Application Note

Genomics

Agilent

Trusted Answers

SureSelect Targeted Panel Paired With iSort Enables Robust Enumeration of Immune Cell-type Abundances in Human Blood Samples

Authors

Aki Nakao[1]
Aaron M. Newman[1,2]
Ash A. Alizadeh[1,2]
Maximilian Diehn[1,2]
Manjula Aliminati[3]
Mistuni Ghosh[3]
Jayati Ghosh[3]
Kristi Stephenson[3]
Ashutosh[3]

1. CiberMed, Inc., California, USA

2. Stanford University, Stanford, California, USA

3. Agilent Technologies Inc., California, USA

Abstract

SureSelect XT HS2 RNA sequencing was combined with CiberMed's iSort digital cytometry solution to enable highly accurate, cost-effective profiling of leukocyte composition from peripheral blood samples. Enrichment of genes in the LM22 signature matrix, a well-established collection of reference profiles for distinguishing 22 human immune subsets, was achieved with two new targeted sequencing panels—Agilent SureSelect CD CiberMed Heme and Agilent SureSelect CD CiberMed Heme + HiRes. Both panels were assessed for their ability to profile leukocyte subsets with CiberMed's iSort Fractions software.

Applied to 36 whole blood samples and seven major leukocyte populations, cell type fractions determined by iSort were highly concordant with ground truth fractions determined by clinical grade standards ($r \geq 0.96$) and exhibited strong reproducibility across technical replicates ($r \geq 0.98$). Furthermore, targeted enrichment using the SureSelect CD CiberMed Heme panel reduced the sequencing requirement by nearly 50-fold compared to whole-transcriptome sequencing, while also improving accuracy. These new panels for digital cytometry are available through the Agilent Community Design program to enable focused, reliable, and high-throughput analysis of cell type composition from peripheral blood samples.

Introduction

Cellular heterogeneity and complex intercellular interactions underlie diverse physiological and pathological states, including various malignancies. Therefore, it is critically important to study the phenotypic and genotypic composition of cell subsets within the diseased milieu. It is also essential to monitor changes in their relative abundances during disease progression and in response to therapy. The importance of studying cellular heterogeneity and immune composition within the tumor microenvironment (TME) is well established.[1-3] Enumerating immune cell composition has prognostic value and holds great promise as a potential predictive biomarker for therapy response.[4,5] Therefore, both traditional and next-generation methodologies are routinely employed to quantify immune cell subsets in research and clinical settings (Table 1).

Traditional cell profiling methods such as flow cytometry, mass cytometry,[6,7] immunohistochemistry (IHC),[8] and immunofluorescence (IF)[9] are powerful tools for quantifying and characterizing immune cell subtypes. However, only a handful of markers can be interrogated by these methods and there is often a trade-off between the number of markers that can be measured and the throughput of the assay. Cytometry by time-of-flight (CyTOF)[10] is destructive to the sample and doesn't enable co-interrogation of cell type fractions and cell type expression profiles across thousands of genes. In more recent years, single-cell RNA sequencing (RNA-Seq) approaches have been embraced as a means of characterizing immune cell composition and gene expression at the single-cell level. However, single-cell RNA-Seq is limited by sample preparation artifacts, including dissociation-induced distortions in cellular composition, and remains cost prohibitive for large-scale cellular profiling.[11]

Given these limitations, deconvolution algorithms for determining cell type abundances from bulk tissue expression profiles have gained traction.[1,2,12-16] These methodologies enable dissection of cell-type-specific signals from bulk sequencing data. Comparative analyses of single-cell RNA-Seq, deconvolution of bulk expression data, and IHC revealed that deconvolution is free from artifacts arising from cell separation and tissue dissociation.[17] Of these deconvolution methods, CIBERSORT and CIBERSORTx, have emerged as robust and accurate tools for determining immune cell type proportions from blood and tissue samples.[3,17-28] In fact, CIBERSORT/x was recently identified as one of the five fastest-growing software tools in the biosciences.[29]

CiberMed further optimized and standardized CIBERSORT/x with proprietary enhancements, validated the improved algorithms with different sample types and platforms. CiberMed currently offers them via two flagship products within the iSort digital cytometry suite (Figure 1). These include iSort Fractions for reliably determining cell subset abundance from bulk tissue or blood expression data and iSort HiRes for inferring cell-type-specific gene expression profiles from bulk tissue or blood expression data. iSort Fractions uses the well-established LM22 leukocyte gene signature matrix to distinguish 22 human hematopoietic cell subsets. This includes seven T cell subsets, naïve and memory B cells, plasma cells, resting and activated NK cells, monocytes, eosinophils, and neutrophils (see Table 2 for the complete list of cell subsets).[3] LM22 has been validated in pure leukocyte subset titrations, blood samples, and tumors from multiple cancer types. Because iSort Fractions determine cell type proportions, cell subsets within LM22 can be further grouped into various parental cell types of interest, including 11 major leukocyte types based on shared lineage.[3]

In this application note, we demonstrate robust, accurate, and reproducible enumeration of immune cell subsets from 36 whole blood samples using iSort Fractions combined with the Agilent SureSelect target enrichment solution (Figure 2). Two panels, SureSelect CD CiberMed Heme and SureSelect CD CiberMed Heme + HiRes, were employed to enrich Agilent SureSelect XT HS2 RNA libraries prepared from these samples (panel design details are provided under Table 3). The SureSelect CD CiberMed Heme panel includes LM22 genes, and in conjunction with CiberMed's iSort Fractions analysis, is intended for determination of immune composition from blood samples. The SureSelect CD CiberMed Heme + HiRes panel includes both LM22 genes and a set of control genes for demonstrating the capabilities of iSort HiRes. The latter is intended to be paired with a custom gene panel covering additional genes of interest.

Across 36 whole blood samples profiled by one or both panels, leukocyte proportions predicted by iSort Fractions showed high concordance with ground truth composition across seven major populations ($r \geq 0.96$; Figure 3). There was also strong reproducibility between sample technical replicates (Figure 4). Because of several proprietary enhancements, we found that iSort Fractions outperforms CIBERSORT/x for the enumeration of immune composition from peripheral blood. Notably, targeted enrichment with the SureSelect CD CiberMed Heme panel reduced the sequencing requirement to only 500 k (250 k × 2) input reads per sample, as opposed to approximately 20 M (10 M × 2) input reads per sample for whole-transcriptome sequencing (Figures 5 to 7). The decrease equates to a nearly 50-fold reduction in sequencing cost-per-sample.

Altogether, these results underscore the accuracy, reliability, and cost-effectiveness of the combined SureSelect/iSort assay for immune cell profiling.

