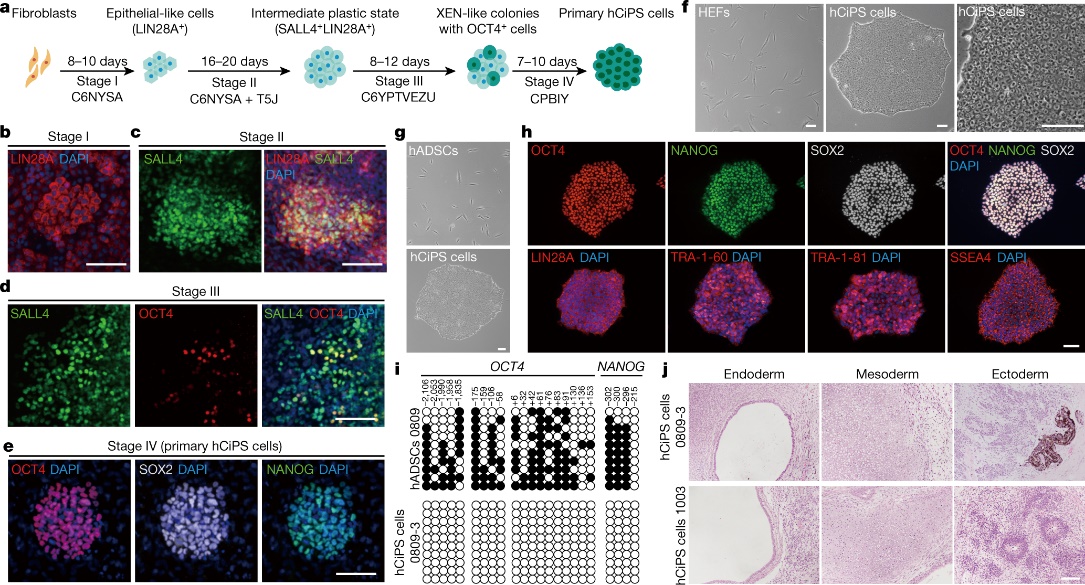
## *8 Abril 2024*

DATABASES

### [1] Guan, J., Wang, G., Wang, J. et al. Chemical reprogramming of human somatic cells to pluripotent stem cells. Nature 605, 325–331 (2022). <https://doi.org/10.1038/s41586-022-04593-5>

#### 1. Conclusiones/Descubrimientos

* Reprogramación química en 4 fases con moléculas no integrables en el genoma y controlables para generar células hCiPSC 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.

#### 2.Datos (tipo, línea celular etc)

Estado plástico intermedio -> fase II.

*scRNA-seq*

* *hEFSs: Fibroblastos embrionarios humanos (0330)*
* *hASFs: Fibroblastos dérmicos humano adulto (0605)*
* ***hADSCs: Células madre mesenquimales derivadas de adipocitos humano adulto (0618, 0809)***
* *H1: Células madre embrionarias*

*scATAC-seq*

* *hEFSs: Fibroblastos embrionarios humanos (0330)*
* ***hADSCs: Células madre mesenquimales derivadas de adipocitos humano adulto (0618)***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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)** |
| hEFSs: 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 |
| hASFs: 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 |
| hADSCs:  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 |
| H1:  Células madre embrionarias | Marca de insignia1 con relleno sólido |  |  |  |  |  |

#### 3.Usan proteínas para validar

---

#### 4.Algortimos de ML/Métodos

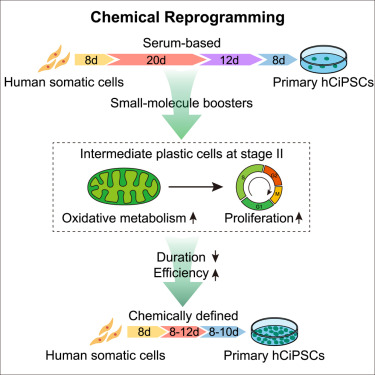
* Waddington-OT: reconstruir trayectorias de reprogramación
* [GO análisis](https://geneontology.org/docs/go-enrichment-analysis/): anota información sobre los genes usando conocimiento experto como PANTHER.
* UMAP

