Nature Reviews Genetics | AOP, published online 9 March 2011;
doi:10.1038/nrg2975

GETTY & IMAGE100

STEM CELLS

Reprogramming's unintended consequences

If induced pluripotent stem cells (iPSCs) are to be used safely and effectively in regenerative medicine and disease research, we need to know the impact of reprogramming on genomic and epigenomic integrity. Three papers that look at point mutations, copy number variations and DNA methylation, respectively, provide lessons for researchers seeking to use iPSCs in these applications.

Gore *et al.* performed exome sequencing on 22 human iPSC lines — generated in seven laboratories by five different methods — and the fibroblast lines from which they derive. They validated 124 point mutations that were fixed in the iPSCs but not the fibroblasts, and from this they predict a mutational load of six coding mutations per iPSC genome. Notably, their study identified many missense mutations that are predicted to alter protein function, and point mutations were enriched in genes that are implicated in cancers. Through ultradeep sequencing, they found that around half of the mutations pre-existed at low levels in the fibroblast populations and the others occurred during or after reprogramming. The authors suggest that mutational load is more likely to be influenced by selection during reprogramming and culturing than by reprogramming being intrinsically mutagenic.

Another aspect of genomic integrity was studied by Hussein and colleagues. Using SNP arrays, they looked at copy number variants (CNVs) in a large number of human embryonic stem cell (ESC) lines, human iPSC lines, and their corresponding fibroblast lines. iPSCs had a higher average number of CNVs than did ESCs, but this number decreased with passaging. The authors showed that CNVs form at a high rate during reprogramming (possibly owing to replication stress), leading to genetic mosaicism in early-passage iPSC populations. During propagation, rapid selection occurs against cells with high numbers of CNVs, causing iPSCs to be more similar to ESCs at later passages. However, the authors caution that it is possible that some CNVs could confer a selective advantage.

Several studies have suggested that the epigenomes of iPSCs and ESCs differ. Lister *et al.* addressed this in detail by using shotgun bisulphite sequencing to map DNA methylation genome-wide at single-base resolution in iPSCs, ESCs, somatic cells and cells differentiated from the pluripotent lines. They found that although the methylomes of iPSCs and ESCs are broadly similar, the five iPSC lines studied had notable differences to ESCs and to each other, and some differences persist after differentiation. The differences in iPSCs were a mixture of methyl marks that represent a 'memory' of the somatic cell type from which the iPSCs were generated and other iPSC-specific changes. Of note are differentially methylated regions that are shared among lines — suggesting that some regions are particularly susceptible to aberrant reprogramming — and megabase-scale regions near centromeres and telomeres that have altered non-CG methylation and are associated with changes in transcription and histone methylation.

As well as underlining the need for careful characterization of iPSC lines, these studies raise intriguing questions about the relative influences of cell division and culture, and of molecular events that are intrinsic to reprogramming.

Mary Muers

ORIGINAL RESEARCH PAPERS Gore, A. *et al.* Somatic coding mutations in human induced pluripotent stem cells. *Nature* **471**, 63–67 (2011) | Hussein, S. M. *et al.* Copy number variation and selection during reprogramming to pluripotency. *Nature* **471**, 58–62 (2011) | Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68–73 (2011)