

Endoderm Differentiation - Single Cell RNA-sequencing

Rivera Ramírez Andrés

Resumen

En este reporte se resume el proyecto en el que se llevó a cabo un análisis de expresión diferencial utilizando los datos del transcriptoma de células madre embrionarias humanas obtenido a través de Single-cell RNA-seq. Las células fueron tratadas con nocodazole o DMSO antes y después de diferenciarse hacia endodermo. Las colonias hPSC fueron tratadas con alguno de estos estímulos a una concentración de 100ng/ml por 16 horas e inducidas hacia el estado de endodermo por tres días.

Objetivo

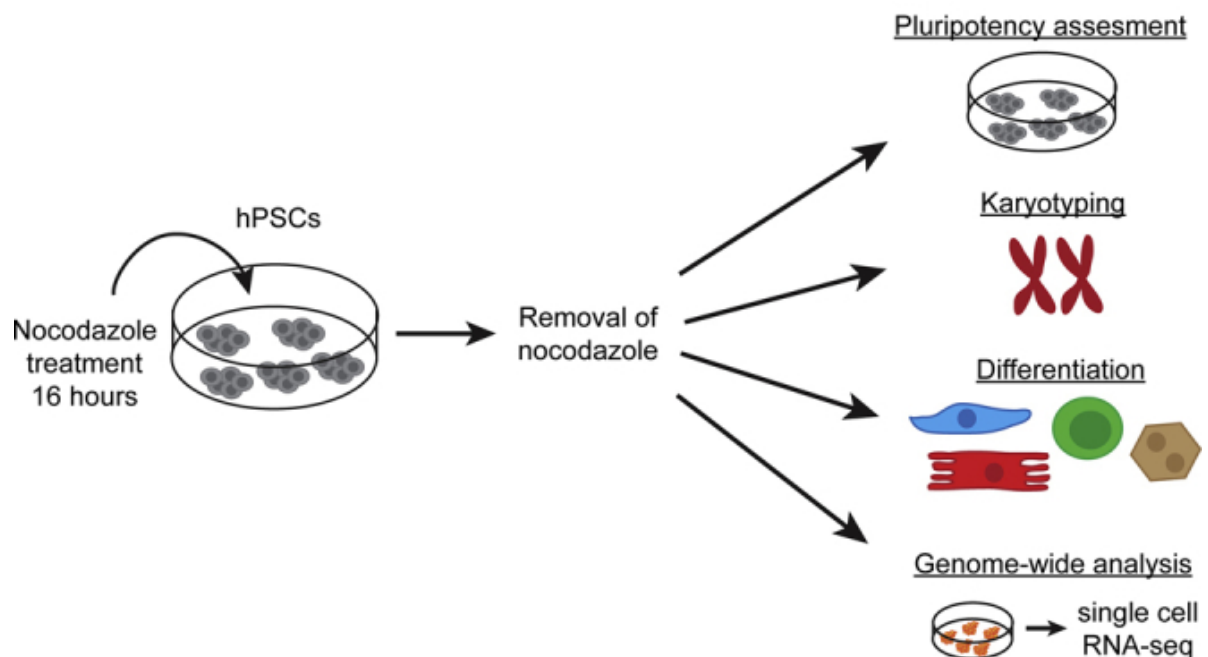
Obtener información acerca de genes diferencialmente expresados en células sometidas a moléculas inhibitoras de la progresión del ciclo celular, con el objetivo de sincronizar grandes cantidades de células, en particular nocodazole y DMSO.

Justificación

Las células madre pluripotentes humanas (hPSCs) son un modelo único para estudiar las difrenciaciones tempranas de las células ya que pueden ser crecidas in vitro de forma indefinida mientras mantienen la capacidad de diferenciarse en tres líneas germinales: endodermo, mesodermo y neuroectodermo. Sin embargo,, los estudios llevados a cabo en estas células se han visto obstaculizados debido a la dificultad de sincronizar de forma exitosa una gran cantidad de células en las distintas etapas del ciclo celular. Por lo tanto es importante desarrollar nuevas técnicas de sincronización celular en hPSCs .

