Back to the stemhood

A Systems Biology approach to identify extracellular factors involved in stemness of human embryonic stem cells

Virginie Mournetas[1,2], Patricia Murray[1], Christopher Sanderson[1], David Fernig[2]

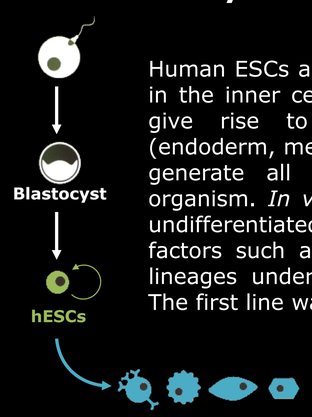
[1]Institute of Translational Medicine, [2]Institute of Integrative Biology, University of Liverpool, Crown Street, L69 7ZB

Human embryonic stem cells (hESCs) are a fundamental tool for understanding human embryonic development and the differentiation mechanisms involved in that process. Moreover, these cells have a huge potential in regenerative medicine for repairing/replacing tissues or organs due to functional deficiencies arising as a consequence of age, disease, damage or congenital defects. However, little is known about the regulation of hESC self-renewal; in particular, the role of cell-microenvironment interactions has been relatively unexplored.

In this project, I aim to characterize novel proteins and/or interactions responsible for stemness maintenance, focusing on the links between the extracellular matrix and the transcriptional network. As a first step, I identified candidate proteins using an *in silico* Systems Biology approach (use of high-throughput experiments, mathematical theories and computing tools to model complex systems). I am now undertaking an *in vitro* study to assess the role of these candidates in maintaining self-renewal and pluripotency of hESCs.

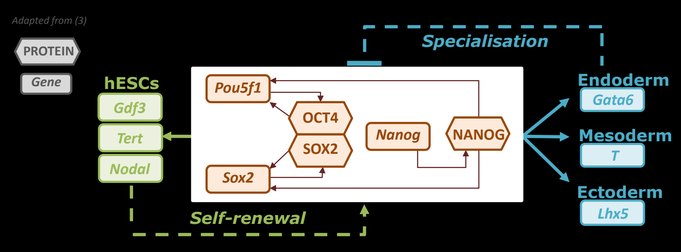
This work should give new clues for a better understanding of the stemness state in order to realise the great potential of hESCs.

Embryonic stem cell culture



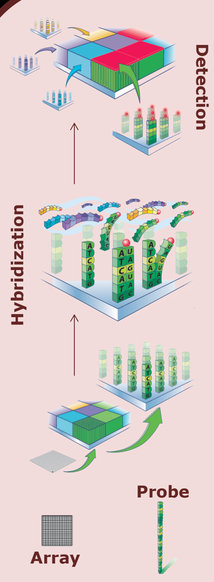
Human ESCs are transiently present *in vivo* in the inner cell mass of blastocysts. They give rise to the three germ layers (endoderm, mesoderm, ectoderm) and thus generate all cell types in the mature organism. *In vitro* they can be maintained undifferentiated by using specific growth factors such as FGF2[1], or committed into lineages under specific culture conditions. The first line was derived in 1998[2].

The ESC core transcriptional network



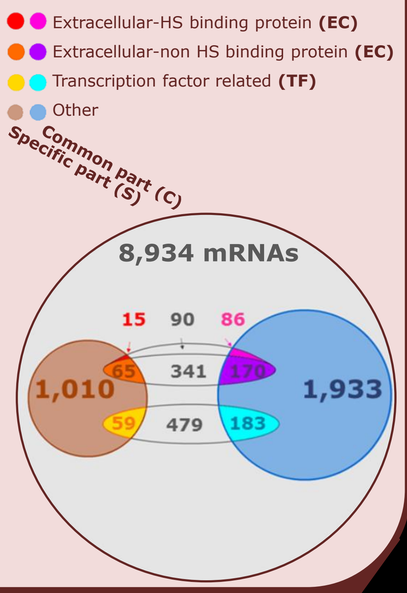
The ESC core transcriptional circuit regulating stemness consists of the three transcription factors OCT4, SOX2 and NANOG. Interactions between them and their target genes are essential for the maintenance of stemness[3] and these form the focus of many studies. However, it is clear that other factors impact on self-renewal and pluripotency, such as cell-cell and cell-matrix interactions, yet these are not as well characterised.

hESC transcriptome



By analysing microarrays (technology used to measure gene expression, *figure* on the left) from ArrayExpress public database[4], transcriptomes (sets of mRNAs) from hESCs and early committed hESCs were established and compared (*figure* on the right).

