APPLICATION NOTE

Tackling the globin problem in single-cell malaria research.

Authors

Barbara Stokes,
Fiona Achcar,
Julie Galbraith,
Matthias Marti

School of Infection and Immunity,
College of Medical, Veterinary &
Life Sciences
University of Glasgow, Scotland

Introduction

Human whole blood is a challenging sample type for transcriptome analysis because most circulating cells are enucleated erythrocytes or immature reticulocytes, whose RNA content is dominated by hemoglobin transcripts. In typical adult samples, HBA1, HBA2, and HBB together can represent ~50% of total polyadenylated RNA extracted from whole blood, while HBD contributes at much lower levels (<1%)[1,2]. These high globin levels can obscure low abundance transcripts and increase apparent duplication rates.

Various approaches have been developed to remove or minimize globin reads. Early probe-based methods used biotinylated oligonucleotides linked to magnetic beads to pull down globin mRNA upstream of library preparation. More recent RNase H depletion protocols achieve comparable globin removal through hybridization of short DNA probes, followed by RNase H digestion of RNA:DNA hybrids.

However, both strategies have important limitations. They require relatively high RNA input, perform best with intact RNA, and show reduced efficiency with low-yield or degraded samples[2,3]. The hybridization and enzymatic steps can fragment molecules and reduce read length, compromising full-length coverage and limiting their suitability for long-read sequencing. Their reliance on pre-hybridization steps also constrains use in workflows that process individual cells or nuclei. Most critically, the depletion reactions act on bulk RNA prior to reverse transcription.

The pre-hybridization step mixes RNA from many cells into a single pool, so single-cell barcodes cannot be maintained. As a result, conventional probe- and RNase H based depletion methods are incompatible with single-cell transcriptomic assays.

This limitation is particularly evident in malaria research, where one of the current challenges is resolving parasite transcriptional activity within individually infected erythrocytes. In uncomplicated *Plasmodium falciparum* infections, most infected red blood cells harbour one parasite per cell, although multiply infected cells can occur under high parasitemia or mixed-species conditions. Recent studies have shown that analysing gene expression at the level of individual *Plasmodium* parasites reveals developmental heterogeneity and host-parasite interactions that bulk RNA-seq cannot resolve[4-7]. In erythrocytes, ~90-95% of total cellular mRNA corresponds to globin RNA[8], masking the comparatively scarce parasite mRNAs.

Here we describe the use of a CRISPR/Cas9 based depletion approach, also known as DASH[9], on infected cells to selectively cleave unwanted hemoglobin sequences while preserving single-cell resolution.

Methods and results

Blood cells (O+) were obtained from the NHS and were leukocyte-depleted prior to parasite infection. Bone marrow aspirates were commercially purchased from CGT Global. Samples (blood or bone marrow) were infected with schizont-stage *Plasmodium falciparum* parasites from the Pf2004/164-tdTomato gametocyte reporter line[10].

Cultures were maintained in a 5% CO_2, 1% O_2, and 94% N_2 gas mixture and grown either in standard growth medium consisting of RPMI 1640 with 25 mM HEPES and 24 mM sodium bicarbonate supplemented with 100 µM hypoxanthine, 10 µg/mL gentamicin, and 10% O+ human serum, or in minimal fatty acid medium (gametocyte-inducing) consisting of RPMI 1640 with 25 mM HEPES and 24 mM sodium bicarbonate supplemented with 100 µM hypoxanthine, 10 µg/mL gentamicin, 0.39% fatty acid-free BSA, 30 µM oleic acid, and 30 µM palmitic acid.

Samples (31,351 cells on average) were processed using the 3′ GEX Next GEM protocol (10x Genomics). For each sample, 10 ng of amplified 10x cDNA was treated with the NEXTFLEX™ Cas9-gRNA Globin Depletion Enzyme (Human; Revvity), which contains 30 unique guide RNAs targeting human HBA1, HBA2, HBB, and HBD. Depleted and undepleted aliquots from each sample were then carried through the standard 10x Genomics library construction protocol according to the manufacturer's instructions. Libraries were pooled and sequenced on an Illumina® NextSeq® 2000 instrument (Figure 1). FASTQ files were aligned to a mixed reference of the human genome and *Plasmodium falciparum* (PlasmoDB v65) using Cell Ranger (v7.0.0). Output matrices were filtered to retain only cells with more than 10 features and 200 UMIs using Seurat v5.1.0, and raw counts were visualized using ggplot2 v3.5.1.