Diseño experimental

Se llevó a cabo Single-cell RNA-seq en células tratadas con nocodazole y DMSO antes y después su diferenciación en endodermo. Las colonias e hPSC fueron tratadas con DMSO o nocodazole a concentraciones de 100ng/ml por 16 horas e inducidas a diferenciarse a endodermo definitivo por tres días. Se colectaron células únicas en condiciones indeferenciadas o en estado de endodermo y después se colocaron en 384 pozillos para llevar a cabo Smart-seq2.



Análisis de expresión diferencial.

El análisis se llevó a cabo utilizando distintos paquetes de Bioconductor.

Se utilizaron datos obtenidos a través de recount3 con el identificador ERP110066, con los cuales se generó un objeto RSE para comenzar con el análisis exploratorio de los datos. Al finalizar el análisis exploratorio despues de establecer los parámetros de calidad deseados se eliminaron ciertas lecturas que no cumplían con los niveles de expresión deseados.

Los datos fueron normalizados y finalmente se generó un modelo estadístico adecuado con los atributos de interés y se realizó el AED utilizando el paquete limma.

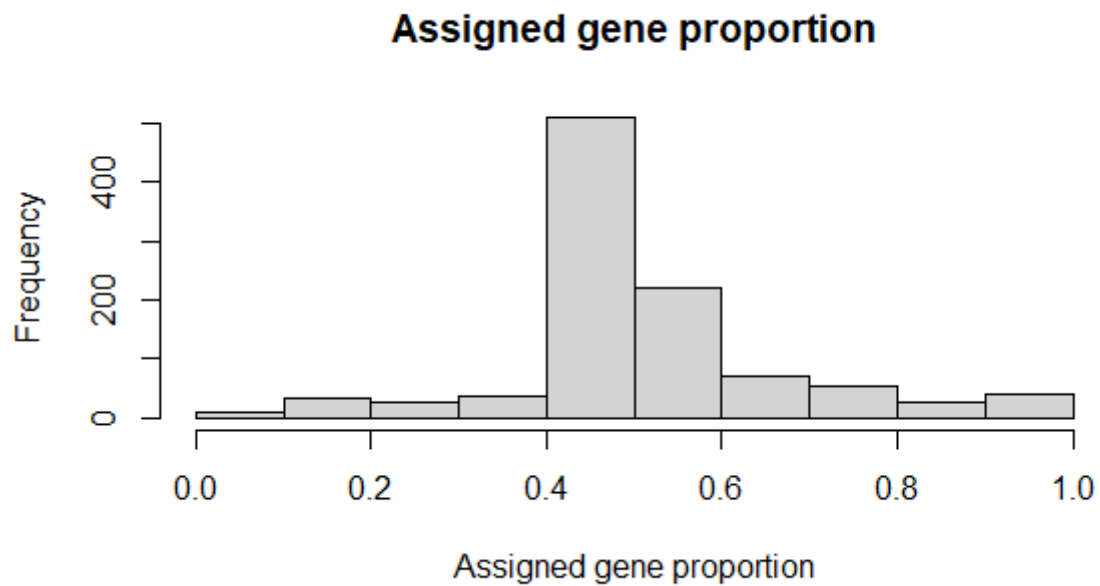
Resultados

Análisis exploratorio

Se comprobó que las células utilizadas para la secuenciación venían todas de una única línea germinal humana, eran células sanguíneas de un hombre adulto y los atributos en los que estaban diferenciadas era las condiciones de crecimiento que poseen dos categorías: cultivo para mantener pluripotencia y diferenciación hacia endodermo, y en el estímulo que recibieron que posee un gran número de categorías de entre las cuales la mas importante es la sincronización.

Parámetros de calidad

Se determinó la proporción de lecturas asignadas a genes para asegurar que las lecturas eran congruentes con el transcriptoma utilizado.