### [2] Liuyang S, Wang G, Wang Y, He H, Lyu Y, Cheng L, Yang Z, Guan J, Fu Y, Zhu J, Zhong X, Sun S, Li C, Wang J, Deng H. Highly efficient and rapid generation of human pluripotent stem cells by chemical reprogramming. Cell Stem Cell. 2023

### Apr 6;30(4):450-459.e9. doi: 10.1016/j.stem.2023.02.008. Epub 2023 Mar 20. PMID: 36944335.

#### 1.Conclusiones/Descubrimientos

* Detección de moléculas booster y optimización sistemática de la reprogramación química para generar un protocolo acortado y eficiente para la generación de hCiPSC (de 50 a 16 días, sin fase XEN, directamente pluripotencia).
* Las primeras células intermedias aumentaron su actividad metabólica de proliferación y fosforilación oxidativa (OXPHOS), pero no se requiere glucólisis para la etapa II de la reprogramación química
* Se sugiere que los patrones y la dinámica metabólica celular son distintos entre la reprogramación química y la reprogramación basada en Yamanaka TF.



Esquema comparación [1] y [2]

#### 2.Datos (tipo, línea celular etc)

Estado plástico intermedio -> final de la fase II.

*RNA-seq, WGBS, CUT & Tag, scRNA-seq, scATAC-seq*

* hEFSs: Fibroblastos embrionarios humanos (0330)
* *hASFs: Fibroblastos dérmicos humano adulto (0605)*
* ***hADSCs: Células madre mesenquimales derivadas de adipocitos humano adulto (0618, 0809)***
* *H1: Células madre embrionarias*
* H9:

#### 3.Usan proteínas para validar

Análisis de repetición corta en tándem (STR) confirmó que las hCiPSC se derivaron de sus células somáticas parentales.

#### Análisis de inmunofluorescencia y RT-qPCR confirmó la expresión de genes clave de pluripotencia

#### 4.Algortimos de ML

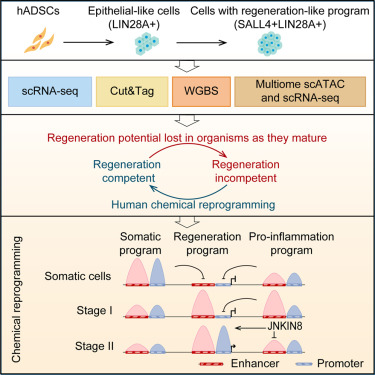
scRNA-seq:

* Read10X del paquete R Seurat.
* Eliminar las células con recuentos totales de UMI <1000, número de genes <500 o porcentaje de recuento de UMI mitocondrial> 20.
* Recuentos de UMI se normalizaron utilizando la función NormalizeData.
* Preproceso estándar de Seurat ejecutando las funciones FindVariableFeatures, ScaleData, RunPCA, FindNeighbors y FindClusters.
* Se eliminaron los grupos con un recuento medio de UMI bajo (<1500).
* Para identificar las células que se convirtieron en hCiPSC con éxito en cada momento, realizamos un análisis de Waddington-OT para construir una trayectoria unicelular utilizando el módulo wot de Python.
* Seguimos el proceso de análisis WOT para calcular la probabilidad de transición hacia hCiPSC de cada celda + anotamos las celdas en la trayectoria correcta de acuerdo con la distribución de esta probabilidad entre los grupos de celdas.
* Función FindMarkers para identificar genes expresados ​​diferencialmente (DEG) entre grupos de células.
* El análisis GO se realizó a esos DEG utilizando el paquete R clusterProfiler

### [3] Wang G, Wang Y, Lyu Y, He H, Liuyang S, Wang J, Sun S, Cheng L, Fu Y, Zhu J, Zhong X, Yang Z, Chen Q, Li C, Guan J, Deng H. Chemical-induced epigenome resetting for regeneration program activation in human cells. Cell Rep. 2023