Table 1. Comparison of cell profiling methods.

	Software-Based Deconvolution	Single-Cell RNA-Seq	Flow Cytometry/ CyTOF	IHC
Throughput	+++++	+	++	+++
Requires tissue dissociation	No	Yes	Yes	No
Artifacts introduced by dissociation	No	Yes	Yes	No
Workflow simplicity	+++++	+	++	++++
Manual data analysis required	No	Yes	Yes	Yes
Number of cell types/cell type resolution	+++++	+++++	+++	+++

Table 2. Twenty-two human hematopoietic cell types in the LM22 signature matrix.

Parent Subsets	Cell Type Description
B cells	B cells naïve
	B cells memory
Plasma cells	Plasma cells
CD8 T cells	T cells CD8
CD4 T cells	T cells CD4 naïve
	T cells CD4 memory resting
	T cells CD4 memory activated
	T cells follicular helper
	T cells regulatory (Tregs)
Gamma delta T cells	T cells gamma delta
NK cells	NK cells resting
	NK cells activated
Monocytes and Macrophages	Monocytes
	Macrophages M0
	Macrophages M1
	Macrophages M2
Dendritic cells	Dendritic cells resting
	Dendritic cells activated
Mast cells	Mast cells resting
	Mast cells activated
Eosinophils	Eosinophils
PMNs	Neutrophils

Table 3. Comparison of panel designs and intended applications.

	SureSelect CD CiberMed Heme	SureSelect CD CiberMed Heme + HiRes
Intended Application	Enumeration of 22 immune cell subsets (Table 1) with iSort Fractions	Enumeration of 22 immune cell subsets (Table 1) with iSort Fractions and confirmation of cell-type-specific gene expression purification with iSort HiRes
Sample Type	Whole blood, PBMCs	Whole blood, PBMCs
Includes LM22 Signature Matrix	Yes	Yes
Number of Targets	547 genes	685 genes
Total Capture Size	1.8 Mb	2.4 Mb
Recommended Min Reads/Sample	500k (250k × 2, 150 bp)	1 M (500k × 2, 150 bp)

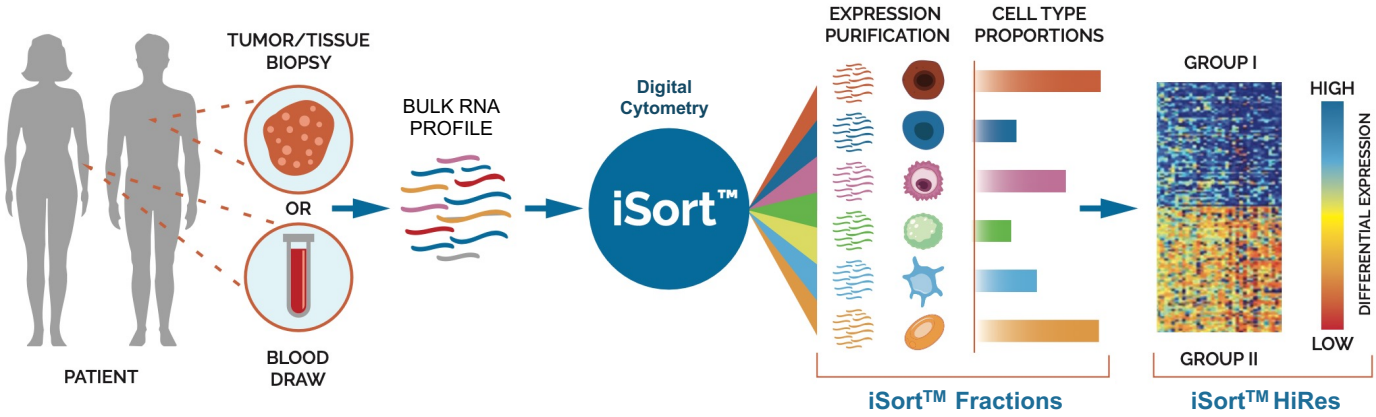

Figure 1. Schematic of iSort algorithms for digital cytometry. Expression purification estimates of cell type proportions (iSort Fractions) and cell-type-specific expression (iSort HiRes).

