MASS SPECTROMETRY

Designer proteases for post-translational modifications

Researchers have engineered a protease specific for a post-translational modification.

Post-translational modifications (PTMs) come in many flavors and broadly influence protein behavior. Mapping the positions of a specific PTM from the single-protein to the genome-wide level represents an important problem in modern biochemistry. One powerful method for such mapping is mass spectrometry (MS).

In a typical MS workflow, proteins are digested by the protease trypsin, which selectively cleaves the C-terminal peptide bonds of Arg and Lys. The mass-to-charge ratio for the resulting peptides is then determined and offset from the expected ratio based on the primary sequence can often be specifically attributed to one or more PTMs. The PTM must occur on a peptide of appropriate size, which is determined by the positions of Arg and Lys within the protein. When Arg and Lys spacing do not yield the correct peptide sizes, coverage gaps can conceal PTMs and hinder mapping. Surprisingly, very few proteases exist as backups for trypsin.

To address this challenge, Brian Paegel and Duc Tran, a postdoc fellow in Paegel's laboratory at Scripps Research Institute in Florida, sought to develop better protease tools for studying modified proteins. According to Paegel, the work was born from conversations he had in 2008 with colleague and coauthor Valerie Cavett about the limitations in mass analysis of proteins, which led him to "start sketching schemes to evolve proteolytic function."

Their idea was to develop proteases that specifically cleave proteins at the site of target PTMs, which would facilitate precise determination of the PTM positions by MS. The notion of such 'designer proteases' was not new, but no such mutants of trypsin exist—possibly because trypsin has many residues involved in binding, making it a challenging target for directed evolution.

The group developed a trypsin mutant that cleaves at citrulline, a PTM implicated in several important epigenetic and immunological functions. To develop this mutant, they

STEM CELLS

STEM CELLS FROM THE YOUNG AND OLD

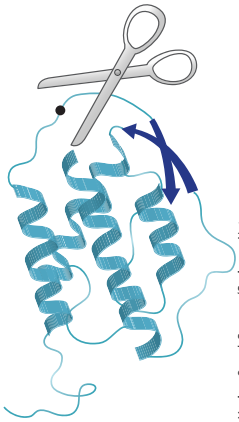
Donor age impacts the epigenetic state and mutational burden of induced pluripotent stem cells.

The reprogramming of adult somatic cells to pluripotent stem cells does not amount to a complete epigenetic reboot, at least not using current protocols. Researchers have found that induced pluripotent stem cells (iPSCs) preserve some epigenetic memory of their donor cells. They can also carry somatic mutations that were present prior to reprogramming.

The derivation of iPSCs from a patient's cells is a potential route for repairing damaged organs or genetically deficient cells without fear of rejection by the immune system. To better understand these cells and to minimize risks for therapy, it is important to know the effects of source tissue and reprogramming technique on iPSCs. Yet the role of donor age has not been systematically examined.

At the Scripps Research Institute in California, Kristin Baldwin, Ali Torkamani and colleagues studied the effect of donor age on iPSCs generated from frozen peripheral blood mononuclear cells (PBMCs). Cells from sixteen individuals ranging in age from 21 to 100 were expanded and enriched for erythroid progenitor cells and reprogrammed using an optimized episomal protocol. The researchers generated and selected three lines with normal karyotypes from each individual, then they profiled DNA methylation on a chip and carried out whole-exome sequencing.

Variation in DNA methylation patterns among the iPSC lines showed a largely linear but relatively weak relationship between donor age and age-dependent methylation patterns established from blood. Cells from older donors appeared to resist the characteristic global demethylation that occurs during reprogramming, particularly at age-associated DNA methylation sites. Across the genome, stem cells derived from older donors exhibited about 5% more methylation than the younger donors.



Marina Corral Spence/Springer Nature

A cartoon depiction of a designer protease cleaving at a PTM.

had to develop a method for identifying protease mutants with the desired activity and high specificity. "We tried and failed at many strategies to identify functional mutants before arriving at the activity-based screening strategy," recalls Paegel. The activity-based screen involved using a home-made bisamide rhodamine probe, which fluoresces only upon cleavage at citrulline and allowed them to identify mutants that were highly specific.

Paegel says they got "a lucky break" when they found a mutation in the protease that improves the expression both *in vitro* and in cells; and they now include this mutation in all their designer proteases. He calls this a stark reminder of Leslie Orgel's second rule, "evolution is cleverer than you are."

After successfully generating the citrulline-specific protease, the team demonstrated that this protease allows them to precisely determine 12 sites of citrullination in the enzyme protein arginine deiminase 4 and 25 sites of citrullination in fibrinogen, two of which were previously unknown. They are currently testing the protease on a proteome-wide scale.

Future directions involve generalizing the directed evolution approach to proteases that target other PTMs. Paegel and his team are interested in developing proteases that recognize PTMs that are labile in MS and are thus difficult to observe by changes in mass-to-charge ratio. Examples include phosphorylations of serine, threonine or sulfated tyrosine. These upcoming tools, as well as the citrulline-specific protease, signal a promising future for precision PTM mapping.

Rita Strack

RESEARCH PAPERS

Tran, D.T. *et al.* Evolution of a mass spectrometry-grade protease with PTM-directed specificity. *Proc. Natl. Acad. Sci. USA* **113**, 14686–14691 (2016).

How this age signature may affect cell function is still unclear. Luckily, iPSCs exhibit memory loss upon continued passaging. The researchers found age-associated differences at the eighth cell culture passage, but adding 20 more passages eliminated these differences.

Exome sequencing led to potentially greater concerns for clinical practice than the results of epigenetic profiling. The number of mutations in genic regions was doubled in iPSCs from donors in their late eighties compared with those in their early twenties. The vast majority of these mutations were likely generated prior to reprogramming, and many may have functional implications. Of the 326 candidate somatic mutations that the researchers detected in total from all individuals, 24 were predicted to knock out genes important for cellular function, implicated in disease or tumor suppression.

Most studies do not use cells from older individuals, as these can be less efficient at reprogramming (fibroblast cells exhibit age-related decline, for example). Additional work, including a sampling of more donors and different source tissues, is needed to confirm these trends in genetic and DNA methylation heterogeneity as a function of age. But at a minimum, the findings of Torkamani, Baldwin and colleagues suggests that multiple lines should be screened for safety before clinical use, in particular for older donors.

The work also has an interesting coda. For the oldest donors who were at least 90 years of age, iPSCs had epigenetic signatures that resembled those derived from donors as young as half their age. The ratio of transversions to transitions revealed that elderly donor cells have undergone fewer divisions than expected—consistent with a model in which very old donors lack rapidly dividing progenitors which accumulate more somatic mutation in middle age.

Tal Nawy

RESEARCH PAPERS

Lo Sardo, V. *et al.* Influence of donor age on induced pluripotent stem cells. *Nat. Biotechnol.* **35**, 69–74 (2017).