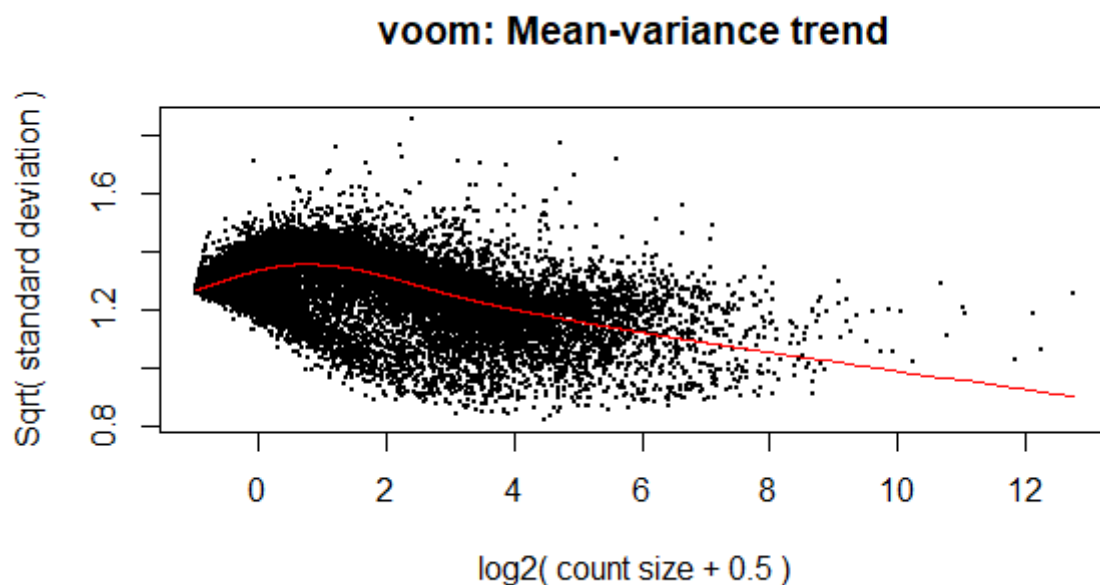
Se eliminaron aquellas lecturas de genes con una expresión por debajo de 0.1, quedando así el 34.89 de los genes,

Análisis de expresión diferencial

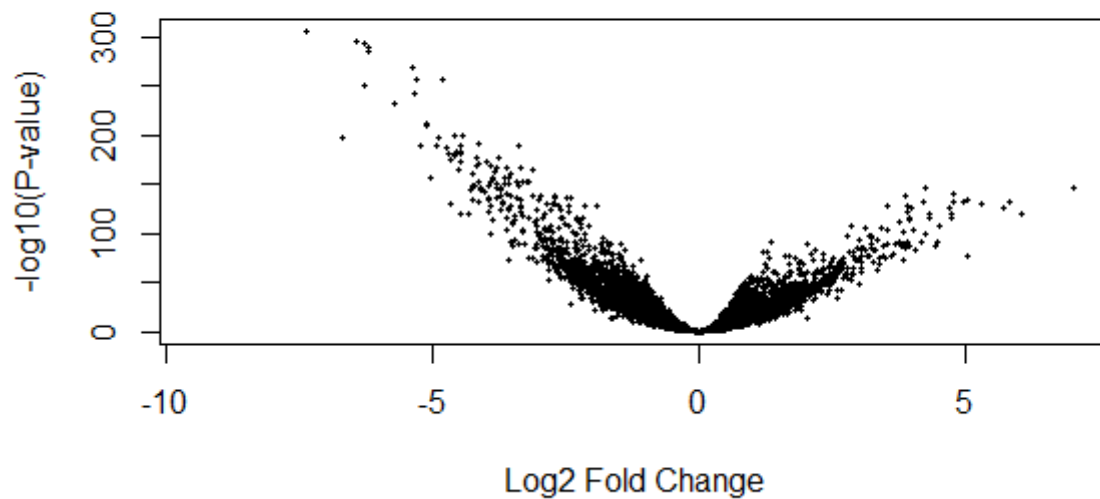
Tras la normalización de los datos se estableció un modelo estadístico para la normalización de los datos utilizando el atributo de condición de crecimiento.

Se realizó el análisis de expresión diferencial seleccionando los genes con un valor P inferior a 0.05.