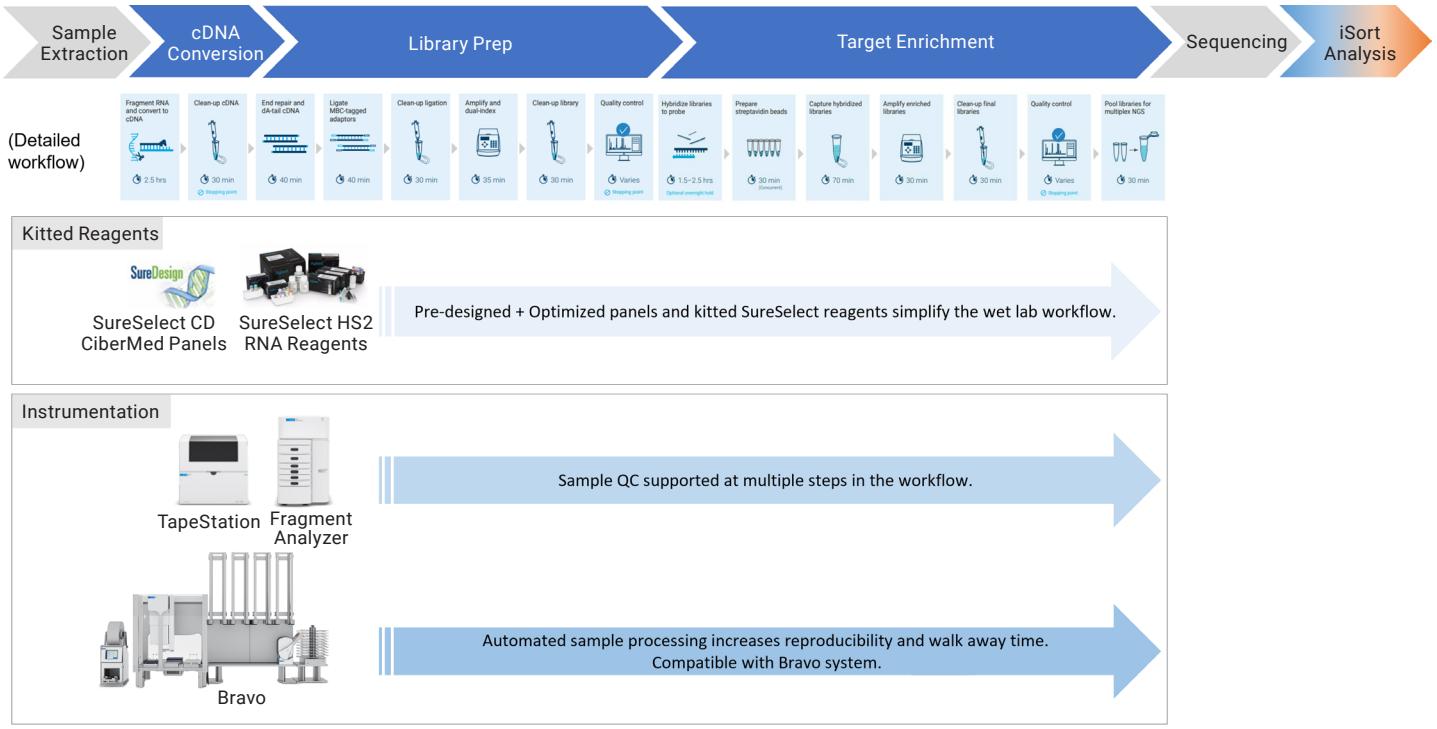

Figure 2. End-to-end workflow supported by kitted reagents and instrumentation from Agilent and an integrated analysis solution from CiberMed.

Materials and methods

Samples

Whole blood samples were freshly collected from 36 healthy donors and split into two fractions. One fraction was immediately processed for complete blood count (CBC) and enumeration of the major leukocyte populations by flow cytometry. The other fraction was stored in PAXgene Blood RNA tubes for subsequent RNA sequencing. A Sysmex system was employed for CBC quantification of neutrophils, lymphocytes, and monocytes. A Becton Dickinson six-color TBNK MultiTest in vitro diagnostic (IVD) assay was employed for enumerating B cells, CD8 T cells, CD4 T cells, and NK cells by flow cytometry. CBC and TBNK data were jointly used as ground truth to assess deconvolution performance from targeted and whole-transcriptome bulk RNA sequencing data.

Library preparation and target enrichment for targeted RNA sequencing

Total RNA was extracted using a PAXgene Blood RNA kit from whole blood samples collected in PAXgene Blood RNA tubes from healthy adult donors. The isolated total RNA was of high quality with RIN values of 8.0 or above.

Library preparation and targeted enrichment were performed using an Agilent Bravo NGS workstation option B following the SureSelect XT HS2 RNA system user guide G9993-90010.31 SureSelect XT HS2 RNA Reagent Kits G9991A and G9991B were employed to generate three technical replicate libraries per sample, each starting with an input of 30 ng total RNA. Agilent SureSelect XT HS2 RNA Target Enrichment kit part number G9994A was employed to enrich blood sample libraries with either SureSelect CD CiberMed Heme (for 12 blood samples) or SureSelect CD CiberMed Heme + HiRes panel (for 36 blood samples).

Agilent SureDesign software was used to create the two community design panels, SureSelect CD CiberMed Heme and SureSelect CD CiberMed Heme + HiRes. The SureSelect CD CiberMed Heme panel covers all 547 genes in the LM22 signature matrix. The SureSelect CD CiberMed Heme + HiRes panel includes an additional 138 control genes for iSort HiRes analysis (Table 3). Manual curation was performed to ensure full coverage of the coding region for every included gene.

Enriched libraries were sequenced as 2 × 150 bp paired end reads on either an Illumina HiSeq4000 or an Illumina NovaSeq 6000 instrument. Each sample was sequenced to approximately 24 M (12 M × 2) total reads to allow performance assessment at varying depths of coverage.

Data processing

As input to iSort, RNA sequencing reads (in FASTQ format) were first summarized to gene expression values in transcripts per million (TPM). While this can be accomplished using any standard mapping/alignment approach, for the data presented here, we used Salmon v1.1[32] to map and quantify RNA-Seq reads. The TPM values were then used as input to iSort Fractions v1.4, which determined the relative fractions of 22 immune subsets (Table 1) in each blood sample. As a comparator, we also evaluated STAR[33] for mapping and RSEM[34] for read summarization. We confirmed that the distinct read alignment and quantification methods didn't significantly impact the performance of iSort Fractions (data not shown).

CiberMed software tools, iSort Fractions, and iSort HiRes, are currently available as docker-containerized tools that can be run locally. iSort Fractions is also available via a user-friendly Web interface run securely via AWS (Amazon Web Services) at https://isort.cibermed.com.

For more details about the iSort digital cytometry suite, please follow this link: https://isort.cibermed.com/iSort_ProductsAndServices.pdf.

Performance evaluation

Concordance was determined by applying Spearman rho, Pearson r, and RMSE to cell type proportions. Mutually exclusive cell types were normalized to 100% per sample to promote a fair comparison.

Read titration

To determine the impact of the number of reads per sample on deconvolution accuracy and reproducibility, paired reads were randomly sampled to 600 k, 200 k, 100 k, 20 k, and 2 k effective reads per sample. Effective reads are defined as the number of reads mapped "on-target" to the genes included in the SureSelect CD CiberMed Heme panel. Input reads are calculated as [minimum effective reads] / [on-target map rate].

Results

Quality control (QC) metrics for targeted RNA-Seq data were generated in this study using the SureSelect CD CiberMed Heme panel and an internal pipeline. All QC data passed internal QC requirements. Comparable results were obtained with data generated by the SureSelect CD CiberMed Heme + HiRes panel (data available upon request).

Accuracy

Figure 3 summarizes the accuracy of iSort Fractions for deconvolving major leukocyte populations in 36 blood samples as compared to orthogonal cell profiling by clinical grade standards. Clear performance gains can be seen for iSort over CIBERSORTx (Figures 3b versus 3a). Gains in performance can also be observed for targeted sequencing with the SureSelect CD CiberMed Heme panels (including the Heme + HiRes panel) over whole-transcriptome sequencing (Figures 3c versus 3b). Both SureSelect panels showed comparable performance (Figures 3c and 3d). While CIBERSORT/x was developed by Stanford University[3,17] and is the most widely used deconvolution tool in academic settings,[30] iSort Fractions is an optimized and standardized version of CIBERSORT/x that can achieve superior performance (Figure 3).

Reproducibility

To assess reproducibility, blood samples were sequenced in triplicate and the concordance of iSort Fractions across technical replicates was evaluated (Figure 4). All pairwise comparisons across the three replicates achieved Pearson correlations of at least 0.98 and Spearman correlations of at least 0.94 (Figure 4), demonstrating strong reproducibility.