### Jun 27;42(6):112547. doi: 10.1016/j.celrep.2023.112547. Epub 2023 May 23. PMID: 37224020.

#### 1.Conclusiones/Descubrimientos

* Se usa el análisis single-cell de transcriptoma para demostrar que la vía de la reprogramación química humana con estado de regeneración es distinta de la de reprogramación mediada por factores de transcripción (OSKM).
* Caracterización del programa genético similar a la regeneración (estado plástico intermedio)
  + **Presencia de células plásticas intermedias al final de la etapa II de la reprogramación química en la trayectoria correcta (R4) que expresaban genes clave del desarrollo, además de la coexpresión de genes relacionados con la regeneración (los cuáles a su vez se relacionan con el desarrollo de extremidades embrionarias)**
  + Mediante Multiome scATAC y expresión génica se detecta el programa genético similar a la regeneración (G3) mediante el estudio de genes que adquirieron una mayor accesibilidad a la cromatina, así como una expresión génica regulada positivamente en las células plásticas intermedias (157 genes)
  + En el programa de regeneración se incluyen genes como SALL4 , MSX1 y HOXB9. También se identificó el programa de genes somáticos (G1) que fue reprimido antes de la activación del programa de genes similar a la regeneración.
  + Gracias a CUT&Tag (analizar cromatina) se mapean cuatro modificaciones de histonas, incluidos los principales marcadores activos promotores y potenciadores H3K4me3 (marca el sitio de inicio de la transcripción de genes activos), H3K4me1 y H3K27ac y el marcador represivo H3K27me3. Sugieren que en la etapa II, las células obtuvieron gradualmente más loci de genes con un estado de cromatina activa a nivel global en comparación con las células somáticas iniciales
  + Las dinámicas de actividad del promotor (región cercana a un gen) y potenciador (sitio de unión a factores de transcripción: activar/desactivar un gen) sugieren la existencia de dos oleadas principales de remodelación del epigenoma, una para el silenciamiento epigenético del programa de fibroblastos (potenciadores de menor actividad) y otra para la activación epigenética de genes implicados en el programa similar a la regeneración (potenciadores de mayor actividad).
  + La mayoría de las regiones promotoras en las células somáticas iniciales permanecieron en un estado represivo. Al final de la etapa II, la mayoría de los genes del programa similar a la regeneración habían aumentado la modificación de H3K4me3 en las regiones promotoras, se activaron el 49,04% de los genes (regulación negativa de la modificación H3K27me3). Esto sugiere que estos genes reprimidos en las células somáticas iniciales se activan mediante la eliminación del marcador de histonas represivas H3K27me3 para activar las regiones promotoras -> regulación negativa de H3K27me3 y la metilación del ADN en las regiones potenciadoras durante la reprogramación química. la familia TCF/LEF se enriqueció como reguladores ascendentes del programa de genes de regeneración
* La activación de un programa similar a la regeneración a través de una vía inversa ocurre durante la pérdida del potencial de regeneración en los organismos durante su maduración
* LEF1 se identifica como un regulador clave para la activación del programa de genes de regeneración. Concretamente, se encuentra regulado negativamente en las células con capacidad regenerativas del estado II. Su sobreexpresión promueve la activación del programa de regeneración.
* Cuando se activa el programa de regeneración se “downregulate” el programa somático (se inhibe vía JNK que lleva a la downregulation de los genes AP-1 y suprime el programa proinflamatorio; se inhiben H3K4me3, H3K4me1 y H3K27ac + decrece la accesibilidad a la cromatina en esas regiones )
* Genes relacionados con la regeneración: HOXA4, CDO1, MSX2, SNCG, IL2, PLBD1y DYDC2
* SEGUIR POR LA DISCUSIÓN.
  + 

#### 2.Datos (tipo, línea celular etc)

Estado plástico intermedio -> reprogramación parcial, final del estado II. “*Al final de la etapa II, detectamos la coexpresión de los genes relacionados con la regeneración*”.

* GO en trayectoria R0 (somática) y R4 (estado plástico intermedio).