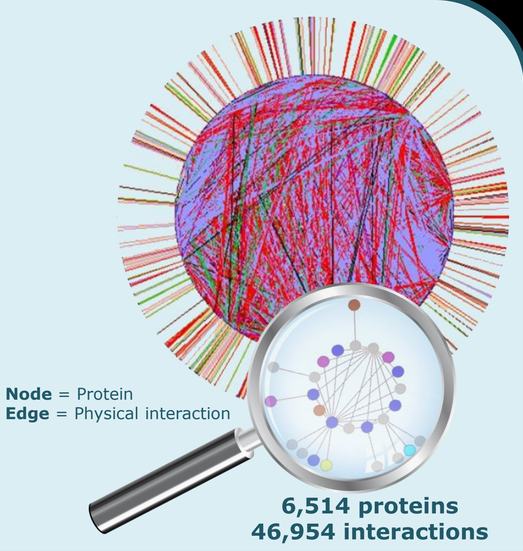
The hESC transcriptome containing 8,934 mRNAs has a specific mRNA set (S-1,010 mRNAs) only expressed in hESCs and a common one (C-1,933 mRNAs) expressed in hESCs and their early derived cells. Extracellular (EC) and transcription factor (TF) parts were found using the Gene Ontology database[5]. The EC part can be split into two sub-sets: heparan sulfate (HS) binding and non binding proteins.



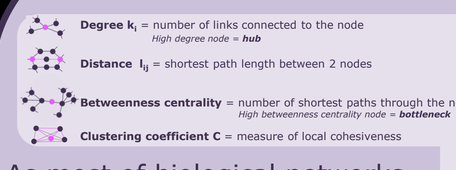
hESC interactome

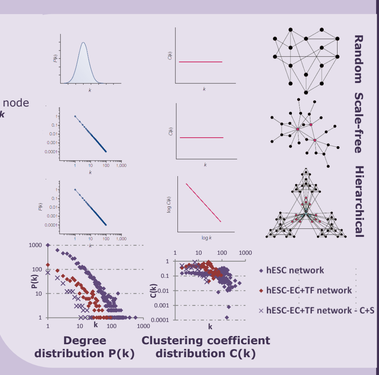
For the purpose of the analysis, we assumed that all mRNAs were translated to proteins. The hESC interactome (set of protein-protein interactions) was deduced from the hESC transcriptome (*panel* on the left) and built using STRING database[6].

It is composed of 6,514 proteins and 46,954 interactions and had been structurally studied (*panel* above) and filtered to identify a short list of 125 candidate proteins potentially involved in maintaining stemness (*panel* on the bottom-left). The filtering focused on the extracellular (EC) matrix molecules and their links with the transcription factor (TF) network.



Network architecture





As most of biological networks, the hESC interactome exhibits a scale-free structure, where the degree distribution P(k) follows a power-law $P(k) \sim k^{-\gamma}$, involving the presence of hubs, and the clustering coefficient distribution C(k) is independent of k meaning there is no inherent presence of modules (*figures* above).

Enrichments

GO Biological Processes term enrichment was used to determine if the EC+TF interactome contained meaningful sub-sets of proteins. Obviously, terms related to EC ('extracellular matrix organization') and TF ('transcription, DNA-dependent') appeared. More interestingly, terms related to the development ('in utero embryonic development'), cell adhesion, cell communication, cell differentiation and cell proliferation were also represented (*graph* on the bottom right).