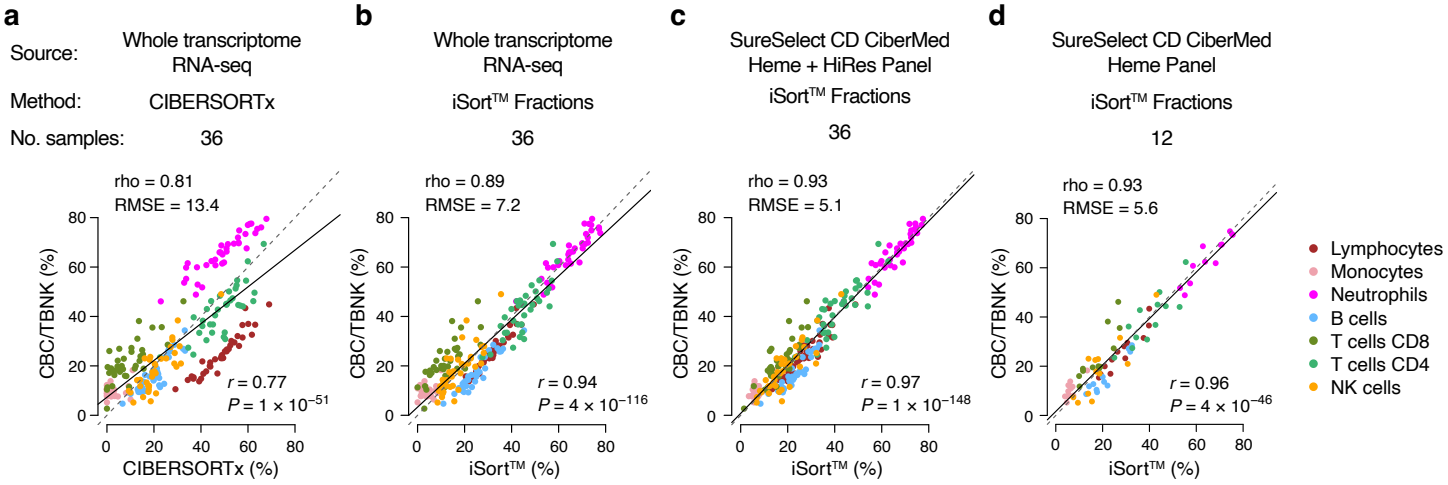

Figure 3. Relative leukocyte abundances from the whole blood of 36 healthy adult donors. Data comparison generated by (a) whole-transcriptome sequencing analyzed by CIBERSORTx, (b) whole-transcriptome sequencing analyzed by iSort Fractions, (c) SureSelect CD CiberMed Heme + HiRes panel analyzed by iSort Fractions, and (d) SureSelect CD CiberMed Heme panel analyzed by iSort Fractions. Each data point represents a sample and is color-coded by cell type. Predicted and ground truth proportions are shown on the x-axis and y-axis, respectively.

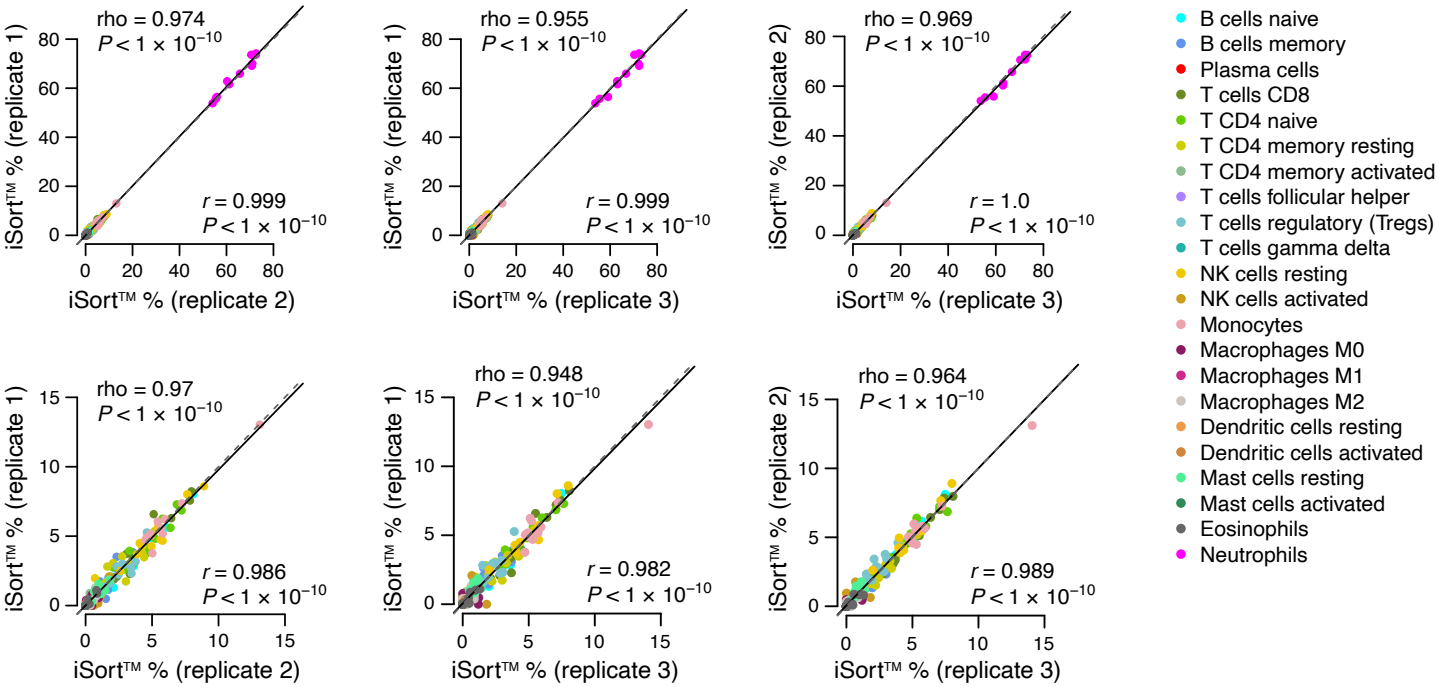

Figure 4. Comparison of iSort Fractions results among three sequencing replicates. Each data point represents a sample and is color-coded by cell type. The first three plots (upper row) show the abundance of all 22 cell types. Neutrophils are excluded in the bottom three scatterplots to expand the 0 to 15% range. Results were calculated using the SureSelect CD CiberMed Heme panel applied to whole blood from 12 healthy adult donors.