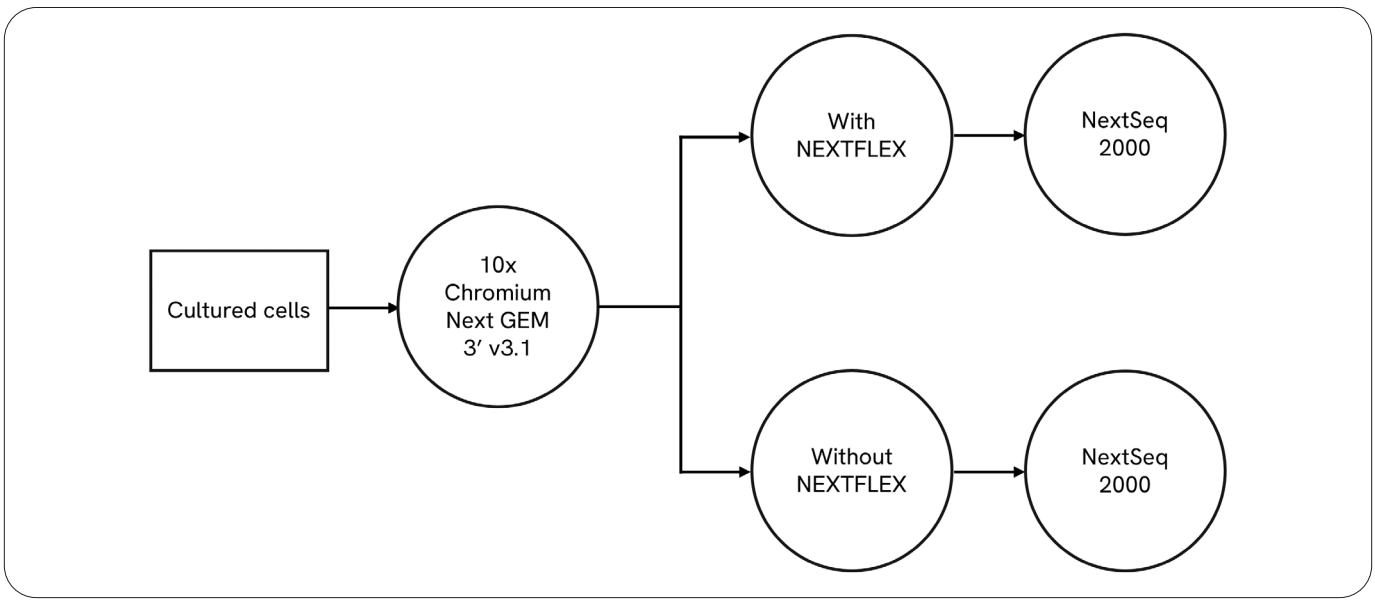

Figure 1: Experimental design used in the pilot study described.

Cultured blood cells show ~95% of UMI counts mapping to globin genes. After depletion this percentage is much lower and ranges from 9–16%, depending on the sample inspected (Figure 2A). In cultured bone marrow we also observe ~95% of UMI counts mapping to globin genes before depletion.

After depletion, the globin UMI fraction ranges from 7–39% (Figure 2B). Depletion efficiency appears to be more variable in bone marrow than in blood cells, although the reason for this has not yet been systematically investigated.