Data analysis

Model design

DRY / *In silico* experiments

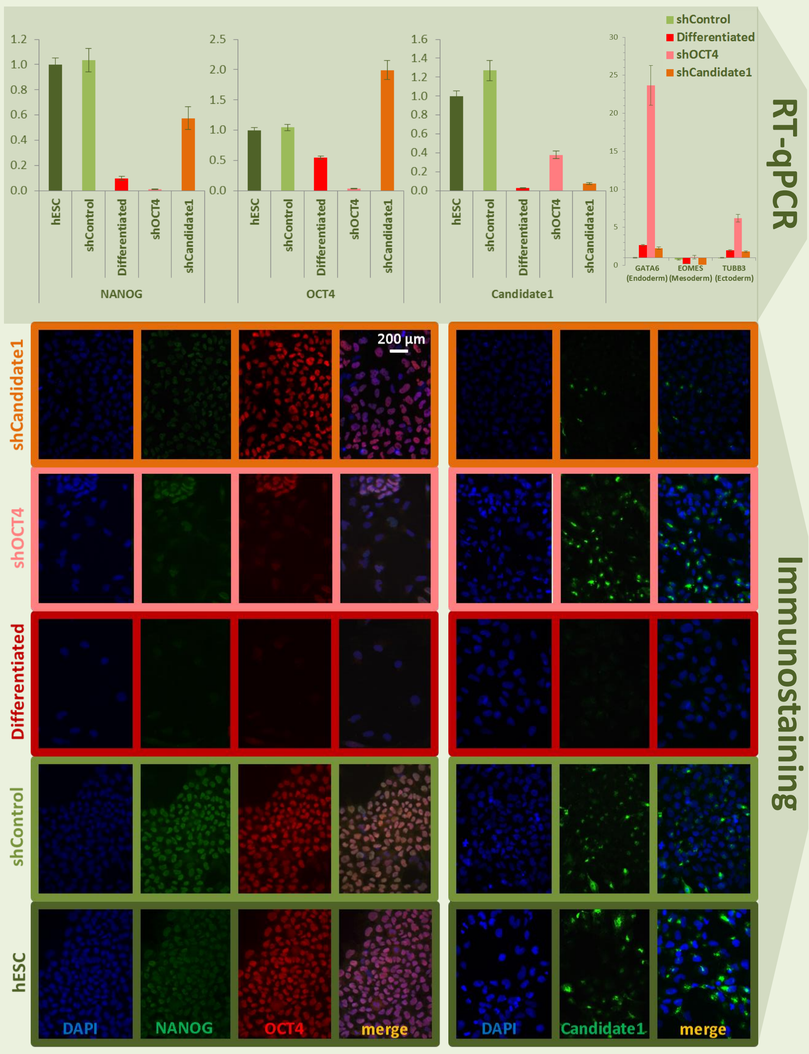
Predictions/ Hypothesis

WET / *In vitro* experiments

Experimental design

The candidates' role in hESC stemness - *in vitro* study

The role of these candidates in maintaining hESC stemness is assessed by an *in vitro* study where the effects of stable knockdown using shRNA on the expression of pluripotency and differentiation markers is investigated at the mRNA level using qPCR and at the protein level using immunofluorescence.



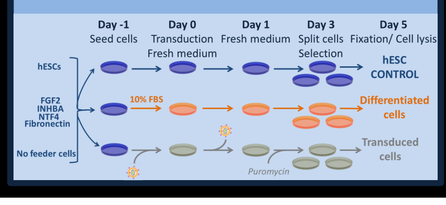
Graphs at the top show the relative gene expression of pluripotency markers NANOG and OCT4, Candidate1 and differentiation markers GATA6, EOMES and TUBB3 in five different conditions.
After 5 days of differentiation in 10% FBS (red) or when OCT4 is knockdown (pink), hESCs loose pluripotent markers and candidate1 expression whereas differentiation markers are upregulated.

The co-immunostaining of NANOG, OCT4 and candidate1 above confirms qPCR results.

The *in vitro* strategy

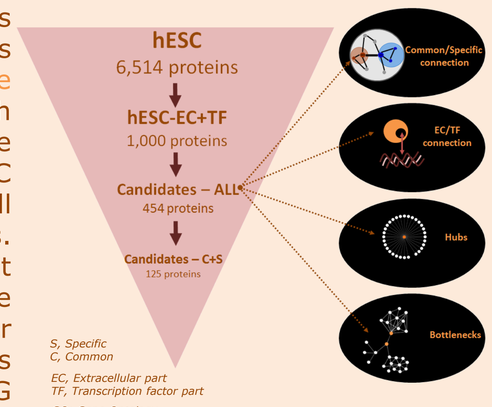
Messenger RNAs of interest were stably knockdown using a small hairpin (sh) RNA lentivirus approach (*figure* below).

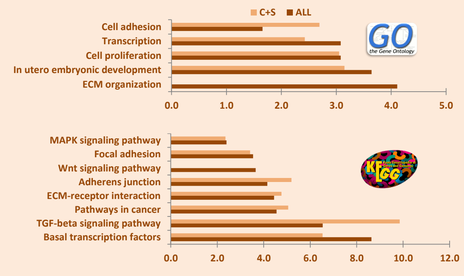
HESCs alone or transduced with a shRNA control were the undifferentiated controls whereas hESCs cultured in 10%FBS or transduced with an shRNA against OCT4 were the differentiated controls (*figure* below and *panel* on the left).



The candidates' role in hESC stemness - *in silico* study

The list of 125 candidates is composed of hubs (highly connected proteins), bottlenecks and nodes in the S/C and EC/TF interfaces (*figure* on the right). Hubs are central proteins due to their high number of interactions and bottlenecks are connections between different processes. The S/C interface reflects the links between the general cell machinery and the features characteristic to hESCs. The EC/TF interface is also essential, since it represents the communication between the genome and the cell's environment, which includes other cells. To determine if this list still contains proteins of interest, GO Biological Processes and KEGG pathway enrichments were processed (*graphs* below).





The list of candidates incorporates proteins involved in development and cell communication. Among these proteins, some are already known to be required for hESC stemness maintenance, either directly such as NODAL[7], FGF2[7] or Activin A[7], or indirectly through signalling pathways such as TGF-β[8] or Wnt[9]. Other proteins are also known for their role in mouse ESC pluripotency but not in hESC yet, such as TCF3[10]. However, for the majority, nothing is known yet in the hESC context.

vmounet@liv.ac.uk
http://bit.ly/IcFl6V
@VMournet

UNIVERSITY OF LIVERPOOL

(1) Xu, Nat Methods, 2005. 2: p. 185-90
(2) Thomson et al., Science, 1998. 282: p. 1145-7
(3) Macarthur et al., Nat Rev Mol Cell Biol, 2009. 10: p. 672-81
(4) http://www.ebi.ac.uk/arrayexpress/
(5) http://www.geneontology.org/
(6) http://string-db.org/
(7) Vallier, J Cell Sci, 2005. 118: p.4495-509
(8) James, Development, 2005. 132: p. 1273-82
(9) Sato, Nat Med, 2004. 10: p. 55-63
(10) Cole, Genes Dev, 2008. 22: p. 746-55