Sequencing requirement

Each blood sample was sequenced to an average of 24 M (12 M × 2) reads for SureSelect CD CiberMed Heme panel, and 84 M (42 M × 2) reads for whole-transcriptome libraries. To determine the minimum number of reads needed per sample without compromising performance, we performed a titration experiment. In this experiment, the number of effective reads per sample was down-sampled to predefined quantities before running iSort Fractions (where "effective reads" denotes on-target reads only) (Figure 5). The SureSelect CD CiberMed Heme panel maintained accurate and stable performance down to 100 k effective reads (Figures 5a, 6, and 7). The panel also achieved a lower median RMSE than whole-transcriptome sequencing for all read quantities evaluated (Figure 5b).

Based on these data, when using the SureSelect CD CiberMed Heme Panel, the total number of required input reads is expected to range from approximately 140 k per sample (assuming a 75% on-target mapping rate) to approximately 500 k per sample (if the mapping rate is uncharacteristically low). Therefore, a conservative projection of 500 k (250 k × 2) total input reads per sample should be sufficient for nearly every application. Since at least 20 M (10 M × 2) input reads are recommended when using data from whole-transcriptome sequencing for iSort Fractions, the SureSelect CD CiberMed Heme panel is expected to reduce the sequencing cost by nearly 50-fold while achieving superior performance.

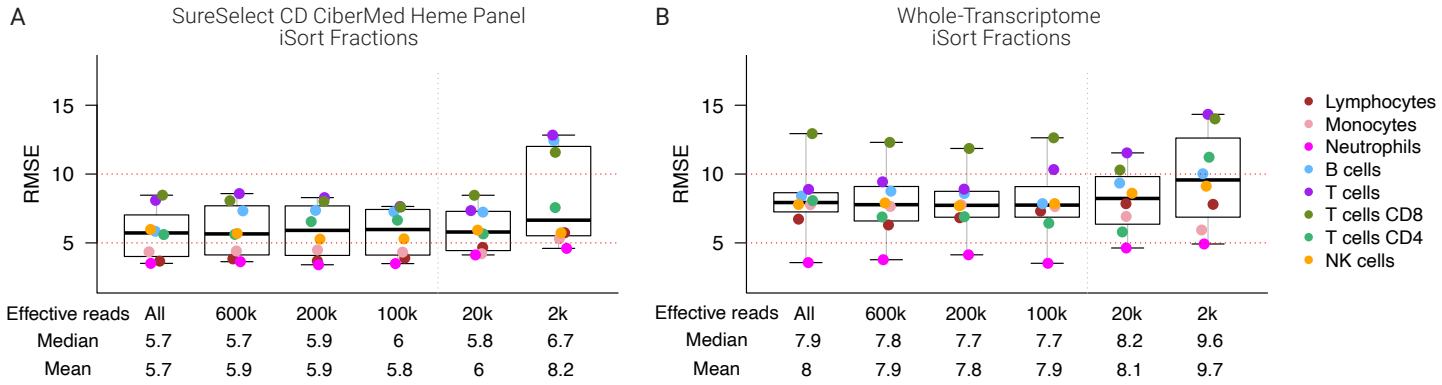

Figure 5. Root mean squared error (RMSE) between leukocyte proportions determined by iSort Fractions and clinical grade cell profiling assays (CBC and TBNK) on whole blood from 12 healthy adult donors. "All" denotes all evaluable reads per sample. The SureSelect CD CiberMed Heme panel targeted sequencing panel (5a) maintained stable performance down to 100 k (50 k × 2) effective ("on-target") reads per sample and outperformed whole-transcriptome sequencing (5b) for all evaluated read quantities.

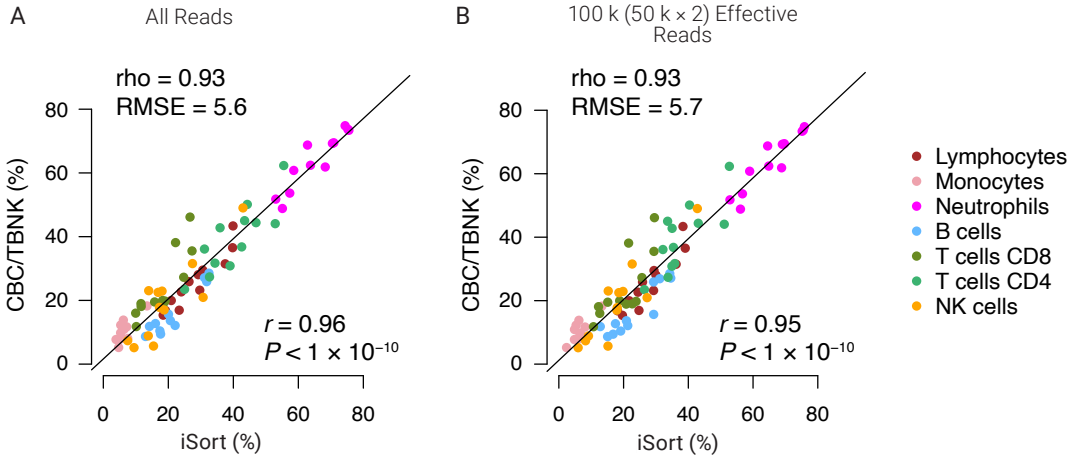

Figure 6. Accuracy of iSort Fractions, comparing "all reads" per sample (no down-sampling) (6a) with 100 k (50 k × 2) effective (that is, on-target) reads per sample (6b). Each data point represents a sample and is color-coded by cell type. Results were calculated using the SureSelect CD CiberMed Heme panel applied to whole blood from 12 healthy adult donors.