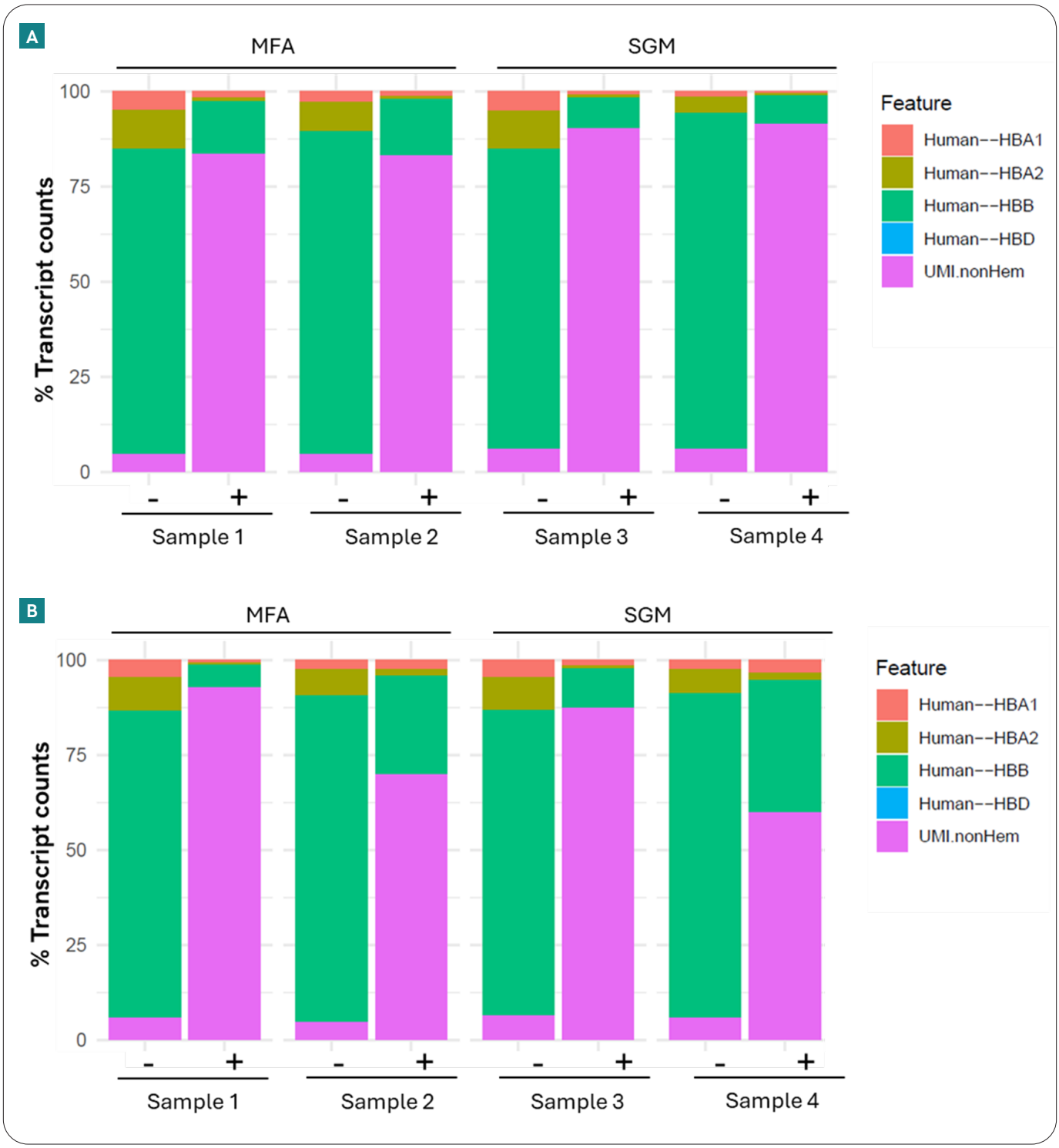

Figure 2: Comparison of detected transcripts in (A) infected blood cells or (B) infected bone marrow cells. Depleted samples (+) show lower percentage of UMI counts mapping to the target globin genes compared to undepleted ones (-). Variability observed does not seem to depend on growth media used in this set of samples. Minimal Fatty Acid Media (MFA). Standard Growth Media (SGM).

Conclusions

Across the experiments performed, we observe greater than 3-fold depletion of the NEXTFLEX Cas9-gRNA globin targets on average, with up to 10-fold depletion in some samples. These results show that adding a Cas9 based library depletion step, applied to amplified 10x cDNA, markedly boosts information density in 10x single-cell workflows, facilitating the detection of rare transcripts, including parasite mRNAs in infected erythrocytes. Although not shown here, similar performance has been observed when applying this approach to malaria-infected human blood samples.

References

1. Choi, I., et al. (2014). Increasing gene discovery and coverage using RNA-seq of globin RNA reduced porcine blood samples. BMC Genomics. 15:954. doi:10.1186/1471-2164-15-954.
2. Jang, J.S., et al. (2020). Comparative evaluation for the globin gene depletion methods for mRNA sequencing using the whole blood-derived total RNAs. BMC Genomics. 21(1):890. doi:10.1186/s12864-020-07304-4.
3. Sheerin, D., et al. (2023). Identification and control for the effects of bioinformatic globin depletion on human RNA-seq differential expression analysis. Sci Rep. 13:1859. doi:10.1038/s41598-023-28218-7.
4. Rosenthal, M.R., Ng, C.L. (2023). High-content imaging as a tool to quantify and characterize malaria parasites. Cell Rep Methods. 3(7):100516. doi:10.1016/j.crmeth.2023.100516.
5. Howick, V.M., et al. (2019). The Malaria Cell Atlas: Single parasite transcriptomes across the complete Plasmodium life cycle. Science. 365(6455):eaaw2619. doi:10.1126/science.aaw2619.
6. Real, E., Mancio-Silva, L. (2022). Single-cell views of the Plasmodium life cycle. Trends Parasitol. 38(9):748–757. doi:10.1016/j.pt.2022.05.009.
7. Hentzschel, F., et al. (2022). Host cell maturation modulates parasite invasion and sexual differentiation in Plasmodium berghei. Sci Adv. 8(17):eabm7348. doi:10.1126/sciadv.abm7348.
8. Chkheidze, A.N., et al. (1999). Assembly of the alpha-globin mRNA stability complex reflects binary interaction between the pyrimidine-rich 3′ untranslated region determinant and poly(C) binding protein alphaCP. Mol Cell Biol. 19(7):4572–81. doi:10.1128/MCB.19.7.4572.
9. Gu, W., et al. (2016). Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. Genome Biol. 17:41. doi:10.1186/s13059-016-0904-5.
10. Brancucci, N.M., et al. (2015). An assay to probe Plasmodium falciparum growth, transmission stage formation and early gametocyte development. Nat Protoc. 10(8):1131–42. doi:10.1038/nprot.2015.072.