Identification of gene co-expression modules associated with long-term engraftment potential hematopoietic stem cells (LT-HSC) by analysis of public transcriptome data

Gustavo Bueno[1], Karina Griesi Oliveira[1]

[1]Hospital Israelita Albert Einstein

Long-term hematopoietic stem cells (LT-HSCs) are essential for the maintenance of adult hematopoiesis due to their self-renewal capacity and multipotent differentiation potential[1,2]. However, culturing these cells remains a challenge, as surface markers commonly used to identify fresh LT-HSCs lose their reliability after cultivation[3,4]. Although xenotransplantation remains the gold standard for functional validation, it is a time-consuming and costly method. Nevertheless, it remains unclear which genes are essential to the identity of LT-HSCs, regardless of their state—fresh or cultured[5]. This study aimed to identify gene co-expression modules associated with LT-HSCs by analyzing public transcriptome datasets using the Weighted Gene Co-expression Network Analysis (WGCNA) approach. Over 40 publicly available studies were screened, and five were selected based on strict criteria, including adequate sample numbers, subpopulation separation, and functional validation through xenotransplantation. Two of these studies included fresh LT-HSCs, while three involved cells cultured under conditions known to result in higher or lower enrichment of LT-HSCs. The selected datasets were carefully curated and organized to ensure quality and suitability for analysis. Subsequently, co-expression networks were individually constructed using the WGCNA approach. The overlap of LT-HSC-associated modules across the selected studies revealed a set of 69 common genes. Among these, HLF, EPCR, and FGD5 had previously been described in the literature for their roles in HSC biology, reinforcing the consistency of our findings[6–9]. In addition, several novel candidate genes were identified that may represent promising targets. To investigate whether these genes were indeed associated with LT-HSC-enriched conditions, their combined expression was summarized using the first principal component (eigengene) and correlated with experimental conditions indicating higher or lower LT-HSC enrichment in independent datasets. This analysis revealed a strong association across all datasets evaluated. These findings suggest that the co-expression pattern of the identified genes can be used to recognize LT-HSCs, even in cultured contexts where traditional surface markers are no longer reliable. This study contributes to advancing our understanding of LT-HSC biology and lays the groundwork for future applications in compound screening, drug discovery, and the improvement of ex vivo expansion protocols.

Key words: Hematopoietic Stem Cells, LT-HSC, Transcriptomics, WGCNA

Funding: FAPESP

References

1. Eaves, C. J. Review Series Hematopoietic stem cells: concepts, definitions, and the new reality. (2015) doi:10.1182/blood-2014.
2. Kumar, S. & Geiger, H. HSC Niche Biology and HSC Expansion Ex Vivo. Trends in Molecular Medicine vol. 23 799–819 Preprint at https://doi.org/10.1016/j.molmed.2017.07.003 (2017).
3. Notta, F. et al. Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. Science (1979) 333, 218–221 (2011).
4. Majeti, R., Park, C. Y. & Weissman, I. L. Identification of a Hierarchy of Multipotent Hematopoietic Progenitors in Human Cord Blood.
5. Wilkinson, A. C., Igarashi, K. J. & Nakauchi, H. Haematopoietic stem cell self-renewal in vivo and ex vivo. Nature Reviews Genetics vol. 21 541–554 Preprint at https://doi.org/10.1038/s41576-020-0241-0 (2020).
6. Belluschi, S. et al. Myelo-lymphoid lineage restriction occurs in the human haematopoietic stem cell compartment before lymphoid-primed multipotent progenitors. Nat Commun 9, (2018).
7. Lehnertz, B. et al. HLF Expression Defines the Human Hematopoietic Stem Cell State. https://kleintools. (2021).
8. Gazit, R. et al. Fgd5 identifies hematopoietic stem cells in the murine bone marrow. Journal of Experimental Medicine 211, 1315–1331 (2014).
9. Fares, I. et al. EPCR expression marks UM171-expanded CD34 1 cord blood stem cells Key Points. (2017) doi:10.1182/blood-2016-11.