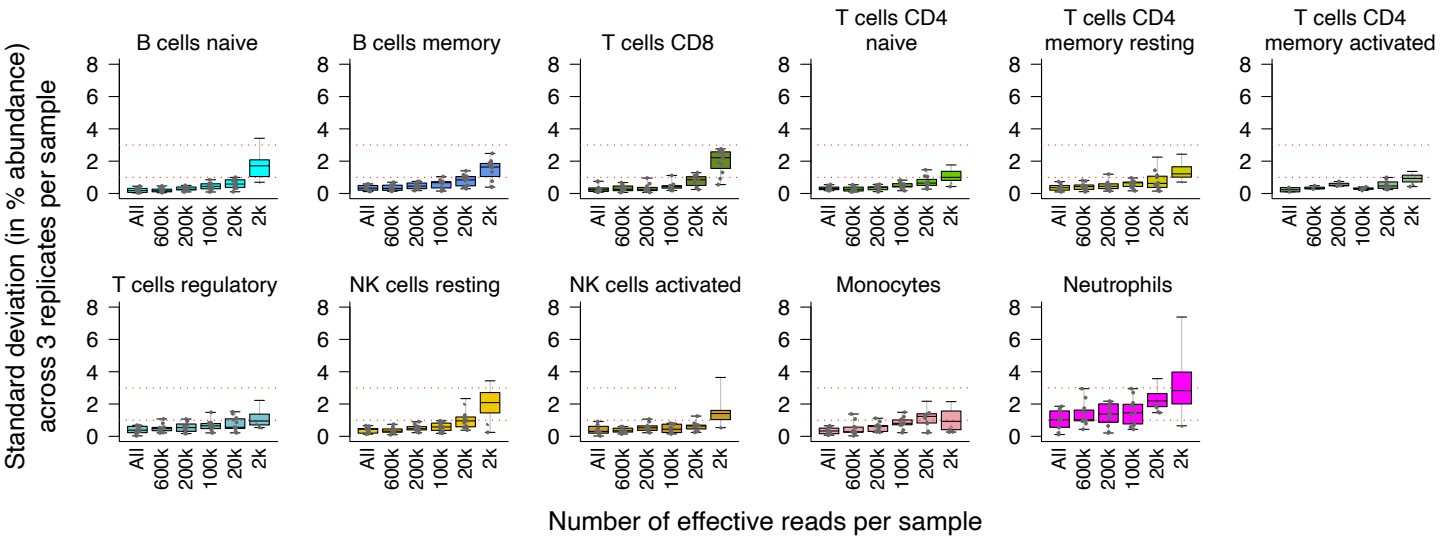

Figure 7. Reproducibility across three sequencing replicates, analyzed by the number of effective reads per sample (that is, the number of reads mapped to SureSelect CD Heme panel genes). Except for neutrophils, all cell types show consistent performance down to 100 k (50 k × 2) effective reads with a mean standard deviation less than one. Neutrophils show a comparable trend but have a slightly larger mean standard deviation owing to a higher mean abundance than other evaluated cell types (approximately 65% versus a mean abundance of <7%). Cell types with a mean abundance of less than 0.1% are not shown. The x-axis shows the number of effective reads per sample. Results were calculated using the SureSelect CD CiberMed Heme panel applied to whole blood from 12 healthy adult donors.

Conclusion

In this application note, we describe the performance characteristics of two SureSelect CD Heme panels for leukocyte profiling with iSort, a state-of-the-art software solution for digital cytometry (CiberMed). There are several applications that would benefit from the enhanced robustness, accuracy, and cost-effectiveness of this new joint assay for "cytometry by sequencing". These include routine and translational blood composition analysis,[35] retrospective characterization of bulk expression data to derive new insights into cellular composition[1], and large-scale validation of sequencing analyses.[36] Additional applications include the general assessment of leukocyte composition under diverse physiological and pathological conditions, all without the need for antibodies, fresh specimens, viable material, or millions of cells.

The utility of the core iSort methodology has been demonstrated in multiple contexts, including immuno-oncology,[37,38] organ transplantation,[39] cardiology,[40] Crohn's disease,[30] and neonatal sepsis.[41] The ability to deconvolve alternative genomic data types, including methylation and proteomic profiles, has also been demonstrated.[19,42] Furthermore, iSort has been analytically validated for whole blood samples and fresh, frozen, and fixed tumor specimens.[27,30] The core algorithm underlying iSort has become the standard methodology for deconvolution in large cancer studies and datasets, such as The Cancer Genome Atlas (TCGA) and is being applied in clinical trials.[43-46] Given the performance gains demonstrated here, the SureSelect/iSort assay for digital cytometry promises to facilitate many exciting and impactful future applications.