*RNA-seq, scRNA-seq, scRNA+ATC-seq multiome data, CUT&Tag*

* **hADCS*:* Células madre mesenquimales derivadas de adipocitos humano adulto (1013). Linea “self derived”, mujer de 52 años.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Fase de reprogramación** | **scRNA-seq** | **CUT&Tag** | **WGBS** | **ScRNA+ATAC seq multiome** |
| Control (hADSC) | Marca de insignia1 con relleno sólido | **¿** | **¿** | **¿** |
| Stage I – Day 0 (Control, somáticas) | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido |
| Stage I - Day 4 | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido |  |  |
| Stage I – Day 8 | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido |
| Stage II – Day 4 | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido |  |  |
| Stage II - Day 8 | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido |  | Marca de insignia1 con relleno sólido |
| Stage II – Day 12 | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido |  |  |
| Stage II – Day 16  (Estado plástico) | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido | Marca de insignia1 con relleno sólido |

**Pre-procesamiento data hADSC -> somáticas, sin reprogramar**

*The clean single-cell RNA-seq reads were mapped to human reference genome hg19 using Cell Ranger v3.1.0: .tar.gz files containing cellranger output (count matrix, barcodes, genes id).*

* Matriz de conteo: 32738 genes x 5404 células**. Tiene el mismo número de genes que en los otros experimentos, podríamos comprobar si son iguales. En el caso de que los genes coinciden, podríamos realizar una comparación entre los experimentos cuando las células llegan a un estado plástico intermedio (o incluso a la pluripotencia) y los genes expresados. De esta forma, confirmaríamos si se expresan los mismos genes aunque los experimentos/reprogramación sea diferente (en la etapa elegida) o el caso contrario y ver qué diferencias hay.**
* Features: genes. Tiene información sobre el nombre de los genes, el tipo de gen y el id.
  + Hay genes con el nombre repetido. Nos fiamos de que son genes diferentes y usamos los ids (son todos diferentes), en un futuro usaremos estos ids para encontrar los nombres reales de los genes.
  + Todos son genes de expresión.
* Barcodes: céulas. Contiene los ids de las células.
  + Todos los barcodes son únicos así que damos por hecho que todas las muestras son únicas.

**Pre-procesamiento data stageII\_day16 -> estado plástico intermedio**

*The clean single-cell RNA-seq reads were mapped to human reference genome hg19 using Cell Ranger v3.1.0: .tar.gz files containing cellranger output (count matrix, barcodes, genes id).*

* Matriz de conteo: 32738 genes x 5404 células**. Tiene el mismo número de genes que en los otros experimentos, podríamos comprobar si son iguales. En el caso de que los genes coinciden, podríamos realizar una comparación entre los experimentos cuando las células llegan a un estado plástico intermedio (o incluso a la pluripotencia) y los genes expresados. De esta forma, confirmaríamos si se expresan los mismos genes aunque los experimentos/reprogramación sea diferente (en la etapa elegida) o el caso contrario y ver qué diferencias hay.**
* Features: genes. Tiene información sobre el nombre de los genes, el tipo de gen y el id.
  + Hay genes con el nombre repetido. Nos fiamos de que son genes diferentes y usamos los ids (son todos diferentes), en un futuro usaremos estos ids para encontrar los nombres reales de los genes.
  + Todos son genes de expresión.
* Barcodes: céulas. Contiene los ids de las células.
  + Todos los barcodes son únicos así que damos por hecho que todas las muestras son únicas.

**Datos preprocesados**

De cada experimento obtengo dos archivos.

* Archivo .csv con metadatos de los genes: contiene tres columnas (gene id, gene name, gene type). Para hacer un merge cuando haya que analizar.
* Archivo .txt.gz con los datos (células x genes). Contiene los conteos de cada gen por célula (lo que será la count matrix), también datos relativos al experimento (son sintéticos, 5 campos: reprogramming\_type, reprogramming\_stage, cell\_state, cell\_line, cell\_type. Si los campos asociados a la reprogramación son False significa que las células son somáticas, control sin reprogramar). Las columnas son los ids de los genes, mientras que las filas son los ids de las células.

Cuando queramos usarlo habrá que crear el campo Y (0 sin reprogramar, 1 reprogramado) y X será la matriz de conteos (quitar los 5 campos sintéticos).

#### 3.Usan proteínas para validar

#### 4.Algortimos de ML