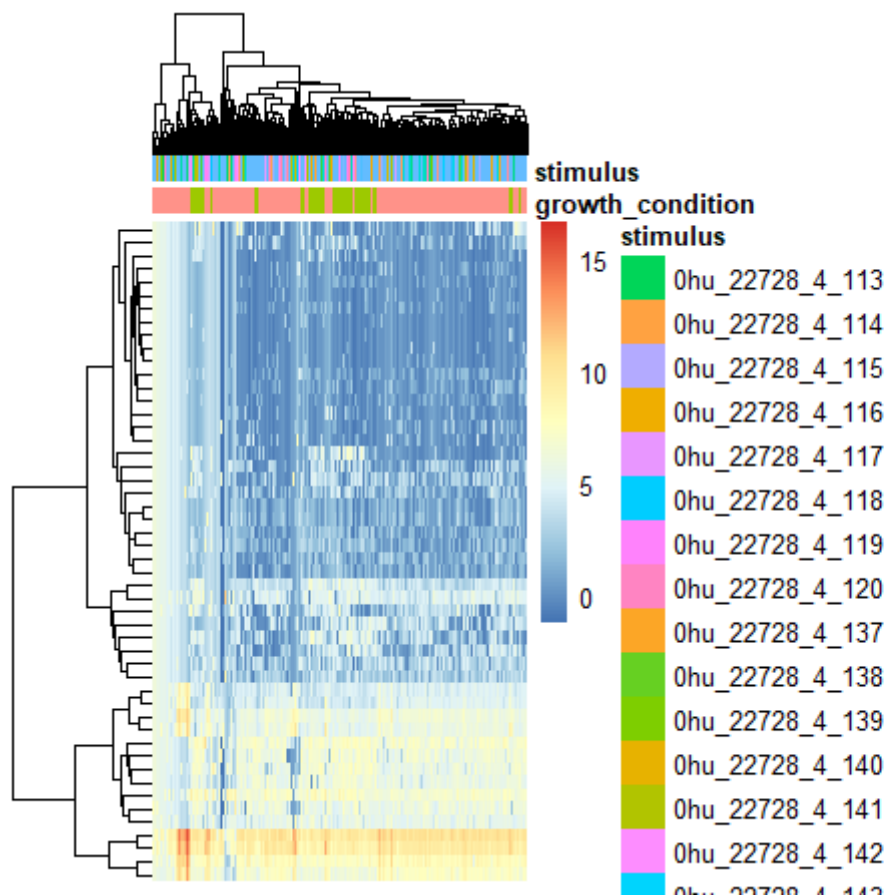
Las siguientes gráficas corresponden a la búsqueda de los DEGs.



En el caso de este volcano plot se puede apreciar que si existen lecturas que corresponden a genes diferencialmente expresados en la parte izquierda superior ya que cuentan con una gran magnitud en su fold change y significancia estadística.



Finalmente, se generó el siguiente heatmap con los valores P ajustados, sin embargo, debido a la gran cantidad de categorías en el atributo "stimulus" no resulta sencillo ver zonas de DEGs.



Conclusiones.

Debido a la estructura del RSE no fue posible obtener DGEs, ya que la cantidad de estímulos utilizados no permite construir un heatmap lo suficientemente explícito como para reconocer áreas de expresión diferenciada específicas y claras. Sería recomendable realizar un estudio con ambas condiciones de crecimiento presentadas en este y únicamente un par de estímulos correspondientes a 16 horas de DMSO y nocodazole para obtener datos que sean mas sencillos de utilizar.

Bibliografía.

- Yiangou L, Grandy RA, Morell CM, Tomaz RA, Osnato A, Kadiwala J, Muraro D, Garcia-Bernardo J, Nakanoh S, Bernard WG, Ortmann D, McCarthy DJ, Simoncic I, Sinha S, Vallier L. Method to Synchronize Cell Cycle of Human Pluripotent Stem Cells without Affecting Their Fundamental Characteristics. Stem Cell Reports. 2019 Jan 8;12(1):165-179. doi: 10.1016/j.stemcr.2018.11.020. Epub 2018 Dec 27. PMID: 30595546; PMCID: PMC6335580.