References

1. Thorsson, V.; Gibbs, D. L.; Brown, S. D.; Wolf, D.; Bortone, D. S.; Yang, T. H. O.; Porta-Pardo, E.; Gao, G. F.; Plaisier, C. L.; Eddy, J. A.; et al. The Immune Landscape of Cancer. *Immunity*. **2018**, 48 (4), P812-830. https://doi.org/10.1016/j.immuni.2018.03.023.
2. Farc, O.; Cristea, V. An Overview of the Tumor Microenvironment, from Cells to Complex Networks (Review). *Exp. Ther. Med.* **2020**, 21 (1). https://doi.org/10.3892/etm.2020.9528.
3. Newman, A. M.; Liu, C. L.; Green, M. R.; Gentles, A. J.; Feng, W.; Xu, Y.; Hoang, C. D.; Diehn, M.; Alizadeh, A. A. Robust Enumeration of Cell Subsets from Tissue Expression Profiles. *Nat. Methods* **2015**, 12 (5), 453–457. https://doi.org/10.1038/nmeth.3337.
4. Gentles, A. J.; Newman, A. M.; Liu, C. L.; Bratman, S. V.; Feng, W.; Kim, D.; Nair, V. S.; Xu, Y.; Khuong, A.; Hoang, C. D.; et al. The Prognostic Landscape of Genes and Infiltrating Immune Cells across Human Cancers. *Nat. Med*. **2015**, 21 (8), 938–945. https://doi.org/10.1038/nm.3909.
5. Bentham, R.; Litchfield, K.; Watkins, T. B. K.; Lim, E. L.; Rosenthal, R.; Martínez-Ruiz, C.; Hiley, C. T.; Bakir, M. A.; Salgado, R.; Moore, D. A.; et al. TRACERx Consortium. Using DNA Sequencing Data to Quantify T Cell Fraction and Therapy Response. *Nature* **2021**, 597 (7877), 555–560. https://doi.org/10.1038/s41586-021-03894-5.
6. Herold, N. C.; Mitra, P. I*mmunophenotyping; StatPearls Publishing*, **2023**.
7. Alfonso, B.-F.; Al-Rubeai, M. Flow Cytometry. In Comprehensive Biotechnology; Elsevier, **2011**; pp 559–578.
8. Dixon, A. R.; Bathany, C.; Tsuei, M.; White, J.; Barald, K. F.; Takayama, S. Recent Developments in Multiplexing Techniques for Immunohistochemistry. *Expert Rev. Mol. Diagn*. **2015**, 15 (9), 1171–1186. https://doi.org/10.1586/14737159.2015.1069182.
9. Rashid, R.; Gaglia, G.; Chen, Y.-A.; Lin, J.-R.; Du, Z.; Maliga, Z.; Schapiro, D.; Yapp, C.; Muhlich, J.; Sokolov, A.; et al. Highly Multiplexed Immunofluorescence Images and Single-Cell Data of Immune Markers in Tonsil and Lung Cancer. *Sci. Data* **2019**, 6 (1), 1–10. https://doi.org/10.1038/s41597-019-0332-y.
10. Palit, S.; Heuser, C.; de Almeida, G. P.; Theis, F. J.; Zielinski, C. E. Meeting the Challenges of High-Dimensional Single-Cell Data Analysis in Immunology. *Front. Immunol.* **2019**, 10. https://doi.org/10.3389/fimmu.2019.01515.
11. Lähnemann, D.; Köster, J.; Szczurek, E.; McCarthy, D. J.; Hicks, S. C.; Robinson, M. D.; Vallejos, C. A.; Campbell, K. R.; Beerenwinkel, N.; Mahfouz, A.; et al. Eleven Grand Challenges in Single-Cell Data Science. *Genome Biol.* **2020**, 21(1). https://doi.org/10.1186/s13059-020-1926-6.
12. Aran, D.; Hu, Z.; Butte, A. J. xCell: Digitally Portraying the Tissue Cellular Heterogeneity Landscape. *Genome Biol.* **2017**, 18(1). https://doi.org/10.1186/s13059-017-1349-1.
13. Racle, J.; Gfeller, D. EPIC: A Tool to Estimate the Proportions of Different Cell Types from Bulk Gene Expression Data. In Bioinformatics for Cancer Immunotherapy; *Springer US*: New York, NY, **2020**; Vol. 2120, pp 233–248.
14. Li, B.; Li, T.; Liu, J. S.; Liu, X. S. Computational Deconvolution of Tumor-Infiltrating Immune Components with Bulk Tumor Gene Expression Data. In Bioinformatics for Cancer Immunotherapy; *Springer US*: New York, NY, **2020**; Vol. 2120, pp 249–262.
15. Becht, E.; Giraldo, N. A.; Lacroix, L.; Buttard, B.; Elarouci, N.; Petitprez, F.; Selves, J.; Laurent-Puig, P.; Sautès-Fridman, C.; Fridman, W. H.; de Reyniès, A. Estimating the Population Abundance of Tissue-Infiltrating Immune and Stromal Cell Populations Using Gene Expression. *Genome Biol*. **2016**, 17 (1). https://doi.org/10.1186/s13059-016-1070-5.
16. Shen-Orr, S. S.; Tibshirani, R.; Butte, A. J. Gene Expression Deconvolution in Linear Space. *Nat. Methods* **2012**, 9 (1), 9–9. https://doi.org/10.1038/nmeth.1831.
17. Newman, A. M.; Steen, C. B.; Liu, C. L.; Gentles, A. J.; Chaudhuri, A. A.; Scherer, F.; Khodadoust, M. S.; Esfahani, M. S.; Luca, B. A.; Steiner, D.; et al. Determining Cell Type Abundance and Expression from Bulk Tissues with Digital Cytometry. *Nat. Biotechnol.* **2019**, 37 (7), 773–782. https://doi.org/10.1038/s41587-019-0114-2.
18. Bionetworks, *S. Synapse*. Synapse.org. https://www.synapse.org/ (accessed 2023-10-30).
19. Nadel, B. B.; Oliva, M.; Shou, B. L.; Mitchell, K.; Ma, F.; Montoya, D. J.; Mouton, A.; Kim-Hellmuth, S.; Stranger, B. E.; Pellegrini, M.; Mangul, S. Systematic Evaluation of Transcriptomics-Based Deconvolution Methods and References Using Thousands of Clinical Samples. *Brief. Bioinform.* **2021**, 22 (6). https://doi.org/10.1093/bib/bbab265.
20. Avila Cobos, F.; Alquicira-Hernandez, J.; Powell, J. E.; Mestdagh, P.; De Preter, K. Benchmarking of Cell Type Deconvolution Pipelines for Transcriptomics Data. *Nat. Commun*. **2020**, 11 (1). https://doi.org/10.1038/s41467-020-19015-1.

