Ultra-Sensitive Duplex Sequencing for Quantifying Multi-Individual Cell Therapy Sub-Population Fractions

Zach Norgaard[1], Jake Higgins[1], Jeffry Yaplee[1], Joe Blake[3], Mary Prieve[3], Fang Yin Lo[1], Colleen Delaney[3], Charles C Valentine[1], Jesse J Salk[1,2]

[1] TwinStrand Biosciences, Seattle, WA
[2] University of Washington, Division of Medical Oncology, Seattle, WA
[3] Deverra Therapeutics, Seattle, WA

BACKGROUND & AIM

Umbilical cord blood (CB) contains CD34+ hematopoietic stem cells and is routinely used as a source of donor stem cells for allogeneic transplant. CB units with insufficient cell dose for transplant largely go unused but are increasingly being repurposed as source material to generate cell therapies and regenerative medicine products. If these small units are pooled to generate a cell therapy product, current techniques are often insufficiently sensitive for quantitative deconvolution of these mixtures to assess the relative expansion of each subpopulation during the manufacturing process or after infusion into patients. A more sensitive and widely applicable assay would accelerate development of these new treatments. Duplex Sequencing (DS) compares both strands of each original DNA molecule to eliminate technical errors and achieve extreme accuracy and sensitivity, with an error rate below one-in-ten-million.

METHODS, RESULTS & CONCLUSION

We designed a hybrid capture panel targeting 277 single nucleotide polymorphisms (SNPs) to distinguish each contributor genome in a complex mixture. From 16 individual CB samples, we prepared 10 synthetic mixtures of up to 8 contributing genomes at fractions ranging from 82% to 0.05%. We performed DS on the individual CB samples and the mixtures, obtaining a total of 23 billion duplex error-corrected bases across all samples and a mean duplex molecular sequencing depth of 33,191x for the mixtures. We identified 80 informative alleles specific to a single CB sample (median 4 per individual, range 1-14). To determine the estimated fraction of a specific CB genome in a mixture, we divided the total number of molecular counts of that individual's informative alleles by the total effective duplex molecular depth at all relevant informative sites. Operators and analysts were blinded to the composition of 7/10 mixtures until after observed genomic fractions were calculated. We observed all contributor genomes at near expected and highly correlated (R=0.99) frequencies in every mixture.

DS provides unprecedented statistical power for mixture deconvolution through a combination of extreme accuracy and the ability to simultaneously capture a nearly unlimited number of independent lineage-defining genetic variants. This same principle can be extended to tracking the abundance of subpopulations in a patient, showing great potential for using DS in allogeneic cell therapies, transplant, and regenerative medicine.

MKT0028 v1.0