21. Le, T.; Aronow, R. A.; Kirshtein, A.; Shahriyari, L. A Review of Digital Cytometry Methods: Estimating the Relative Abundance of Cell Types in a Bulk of Cells. *Brief. Bioinform.* **2021**, 22 (4). https://doi.org/10.1093/bib/bbaa219.
22. Chen, B.; Khodadoust, M. S.; Liu, C. L.; Newman, A. M.; Alizadeh, A. A. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. In Methods in Molecular Biology; *Springer New York:* New York, NY, **2018**; Vol. 1711, pp 243–259.
23. Sutton, G. J.; Poppe, D.; Simmons, R. K.; Walsh, K.; Nawaz, U.; Lister, R.; Gagnon-Bartsch, J. A.; Voineagu, I. Comprehensive Evaluation of Deconvolution Methods for Human Brain Gene Expression. *Nat. Commun.* **2022**, 13 (1), 1–18. https://doi.org/10.1038/s41467-022-28655-4.
24. Breen, M. S.; Ozcan, S.; Ramsey, J. M.; Wang, Z.; Ma'ayan, A.; Rustogi, N.; Gottschalk, M. G.; Webster, M. J.; Weickert, C. S.; Buxbaum, J. D.; Bahn, S. Temporal Proteomic Profiling of Postnatal Human Cortical Development. *Transl. Psychiatry* **2018**, 8 (1), 1–14. https://doi.org/10.1038/s41398-018-0306-4.
25. Li, L.; Shen, L.; Ma, J.; Zhou, Q.; Li, M.; Wu, H.; Wei, M.; Zhang, D.; Wang, T.; Qin, S.; Xing, T. Evaluating Distribution and Prognostic Value of New Tumor-Infiltrating Lymphocytes in HCC Based on a scRNA-Seq Study with CIBERSORTx. *Front. Med.* (Lausanne) **2020**, 7. https://doi.org/10.3389/fmed.2020.00451.
26. Newman, A.M.; Gulati, S.G.; Clarke, M.F.; Sikandar, S.S. Methods Utilizing Single Cell Genetic Data for Cell Population Analysis and Applications Thereof. US202003700112 A1, November 26, **2020**.
27. Newman, A. M.; Nakao, A.; Li, K.; Liu, C.-L.; Mathi, K.; Sigal, N.; Maecker, H.; Diehn, M.; Alizadeh, A. A. Analytical Validation of Digital Cytometry (iSort) for Leukocyte Enumeration Using Stored Blood. J. *Clin. Oncol.* **2020**, 38 (15_suppl), 3542–3542. https://doi.org/10.1200/jco.2020.38.15_suppl.3542.
28. Chakravarthy, A.; Furness, A.; Joshi, K.; Ghorani, E.; Ford, K.; Ward, M. J.; King, E. V.; Lechner, M.; Marafioti, T.; Quezada, S. A.; Thomas, G. J.; Feber, A.; Fenton, T. R. Pan-Cancer Deconvolution of Tumour Composition Using DNA Methylation. *Nat. Commun.* **2018**, 9 (1). https://doi.org/10.1038/s41467-018-05570-1.
29. Hutson, M. Hunting for the Best Bioscience Software Tool? Check This Database. *Nature* **2023**. https://doi.org/10.1038/d41586-023-00053-w.
30. Chen, H.; Chen, C.; Yuan, X.; Xu, W.; Yang, M.-Q.; Li, Q.; Shen, Z.; Yin, L. Identification of Immune Cell Landscape and Construction of a Novel Diagnostic Nomogram for Crohn's Disease. *Front. Genet.* **2020**, 11. https://doi.org/10.3389/fgene.2020.00423.
31. A, V. Agilent.com. https://www.agilent.com/cs/library/usermanuals/public/G9993-90010.pdf (accessed 2023-10-30).
32. Patro, R.; Duggal, G.; Love, M. I.; Irizarry, R. A.; Kingsford, C. Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression. *Nat. Methods* **2017**, 14 (4), 417–419. https://doi.org/10.1038/nmeth.4197.
33. Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, 29 (1), 15–21. https://doi.org/10.1093/bioinformatics/bts635.
34. Li, B.; Dewey, C. N. RSEM: Accurate Transcript Quantification from RNA-Seq Data with or without a Reference Genome. *BMC Bioinformatics* **2011**, 12 (1). https://doi.org/10.1186/1471-2105-12-323.
35. Orange, D. E.; Yao, V.; Sawicka, K.; Fak, J.; Frank, M. O.; Parveen, S.; Blachere, N. E.; Hale, C.; Zhang, F.; Raychaudhuri, S.; et al. RNA Identification of PRIME Cells Predicting Rheumatoid Arthritis Flares. *N. Engl. J. Med.* **2020**, 383 (3), 218–228. https://doi.org/10.1056/nejmoa2004114.
36. Gohil, S. H.; Iorgulescu, J. B.; Braun, D. A.; Keskin, D. B.; Livak, K. J. Applying High-Dimensional Single-Cell Technologies to the Analysis of Cancer Immunotherapy. *Nat. Rev. Clin. Oncol.* **2021**, 18 (4), 244–256. https://doi.org/10.1038/s41571-020-00449-x.
37. Nabet, B. Y.; Esfahani, M. S.; Moding, E. J.; Hamilton, E. G.; Chabon, J. J.; Rizvi, H.; Steen, C. B.; Chaudhuri, A. A.; Liu, C. L.; Hui, A. B.; et al. Noninvasive Early Identification of Therapeutic Benefit from Immune Checkpoint *Inhibition. Cell* **2020**, 183 (2), 363-376.e13. https://doi.org/10.1016/j.cell.2020.09.001.
38. Lozano, A. X.; Chaudhuri, A. A.; Nene, A.; Bacchiocchi, A.; Earland, N.; Vesely, M. D.; Usmani, A.; Turner, B. E.; Steen, C. B.; Luca, B. A.; et al. T Cell Characteristics Associated with Toxicity to Immune Checkpoint Blockade in Patients with Melanoma. *Nat. Med.* **2022**, 28 (2), 353–362. https://doi.org/10.1038/s41591-021-01623-z.

39. Huang, S.; Chen, H.; Guo, Z.; He, X. A Comprehensive Cibersort Study and Gene-Expression Based Transcriptomic Analysis on Patterns of Immune Infiltration in Ischemia-Reperfusion Injury Livers Post Liver Transplantation. *Transplantation* **2020**, 104 (S3), S195–S195. https://doi.org/10.1097/01.tp.0000699368.56252.17.

40. Kawada JI, Takeuchi S, Imai H, Okumura T, Horiba K, Suzuki T, Torii Y, Yasuda K, Imanaka-Yoshida K, Ito Y. Immune cell infiltration landscapes in pediatric acute myocarditis analyzed by CIBERSORT. J Cardiol. **2021**, 77(2), 174-178. https://doi.org/10.1016/j.jjcc.2020.08.004.

41. Jiang, Z.; Luo, Y.; Wei, L.; Gu, R.; Zhang, X.; Zhou, Y.; Zhang, S. Bioinformatic Analysis and Machine Learning Methods in Neonatal Sepsis: Identification of Biomarkers and Immune Infiltration. *Biomedicines* **2023**, 11 (7), 1853. https://doi.org/10.3390/biomedicines11071853.

42. Voss, M. H.; Buros Novik, J.; Hellmann, M. D.; Ball, M.; Hakimi, A. A.; Miao, D.; Margolis, C.; Horak, C.; Wind-Rotolo, M.; De Velasco, G.; et al. Correlation of Degree of Tumor Immune Infiltration and Insertion-and-Deletion (Indel) Burden with Outcome on Programmed Death 1 (PD1) Therapy in Advanced Renal Cell Cancer (RCC). *J. Clin. Oncol.* **2018**, 36 (15_suppl), 4518–4518. https://doi.org/10.1200/jco.2018.36.15_suppl.4518.

43. Metzger Filho, O.; Stover, D. G.; Asad, S.; Ansell, P. J.; Watson, M.; Loibl, S.; Geyer, C. E.; O'Shaughnessy, J.; Untch, M.; Rugo, H. S.; et al. Immunophenotype and Proliferation to Predict for Response to Neoadjuvant Chemotherapy in TNBC: Results from BrighTNess Phase III Study. *J. Clin. Oncol.* **2019**, 37 (15_suppl), 510–510. https://doi.org/10.1200/jco.2019.37.15_suppl.510.

44. CTG labs - NCBI. Clinicaltrials.gov. https://clinicaltrials.gov/study/NCT03979508 (accessed 2023-10-30).

45. Clinicaltrials.gov. https://classic.clinicaltrials.gov/ProvidedDocs/10/NCT03955510/Prot_000.pdf (accessed 2023-10-30).

46. Medjebar, S.; Richard, C.; Fumet, J.-D.; Malo, J.; Elkrief, A.; Blais, N.; Tehfe, M.; Florescu, M.; Boidot, R.; Truntzer, C.; et al. Angiotensin-Converting Enzyme Inhibitor Prescription Is Associated with Decreased Progression-Free Survival (PFS) and Overall Survival (OS) in Patients with Lung Cancers Treated with PD-1/PD-L1 Immune Checkpoint Blockers. *J. Clin. Oncol.* **2019**, 37 (15_suppl), e20512–e20512. https://doi.org/10.1200/jco.2019.37.15_suppl.

www.agilent.com

PR7001-1756

Published in the USA, March 15, 2024
5994-6964EN

Agilent
Trusted Answers